Direito e Inteligência Artificial

Coordenação: Maria Raquel Guimarães, Rute Teixeira Pedro

Organização:
Fernanda de Araujo Meirelles Magalhães
Luísa Eckenroth Moreira
Tiago Morais Rocha

ALMEDINA

Direito e Inteligência Artificial

2023

Maria Raquel Guimarães
Rute Teixeira Pedro
Coordenação

Fernanda de Araujo Meirelles Magalhães
Luísa Eckenroth Moreira
Tiago Morais Rocha
Organização

ALMEDINA

DIREITO E INTELIGÊNCIA ARTIFICIAL
COORDENAÇÃO
Maria Raquel Guimarães • Rute Teixeira Pedro
EDITOR
EDIÇÕES ALMEDINA, S.A.
Avenida Emídio Navarro, 81, 3D
3000-151 Coimbra
Tel.: 239 851 904 · Fax: 239 851 901
www.almedina.net · editora@almedina.net
DESIGN DE CAPA
EDIÇÕES ALMEDINA, S.A.
PRÉ-IMPRESSÃO
João Jegundo
IMPRESSÃO E ACABAMENTO

Outubro, 2023

ISBN
978-989-40-1434-8

ALMEDINA | GRUPOALMEDINA

[Publicação do Projecto "It's a wonderful (digital) world": O direito numa sociedade digital e tecnológica (CIJ)]
Este trabalho foi desenvolvido com o apoio da Fundação para a Ciência e a Tecnologia – UIDB/00443/2020 (Centro de Investigação Jurídica)

SUMÁRIO

NOTAS INTRODUTÓRIAS

A obra que agora se publica surge na sequência da realização, nos dias 12 e 13 de Maio de 2022, na Faculdade de Direito da Universidade do Porto, do Congresso Internacional "Direito e Inteligência Artificial". O Encontro científico, que congregou investigadores de várias nacionalidades, foi organizado pelo CIJ – Centro de Investigação Jurídica (então CIJE – Centro de Investigação Jurídico-Económica) da Faculdade de Direito da Universidade do Porto, no âmbito do Projeto de Investigação "It's a wonderful (digital) world: O direito numa sociedade digital e tecnológica" [apoio da Fundação para a Ciência e a Tecnologia – UIDB/00443/2020 (Centro de Investigação Jurídica)], com a colaboração de colegas da Universidade de Cantábria, no âmbito do Projeto de Investigação "La inteligencia artificial jurídica" [RTI2018-096601-B-100 (MCIU/AEI/FEDER, UE)], del Programa Estatal de I+D+i Orientada a los Retos de la Sociedad, financiado pelo Ministerio de Ciencia e Innovación e a Agencia Estatal de Investigación espanhóis.

No evento realizado propúnhamo-nos refletir, transversalmente, sobre os múltiplos desafios que a Inteligência Artificial coloca ao Direito, tendo a reflexão ocorrido durante dois dias desdobrando-se em vários painéis dedicados a áreas temáticas variadas. A discussão foi enriquecida não só pelas intervenções dos oradores convidados, mas também pela participação de oradores escolhidos no âmbito de uma "call for abstracts" muito concorrida.

A presente obra faz jus à natureza que se imprimiu ao encontro científico e à qualidade da reflexão levada a cabo, contendo contributos que versam múltiplas temáticas organizadas em seis partes distintas sob os seguintes títulos: I – O admirável mundo novo da IA: desenvolvimentos nacionais e europeus; II – IA: uma revolução processual em curso?; III – Personalidade e Pessoa numa sociedade digital; IV – Questões de res-

ponsabilidade civil; Painel V – IA, contratos e consumo e VI – Mercado e empresa num mundo computacional.

Aproveita-se esta oportunidade para agradecer a todos os oradores intervenientes no Congresso "Direito e Inteligência Artificial" e muito em especial a todos os que deram o seu contributo escrito para a construção desta publicação.

Um agradecimento é também dirigido aos elementos que compuseram a comissão que se dedicou à organização do evento científico e desta publicação, sendo devido um agradecimento especial à Drª Fernanda de Araujo Meirelles Magalhães.

Porto e FDUP, Abril de 2023

Maria Raquel Guimarães
Rute Teixeira Pedro

I
O ADMIRÁVEL MUNDO NOVO DA IA: DESENVOLVIMENTOS NACIONAIS E EUROPEUS

Algumas notas sobre direitos fundamentais, transformação digital e inteligência artificial

Some remarks on fundamental rights, digital transformation, and artificial intelligence

Anabela Costa Leão*

RESUMO: A necessidade de uma abordagem de direitos fundamentais à transformação digital e à regulação das tecnologias digitais e da inteligência artificial (IA) tem sido sustentada. Neste texto, identificam-se algumas linhas para reflexão sobre a proteção e promoção dos direitos fundamentais no ambiente digital. Após constatar o impacto simultaneamente positivo e negativo das tecnologias digitais e da IA na proteção e exercício de direitos fundamentais, discute-se a afirmação de "novos direitos" no espaço digital, bem como específicas questões de igualdade e não discriminação e de proteção da vulnerabilidade. Por último, identifica-se, a par da complexidade (da matéria e da regulação), a necessidade de uma regulação constitucional.

PALAVRAS-CHAVE: Direitos fundamentais; igualdade; constitucionalismo; inteligência artificial; tecnologia.

* Investigadora do Centro de Investigação Jurídica da Faculdade de Direito da Universidade do Porto (CIJ-FDUP) e Professora Auxiliar da FDUP. Faculdade de Direito da Universidade do Porto. aleao@direito.up.pt

ABSTRACT: A fundamental rights approach to digital transformation, digital technologies, and artificial intelligence (AI) has been emphasized. This paper aims to discuss the protection and promotion of fundamental rights in the digital space. It addresses the positive and negative impacts of digital technologies and AI in fundamental rights, the need for "new rights" in the digital space and specific issues of equality and nondiscrimination and protection of vulnerable groups and persons. Finally, it discusses the need for constitutional regulation.

KEYWORDS: Fundamental rights; equality; constitutionalism; artificial intelligence; technology.

1. Enquadramento[1]

A necessidade de uma abordagem de direitos fundamentais e humanos à transformação digital e à regulação das tecnologias digitais e da inteligência artificial (IA)[2] tem sido sustentada, *inter alia,* pela União Europeia e pelas Nações Unidas, afirmando de forma clara que os direitos fundamentais e

[1] Este texto corresponde, com alguns aprofundamentos, ao suporte escrito da intervenção com tema "Direitos fundamentais e Inteligência Artificial: dimensões de inclusão e exclusão", proferida no *Congresso Internacional "Direito e Inteligência Artificial",* org. FDUP e CIJ-FDUP, que decorreu na Faculdade de Direito da Universidade do Porto (Portugal) a 12 e 13 de maio de 2022.

[2] Não existindo uma definição universalmente aceite de IA, utiliza-se aqui o conceito proposto pelo Grupo de Peritos de Alto Nível em Inteligência Artificial da Comissão Europeia: "O conceito de inteligência artificial (IA) aplica-se a sistemas que apresentam um comportamento inteligente, analisando o seu ambiente e tomando medidas – com um determinado nível de autonomia – para atingir objetivos específicos. Os sistemas baseados em inteligência artificial podem estar puramente confinados ao "software", atuando no mundo virtual (por exemplo: assistentes de voz, programas de análise de imagens, motores de busca, sistemas de reconhecimento facial e de discurso), ou podem estar integrados em dispositivos físicos (por exemplo: robôs avançados, automóveis autónomos, veículos aéreos não tripulados ou aplicações da Internet das coisas)"- *apud* Agência de Direitos Fundamentais da UE, *Preparar o futuro: Inteligência Artificial e Direitos Fundamentais. Síntese,* 2021, p. 1. Disponível em https://fra.europa.eu/sites/default/files/fra_uploads/fra-2021-artificial-intelligence-summary_pt.pdf [04/11/2022].

humanos se aplicam tanto *online* como *offline*[3]. Neste sentido vejam-se, por exemplo, a *Declaração de Lisboa – Democracia Digital com propósito* (2021)[4], apelando ao reforço da dimensão humana do ecossistema digital, ou o *Roteiro para a Cooperação Digital* apresentado pelo Secretário-Geral da ONU (2020) em linha com os *Objetivos do Desenvolvimento Sustentável*[5] .

Neste texto procura dar-se brevíssima nota de algumas questões suscitadas pela adoção de uma abordagem de direitos fundamentais em sentido amplo (incluindo direitos fundamentais constitucionais e direitos humanos)[6] ao digital e à IA, tendo presente que também neste domínio se manifesta a necessidade de articulação entre diferentes níveis de proteção, internos e internacionais, de direitos, em necessário diálogo e frequente convergência[7], e a consolidação dos

[3] The United Nations Secretary-General's roadmap for digital cooperation "Ensuring the protection of human rights", disponível em https://www.un.org/techenvoy/content/digital-human-rights [04/11/2022].

[4] *Declaração de Lisboa – Democracia Digital com propósito* (2021) https://www.lisbondeclaration.eu

[5] United Nations, *Report of the Secretary-General. Road map for digital cooperation,* June 2020. Disponível em https://www.un.org/en/content/digital-cooperation-roadmap/assets/pdf/Roadmap_for_Digital_Cooperation_EN.pdf [04/11/2022].

[6] Tradicionalmente distingue-se entre direitos fundamentais e direitos humanos, os primeiros referindo-se à esfera estadual constitucional de proteção e os últimos reservados à esfera jus-internacional. Ora, como ensina José Carlos Vieira de Andrade, *Os direitos fundamentais na Constituição portuguesa de 1976,* Coimbra: Almedina, 2021, p. 15, os direitos a que chamamos direitos fundamentais podem ser vistos sob diferentes perspetivas: como direitos naturais de todas as pessoas, independentemente dos tempos e dos lugares (perspetiva filosófica ou jusnaturalista), como os direitos mais importantes das pessoas, num determinado tempo e lugar, i. é, num Estado concreto ou numa comunidade de Estados concreta (perspetiva estadual ou constitucional) e como direitos de todas as pessoas num certo tempo em todos os lugares ou, pelo menos, em grandes regiões do mundo (perspetiva universalista ou internacionalista). E, com efeito, verifica-se que a expressão "direitos fundamentais" é hoje utilizada para além do domínio das constituições nacionais (*e.g.* na União Europeia), no contexto de uma "compreensão plural" e "multinível" do constitucionalismo, *vd. inter alia* Suzana Tavares da Silva, *Direitos Fundamentais na Arena Global,* Coimbra: Imprensa Universidade Coimbra, 2011, pp. 9 ss. Sem prejuízo da utilidade da distinção entre direitos fundamentais e direitos humanos, ao menos no que diz respeito às diferentes estruturas de proteção (como se escreveu em Anabela Costa Leão, "A Carta dos Direitos Fundamentais da União Europeia", *in RFDUP,* Ano III, 2006, p. 41 ss.), utiliza-se neste texto um sentido amplo de direitos fundamentais, reservando o sentido estrito da expressão para a proteção conferida por uma constituição estadual. Para mais desenvolvimentos, remete-se para Suzana Tavares da Silva, *Direitos Fundamentais na Arena Global,* cit., e, para uma defesa da distinção entre direitos humanos e direitos fundamentais, para José de Melo Alexandrino, *O discurso dos direitos,* Coimbra: Coimbra Editora/Wolters Kluwer, 2011, p. 179 ss. e p. 205 ss.

[7] Convocam-se aqui a chamada "proteção multinível" dos direitos humanos e fundamentais e o "diálogo" entre níveis universais, regionais – com ênfase, no contexto europeu, para o Direito da União Europeia e para o sistema de proteção da Convenção Europeia de Direitos Humanos (CEDH)

direitos fundamentais como dimensão essencial de um Direito global – para alguns mesmo de um *Direito constitucional global*[8] – em construção.

Considerando a natureza evolutiva e aberta da proteção de direitos fundamentais, afina-se a proteção dos direitos das pessoas em *contextos específicos*, neste caso, o ambiente digital, e face a *desafios específicos,* como os que resultam das tecnologias digitais e da IA[9]. Para Custers, no cruzamento entre o Direito e o desenvolvimento das tecnologias digitais colocam-se diversos tipos de questões de proteção de direitos, sejam as relativas a violações de direitos ou aos conflitos de direitos existentes resultantes (do uso) de novas tecnologias, sejam novas questões resultantes (do uso) de novas tecnologias, para as quais "ainda não existem direitos"[10].

Os direitos fundamentais têm de ser compreendidos à luz do desenvolvimento científico e tecnológico, da globalização e da sociedade de informação e de "risco"[11]. Tomando de Vieira de Andrade[12] a proposta de interpretação atual dos direitos fundamentais a partir da trilogia *segurança, diversidade, solidariedade,* todos estes vetores parecem especialmente relevantes no contexto digital. A *segurança* é hoje também a segurança digital ou a garantia de um ambiente em linha protegido e seguro, seja quanto à privacidade, seja quanto à proteção contra a cibercriminalidade. Do mesmo passo, a *garantia da diversidade,* que é uma característica do sistema de proteção de direitos fundamentais, manifestada no pluralismo ideológico mas igualmente na diversidade cultural e linguística, ganha também novos contornos no ambiente digital, com a *Declaração Europeia sobre os direitos e princípios digitais para a década digital* (2022)[13]

– e constitucionais de proteção de direitos. Sobre a questão, *inter alia, vd.* Suzana Tavares da Silva, *Direitos Fundamentais na Arena Global,* cit.

[8] Discutindo o conceito e as perspetivas de constitucionalismo global, Ángel Aday Giménez Alemán, "El constitucionalismo global: ¿neologismo necesario o mera cacofonía?", *in Revista Española de Derecho Constitucional*, 117, septiembre-diciembre, 2019, pp. 139-166. DOI: https://doi.org/10.18042/cepc/redc.117.05

[9] Toma-se aqui de empréstimo o conceito de especificação de Gregorio Peces-Barba Martinez, *Curso de Derechos Fundamentales. Teoria General,* Madrid, Universidad Carlos III, 1999, pp. 154 ss., que identifica quatro processos ou linhas de evolução dos direitos fundamentais: i) positivação, ii) generalização, iii) internacionalização e iv) especificação (este, baseado em Norberto Bobbio).

[10] Bart Custers, "New digital rights: Imagining additional fundamental rights for the digital era", *in Computer Law & Security Review,* Volume 44, 2022,105636, p. 1, https://doi.org/10.1016/j.clsr.2021.105636.

[11] Assim ensina José Carlos Vieira de Andrade, *Os direitos fundamentais...*, cit., p. 59-61.

[12] José Carlos Vieira de Andrade, *Os direitos fundamentais...*, cit., p. 61.

[13] COM (2022) 27 final, *Comunicação da Comissão que estabelece uma Declaração Europeia sobre os direitos e princípios digitais para a década digital,* disponível em https://digital-strategy.ec.europa.eu/en/library/declaration-european-digital-rights-and-principles

a estabelecer que "todas as pessoas devem ter acesso a um ambiente em linha fiável, diversificado e multilingue" (*vd.* art. 22º da Carta dos Direitos Fundamentais da União Europeia, doravante CDFUE) e a vincar que "o acesso a conteúdos diversificados contribui para um debate público pluralista e deve permitir que todos participem no processo democrático" (Cap. IV). A *solidariedade e inclusão*, por seu turno, devem ser preservadas e aprofundadas na transformação digital. Como se lê no suprarreferido *Roteiro para a Cooperação Digital*, as tecnologias digitais não existem no vazio[14] e o comprovado acesso desigual às tecnologias tende a refletir e a agravar as desigualdades económicas, sociais e culturais já existentes, tornando-se necessário agir no sentido da "inclusão digital"[15].

Os direitos fundamentais integram, juntamente com a limitação e separação de poderes, o núcleo da ideia de constituição. Além desta dimensão constitucional, é necessário atender aos impactos do digital e da IA na organização das sociedades e, em especial, no funcionamento e na qualidade das democracias constitucionais[16], tendo ainda em vista as funções que as constituições são chamadas a desempenhar, designadamente de limitação do poder e de garantia do pluralismo, e a sua capacidade reguladora perante o digital (e que não pode, hoje, deixar de convocar dimensões transnacionais)[17].

O desenvolvimento da tecnologia digital e o seu "impacto disruptivo" nas nossas sociedades levam mesmo Edoardo Celeste a afirmar que vivemos um "novo momento constitucional"[18]. O autor adota o conceito de "constitucionalismo digital" – sobre o qual não há consenso na doutrina[19] – para identificar a ideologia (que toma em sentido neutro, ou seja, como "conjunto estruturado de valores e ideais") "que visa estabelecer e garantir a existência de um enqua-

[14] *Roadmap for digital cooperation,* cit., p. 2.

[15] *Roadmap for digital cooperation,* cit., pp. 10-11.

[16] Vd. Oreste Pollicino, e Giovanni De Gregorio, "Constitutional Law in the Algorithmic Society", *in* H. Micklitz, O. Pollicino, A. Reichman, A. Simoncini, G. Sartor e G. De Gregorio (Eds.), *Constitutional Challenges in the Algorithmic Society*, Cambridge: Cambridge University Press, 2021, pp. 3-24, pp. 4 ss. Sobre o que designa como "capitalismo da vigilância", antidemocrático e anti igualitário, correspondente ao "triunfo de um poder instrumentarista", *vd.* Shoshana Zuboff, *A era do capitalismo da vigilância. A disputa por um futuro humano na nova fronteira do poder,* Relógio D'Água, 2019, *passim,* em especial, p. 11(noção) e pp. 576 ss.

[17] Assim, Oreste Pollicino e Giovanni De Gregorio, "Constitutional Law in the Algorithmic Society", cit., *maxime* pp. 5 ss.

[18] Edoardo Celeste, "Digital constitutionalism: a new systematic theorisation", *in International Review of Law, Computers & Technology,* 33, 1, 2019, pp. 76-99, 78 ss. DOI: 10.1080/13600869.2019.1562604.

[19] Apresentando e discutindo diferentes perspetivas, Edoardo Celeste, "Digital constitutionalism...", cit.

dramento normativo para a proteção dos direitos fundamentais e do balanceamento de poderes no ambiente digital" [20]. Trata-se, portanto, para o autor, de uma "declinação do moderno constitucionalismo", ou seja, de perceber de que forma a tecnologia digital afeta o "equilíbrio constitucional" resultante da proteção dos direitos fundamentais e o balanceamento de poderes, que designa por funções básicas do direito constitucional. Para Celeste, a tecnologia digital altera o equilíbrio constitucional pois amplia, simultaneamente, as possibilidades de exercício dos direitos fundamentais e o risco de ameaças a esses mesmos direitos (em virtude da expansão da possibilidade de transmitir informação), do mesmo passo que afeta o equilíbrio de poderes dada a centralidade desempenhada pelos atores privados, designadamente corporações, no ambiente digital, com poder para regular o uso das tecnologias digitais pelos indivíduos e cuja atuação pode consubstanciar violação de direitos fundamentais de particulares, deslocando a questão do equilíbrio de poderes da relação indivíduo-Estado[21]. Caberá ao "constitucionalismo digital" que propõe proteger direitos fundamentais e limitar tanto atores públicos como privados, já que tanto uns como outros podem violar direitos fundamentais[22], o que se concretiza num *processo* de produção de normas com esse preciso objetivo, que designa por "constitucionalização do ambiente digital" [23].

Assim, apesar de este texto se ocupar do impacto das tecnologias digitais nos direitos fundamentais (*lato sensu*), as dimensões constitucionais da transformação digital são mais amplas e variadas e estão, naturalmente, relacionadas entre si. A isto se voltará *infra,* na parte final do texto.

2. Direitos fundamentais e (em) ambiente digital

De seguida, procura-se identificar algumas linhas para reflexão sobre a proteção e promoção dos direitos fundamentais no ambiente digital.

2.1. Impactos (positivos e negativos) da "transição digital" no exercício de direitos fundamentais

O uso de tecnologias e o desenvolvimento digital pode afetar diversos direitos e de diversas formas[24], podendo identificar-se impactos positivos e negativos.

[20] Edoardo Celeste, "Digital constitutionalism...", cit., pp. 88-89.

[21] Edoardo Celeste, "Digital constitutionalism...", cit., p. 78 ss.

[22] Edoardo Celeste, "Digital constitutionalism...", cit., p. 89.

[23] Aludindo a um processo de generalização e re-especificação, Edoardo Celeste, "Digital constitutionalism...", cit., p. 89.

[24] Neste sentido, para o ambiente digital, Bart Custers, "New digital rights", cit. Exemplificativamente, um estudo do Conselho da Europa de 2017 sobre algoritmos e direitos humanos mostra

O espaço digital fornece um novo contexto de efetivação e suscita novas questões de proteção de direitos fundamentais, como o direito à identidade pessoal, o direito à proteção de dados pessoais ou o direito à proteção da vida privada, e de princípios com eles relacionados, como o princípio da igualdade. Este impacto não se faz sentir apenas nos chamados "direitos de primeira geração", mas igualmente em direitos sociais e culturais, como a educação ou a fruição cultural, e na própria forma como o Estado desempenha as suas tarefas de concretização de direitos.

Se a tecnologia digital e a IA representam uma oportunidade, enquanto contexto de aprofundamento e efetivação de direitos fundamentais, trazem igualmente riscos e desafios para a segurança e para a proteção destes direitos.

Tome-se como exemplo o direito à informação. O ambiente digital expande as possibilidades do já existente direito a informar, ser informado e se informar[25]. Ao mesmo tempo, amplia a possibilidade de definir perfis de utilizador e influenciar comportamentos através de ferramentas de IA[26], bem como amplia os riscos associados à chamada "desinformação[27], discutindo-se neste contexto a necessidade e os limites da moderação de conteúdos *online* (designadamente com o auxílio de ferramentas de IA)[28] e de uma proteção contra a desinformação[29], dada a ligação entre democracia, pluralismo e liberdade de expressão[30].

o impacto da IA em vários direitos, dos direitos sociais à liberdade de expressão e à participação política, bem como no plano da igualdade, vd. Council of Europe, *Algorithms and Human Rights. Study on the human rights dimensions of automated data processing techniques and possible regulatory implications,* Council of Europe Study, DGI (2017) 12, 2018. Disponível em https://rm.coe.int/algorithms-and-human-rights-en-rev/16807956b5. [04/11/2022].

[25] Consagrado, designadamente, no art. 37º CRP, no art. 10º CEDH e no art. 11º CDFUE.

[26] Oreste Pollicino e Giovanni De Gregorio, "Constitutional Law in the Algorithmic Society", cit., pp. 6 ss.

[27] Sobre a questão, Edoardo Celeste, "Digital constitutionalism…", cit. pp. 78 ss.

[28] Oreste Pollicino e Giovanni De Gregorio, "Constitutional Law in the Algorithmic Society", cit., pp. 8 ss.

[29] Sobre a questão, *inter alia,* Oreste Pollicino e Giovanni De Gregorio, "Constitutional Law in the Algorithmic Society", cit., pp. 7 ss. Entre nós, a *Carta Portuguesa de Direitos Humanos na Era Digital,* aprovada pela Lei nº 27/2021, de 17 de maio, consagrou no seu artigo 6º um "direito à proteção contra a desinformação", nos termos de cujo nº 1 cabia ao Estado *"assegurar o cumprimento em Portugal do Plano Europeu de Ação contra a Desinformação, por forma a proteger a sociedade contra pessoas singulares ou coletivas, de jure ou de facto, que produzam, reproduzam ou difundam narrativa considerada desinformação",* conceito que se procurava definir nos nºs 2 a 4: *"Considera-se desinformação toda a narrativa comprovadamente falsa ou enganadora criada, apresentada e divulgada para obter vantagens económicas ou para enganar deliberadamente o público, e que seja suscetível de causar um prejuízo público,*

Esta dupla dimensão é por demais enfatizada. Como se escreve no *Livro Branco sobre a inteligência artificial* (2020)[31] da Comissão Europeia, embora o recurso à IA possa tornar os produtos e os processos mais seguros, pode

nomeadamente ameaça aos processos políticos democráticos, aos processos de elaboração de políticas públicas e a bens públicos", sendo que *"considera-se, designadamente, informação comprovadamente falsa ou enganadora a utilização de textos ou vídeos manipulados ou fabricados, bem como as práticas para inundar as caixas de correio eletrónico e o uso de redes de seguidores fictícios"*, e excluindo *"os meros erros na comunicação de informações, bem como as sátiras ou paródias"*. O artigo estabelecia ainda (nº 5) um direito de queixa à Entidade Reguladora para a Comunicação Social "*contra as entidades que pratiquem os atos previstos no presente artigo, sendo aplicáveis os meios de ação referidos no artigo 21º e o disposto na Lei nº 53/2005, de 8 de novembro, relativamente aos procedimentos de queixa e deliberação e ao regime sancionatório"* e incumbia o Estado de apoiar *"a criação de estruturas de verificação de factos por órgãos de comunicação social devidamente registados e incentiva a atribuição de selos de qualidade por entidades fidedignas dotadas do estatuto de utilidade pública"* (nº6).

O Presidente da República entendeu suscitar, em 28 de julho de 2021, nos termos da alínea a) do nº 2 do art. 281º da CRP, a fiscalização abstrata sucessiva das normas do artigo 6º da Lei, por violação da liberdade de expressão (art. 37º CRP) e incumprimento dos requisitos do regime material e orgânico dos direitos, liberdades e garantias, em especial do regime das restrições e do princípio da proporcionalidade (artigo 18º CRP), bem como do princípio do Estado de Direito (art. 2º), designadamente por se entender em causa o princípio da densidade constitucional suficiente da lei restritiva de direitos liberdades, e garantias (considerando 7º do Pedido, disponível em www.presidencia.pt). Em 2022, a Provedora de Justiça requereu igualmente a fiscalização abstrata sucessiva dos nºs 5 e 6 da referida Lei, "por constituírem uma restrição injustificada e desproporcionada" (nº 2 do artigo 18º CRP) da liberdade de expressão e informação (artigo 37º CRP) e, quanto ao nº 6 do artigo 6º, da liberdade de imprensa (primeira parte do nº 4 do artigo 38º da CRP) (pedido disponível em www.provedor-jus.pt).

Entretanto, sem que o Tribunal Constitucional chegasse a pronunciar-se, o artigo 6º foi quase totalmente revogado pelo artigo 2º da Lei nº 15/2022, de 11 de agosto, passando a dispor, sob a mesma epígrafe, que *"O Estado assegura o cumprimento em Portugal do Plano Europeu de Ação contra a Desinformação, por forma a proteger a sociedade contra pessoas singulares ou coletivas, de jure ou de facto, que produzam, reproduzam ou difundam narrativa considerada desinformação"*. O procedimento legislativo parlamentar pode ser consultado em https://www.parlamento.pt/Legislacao/Paginas/Educacao_Carta-Portuguesa-de-Direitos-Humanos-na-Era-Digital.aspx .

Sobre a questão, *inter alia*, José de Melo Alexandrino, *Dez breves apontamentos sobre a Carta Portuguesa de Diretos Humanos na Era Digital*, ICJP/CIDP, 2021, p. 1, Domingos Soares Farinho, "The Portuguese Charter of Human Rights in the Digital Age: a legal appraisal", *in Revista Española de la Transparencia*, nº 13, 2º semestre jul-dec. 2021, pp. 85-105. DOI: https://doi.org/10.51915/ret.191 e, ainda, os Pareceres nº 2020/116, nº 2020/117 e nº 2022/56 da CNPD – Comissão Nacional de Proteção de Dados (disponíveis em www.cnpd.pt) sobre a dificuldade de conciliação da liberdade de expressão e opinião e a proteção contra a desinformação.

[30] Sobre a questão, *inter alia*, Oreste Pollicino e Giovanni De Gregorio, "Constitutional Law in the Algorithmic Society", cit., pp. 7 ss.

[31] COM (2020)65final, 19.02.2020, *Livro Branco sobre a inteligência artificial-uma abordagem europeia virada para a excelência e a confiança*. Disponível em https://eur-lex.europa.eu/legal-content/PT/TXT/?uri=CELEX:52020DC0065 [04/11/2022].

ter impactos negativos materiais e imateriais, designadamente associados à perda de privacidade, às limitações à liberdade de expressão ou à discriminação, bem como estar associada a riscos de lesão de direitos fundamentais e para a segurança (cibersegurança, aplicações de IA em infraestruturas críticas ou utilização maliciosa de IA)[32]. O mesmo é reiterado na *Proposta de Regulamento sobre Inteligência Artificial* apresentada em 2021[33], em cujo Considerando 15 se lê que "[a]lém das inúmeras utilizações benéficas da inteligência artificial, essa tecnologia pode ser utilizada indevidamente e conceder instrumentos novos e poderosos para práticas manipuladoras, exploratórias e de controlo social. Essas práticas são particularmente prejudiciais e devem ser proibidas, pois desrespeitam valores da União, como a dignidade do ser humano, a liberdade, a igualdade, a democracia e o Estado de Direito, bem como direitos fundamentais da União, incluindo o direito à não discriminação, à proteção de dados pessoais e à privacidade, e os direitos das crianças".

Configuram-se também, nota Custers[34], novas possibilidades de conflito entre direitos, a requerer operações de concordância prática e de restrição de direitos guiadas pelo princípio da proporcionalidade. Os exemplos de conflitos são vários e, note-se, podem configurar-se de forma subjetiva e/ou objetiva, como conflitos entre direitos de diferentes sujeitos ou entre direitos e bens jurídicos como, por exemplo, a segurança. São exemplos a colisão entre a proteção de dados pessoais e a segurança, visível por exemplo na questão do acesso a dados pessoais para efeitos de investigação criminal[35] ou o conflito entre a proteção de dados pessoais e a autonomia privada e a liberdade empresarial dos operadores[36].

A relevância de uma abordagem de direitos fundamentais não diz, assim e apenas, respeito à determinação de novos direitos ou à reconfiguração do âmbito de proteção de direitos menos novos, mas igualmente à possibilidade e necessidade de mobilizar toda uma dogmática voltada para a proteção de direitos e para a resolução de conflitos de direitos ou entre direitos e outros bens jurídicos, designadamente através do princípio do caráter restritivo das

[32] COM (2020)65final, *Livro Branco sobre a inteligência artificial*, cit., pp. 11 ss.

[33] COM (2021) 206 final 2021/0106(COD), *Proposta de Regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial (Regulamento Inteligência Artificial) e altera determinados atos legislativos da União* (doravante, *Proposta de Regulamento Inteligência Artificial).*

[34] Bart Custers, "New digital rights", cit., p. 3 ss.

[35] Bart Custers, "New digital rights", cit., p. 3 ss.

[36] *vd. Livro Branco* e *Proposta de Regulamento Inteligência Artificial*, cit., Ponto 3.5.

restrições e, muito particularmente, do *princípio da proporcionalidade*[37], que reivindica, também ele, uma espécie de estatuto de princípio constitucional global[38].

A proteção de direitos fundamentais determinará, assim, neste como noutros domínios, limites à utilização das tecnologias existentes. O elenco de práticas de IA proibidas do artigo 5º da *Proposta de Regulamento da Inteligência Artificial*[39] ilustra esta função de limite dos direitos fundamentais. Mas os deveres de proteção dos direitos fundamentais projetam-se também na própria definição do regime aplicável e do estabelecimento de garantias. Como resulta da *Proposta de Regulamento sobre Inteligência Artificial*, procura-se assegurar um nível elevado de proteção dos direitos fundamentais, estabelecendo obrigações de testagem *ex ante*, de gestão de riscos e um princípio de supervisão humana, complementados por garantias de reação contra violações dos seus direitos[40].

2.2. A natureza evolutiva da proteção e a questão dos "novos direitos"

A proteção dos direitos fundamentais (*lato sensu*) é dinâmica e evolutiva, providenciando novas respostas para novos desafios de proteção da pessoa e da sua dignidade[41]. A enumeração de novos direitos ou de novas dimensões de direitos fundamentais pode ocorrer através da aprovação ou atualização de catálogos de direitos, e.*g.* cartas de direitos, constituições ou convenções, ou da interpretação jurisprudencial, concretizando dimensões de jusfundamentalidade[42]. Contudo, mesmo considerando a previsão constitucional de

[37] Tal como resulta dos nºs 2 e 3 do art. 18º da CRP, do art. 52º da CDFUE e da CEDH, designadamente do nº 2 do art. 8º (proteção da vida privada e familiar).

[38] Sobre a questão, remetendo-se para os autores aí referidos, Anabela Costa Leão, "O Princípio da Proporcionalidade e os seus críticos", *in* AA. VV., *O Princípio da Proporcionalidade. XIII Encontro de Professores de Direito Público,* Coimbra: Instituto Jurídico, 2021, pp. 127 ss., www.doi.org/10.47907/clq2021_2a7.

[39] *Proposta de Regulamento Inteligência Artificial,* cit., Ponto 3.5. Sobre a "pirâmide de risco" em que assenta a *Proposta de Regulamento Inteligência Artificial, vd.* Dimitar Lilkov, "Regulating artificial intelligence in the EU: A risky game", *in European View,* 20(2), 2021, pp. 166–174, pp. 168 ss. DOI: https://doi.org/10.1177/17816858211059248

[40] *Proposta de Regulamento Inteligência Artificial,* cit., Ponto 3.5.

[41] A que não é alheio o risco, para o qual alerta de há muito José Casalta Nabais, de a "panjusfundamentalização" levar à banalização dos direitos fundamentais, *vd.* "Algumas reflexões críticas sobre os direitos fundamentais", in *Por uma liberdade com responsabilidade*, Coimbra: Coimbra Editora, 2007, p. 103 ss.

[42] O que dependerá da forma como está configurado cada sistema de proteção, *vd.* por exemplo entre nós o mecanismo da cláusula aberta do nº 1 do art. 16º da CRP, que permite o reconhecimento de direitos fundamentais consagrados pelo legislador ordinário. Sobre a questão dos "direitos

direitos fundamentais e a interpretação das normas constitucionais que os consagram, pode suscitar-se a questão dos limites a uma interpretação atualista e da repartição de tarefas entre o legislador – neste caso, o legislador da revisão constitucional – e o juiz no aprofundamento da resposta constitucional aos novos desafios da era digital, mantendo a "reflexividade da constituição" (no sentido proposto por J. J. Gomes Canotilho), questão para que alerta Raquel Brízida de Castro[43].

Alguns dos direitos que vêm sendo formulados no contexto digital podem ser reconduzidos ao âmbito de proteção de direitos fundamentais já existentes, como a proteção de dados pessoais ou a reserva da vida privada, enquanto densificações destes. Contudo, poderá discutir-se se e em que medida a proteção conferida pelos direitos humanos e fundamentais existentes é suficiente e/ou deverá (e em que termos) ser reforçada pela especificação normativa de novos direitos relacionados com a transformação digital ou, mais especificamente, com a IA, sejam direitos positivos ou direitos de defesa, designadamente nos textos constitucionais[44].

Um dos argumentos utilizados a favor da especificação e consagração de novos direitos é a incerteza e a insegurança sobre a proteção existente, indesejável à luz do princípio do Estado de Direito (*rule of law*)[45].

Como Custers nota, muitos dos textos que protegem direitos fundamentais foram elaborados no e para um mundo completamente diferente do mundo atual (sendo, muitos deles, anteriores ao digital, à *internet* ou à IA), ou para compreensões muito diferentes das atuais sobre os direitos consagrados[46]. Este argumento deve, contudo, ser relativizado – como o autor também nota, a abertura das normas de direitos fundamentais permite, através da interpreta-

não enumerados", *vd.* Cristina Queiroz, *Direitos Fundamentais,* Coimbra: Coimbra Editora/ Wolters Kluwer, 2010.

[43] Raquel Brízida de Castro, "Constituição e ciberespaço: argumentos para um 'direito constitucional do inimigo'? "*in Ciberlaw,* nº 1, 2016.

[44] Foi recentemente notícia a previsão específica dos chamados "neuro direitos" na Constituição chilena através da Lei 21.383 (https://bcn.cl/2scpd). Sobre a questão, veja-se Cinthia Obladen de Almendra Freitas, "Neurodireitos: O exemplo do Chile e a regulação das neurotecnologias", *IberICONnect,* 8/02/2022. Disponível em: https://www.ibericonnect.blog/2022/02/neurodireitos-o-exemplo-do-chile-e-a-regulacao-das-neurotecnologias/ [04/11/2022].

[45] Sobre este argumento, a que se voltará *infra,* Bart Custers, "New digital rights", cit., pp. 4 ss. e Oreste Pollicino e Giovanni De Gregorio, "Constitutional Law in the Algorithmic Society", cit., pp. 9 ss.

[46] Assim, e dando o exemplo do entendimento sobre a privacidade e o chamado "paradoxo da privacidade", Bart Custers, "New digital rights", cit., p. 4.

ção, a sua aplicação a situações e contextos muito diversificados sem perder o seu valor jusfundamental[47]. E, com efeito, a interpretação atualista, dinâmica e evolutiva das normas de direitos fundamentais (que tem, contudo, limites) permite estender a sua proteção a "novas realidades" resultantes da evolução científica e tecnológica – veja-se o exemplo da CEDH[48].

Para Custers, perante lacunas ou incertezas na proteção conferida por direitos já consagrados, a par das mudanças nas conceções subjacentes a esses direitos, será de refletir sobre a necessidade de especificar direitos fundamentais adicionais para a "era digital", bem como de procurar determinar quais serão esses direitos[49]. Neste contexto, propõe-se lançar a discussão sobre alguns "novos direitos digitais" , formulados em termos abstratos[50] e, na realidade, mais novos uns do que outros[51]: o "direito a estar *offline*"[52], o "direito de aceso à internet"[53], o "direito a não saber"[55], o direito (reforçado) "a mudar

[47] Bart Custers, "New digital rights", cit.,

[48] Veja-se, a propósito dos desenvolvimentos científicos e tecnológicos nas áreas das ciências da vida, o recurso, pelo Tribunal Europeu de Direitos Humanos, à interpretação da CEDH como "instrumento vivo" e o recurso ao art. 8º CEDH (reserva da vida privada e familiar) para "revelar" novos direitos no âmbito da biomedicina, por exemplo relativos à procriação medicamente assistida, vd. Anabela Costa Leão, "O contributo do Tribunal Europeu dos Direitos do Homem" *in* Luísa Neto e Rute Teixeira Pedro, R., *Debatendo a procriação medicamente assistida*. (pp. 23-40). Centro de Investigação Jurídico-Económica: Faculdade de Direito da Universidade do Porto, 2018, pp. 23-40, online em www.direito.up.pt (e-books) e, para maior desenvolvimento, Alasdair Mowbray, "Between the will of the Contracting Parties and the needs of today: Extending the scope of Convention rights and freedoms beyond what could have been foreseen by the drafters of the ECHR", *in* Eva Brems e Janneke Gerards (Eds.), *Shaping Rights in the ECHR: The Role of the European Court of Human Rights in Determining the Scope of Human Rights* , 2014, Cambridge: Cambridge University Press, pp. 17-37). DOI:10.1017/CBO9781107337923.003

[49] Bart Custers, "New digital rights", cit., pp. 4 ss., em especial p. 5 ss.

[50] Bart Custers, "New digital rights", cit., p. 5, faz algumas ressalvas: trata-se sobretudo de lançar a discussão sobre a necessidade e viabilidade destes "novos" direitos, mais do que afirmar desde já a sua existência; os direitos não se encontram elaborados com detalhe; não se trata de um catálogo exaustivo e a "novidade" destes direitos não significa que nunca tenham sido mencionados, mas apenas que ainda não foram amplamente incorporados nas diferentes ordens jurídicas ou, pelo menos, não no contexto digital.

[51] Bart Custers, "New digital rights", cit., pp. 6 ss.

[52] Bart Custers, "New digital rights", cit., p. 6 ss. Se, como se verá *infra,* nalguns casos a recusa em utilizar serviços digitais pode resultar da falta de meios para o fazer, poderá corresponder a uma *opção* por *não estar ligado*.

[53] A defesa de um direito humano de acesso à internet foi já sugerida de forma não vinculativa pelas Nações Unidas em 2016 e é sustentada por vários autores, designadamente por Merten Reglitz, para quem o direito de acesso livre à internet, gratuito para quem não tenha meios económicos, deve ser reconhecido como direito humano universal dado o facto de ele ser necessário para possibilitar o exercício de outros direitos fundamentais, sendo não um "luxo", mas um requisito mínimo

de opinião"[55], o "direito a começar de novo com um registo digital limpo"[56], o "direito a prazos de validade para dados", o "direito a saber o valor dos nossos dados", bem como os direitos "a um ambiente digital limpo" e "seguro" e o "direito à educação digital". Direitos estes que, note-se, têm como contrapartida deveres ou obrigações estaduais, por exemplo deveres de proteção e promoção da segurança[57]. Destes direitos, Custers nota que alguns já estão normativamente consagrados nalguns ordenamentos, como o direito a estar desligado ou o direito de acesso à Internet[58], ao passo que outros não, sendo sobretudo propostas doutrinais, e outros ainda são extensões de direitos não-digitais já existentes[59], *e.g.* os direitos a um ambiente digital limpo e seguro e o direito à educação digital, que são extensões dos direitos ao ambiente, à segurança e à educação.

No que toca especificamente à IA, a transparência – ou, por outra, a *opacidade* – é problemática do ponto de vista dos direitos fundamentais[60], avultando aqui especificamente a garantia de direitos de informação sobre a utilização de

para que as pessoas possam levar "vidas decentes", vd. Merten Reglitz, "The human right to free internet access", *in Journal of Applied Philosophy*, Vol. 37, No. 2, 2020, pp. 314 ss. DOI: 10.1111/japp.12395. Note-se que não se trata apenas de um direito a não ser privado de acesso à internet, por exemplo pela interrupção do acesso, vd. art. 5º da Carta Portuguesa de Direitos Humanos na Era Digital, mas da garantia de condições de acesso efetivo, igualdade e neutralidade – vd. Bart Custers, "New digital rights", cit., pp. 7 ss.

[54] Segundo Bart Custers, "New digital rights", cit., p. 7, trata-se do oposto do direito a saber, traduzindo-se, por exemplo, no direito a não conhecer as previsões obtidas por recurso a *big data* sobre aspetos da sua vida, como a probabilidade de contrair uma doença ou de se divorciar.

[55] O autor, "New digital rights", cit., p.9, ressalva que este direito poderia porventura enquadrar-se na liberdade de pensamento ou de expressão, mas alvitra que o desenvolvimento tecnológico atual e o facto de o utilizador se encontrar frequentemente em "*filter bubbles*" pode exigir um direito reforçado a mudar de opinião.

[56] Como nota Bart Custers, "New digital rights", cit., p. 9, já parcialmente ensaiado no direito ao esquecimento do Regulamento Geral de Proteção de Dados (aprovado pelo Regulamento UE nº 679/2016, de 27 de abril), mas com um alcance mais alargado.

[57] Bart Custers, "New digital rights", cit., pp. 9-10.

[58] Veja-se o exemplo do nº 2 do artigo 5ºA da Constituição da Grécia, referido por Bart Custers, "New digital rights", cit., p. 4 que estabelece "*2. All persons have the right to participate in the Information Society. Facilitation of access to electronically transmitted information, as well as of the production, exchange, and diffusion thereof, constitutes an obligation of the State, always in observance of the guarantees of articles 9, 9A and 19.*" (Constituição da Grécia de 1975, revista em 2019, versão em língua inglesa disponível e consultada em https://www.hellenicparliament.gr)

[59] Para uma visão global, Bart Custers, "New digital rights", cit., p. 12, Table 1.

[60] Oreste Pollicino e Giovanni De Gregorio, "Constitutional Law in the Algorithmic Society", cit., pp. 3-4.

tecnologias de IA[61], frequentemente invisível para o utilizador[62], e do direito a não ser submetido a uma decisão baseada apenas em processamento automatizado[63]. O facto de as pessoas não terem conhecimento do recurso a IA não só afeta as suas escolhas e a sua capacidade de autodeterminação como pode também afetar, depois, a sua capacidade de reação face a um uso ilícito dessas tecnologias, prejudicando o seu direito a uma tutela jurisdicional efetiva, tal como é, por exemplo, acentuado no Livro Branco da Comissão Europeia[64].

A proteção destes direitos pode fazer-se por diversas vias, no contexto da proteção multinível suprarreferida, com graus de eficácia e mecanismos de proteção – é preciso dizê-lo – diferenciados. Acresce que, porque se encontram no centro da regulação jurídica digital, os direitos fundamentais são aptos a motivar convergências na proteção, mas igualmente a revelar dissonâncias, as mesmas que em geral já podem resultar da aplicação multinível dos direitos fundamentais, exigindo operações de interpretação e aplicação jurídica dos direitos previstos em diferentes catálogos pelos operadores jurídicos e órgãos judiciais – incluindo "diálogos interjurisdicionais", na expressão do Tribunal Constitucional português[65].

[61] Vejam-se a alínea f) do nº 2) do artigo 13º e a alínea g) do nº 2) do artigo 14º do Regulamento Geral de Proteção de Dados e os artigos 13º e 52º da *Proposta de Regulamento sobre Inteligência Artificial* sobre transparência e prestação de informação sobre IA aos utilizadores.

[62] Sobre a questão, Cinthia Obladen de Almendra Freitas, Neurodireitos, cit.

[63] Oreste Pollicino e Giovanni De Gregorio, "Constitutional Law in the Algorithmic Society", cit., pp. 20 ss. Veja-se o artigo 22º do Regulamento Geral de Proteção de Dados da União Europeia (Regulamento UE 2016/679 do Parlamento Europeu e do Conselho de 27 de abril de 2016). Também a já mencionada Carta Portuguesa dos Direitos Humanos na Era Digital, no seu art. 9º, sobre uso da inteligência artificial e robots, estabelece que *"1 – A utilização da inteligência artificial deve ser orientada pelo respeito dos direitos fundamentais, garantindo um justo equilíbrio entre os princípios da explicabilidade, da segurança, da transparência e da responsabilidade, que atenda às circunstâncias de cada caso concreto e estabeleça processos destinados a evitar quaisquer preconceitos e formas de discriminação. 2 – As decisões com impacto significativo na esfera dos destinatários que sejam tomadas mediante o uso de algoritmos devem ser comunicadas aos interessados, sendo suscetíveis de recurso e auditáveis, nos termos previstos na lei."*

[64] *Livro Branco sobre a inteligência artificial*, cit. Aí se lê, p. 13, "As características específicas de muitas tecnologias de IA, incluindo a opacidade ("efeito de caixa negra"), a complexidade, a imprevisibilidade e o comportamento parcialmente autónomo, podem dificultar a verificação do cumprimento e prejudicar a aplicação efetiva das regras do direito da UE em vigor destinadas a proteger os direitos fundamentais. As autoridades responsáveis pela aplicação da lei e as pessoas afetadas podem não dispor dos meios necessários para verificar a forma como foi tomada uma determinada decisão com o envolvimento da IA e, por conseguinte, se as regras pertinentes foram respeitadas. As pessoas singulares e coletivas podem deparar-se com dificuldades no acesso efetivo à justiça em situações em que tais decisões as possam afetar negativamente".

[65] Veja-se a propósito o Acórdão nº 268/2022 do Tribunal Constitucional. ("Metadados"), em especial o Ponto, 8, onde se lê "Por estas razões, situando-se as normas fiscalizadas no domínio

A CRP fornece um bom exemplo de abertura do catálogo de direitos fundamentais a novas realidades científicas e tecnológicas[66]. Do texto constitucional resulta, desde logo, a proteção da reserva da vida privada e a proteção de dados pessoais (vd. nº 1 do artigo 26º e artigo 35º CRP)[67]. A Constituição portuguesa foi pioneira na consagração da proteção de dados face à informática, dedicando um artigo à "utilização da informática" e consagrando

de aplicação do direito da União Europeia, a interpretação dos parâmetros constitucionais a que as regras em crise se submetem tem em conta o sentido das normas europeias, procurando-se estabelecer a interpretação mais próxima do direito europeu. É, aliás, o que a requerente sustenta nos artigos 42º a 45º do pedido, solicitando ao Tribunal Constitucional que interprete os parâmetros da Constituição portuguesa à luz da Carta. E foi justamente o que o Tribunal Constitucional concluiu no Acórdão nº 464/2019: *"por força das normas do artigo 8º da Constituição que estabelecem a relevância do Direito Internacional e do Direito da União na ordem jurídica interna e, também, da cláusula aberta no domínio dos direitos fundamentais consagrada no artigo 16º da Constituição, este Tribunal não pode deixar de considerar os direitos fundamentais consagrados na CDFUE e na referida Convenção, devendo igualmente ter em conta, numa perspetiva de diálogo interjurisdicional, a interpretação que dos mesmos tem vindo a ser feita pelas instâncias competentes para a sua aplicação, nomeadamente o Tribunal de Justiça da União Europeia ("TJUE") e o Tribunal Europeu dos Direitos Humanos ("TEDH")"*. Disponível em www.tribunalconstitucional.pt

[66] E não apenas no domínio do digital, vd. n,º 3 do art. 26º e alínea e) do nº 2 do art. 67º CRP, que se referem ao desenvolvimento tecnológico na genética e procriação assistida. Acresce que, no sistema constitucional português, a proteção constitucional de novos direitos fundamentais (em sentido estrito) tem de ser compreendida no contexto da abertura constitucional ao direito internacional e europeu (artigo 8º CRP) e da cláusula aberta de direitos fundamentais (nº 1 do artigo 16º CRP).

[67] Sobre a distinção e relação entre ambos, v. Luísa Neto, *Novos direitos ou novo(s) objetos para o Direito?* U. Porto, 2010, pp. 65 ss. e Catarina Sarmento e Castro, "40 anos de "utilização da informática" – o art. 35º da CRP", *in e-publica*, nº 3, 2016, pp. 43 ss. Entre outra jurisprudência, refira-se que, recentemente, o Tribunal Constitucional foi (novamente) chamado a pronunciar-se, através do Acórdão nº 268/2022, sobre os chamados "Metadados", na sequência de um pedido de fiscalização sucessiva abstrata da Provedora de Justiça (interposto ao abrigo da alínea *d)* do nº 2 do artigo 281º da CRP) da constitucionalidade das normas constantes dos artigos 4º, 6º e 9º da Lei nº 32/2008, de 17 de julho, por violarem o princípio da proporcionalidade na restrição dos direitos à reserva da intimidade da vida privada e familiar (nº 1 do artigo 26º CRP), ao sigilo das comunicações (nº 1 do artigo 34º CRP) e a uma tutela jurisdicional efetiva (nº 1 do artigo 20º CRP), tendo o Tribunal decidido "declarar a inconstitucionalidade, com força obrigatória geral, da norma constante do artigo 4º da Lei nº 32/2008, de 17 de julho, conjugada com o artigo 6º da mesma lei, por violação do disposto nos números 1 e 4 do artigo 35º e do nº 1 do artigo 26º, em conjugação com o nº 2 do artigo nº 18º, todos da Constituição", bem como "declarar a inconstitucionalidade, com força obrigatória geral, da norma do artigo 9º da Lei nº 32/2008, de 17 de julho, relativa à transmissão de dados armazenados às autoridades competentes para investigação, deteção e repressão de crimes graves, na parte em que não prevê uma notificação ao visado de que os dados conservados foram acedidos pelas autoridades de investigação criminal, a partir do momento em que tal comunicação não seja suscetível de comprometer as investigações nem a vida ou integridade física de terceiros, por violação do disposto no nº 1 do artigo 35º e do nº 1 do artigo 20º, em conjugação com o nº 2 do artigo 18º, todos da Constituição".

um direito fundamental à autodeterminação informativa[68]. O artigo, que foi sofrendo alterações em sede de revisão constitucional[69], mostrou abertura à evolução tecnológica e acolheu na revisão constitucional de 1997 a garantia de "livre acesso às redes informáticas de uso público"[70], ou seja, um direito fundamental à internet[71]. Como escreve Catarina Sarmento e Castro, este direito apresenta-se, simultaneamente, como "direito de acesso à rede" e como "direito de utilização das funcionalidades" correspondentes, designadamente mediante o uso do correio eletrónico e das redes sociais[72]. Para a autora, o direito de acesso à internet assume-se como direito fundamental instrumental ao permitir novas formas de exercício direitos já existentes, como a liberdade de expressão ou informação, mas é também um direito com conteúdo próprio, enquanto "direito de interação e de participação na estrutura social em rede, de relacionamento digital" ou "direito de integrar a sociedade digital, através da inserção e da interação na infraestrutura tecnológica, humana e social, que é a rede", ligado ao direito ao desenvolvimento da personalidade[73].

[68] Como nota Catarina Sarmento e Castro, "40 anos de "utilização da informática", cit., p. 44.

[69] Já depois da Conferência que esteve na base deste texto, foram apresentados projetos de revisão constitucional dando-se início ao procedimento de revisão constitucional (art. 284º ss. CRP) que, a concretizar-se, será a 8ª revisão do texto constitucional. Entre as alterações previstas em alguns projetos de revisão, estão precisamente alterações aos artigos 34º e 35º para regulação do acesso aos chamados metadados e introdução de um direito ao apagamento de dados pessoais (vd. app.paralmento.pt).

[70] Nos termos do nº 6 do art. 35º, então introduzido, a CRP estabelece que "A todos é garantido livre acesso às redes informáticas de uso público, definindo a lei o regime aplicável aos fluxos de dados transfronteiras e as formas adequadas de proteção de dados pessoais e de outros cuja salvaguarda se justifique por razões de interesse nacional". Recorde-se ainda que, desde a RC de 1997 a CRP, no art. 26º, a par naturalmente da proteção conferida pelos demais direitos fundamentais (arts. 13º e 35º), já estabelece um "*direito à proteção legal contra quaisquer formas de discriminação*" (nº 1) e que "*A lei estabelecerá garantias efetivas contra a obtenção e utilização abusivas, ou contrárias à dignidade humana, de informações relativas às pessoas e famílias*" (nº 2).

[71] Assim entende Catarina Sarmento e Castro, "40 anos de "utilização da informática",cit., p. 49, que nota que assim se ampliou o objeto material do art. 35º CRP.

[72] Catarina Sarmento e Castro, "40 anos de "utilização da informática",cit., p. 52. Citando a autora, "[o] direito consagrado no nº 6 do artigo 35º deve ser encarado nas suas múltiplas facetas, por um lado, como direito de acesso à rede, às infraestruturas tecnológicas que tornam possível os serviços de comunicação, mas também, por outro, como direito de utilização das funcionalidades comunicação, de informação, de conhecimento, de participação, de interação, de partilha e de relacionamento digital da Internet – designadamente mediante uso do correio eletrónico, de redes sociais, de motores de busca e da *World Wide Web* em geral. É o direito de acesso a um novo caminho ou espaço de comunicação, e o direito de existir nele".

[73] Catarina Sarmento e Castro, "40 anos de "utilização da informática", cit., p. 53.

Não obstante as vantagens de uma especificação dos direitos humanos e fundamentais no contexto digital, são também aqui pertinentes as advertências no sentido de que mais catálogos de direitos não significam necessariamente mais proteção, dadas as inconsistências e dúvidas de interpretação e aplicação que muitas vezes resultam da multiplicação de instrumentos de proteção e, consequentemente, insegurança jurídica. Discussão essa que, entre nós, teve lugar com a aprovação da *Carta Portuguesa dos Direitos Humanos na Era Digital* pela Lei nº 27/2021, de 17 de maio, da Assembleia da República[74].

Finalmente, à complexidade ou multidimensionalidade da proteção dos "direitos digitais" corresponderá a diversidade de deveres e tarefas estaduais que daí resultam também para os poderes públicos, desde obrigações negativas de respeito a deveres positivos de regulação e definição de políticas públicas capazes de tornar efetivo o acesso ao ambiente digital em condições de igualdade e no respeito pelos direitos fundamentais, bem como deveres de proteção. Deveres que recaem sobre o Estado legislador, mas também sobre o Estado administrador ou, mais amplamente, sobre os poderes públicos estaduais e supraestaduais no desempenho de funções legislativas ou administrativas[75], sendo neste contexto relevante a chamada "reserva do possível", aqui também "reserva do tecnologicamente possível"[76]. A este propósito, e ainda que este texto não se ocupe especificamente dessa dimensão, cumpre notar o impacto da digitalização no desempenho das tarefas estaduais em

[74] José de Melo Alexandrino, escrevendo face à versão originária da lei, sustenta que a Carta é "redundante, quer relativamente aos direitos fundamentais constitucionalmente consagrados, quer relativamente aos direitos humanos reconhecidos internacionalmente, quer relativamente aos direitos humanos fundamentais reconhecidos no âmbito da União Europeia, acrescentando assim insegurança jurídica" e que colide em diversos pontos com o Direito da UE e a jurisprudência do TJUE, vd. José de Melo Alexandrino, *Dez breves apontamentos* ..., cit., p. 2. Disponível online em https://www.icjp.pt/sites/default/files/papers/dez_breves_apontamentos_sobre_a_carta_portuguesa.pdf [04/11/2022]. Em sentido idêntico, Domingos Soares Farinho, "The Portuguese Charter of Human Rights...", cit.

[75] O artigo 3º da *Carta Portuguesa de DH na era digital*, cit., é a esse nível exemplificativo, estabelecendo um conjunto de tarefas que vão desde promover o uso autónomo e responsável da internet e o livre aceso às tecnologias ao dever de criação de condições efetivas de acesso ao ambiente digital, com uma dimensão positiva, como tarifas sociais de acesso, pontos de acesso gratuito e garantia de conectividade de qualidade a preço acessível, ou desenvolvimento de ações de capacitação para o digital.

[76] Sobre este conceito, Carla Amado Gomes, "Estado Social e concretização de Direitos Fundamentais na era tecnológica", *in Revista da FDUP,* 7, 2010, pp. 19 ss. e Luísa Neto, "O princípio da proteção da confiança em tempo de crise", *in Direito Administrativo,* CEJ, 2014, pp. 77 ss., p. 83, em nota, disponível em https://cej.justica.gov.pt/

geral[77] e administrativas em especial[78], com ganhos de celeridade e eficiência, mas igualmente com os riscos de opacidade, segurança e discriminação já identificados[79].

2.3. Igualdade e vulnerabilidade(s)

Um outro eixo relevante de análise da relação entre transformação digital e direitos fundamentais é o princípio da igualdade e a proteção contra a discriminação, que atualmente recebem grande atenção[80].

Às formas de discriminação já identificadas[81] juntam-se novas problemáticas como a discriminação por algoritmos, para a qual alertou, designadamente, a Agência de Direitos Fundamentais da União Europeia no seu relatório *Preparar o Futuro. Inteligência Artificial e Direitos Fundamentais* (2021)[82]. À partida, os algoritmos são neutros[83]. Contudo, como faz notar a Agência de

[77] Entre nós e entre diversos exemplos, *vd. Plano de Ação para a Transição Digital,* aprovado através da Resolução do Conselho de Ministros nº 30/2020 de 21 de abril, em especial o *Pilar III* – Digitalização do Estado. Por exemplo, discutindo o impacto da IA na função jurisdicional e a possibilidade de um juiz-máquina ou, mais especificamente, de um "juiz IA", *vd.* Andrew C. Michaels, "Artificial Intelligence, Legal Change, and Separation of Powers", in *University of Cincinnati Law Review*, 88, 2020, pp. 1083 ss., disponível em https://scholarship.law.uc.edu/uclr/vol88/iss4/4

[78] Entre nós e entre diversos exemplos, *vd. Plano de Ação para a Transição Digital,* aprovado através da Resolução do Conselho de Ministros nº 30/2020 de 21 de abril, em especial o *Pilar III* – Digitalização do Estado.

[79] Por exemplo, Oreste Pollicino e Giovanni De Gregorio, "Constitutional Law in the Algorithmic Society", cit., p. 6, notam que os modelos preditivos podem auxiliar as administrações públicas a fornecer serviços públicos mais eficientes e a poupar recursos, mas também podem favorecer decisões discriminatórias. Já Aziz Z. Huq nota que, mesmo que a IA e o "*machine learning*" sejam benéficas do ponto de vista do desempenho das funções estaduais e da realização dos direitos fundamentais, é plausível recear que ao aumento dos poderes estaduais corresponda uma capacidade limitada dos sujeitos regulados – aqui incluído o público em geral mas também os operadores do sistema jurídico – para perceber e desafiar o exercício de tal poder, sendo expectável que a adoção pelas entidades públicas de ferramentas de previsão e inferência aumentem a dificuldade de os cidadãos monitorizarem e responderem a essas atividades, à medida que o seu âmbito também se alarga, *vd.* Aziz Z. Huq, "Constitutional rights in the machine-learning state", *in Cornell Law Review,* vol. 105, 2020, pp.1875 ss., *maxime* p. 1883.

[80] Bart Custers, "New digital rights", cit., p. 3.

[81] Para uma visão geral, *vd.* Fernando Rey Martinez e Luísa Neto (org.), *Direito Antidiscriminatório,* Lisboa: AAFDL, 2021.

[82] Fundamental Rights Agency, Relatório *Preparar o Futuro. Inteligência Artificial e Direitos Fundamentais. Síntese,* 2021. Disponível em https://fra.europa.eu/pt/publication/2021/preparar-o-futuro-inteligencia-artificial-e-direitos-fundamentais-sintese

[83] Oreste Pollicino e Giovanni De Gregorio, "Constitutional Law in the Algorithmic Societ", cit., p. 4.

Direitos Fundamentais da União[84], o recurso a algoritmos pode servir para veicular estereótipos ou intensificar estereótipos já existentes, e tal pode ocorrer devido a falhas na conceção do algoritmo ou, já depois disso, em resultado da aprendizagem realizada pela máquina. Por isso a questão da conceção ou do desenho do algoritmo, e da supervisão humana destes processos, é central[85].

Da mesma forma que a digitalização promove novas formas de inclusão e acesso, aprofunda ou cria também dimensões de exclusão[86]. São disso exemplo as questões de (des)igualdade – por razões económicas, geográficas, de género[87], idade[88] ou outras – no acesso aos recursos que tornam possível a transição digital e no acesso aos serviços digitais (especialmente quando certos serviços deixarem de ser realizados senão de forma digital), e também as desigualdades consequentes no exercício de direitos fundamentais, atendendo à função instrumental desempenhada pelas tecnologias digitais. Estes fatores podem surgir associados, gerando discriminações múltiplas ou interseccionais[89].

Tomando como exemplo o sucedido na educação e ensino durante a pandemia, verificamos que se o acesso às tecnologias permitiu a continuação da

[84] Fundamental Rights Agency, Relatório *Preparar o Futuro...*, cit.

[85] Como se escreveu no *Livro Branco,* cit., p. 12-13*:* "Os preconceitos e a discriminação são riscos inerentes a qualquer atividade económica ou societal. A tomada de decisões humanas não é imune a erros e preconceitos. No entanto, o mesmo preconceito no âmbito da IA pode ter um efeito muito maior, afetando e discriminando muitas pessoas sem os mecanismos de controlo social que regem o comportamento humano. Tal pode também acontecer quando o sistema de IA "aprende" durante o seu funcionamento. Nesses casos, em que o resultado não poderia ter sido evitado ou previsto na fase de conceção, os riscos não resultarão de uma falha na conceção original do sistema, mas sim do impacto prático das correlações ou padrões que o sistema identifica num grande conjunto de dados. (...) Embora a legislação da UE continue, em princípio, a ser plenamente aplicável independentemente do envolvimento da IA, é importante avaliar se pode ser aplicada adequadamente para fazer face aos riscos que os sistemas de IA criam ou se é necessário ajustar determinados instrumentos jurídicos."

[86] Sobre estas, *Roadmap for digital cooperation,* cit., *passim,* pp. 10 ss.

[87] Entre outros, veja-se Argelia Queralt Jiménez, "El mundo digital: un nuevo ámbito de discriminación entre hombres y mujeres", *IberICONnect,* 18/02/2022. Disponível em: https://www.ibericonnect.blog/2022/02/el-mundo-digital-un-nuevo-ambito-de-discriminacion-entre-hombres-y-mujeres/ [21/10/2022], que identifica como manifestações de discriminação entre homens e mulheres no ambiente digital a desigualdade no acesso e uso de novos instrumentos digitais e tecnológicos, especialmente a internet, o diferente impacto das vozes femininas nas redes e a violência contra as mulheres através do digital. Vejam-se ainda os dados no *Road map for digital cooperation,* cit., p.10 ss.

[88] A este propósito, veja-se o "caso Carlos San Juan", ocorrido em Espanha em 2022, e descrito em https://www.nytimes.com/es/2022/03/25/espanol/espana-carlos-san-juan-bancos.html

[89] Sobre estas, *Roadmap for digital cooperation,* cit., *passim,* pp. 10 ss.

prestação educativa a crianças confinadas, expôs também problemas de igualdade de acesso aos recursos (*e.g. Internet*) e de capacidade efetiva para a utilização do digital dos atores educativos[90], expondo desafios de concretização do princípio da igualdade no contexto digital.

Como resulta do *Roteiro para a Cooperação Digital da ONU*[91], a inclusão digital implica a definição de condições que permitam a toda e cada pessoa participar no mundo digital, o que inclui não apenas a conectividade e o acesso a equipamentos, mas também a acessibilidade em termos financeiros e a capacitação para a utilização de ferramentas digitais. O *Roteiro* recomenda ainda que, assente no pressuposto de que todos devem ter iguais oportunidades de empoderamento através das tecnologias de informação e comunicação (*ICT*), a promoção da acessibilidade efetiva se faça através de medidas inclusivas (*e.g. design* inclusivo) e leve em conta fatores como a interseccionalidade e as barreiras linguísticas e estruturais existentes[92].

Especial referência é devida às pessoas e grupos vulneráveis[93]. A necessidade de proteção de pessoas e grupos vulneráveis perante a transição digital, seja no acesso ao ambiente digital, seja na experiência e participação no ambiente digital, vem sendo acentuada, designadamente a proteção de crianças e os jovens ou dos portadores de deficiência (Cap. V da *Declaração Europeia sobre os direitos e princípios digitais para a década digital* (2022)[94]. Acresce que o recurso a meios de IA, veiculando ou reforçando estereótipos, pode reforçar vulnerabilidades já existentes[95].

2.4. Da complexidade à necessidade da regulação constitucional

A velocidade do desenvolvimento das tecnologias, por um lado, e a incerteza e riscos associados, desde logo para os direitos fundamentais e para os valores democráticos, por outro, tornam a regulação dos avanços tecnológicos, equilibrando inovação e riscos, difícil e complexa[96]. Ultrapassado o óbice

[90] *Vd.,* entre muitos contributos, mas refletindo especificamente sobre a situação portuguesa, Provedor de Justiça, *Cadernos da Pandemia. Educação em tempos de emergência,* 2021, disponível online em https://www.provedor-jus.pt/documentos/educ_cadernos_pandemia_2021_web.pdf

[91] *Roadmap for digital cooperation,* cit., *passim,* em especial pp. 5 ss.

[92] *Roadmap for digital cooperation,* cit., *passim,* em especial pp. 10-11.

[93] Sobre o conceito, com referências várias, veja-se Anabela Costa Leão, "O Estado perante a vulnerabilidade", *in Oñati Socio-Legal Series,* 2022. Disponível em: https://opo.iisj.net/index.php/osls/article/view/1326

[94] Vejam-se ainda os Considerandos 16 e 17 da *Proposta de Regulamento Inteligência Artificial*, cit.

[95] Veja-se ainda o Considerando 17 da *Proposta de Regulamento Inteligência Artificial*, cit.

[96] Assim, Oreste Pollicino e Giovanni De Gregorio, "Constitutional Law in the Algorithmic Societ", cit., pp. 5 e 9 ss. Os autores, ob. cit., p. 5, escrevem que deixar sem garantias

da impossibilidade de regulação estadual (e constitucional) do digital e da IA[97], subsiste a questão de saber qual o melhor tipo de regulação[98], ou seja, de determinar *quem* deve regular e *como* pode regular (*e.g.* heteroregulação, autoregulação e/ou corregulação, recurso a *hard* e/ou *soft law*).

Para Dimitar Lilkov, a nova *Proposta de Regulamento Inteligência Artificial* parece traduzir precisamente um afastamento, por parte da União Europeia, da abordagem *soft* que vem sendo utilizada e uma tomada de posição no sentido de que não basta a afirmação de princípios éticos ou mecanismos de autorregulação, sendo necessário o estabelecimento de um quadro normativo vinculativo capaz de fornecer segurança jurídica acautelando direitos fundamentais[99].

De forma particular, atendendo ao objeto da análise empreendida neste texto, trata-se de determinar *quem* e *como* assegura a função de proteção normativa dos direitos fundamentais no espaço digital e se e em que medida se pode alargar a regulação constitucional a estas novas realidades, sendo a via da regulação "constitucional" ensaiada por vários caminhos[100] que apelam a

nem controlos democráticos a regulação das tecnologias algorítmicas pode levar as sociedades ao tecno-determinismo e à marginalização dos atores públicos enquanto garantes dos direitos fundamentais e dos valores democráticos. Para os autores, a tecnologia deve, não comandar, mas estar ao serviço da humanidade, sob pena de um enfraquecimento gradual dos valores democráticos constitucionais em nome da inovação.

97 Se inicialmente a Internet e a transição digital foram vistas como um poderoso desafio à soberania estadual, a regulação da Internet pelos Estados, dos democráticos aos autoritários, é hoje uma realidade, dispondo os Estados de diferentes vias e mecanismos para o fazer, vd. Oreste Pollicino e Giovanni De Gregorio, "Constitutional Law in the Algorithmic Society", cit., pp. 10 ss. e Julia Pohle e Thorsten Thiel, T. "Digital sovereignty", *in Internet Policy Review*, 9, 4, 2020, DOI: 10.14763/2020.4.1532.

98 Sobre a questão, Oreste Pollicino e Giovanni De Gregorio, "Constitutional Law in the Algorithmic Society", cit., pp. 12 ss. e, também, Edoardo Celeste, "Digital constitutionalism...", cit., pp. 78 ss. Pollicino e De Gregorio notam que, se o *soft law* pode parecer a via mais adequada para favorecer o desenvolvimento tecnológico, pode nalguns casos debilitar a proteção de direitos fundamentais, e que se as democracias constitucionais tendem a preferir formas de regulação liberais, dando primazia a mecanismo de autorregulação, nalguns domínios será desejável o desenvolvimento de modelos de corregulação, ob. cit., pp. 12 ss. Sobre estes, veja-se também o Pedido de Fiscalização da constitucionalidade das normas constantes dos números 5 e 6 do artigo 6º, da Lei nº 27/2021, de 17 de maio, que aprova a Carta Portuguesa de Direitos Humanos na Era Digital, Provedoria de Justiça, 18 de maio de 2022, disponível em https://www.provedor-jus.pt/documentos/fiscalizacao-da-constitucionalidade-das-normas-constantes-dos-numeros-5-e-6-do-artigo-6-o-da-lei-n-o-27-2021-de-17-de-maio-que-aprova-a-carta-portuguesa-de-direitos-humanos-na-era-digital/ [03/11/2022].

99 Dimitar Lilkov, "Regulating artificial intelligence in the EU...", cit., pp. 166 ss.

100 Por todos, *vd.* Edoardo Celeste, "Digital constitutionalism...", cit., pp. 82 ss.

conceções mais ou menos amplas de constitucionalismo e da relação entre "constituição" e "Estado" [101].

A perspetiva constitucional, na aceção vista *supra,* exige a proteção dos direitos fundamentais e a limitação do poder. Se é certo que as constituições têm por função (e garantia de legitimidade) defender os indivíduos face ao poder, é necessário compreender de que "poder" se trata. No ambiente digital, a par dos atores públicos surgem relevantes atores privados e as necessidades de proteção de direitos fazem-se sentir, não só – e já não, porventura, sobretudo – face ao Estado, mas também face a entidades privadas, designadamente aos grandes operadores de dados[102], "poderes privados" cuja subtração a limites constitucionais se afigura problemática[103].

Pode, nesta sede, aludir-se à possibilidade de irradiação das normas constitucionais sobre o Direito Privado. Cumpre aqui referir as chamadas teorias da vinculação dos particulares aos direitos fundamentais nas relações entre si[104] para obrigar estes atores privados a respeitar os direitos fundamentais

[101] Pode entender-se que se trata de questões que relevam da autonomia privada e da liberdade económica e que, por conseguinte, terão no direito privado, designadamente no direito comercial e no direito da concorrência, a sua sede própria de regulação, vd. Oreste Pollicino e Giovanni De Gregorio, "Constitutional Law in the Algorithmic Society", cit., pp. 14 ss. Sobre a questão de caber ao direito privado uma função constitucional, Edoardo Celeste, "Digital constitutionalism...", cit., p. 82. Contudo, a via da regulação constitucional permite pôr a enfase na proteção das pessoas, através dos direitos fundamentais e da limitação do poder, e na salvaguarda de valores e princípios fundamentais, vd. Oreste Pollicino e Giovanni De Gregorio, "Constitutional Law in the Algorithmic Society", cit., pp. 14 ss.; Edoardo Celeste, "Digital constitutionalism...", cit., pp. 78 ss.

[102] Como se lê em Council of Europe, *Algorithms and Human Rights,* cit., p. 33, "*(...) the increasing use of automation and algorithmic decision-making in all spheres of public and private life is threatening to disrupt the very concept of human rights as protective shields against state interference. The traditional asymmetry of power and information between state structures and human beings is shifting towards an asymmetry of power and information between operators of algorithms (who may be public or private) and those who are acted upon and governed".*

[103] Vd. por exemplo, Oreste Pollicino e Giovanni De Gregorio, "Constitutional Law in the Algorithmic Society", cit., p. 14 ss. Os autores notam, a pp. 16, que deixar sem regulação pública a atividade de plataformas online corresponde a consolidar áreas de poder privado (uma espécie de ambiente para-legal competindo com as autoridades públicas) nas quais os cidadãos não estão representados na definição das normas nem na tomada das decisões, o que é problemático do ponto de vista do princípio democrático.

[104] Aludimos à problemática da "eficácia horizontal", "eficácia ou validade dos direitos fundamentais nas relações entre particulares" ou "vinculação das entidades privadas aos direitos fundamentais" que tem expressão no nº 1 do artigo 18º da CRP. Esta questão tem sido abordada na teoria sob diversas perspetivas, mais ou menos tendentes à compatibilização entre direitos fundamentais e autonomia privada. Sobre estas teorias, entre nós, José Carlos Vieira de Andrade, *Os direitos*

dos utilizadores, podendo, contudo, eles próprios invocar direitos frequentemente protegidos como fundamentais, *e.g.* a autonomia privada e a liberdade de iniciativa económica[105], solucionáveis por recurso à dogmática de resolução de conflitos de direitos fundamentais, a que já se aludiu. A existência de relações – como a que se estabelece entre um utilizador e uma plataforma online –marcadas por uma "desigualdade posicional"[106] não pode deixar de ser tida em conta em sede de regulação e de harmonização dos direitos em conflito[107]. Pode, de outra perspetiva, sustentar-se a aplicação a atores privados de princípios constitucionais fundamentais, como o princípio do Estado de Direito (*rule of law*)[108].

fundamentais..., cit., pp. 228 ss. Oreste Pollicino e Giovanni De Gregorio, "Constitutional Law in the Algorithmic Society", cit., *maxime* pp. 14 ss.

[105] A este propósito, veja-se a *Proposta de Regulamento sobre Inteligência Artificial* apresentada em 2021, na qual se pode ler "A presente proposta impõe algumas restrições à liberdade de empresa (artigo 16º) e à liberdade das artes e das ciências (artigo 13º), a fim de assegurar o cumprimento de razões imperativas de reconhecido interesse público, como a saúde, a segurança, a defesa dos consumidores e a proteção de outros direitos fundamentais ("inovação responsável") em caso de desenvolvimento e utilização de tecnologia de IA de risco elevado. Essas restrições são proporcionadas e limitadas ao mínimo necessário para prevenir e atenuar riscos de segurança graves e possíveis violações dos direitos fundamentais.

[106] Em geral, sobre a relevância de se estar ou não perante sujeitos privados em situação de poder e sobre as chamadas "relações privadas de poder", José Carlos Vieira de Andrade, *Os direitos fundamentais...,* cit., p. 228 ss., *maxime* pp. 240 ss.

[107] Contudo, como notam Pollicino e De Gregorio (seguindo Balkin), não se trata apenas da relação entre Estado e plataformas, ou entre plataformas e indivíduos, mas de uma relação tripolar entre Estado, plataformas e indivíduos, pelo que, para os autores, será necessário ensaiar uma abordagem alternativa à discussão sobre a eficácia vertical ou horizontal dos direitos fundamentais que passa pelo reconhecimento de novos direitos, vd. Oreste Pollicino e Giovanni De Gregorio, "Constitutional Law in the Algorithmic Society", cit., pp. 20 ss.

[108] Discutindo e aplicando ao contexto digital, Edoardo Celeste, "Digital constitutionalism...", cit., pp. 82 ss. Veja-se ainda Nicolas Suzor, "Digital Constitutionalism: Using the Rule of Law to Evaluate the Legitimacy of Governance by Platforms", in *Social Media + Society*, *4*(3), 2018. Para o autor, *op. cit.*, p. 2: *"The rule of law framework provides a lens through which to evaluate the legitimacy of online governance and therefore to begin to articulate what limits societies should impose on the autonomy of platforms. For the governance of platforms to be legitimate according to rule of law values, we should expect certain basic procedural safeguards. First, decisions must be made according to a set of rules, and not in a way that is arbitrary or capricious. Second, these rules must be clear, well-understood, and relatively stable, and they must be applied equally and consistently. Third, there must be adequate due process safeguards, including an explanation of why a particular decision was made and some form of an appeals process that allows for the independent review and fair resolution of disputes. These are the fundamental minimum procedural standards for a system of governance to be legitimate, and platforms currently perform very poorly on these measures. I argue that the extent of influence that major platforms and other digital intermediaries have over social life implies that we should seek to hold them to account against these values."*

Pollicino e De Gregorio[109] sustentam a necessidade de uma regulação constitucional assente na defesa dos direitos fundamentais. Sustentam a necessidade de reconhecimento e proteção de um novo conjunto de direitos, não apenas face aos poderes públicos, mas também face a particulares, destinados a reforçar a transparência e a responsabilização ("*accountability*") das entidades privadas ("*a digital habeas corpus of substantive and procedural rights*" [110]). Para os autores, assistimos a uma mudança de paradigma (e a um novo "*pactum subjectionis*" do mundo digital[111]) que exige o reconhecimento de novos direitos substanciais (como o direito a que as decisões que nos dizem respeito sejam tomadas por pessoas e não apenas por máquinas ou o direito à explicação de decisões automatizadas[112]) mas também garantias de participação e defesa e procedimentos justos que permitam aos particulares fazer valer os seus direitos também perante entidades privadas, como contrapartida dos ganhos obtidos pelas plataformas e como forma de combater a opacidade e a fragilidade a que aquelas sujeitam os indivíduos[113].

Mais amplamente, para Edoardo Celeste (v. *supra*) está em curso uma "constitucionalização do ambiente digital" [114]. O constitucionalismo digital – que não se circunscreve ao constitucionalismo estadual – fornece o conjunto de valores e ideias que permeia e guia o processo de progressiva constituciona-

[109] Oreste Pollicino e Giovanni De Gregorio, "Constitutional Law in the Algorithmic Society", cit., pp. 20 ss.

[110] Oreste Pollicino e Giovanni De Gregorio, "Constitutional Law in the Algorithmic Society", cit., p. 20.

[111] Oreste Pollicino e Giovanni De Gregorio, "Constitutional Law in the Algorithmic Society", cit., p. 22.

[112] Para os autores, Oreste Pollicino e Giovanni De Gregorio, "Constitutional Law in the Algorithmic Societ", cit., p. 21, "*The right to explanation is only just one of the new rights that could contribute to mitigating the lack of fairness, transparency, and accountability in automated decision-making. Indeed, together with the right to obtain information on the way their data are being processed, individuals should also rely on a right to easy access (right to accessibility) and on a right to obtain translation from the language of technology to the language of human beings. While the former is meant as the right to be provided with the possibility to interact with algorithms and digital platforms implementing the use thereof, the latter requires the use of simple, clear, and understandable information and allows users not only to rely on, for example, the reasons for the removal of online content, but also to better exercise their rights before a judicial or administrative body*".

[113] Assim, Oreste Pollicino e Giovanni De Gregorio, "Constitutional Law in the Algorithmic Societ", cit., p. 20 ss. Para os autores, op. cit., p. 23, esta entendimento está refletido na Proposta de Regulamento do Parlamento Europeu e do Conselho relativo à disputabilidade e equidade dos mercados no setor digital (Regulamento Mercados Digitais) COM (2020) 842 final (entretanto, o Regulamento Mercados Digitais entrou em vigor em 1 de novembro de 2022).

[114] Aludindo a um processo de generalização e re-especificação, Edoardo Celeste, "Digital constitutionalism...", cit., p. 89.

lização em curso. Este está a decorrer de forma não homogénea, antes se traduzindo na emergência de diferentes tipos de "respostas constitucionais"[115]. Estas respostas de tipo constitucional englobam as respostas produzidas no plano nacional, *e.g.* através de normas constitucionais ou de decisões de tribunais supremos ou constitucionais, no plano transnacional "estadocêntrico", *e.g.* através de normas ou de decisões jurisdicionais adotadas no âmbito de organizações internacionais como a União Europeia ou Conselho da Europa, e no plano transnacional não referido ao Estado e, por conseguinte, sem a vinculatividade jurídica que caracteriza as respostas das categorias anteriores[116]. Aqui se incluem, para o autor, as cartas de direitos da Internet[117], as decisões de mecanismos de resolução de litígios relacionados com o ambiente digital e as normas adotadas pelos operadores comerciais como termos de uso ou de serviço ou normas corporativas obrigatórias. Para Edoardo Celeste, estas respostas, incluindo de fonte privada, merecem o qualificativo de "constitucionais" porque orientadas para a amplificação ou para a proteção de direitos fundamentais ou para o reequilíbrio de poderes em ambiente digital[118].

3. Conclusão

As tecnologias digitais e a IA podem simultaneamente ter efeitos positivos e negativos na proteção e efetivação de direitos fundamentais e, em última análise, na proteção da dignidade da pessoa humana, permitindo novas formas de inclusão, mas também de exclusão.

Sendo os direitos fundamentais limites ao desenvolvimento e ao aproveitamento tecnológico, a adoção de uma abordagem de direitos fundamentais exige repensar o âmbito e as exigências de proteção de direitos já existentes, mas também afirmar e proteger novos direitos, tendo presente que as ameaças aos direitos fundamentais provêm quer de entidades públicas quer de entidades privadas, designadamente de atores privados poderosos como plataformas digitais.

A regulação da esfera digital e da atividade destes atores privados é reclamada partindo de diversas perspetivas que utilizam a "linguagem constitucional" de forma mais ou menos ampla para articular diferentes instâncias e diferentes níveis de proteção das pessoas e dos seus direitos, apelando a uma expansão do campo constitucional.

[115] Edoardo Celeste, "Digital constitutionalism...", cit., p. 87 ss.

[116] Edoardo Celeste, "Digital constitutionalism...", cit., p. 90 ss.

[117] Como a *Charter of Human Rights and Principles for the Internet* de 2011 da IRPC (Internet Rights and Principles Coalition), disponível em https://internetrightsandprinciples.org/charter/

[118] Edoardo Celeste, "Digital constitutionalism...", cit., p. 87 ss.

Por isso, também neste domínio se fazem sentir as exigências de aprofundamento de uma regulação constitucional global, com as promessas e as limitações que sempre comporta.

Bibliografia

Alexandrino, José de Melo, *O discurso dos direitos,* Coimbra: Coimbra Editora/Wolters Kluwer, 2011.

Alexandrino, José de Melo, *Dez breves apontamentos sobre a Carta Portuguesa de Diretos Humanos na Era Digital*, ICJP/CIDP, 2021. Disponível online em https://www.icjp.pt/sites/default/files/papers/dez_breves_apontamentos_sobre_a_carta_portuguesa.pdf

Almendra Freitas, Cinthia Obladen, "Neurodireitos: O exemplo do Chile e a regulação das neurotecnologias", in *IberICONnect,* 08/02/2022. Disponível em: https://www.ibericonnect.blog/2022/02/neurodireitos-o-exemplo-do-chile-e-a-regulacao--das-neurotecnologias/ [04/11/2022].

Andrade, José Carlos Vieira de, Os *direitos fundamentais na Constituição portuguesa de 1976,* Coimbra: Almedina, 2021.

Castro, Catarina Sarmento e, "40 anos de "utilização da informática" – o art. 35º da CRP", *in e-publica*, 3, 2016.

Castro, Raquel Brízida de Castro, "Constituição e ciberespaço: argumentos para um 'direito constitucional do inimigo'?", *in Ciberlaw,* 1, 2016.

Custers, Bart, "New digital rights: Imagining additional fundamental rights for the digital era", *in Computer Law & Security Review,* 44, 2022, 105636, https://doi.org/10.1016/j.clsr.2021.105636.

Farinho, Domingos Soares, "The Portuguese Charter of Human Rights in the Digital Age: a legal appraisal", *in Revista Española de la Transparencia*, n. º 13, 2º semestre jul--dec. 2021, pp. 85-105. DOI: https://doi.org/10.51915/ret.191

Gomes, Carla Amado, "Estado Social e concretização de Direitos Fundamentais na era tecnológica", *in Revista da FDUP,* VII, 2010, pp. 19 ss.

Huq, Aziz Z., "Constitutional rights in the machine-learning state", *in Cornell Law Review,* vol. 105, 2020, p.1875 ss.

Leão, Anabela Costa, "A Carta dos Direitos Fundamentais da União Europeia", *in RFDUP,* Ano III, 2006, p. 41 ss.

Leão, Anabela Costa, "O contributo do Tribunal Europeu dos Direitos do Homem" *in* Luísa Neto e Rute Teixeira Pedro, *Debatendo a procriação medicamente assistida,* Porto, CIJE--FDUP, 2017, pp. 23-40, online em https://www.up.pt/press/books/978-989-746-154-5

Leão, Anabela Costa, "O Princípio da Proporcionalidade e os seus críticos", *in* AA. VV., *O Princípio da Proporcionalidade. XIII Encontro de Professores de Direito Público,* Coimbra: Instituto Jurídico, 2021, pp. 127 ss., www.doi.org/10.47907/clq2021_2a7

Lilkov, Dimitar, "Regulating artificial intelligence in the EU: A risky game", *in European View,* 20(2), 2021, pp. 166–174. DOI: https://doi.org/10.1177/17816858211059248

Michaels, Andrew C., "Artificial Intelligence, Legal Change, and Separation of Powers", in *University of Cincinnati Law Review*, 88, 2020, pp. 1083 ss., disponível em https://scholarship.law.uc.edu/uclr/vol88/iss4/4

MOWBRAY, ALASDAIR, "Between the will of the Contracting Parties and the needs of today: Extending the scope of Convention rights and freedoms beyond what could have been foreseen by the drafters of the ECHR", in Eva Brems e Janeke Gerards (Eds.), *Shaping Rights in the ECHR: The Role of the European Court of Human Rights in Determining the Scope of Human Rights*, Cambridge: Cambridge University Press, 2014, pp. 17-37. DOI:10.1017/CBO9781107337923.003

NABAIS, JOSÉ CASALTA, "Algumas reflexões críticas sobre os direitos fundamentais", in *Por uma liberdade com responsabilidade*, Coimbra: Coimbra Editora, 2007, p. 103 ss.

NETO, LUÍSA, *Novos direitos ou novo(s) objetos para o Direito?* U.Porto, 2010.

NETO, LUÍSA, "O princípio da proteção da confiança em tempo de crise", *in Direito Administrativo,* CEJ, 2014, pp. 77 ss., disponível em https://cej.justica.gov.pt/

PECES-BARBA MARTINEZ, GREGORIO, *Curso de Derechos Fundamentales. Teoria General,* Madrid: Universidad Carlos III, 1999.

POHLE, JULIA E THIEL, THORSTEN, "Digital sovereignty", *in Internet Policy Review,* 9, 4, 2020. DOI: https://doi.org/10.14763/2020.4.1532

POLLICINO, ORESTE E DE GREGORIO, GIOVANNI, "Constitutional Law in the Algorithmic Society", *in* H. Micklitz, O. Pollicino, A. Reichman, A. Simoncini, G. Sartor e G. De Gregorio (Eds.), *Constitutional Challenges in the Algorithmic Society*, Cambridge: Cambridge University Press, 2021, pp. 3-24. DOI:10.1017/9781108914857.002

QUEIROZ, CRISTINA, *Direitos Fundamentais,* Coimbra: Coimbra Editora/ Wolters Kluwer, 2010.

QUERALT JIMÉNEZ, ARGELIA, "El mundo digital: un nuevo ámbito de discriminación entre hombres y mujeres", *IberICONnect,* 18/02/2022. Disponível em: https://www.ibericonnect.blog/2022/02/el-mundo-digital-un-nuevo-ambito-de-discriminacion--entre-hombres-y-mujeres/ [21/10/2022].

REGLITZ, MERTEN, "The human right to free internet access", *in Journal of Applied Philosophy*, 37, 2, 2020, p. 314 ss.

REY MARTINEZ, FERNANDO E NETO, LUÍSA (org.), *Direito Antidiscriminatório,* Lisboa: AAFDL, 2021.

SILVA, SUZANA TAVARES DA, *Direitos Fundamentais na Arena Global,* Coimbra: Imprensa Universidade Coimbra, 2011.

SUZOR, NICOLAS, "Digital Constitutionalism: Using the Rule of Law to Evaluate the Legitimacy of Governance by Platforms", in *Social Media + Society*, *4* (3), 2018. https://doi.org/10.1177/2056305118787812

A Estratégia Europeia para a Inteligência Artificial

The European Artificial Intelligence Strategy

Graça Enes*

resumo: A Inteligência Artificial está a revolucionar a indústria, a economia e a sociedade. Os benefícios potenciais são imensos. Os riscos para os direitos fundamentais e para os princípios de um estado-de-direito democrático são inquestionáveis. A União Europeia pretende superar o atraso em relação aos protagonistas norte-americanos e chineses e alcançar a dianteira daquela revolução, assegurando que ela se desenvolve no respeito pelos valores, princípios e direitos fundamentais europeus. Com esse objetivo a Comissão Europeia avançou com a «Estratégia Europeia para a Inteligência Artificial», uma iniciativa política formada por vários instrumentos: programas de I&D, ações de coordenação *multistakeholder* e propostas de regulação normativa da IA, em especial o «Regulamento IA» e a «Diretiva responsabilidade IA», com o objetivo de assegurar uma IA de excelência e confiança, com respeito dos valores, princípios e direitos fundamentais da UE. É uma iniciativa ambiciosa, abrangente, que permitirá à EU posicionar-se na dianteira global. Porém, não está isenta de críticas e insuficiências. Depois de abordar a complexidade inerente à definição da IA (2), percorre-se a elaboração da política europeia para a IA (3) e as propostas de regulação (4).

palavras-chave: União Europeia, Inteligência Artificial (IA), «*risk approach*», harmonização legislativa, «Regulamento IA», «Diretiva responsabilidade IA»

* Investigadora do CIJ. Faculdade de Direito da Universidade do Porto. gef@direito.up.pt.
Este texto é resultado do desenvolvimento do projeto do Módulo Jean Monnet DigEUCit, "A Digital Europe for Citizens. Constitutional and Polcymaking Challenges".
0000-0002-9461-5705 Ciência ID – A614-3C58-C3F5

ABSTRACT: Artificial Intelligence is revolutionizing industry, economy and society. The potential benefits are immense. The risks to fundamental rights and the principles of a democratic rule of law are unquestionable. The European Union intends to overcome the delay in relation to the North American and Chinese protagonists and to reach the forefront of that revolution, ensuring that it develops with respect for European values, principles and fundamental rights. With this objective in mind, the European Commission put forward the «European Strategy for Artificial Intelligence», a political initiative made up of several instruments: R&D programs, multistakeholder coordination actions and proposals for normative regulation of AI, in particular the «AI Regulation» and the «AI Liability Directive». It is an ambitious, comprehensive initiative that will allow the EU to position itself at the global forefront. However, it is not without criticism and shortcomings. After addressing the complexity inherent in the definition of AI (2), the elaboration of the European policy for AI (3) and the proposals for regulation (4) are covered.

KEYWORDS: European Union, Artificial Intelligence (AI), «risk approach», legislative harmonization, «AI Regulation», «AI Liability Directive»

"... a ideia de que a actividade mental, dos seus aspectos mais simples aos mais sublimes, requer um cérebro e um corpo propriamente dito tornou-se notoriamente inescapável."

ANTÓNIO R. DAMÁSIO, *O Erro de Descartes. Emoção, Razão e Cérebro Humano*, 11ª ed., Mem Martins, Publicações Europa-América, p. 18

"*The development of full artificial intelligence could spell the end of the human race.*"

STEPHEN HAWKING, *BBC interview*, 2.12.2014, *in* https://www.bbc.com/news/technology-30290540 (31.10.2022)

1. Introdução

Se a inteligência é um conceito que continua envolto em ambiguidade, o conceito de inteligência artificial (de ora em diante, também IA)[1] e até a sua existência continuam disputadas[2]. Não obstante, sistemas e mecanismos dotados de capacidades que mimetizam capacidades de aprendizagem e decisão da mente humana, com vantagens acrescidas de eficiência e rapidez, são inegáveis e estão cada vez mais disseminados em todas as áreas da indústria e serviços, privados e públicos, e na sociedade em geral. Uma revolução tecnológica, económica, social e cultural está em curso[3], com um potencial impacto semelhante à energia do vapor ou da eletricidade, já designada de 4ª revolução industrial[4].

As oportunidades são numerosas e os benefícios finais são auspiciosos: menos poluição, menos mortes na estrada, melhores cuidados médicos, maior acessibilidade para as pessoas com deficiência e os idosos, melhor educação, mais formas de levar os cidadãos a participarem nos processos democráticos, adjudicações mais rápidas, maior eficácia na luta contra o terrorismo e a criminalidade, tanto em linha como fora de linha, o reforço da cibersegurança, etc. As melhorias de rendimento e redução de custos nas atividades económicas são transversais[5]. Os riscos também são significativos: a potencial exposi-

[1] A expressão IA foi cunhada por John McCarthy, em 1958.

[2] Pelo menos na atualidade, a generalidade dos sistemas disponíveis corresponde mais ao conceito de "Inteligência Aumentada", em que os sistemas tecnológicos complementam e permitem melhorar as capacidades humanas. Este último conceito foi cunhado por Douglas Engelbert, em 1962. Sobre a relação entre Inteligência Artificial (na sigla inglesa, «*AI*») e Inteligência Aumentada (na sigla inglesa, «*IA*»), *vide* Hossein Hassani *et al*, "Artificial Intelligence (AI) or Intelligence Augmentation (IA): What Is the Future?", *AI 2020, 1*, 143–155, doi:10.3390/ai1020008, *in* https://www.researchgate.net/publication/340594734_Artificial_Intelligence_AI_or_Intelligence_Augmentation_IA_What_Is_the_Future (31.10.2022).

[3] Uma exposição do impacto da IA pode encontrar-se em OECD, *Artificial Intelligence in Society*, Paris, OECD Publishing, 2019, *in* https://read.oecd-ilibrary.org/science-and-technology/artificial-intelligence-in-society_eedfee77-en#page13 (31.10.2022).

[4] Ben Goertzel, criador da robot *Sofia*, apresentou uma «*nanny AI*». "Should Humanity Build a Global AI Nanny to Delay the Singularity Until It's Better Understood?", *Journal of Consciousness Studies*, January 2012 19(1-2), pp. 96-111, *in* https://www.researchgate.net/publication/233497378_Should_Humanity_Build_a_Global_AI_Nanny_to_Delay_the_Singularity_Until_It%27s_Better_Understood (31.10.2022). Em 2022, Black Lemoine publicou excertos de uma alegada conversa com um sistema *LaMDA*, afirmando que este poderia ter alcançado a autoconsciência, pois afirmou sentir tristeza e alegria e intitulou-se mesmo como «pessoa». *Is LaMDA Sentient? – an Interview*, 11.06.2022, *in* https://cajundiscordian.medium.com/is-lamda-sentient-an-interview-ea64d916d917 (31.10.2022).

[5] Uma apresentação global é apresentada por Tara Balakrishnan, Michael Chui e Bryce Hall, *McKinsey Global Survey The State of AI 2020*, de 17.11.2020, *in* https://www.mckinsey.com/capabilities/quantumblack/our-insights/global-survey-the-state-of-ai-in-2020 (31.10.2022).

ção de pessoas a erros e manipulações que podem afetar e corroer os direitos fundamentais[6] (*v.g.* a privacidade, a não discriminação, as liberdades)[7], a segurança, a responsabilidade civil e criminal, bem como os processos democráticos[8], o acesso e a administração da justiça, etc.[9].

Na Europa, a realidade acompanha o resto do mundo, ainda que nem sempre na vanguarda. Se 42% das empresas usa habitualmente uma tecnologia de IA (25% utilizam duas, com uma disparidade entre as grandes empresas – 39% – e as PME – 21%), número superior ao do Japão (30%) e dos Estados Unidos (28%)[10], o desenvolvimento e produção ficam aquém dos protagonistas globais, EUA e China. As principais barreiras apontadas são a insuficiência de mão-de-obra especializada (57%) e o custo (52%). Nos Estados Unidos, o volume do investimento privado em IA foi calculado, em 2016, entre € 12,1 e 18,6 mil milhões; na Ásia o número ascendeu a € 6,5 a 9,7 mil milhões, enquanto na Europa se ficou num montante estimado entre € 2,4 e 3,2 mil milhões. Entre 2019 e 2021, o investimento privado global passou de $US 46 mil milhões para $US 93 mil milhões[11]. O domínio da robótica é aquele em

[6] A Agência Europeia para os Direitos Fundamentais publicou, em 2020, o Relatório *Getting the future right – Artificial intelligence and fundamental rights*, *in* https://fra.europa.eu/en/publication/2020/artificial-intelligence-and-fundamental-rights (31.10.2022).

[7] O Conselho da Europa apresentou um estudo sobre o impacto dos sistemas algorítmicos nos direitos humanos. *Algorithms and Human Rights. Study on the human rights dimensions of automated data processing techniques (in particular algorithms) and possible regulatory implications*, Strasbourg, março 2018, *in* https://rm.coe.int/algorithms-and-human-rights-en-rev/16807956b5 (31.10.2022).

[8] O espaço digital é ambivalente. Por um lado, acarreta o risco de desenvolvimento de uma «*Private AIgocracy*», em detrimento da democracia e cidadania. Por outro lado, pode proporcionar vias para o desenvolvimento de um «*Super-collaborative Government*». Estes cenários foram apresentados no projeto do Joint Research Center da Comissão Europeia, *The Future of Government 2030+. A Citizen Centric Perspective on New Government Models*, 2017-2019 (o Projeto, o Relatório final e as Recomendações podem ser encontrados em https://blogs.ec.europa.eu/eupolicylab/futurgov/ (31.10.2022).

[9] Para uma exposição sucinta, mas clara, dessas oportunidades e preocupações, *vide* Grupo Independente de Peritos de Alto-Nível sobre a Inteligência Artificial, *Orientações Éticas para uma IA de Confiança*, Bruxelas, 2019, pp. 41-45, *in* https://op.europa.eu/en/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1/language-pt/format-PDF (31.10.2022).

[10] European Commission, *European enterprise survey on the use of technologies based on artificial intelligence*, Luxemburgo, 2020, *in* https://op.europa.eu/en/publication-detail/-/publication/f089bbae-f0b0-11ea-991b-01aa75ed71a1 (31.10.2022).

[11] Um bom exemplo é a quase total ausência da Europa no desenvolvimento de «sistemas de IA de utilização geral». Future of Life Institute, *Emerging Non-European Monopolies in the Global AI Market*, November 2022, *in* https://futureoflife.org/wp-content/uploads/2022/11/Emerging_Non-

que a Europa está mais avançada, contando com três dos maiores fabricantes mundiais (*KUKA*[12], *ABB* e *Comau*).

A política e o direito não são alheios à IA. Para a Comissão Europeia, "[o] modo como abordarmos a IA definirá o mundo em que vivemos (...) [e] é necessário um quadro europeu sólido"[13], capaz de assegurar que o desenvolvimento e utilização da IA respeite os valores da União e os direitos fundamentais, bem como princípios éticos[14], em especial a responsabilidade e a transparência[15]. É imperativo que os responsáveis políticos assumam a sua responsabilidade de defender os valores, princípios e direitos constitucionais e democráticos de um inefável horizonte de "tecno-determinismo"[16]. A UE tem competências para tal, seja para a adoção de legislação harmonizada para o mercado interno (artigos 4º, nº 2, al. a) TFUE), seja para o desenvolvimento da concorrência da indústria (artigo 173º TFUE), seja a promoção da I&D, incluindo para fomentar a competitividade da indústria (artigos 179º a 190º TFUE).

Desde 2004, a IA foi incluída nos programas de I&D europeus. Já em 2016, a Comissão Europeia integrou a IA na *Estratégia para a digitalização da indústria* [COM(2016) 180 final] e na *Estratégia de Política Industrial renovada da UE* [COM(2017) 479 final]. Por sua vez, os EM assumiram a IA como uma prioridade e, em 10 de abril de 2018 («Dia Digital»), 24 EM[17] e a Noruega firmaram uma Declaração de Cooperação[18]. No período 2014-2017, foram inves-

European_Monopolies_in_the_Global_AI_Market.pdf (4.11.2022). Dois projetos emergem: o *LEAM* (Large European AI Models) e o *Aleph Alpha*.

[12] Desde 2016, é dominada pela empresa chinesa *Midea Group*.

[13] Comunicação da Comissão ao Parlamento Europeu, ao Conselho Europeu, ao Conselho, ao Comité Económico e Social Europeu e ao Comité das Regiões *Inteligência artificial para a Europa*, Bruxelas, 25.4.2018, COM(2018) 237 final, p. 2, *in* https://eur-lex.europa.eu/legal-content/PT/TXT/HTML/?uri=CELEX:52018DC0237&from=EN (31.10.2022).

[14] Sobre a relação entre a IA, direitos humanos e princípios éticos, *vide* Giovanni Sartor, "Artificial intelligence and human rights: Between law and ethics", *Maastricht Journal of European and Comparative Law*, 27 (2020), issue 6, pp. 705-719.

[15] Comissão Europeia, COM 2018 237, *op. cit.*, p. 3.

[16] Oreste Pollicino and Giovanni De Gregorio, "Constitutional Law in the Algorithmic Society", *in* Hans-W. Micklitz *et al*, *Constitutional Challenges in the Algorithmic Society*, Cambridge, Cambridge University Press, pp. 3-24, p. 5, *in* https://www.cambridge.org/core/books/constitutional-challenges-in-the-algorithmic-society/constitutional-law-in-the-algorithmic-society/969E0889109C8092AD0AB57019E507E3 (31.10.2022).

[17] Bélgica, Bulgária, República Checa, Dinamarca, Alemanha, Estónia, Irlanda, Espanha, França, Itália, Letónia, Lituânia, Luxemburgo, Hungria, Malta, Países Baixos, Áustria, Polónia, Portugal, Eslovénia, Eslováquia, Finlândia, Suécia e Reino Unido.

[18] *Declaration Cooperation on Artificial Intelligence*, *in* https://ec.europa.eu/jrc/communities/en/node/1286/document/eu-declaration-cooperation-artificial-intelligence (31.10.2022).

tidos cerca de € 1 100 milhões na investigação e inovação ligadas à IA, no âmbito do programa de investigação e inovação Horizonte 2020, incluindo nos domínios dos megadados («*Big data*»), da saúde, da reabilitação, dos transportes e da exploração espacial. Entre 2017 e 2019, o financiamento da UE para a investigação e a inovação em matéria de IA aumentou c. de 70% e foi de € 1,5 mil milhões. Em 2021, foi criada a «Empresa Comum das Tecnologias Digitais Essenciais» (Parceria Público-Privada que reúne a UE, os EM e três associações industriais – EPoSS, AENEAS, INSIDE), a Iniciativa Europeia de Processadores (consórcio que reúne 30 parceiros europeus, entre os quais a FCT e o Instituto Superior Técnico) ou o *FET Proactive Project* (para estimular a emergência e afirmação de novos paradigmas tecnológicos).

Obviamente, a IA não está isenta do respeito do Direito e dos direitos, ainda que a efetivação daquele e a salvaguarda destes possa não estar assegurada, por desadequação dos instrumentos normativos e institucionais vigentes[19]; em especial, os principais problemas relacionam-se com a potencial opacidade e dificuldades de *accountability*, e logo de responsabilização, dos sistemas de IA. A necessidade de criação de meios específicos de regulação, sólida e flexível, torna-se uma evidência[20], ainda que seja difícil apurar qual a intensidade da intervenção regulatória que não dificulte a inovação.

A intervenção regulatória centra-se na gestão do risco, entre a mitigação e a eliminação, de acordo com o tipo de risco em causa[21]. A estratégia

[19] O quadro normativo geral da União inclui os valores do Artigo 2º TUE, que vinculam a UE e os EM, a CDF e os TUE e TFUE, bem como, no plano do direito secundário, múltiplos atos: as «Diretivas não discriminação» (2000/43, 2000/78, 2004/113, 2006/54); as «Diretivas proteção dos consumidores» (2005/29, 2011/83/CE); os atos que estabelecem os padrões de segurança dos produtos e de proteção dos consumidores (a «Diretiva Máquinas», a «Diretiva Equipamentos de Rádio», a «Diretiva Segurança Geral dos Produtos», bem como as regras de segurança específicas, por exemplo, para brinquedos ou dispositivos médicos); os atos de proteção dos dados pessoais («RGPD»; Reg. 2018/1725; Diretiva 2016/680); as regras de concorrência; a legislação de supervisão financeira; a «Diretiva comércio eletrónico». Grande parte destes atos encontra-se em processo de revisão legislativa. O Comité Europeu para a Proteção de Dados (CEPD) e a Agência Europeia para a Proteção de Dados (AEPD) fazem uma crítica exaustiva da interação da proposta de «Regulamento IA» com a legislação de proteção de dados, que consideram insuficiente e incoerente. Parecer conjunto 5/2021 do CEPD e da AEPD sobre a proposta de regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial (Regulamento Inteligência Artificial), que é, em geral, bastante crítico, *in* https://edpb.europa.eu/system/files/2021-10/edpb-edps_joint_opinion_ai_regulation_pt.pdf (31.10.2022).

[20] A proposta de «Regulamento IA» foi acompanhada por uma proposta de «Regulamento Máquinas».

[21] Todos os sistemas de IA comportam algum risco, mas este varia entre um risco mínimo, em que não é necessária uma regulamentação, e um risco inaceitável, que impõe a proibição de um

europeia distingue 4 níveis de risco: inaceitável, elevado, limitado e mínimo. A abordagem regulatória declarada pela Comissão pretende alcançar objetivos ambiciosos e potencialmente contraditórios: assegurar a concretização dos benefícios da IA, protegendo a inovação, e, simultaneamente, eliminar ou, pelo menos, mitigar os riscos e efeitos negativos, garantindo uma IA de «face humana». As propostas apresentadas no âmbito da «Estratégia Europeia para a IA»[22] vão nesse sentido?

Esta iniciativa política estruturou-se através de múltiplos os documentos apresentados pela Comissão Europeia desde 2018: à Comunicação *Inteligência artificial para a Europa* [COM(2018) 237 final], seguiu-se, ainda em 2018, o *Plano Coordenado para a Inteligência Artificial* [COM(2018) 795]; em 2019, foi apresentada a Comunicação *Aumentar a confiança numa inteligência artificial centrada no ser humano* [COM(2019) 168 final]; já em 2020, o *Livro Branco sobre a inteligência artificial – Uma abordagem europeia virada para a excelência e a confiança* [COM(2020) 65 final][23]; em 2021, após a Comunicação "Fomentar uma abordagem europeia da inteligência artificial" [COM(2021) 205 final] (em anexo encontra-se a "Revisão do Plano Coordenado"), a Comissão propôs o *Regulamento que estabelece regras harmonizadas em matéria de Inteligência Artificial* («Regulamento IA») [COM(2021) 206 final] (em curso, ainda sem posição do PE nem do Conselho); finalmente, em setembro de 2022, a Comissão apresentou a proposta de *Diretiva sobre a responsabilidade na AI* [COM(2022) 496 final] (em simultâneo com a proposta de *Diretiva sobre responsabilidade por produtos defeituosos* [COM(2022) 495 final]. Na esfera internacional, em agosto de 2022, a Comissão apresentou uma Recomendação[24] para uma Decisão do Conselho a autorizar a abertura de negociações no âmbito do Conselho da Europa para uma Convenção sobre IA, direitos humanos e o estado-de-direito[25].

sistema de IA. A Comissão de Ética dos Dados da Alemanha apelou à criação de um sistema de regulamentação de cinco níveis baseado nos diferentes níveis de risco.

[22] *In* https://digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence (31.10.2022).

[23] Ainda em 2020, a Comissão apresentou o *Relatório sobre as implicações em matéria de segurança e de responsabilidade decorrentes da inteligência artificial, da Internet das coisas e da robótica*, de 19.02.2020, COM(2020) 64, *in* https://eur-lex.europa.eu/legal-content/PT/TXT/HTML/?uri=CELEX:52020DC0064&from=en (31.10.2022).

[24] Comissão Europeia, *Recomendação de Decisão do Conselho que autoriza a abertura de negociações, em nome da União Europeia, tendo em vista uma convenção do Conselho da Europa sobre inteligência artificial, direitos humanos, democracia e Estado de direito*, Bruxelas, 18.8.2022, COM(2022) 414 final, *in* https://ec.europa.eu/transparency/documents-register/detail?ref=COM(2022)414&lang=en (31.10.2022).

[25] O Comité para a Inteligência Artificial, instituído no Conselho da Europa, em 2021, recebeu, em 30 de junho de 2022, a incumbência de elaborar uma convenção-quadro sobre IA.

Além desses, não podem ser ignorados documentos conexos, em especial o do Grupo Independente de Peritos de Alto-Nível, *Orientações Éticas para uma IA de confiança*. Múltiplos estudos avaliaram o impacto da IA em numerosos setores na UE[26]. O *Study to Support an Impact Assessment of Regulatory Requirements for Artificial Intelligence in Europe. Final Report*, publicado pela Comissão em abril de 2021[27], apresenta uma descrição exaustiva dos problemas de violação de direitos e princípios fundamentais já comprovados em sistemas de IA e contém uma exposição comparativa das abordagens regulatórias em vigor ou em estudo nos principais parceiros internacionais.

Além disso, deve compreender-se esta estratégia política em conjunto com a extensa reforma no mercado interno, incluindo as propostas de um novo «Regulamento Máquinas» [COM(2021) 202 final][28], atualizando as regras relativas à segurança, e um novo «Regulamento Geral de Segurança de produtos» [COM(2021) 346 final][29]. Não se pode ignorar ainda o «Regulamento dos Serviços Digitais»[30] e o «Regulamento dos Mercados Digitais»[31], bem como as iniciativas legislativas relativas aos dados[32] ou à cibersegurança[33], todos relevantes na construção do ecossistema digital europeu[34].

[26] Sobre aqueles que foram efetuados dentro da UE, pela Comissão, pelo PE e outros organismos, como a Autoridade Europeia para a Proteção de Dados (de ora em diante, também AEPD) veja-se as referências indicadas na Resolução do PE de 3 de maio de 2022. O Grupo de Peritos Independente também apresentou análises setoriais. *Sectoral Considerations on the Policy and Investment Recommendations for Trustworthy Artificial Intelligence*, *in* https://futurium.ec.europa.eu/en/european-ai-alliance/document/ai-hleg-sectoral-considerations-policy-and-investment-recommendations-trustworthy-ai (31.10.2022).

[27] *In* https://op.europa.eu/en/publication-detail/-/publication/55538b70-a638-11eb-9585-01aa75ed71a1 (31.10.2022).

[28] https://eur-lex.europa.eu/legal-content/PT/TXT/HTML/?uri=CELEX:52021PC0202&from=EN.

[29] https://eur-lex.europa.eu/legal-content/PT/TXT/HTML/?uri=CELEX:52021PC0346&from=EN.

[30] Regulamento (UE) 2022/2065 do Parlamento Europeu e do Conselho de 19 de outubro de 2022 relativo a um mercado único para os serviços digitais, JO L 277, 27.10.2022, p. 1. Este ato será aplicável a partir de 17 de fevereiro de 2024.

[31] Regulamento (UE) 2022//1925 do Parlamento Europeu e do Conselho de 14 de setembro de 2022 relativo à disputabilidade e equidade dos mercados no setor digital e que altera as Diretivas (UE) 2019/1937 e (UE) 2020/1828, JO L 265, 12.10.2022, p. 1. Este ato será aplicável a partir de 2 de maio de 2023.

[32] Regulamento (EU) 2022/868 do Parlamento Europeu e do Conselho de 30 de maio de 2022 relativo à governação europeia de dados e que altera o Regulamento (UE) 2018/1724 (Regulamento Governação de Dados), JO L 152, 3.6.2022, p. 1.

[33] Proposta de Regulamento do Parlamento Europeu e do Conselho relativo aos requisitos horizontais de cibersegurança dos produtos com elementos digitais e que altera o Regulamento (UE) 2019/1020, Bruxelas, 15.09.2022, COM(2022) 454 final, *in* https://eur-lex.europa.eu/resource.html?uri=cellar:864f472b-34e9-11ed-9c68-01aa75ed71a1.0008.02/DOC_1&format=PDF (31.10.2022).

2. A Definição de Inteligência Artificial

Numa abordagem global e integrada da IA, impõe-se, antes de mais, delimitar e compreender o objeto de intervenção. No entanto, essa é uma missão praticamente impossível, pois não há um consenso sobre o conceito de IA[35], nem sobre os domínios e subdomínios nela integrados. É difícil recortar uma definição precisa num terreno em que a evolução tecnológica é permanente. Há que optar por uma abordagem pragmática, que recorte e defina um conceito operatório útil para efeitos regulatórios e que assegure a certeza jurídica. Por outro lado, é fundamental assegurar que seja adaptável a inovações tecnológicas em investigação ou a surgir no futuro. Finalmente, seja a definição, seja a classificação dos sistemas de IA tem de ser «*human centric*», ou seja, adequada para alcançar objetivos e necessidades humanas e do planeta[36].

Elementos fundamentais da IA são os dados, a «matéria-prima» dos sistemas de IA, e os algoritmos, a equação lógica de processamento dos dados. Os traços funcionais essenciais são a perceção do entorno, o processamento da informação e autonomia de decisão[37] para a prossecução de objetivos definidos. As funções que os sistemas de IA podem desempenhar vão da resolução e otimização até à atividade criativa, passando pela automação e comunicação e podem ser utilizados nos sistemas de produção, comercialização, administração e gestão, logística, recursos humanos, em todos os domínios económicos, sociais e culturais[38]. A taxonomia[39] de domínios e subdomínios abrangidos

[34] O Parlamento Europeu aprovou várias Resoluções sobre a IA, entre as quais se salienta a Resolução de 20.10.2020, que contém recomendações à Comissão sobre o regime relativo aos aspetos éticos da inteligência artificial, da robótica e das tecnologias conexas. O Conselho, em 21.10.2020, deu uma especial atenção à relação entre a IA e a Carta dos Direitos Fundamentais. *Conclusões sobre a Carta dos Dreitos Fundamentais no contexto da inteligência artificial e da transformação digital*, 11481/20, *in* https://data.consilium.europa.eu/doc/document/ST-11481-2020-INIT/pt/pdf (31.10.2022).

[35] A cautela recomenda-se. Ainda assim, aponta-se a distinção entre IA «estreita» e IA «geral» e as quatro categorias de IA reconhecidas («reativa», «memória limitada», «teoria da mente», «autoconsciente»). Recorde-se até as perspetivas provocadoras do reconhecimento de uma nova forma de «Vida 3.0», pós-humana (Max Tegmark, *Life 3.0. Ser-se Humano na Era da Inteligência Artificial*, (trad. port. de J. Van Zeller), Alfragide, 2019, p. 83), bem como a «*singularity*» de John von Neumann e Vernor Vinge para a explosão de «super-inteligência» que superaria a «era humana».

[36] Esta perspetiva está subjacente ao quadro de classificação dos sistemas de IA elaborado pela OCDE. "OECD Framework for the classification of AI systems", OECD Digital Economy Papers, February 2022, No. 323, *in* https://www.oecd-ilibrary.org/science-and-technology/oecd-framework-for-the-classification-of-ai-systems_cb6d9eca-en (31 de outubro de 2022).

[37] Estas características e as suas implicações são salientadas pela Comissão no *Relatório sobre as implicações em matéria de segurança e de responsabilidade decorrentes da inteligência artificial* (cit.).

[38] Recentemente foi publicado um glossário para uma IA centrada no ser humano. Marina Estevez Almenzar *et al*, *Glossary of human-centred artificial intelligence*, Publications Office of

pode distinguir domínios centrais e domínios transversais. Entre os primeiros, encontra-se o raciocínio, o planeamento, a aprendizagem, a comunicação e a perceção; entre os segundos, conta-se a integração e interação, os serviços de IA e a ética e filosofia da IA. A cadeia de intervenientes é extensa, desde o fabricante, passando pelo fornecedor e o «treinador» dos dados, pelo programador, até ao utilizador.

Essencial nos sistemas de IA é a respetiva autonomia. Tal, no entanto, não significa ausência completa do ser humano e como princípio fundamental da conceção dos sistemas de IA deveria estar a «lealdade» para com o ser humano. As possíveis interações entre o sistema e o ser humano podem assumir várias formas, num espectro que vai da interação permanente de uma pessoa com o sistema de IA («*human in the loop*» -«HITL»), dando *feedback* sobre os desenvolvimentos do sistema, até à total ausência de um ser humano («*human out of the loop*» – «HOOTL»), incluindo sobre a efetivação do resultado, passando por uma intervenção sob a forma de validação ou eliminação do resultado («*human on the loop*» – «HOTL» – e «*human in command*»).

Vamos percorrer apenas as noções avançadas em documentos apresentados no âmbito da UE[40].

Na Comunicação COM(2018) 237 final, a Comissão Europeia apresenta um conceito vago e tautológico, segundo o qual *"[o] conceito de inteligência artificial (IA) aplica-se a sistemas que apresentam um comportamento inteligente, analisando o seu ambiente e tomando medidas – com um determinado nível de autonomia – para atingir objetivos específicos"*[41].

Uma definição mais rigorosa foi dada pelo Grupo Independente de Peritos de Alto Nível sobre a IA (de ora em diante, também Grupo de Peritos),

the European Union, Luxembourg, 2022, doi:10.2760/860665, JRC129614, *in* https://publications.jrc.ec.europa.eu/repository/handle/JRC129614 (31.10.2022).

[39] Seguimos a proposta de Sofia Samoili *et al*, *AI Watch. Defining Artificial Intelligence 2.0*, EUR 30873 EN, Publications Office of the European Union, Luxembourg, 2021, doi:10.2760/019901, JRC126426, *in* https://publications.jrc.ec.europa.eu/repository/handle/JRC126426 (31.10.2022).

[40] No âmbito dos parâmetros de estandardização, veja-se a definição de sistema de IA em ISO/IEC 22989:2022(en) Information technology – Artificial intelligence – Artificial intelligence concepts and terminology: "*[an] engineered system that generates outputs such as content, forecasts, recommendations or decisions for a given set of human-defined objectives*", *in* https://www.iso.org/obp/ui/#iso:std:iso-iec:22989:ed-1:v1:en (31.10.2022). A OCDE apresentou uma definição em 2019 ("Scoping the OECD AI principles: Deliberations of the Expert Group on Artificial Intelligence at the OECD (AIGO)", *OECD Digital Economy Paper*s, November 2019 Nº 291, *in* https://doi.org/10.1787/d62f618a-en (31.10.2022). Por sua vez, na Recomendação aprovada em 22.05.2019, a OCDE salienta a importância de ter em consideração todo o ciclo de vida dos sistemas de IA, *in* https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449 (31.10.2022).

[41] Comissão Europeia, COM(2018) 237 final, cit., p. 1.

segundo a qual "*[o]s sistemas de inteligência artificial (IA) são sistemas de software (e eventualmente também de hardware) concebidos por seres humanos, que, tendo recebido um objetivo complexo, atuam na dimensão física ou digital percecionando o seu ambiente mediante a aquisição de dados, interpretando os dados estruturados ou não estruturados recolhidos, raciocinando sobre o conhecimento ou processando as informações resultantes desses dados e decidindo as melhores ações a adotar para atingir o objetivo estabelecido. Os sistemas de IA podem utilizar regras simbólicas ou aprender um modelo numérico, bem como adaptar o seu comportamento mediante uma análise do modo como o ambiente foi afetado pelas suas ações anteriores.*"[42] Esta definição é retirada de um outro documento do Grupo Independente de Peritos que tem a definição de IA como objeto de análise[43].

Os sistemas de IA baseiam-se em diversas abordagens e técnicas construídas por uma disciplina científica. Entre estas encontram-se "*a aprendizagem automática (de que a aprendizagem profunda e a aprendizagem por reforço são exemplos específicos), o raciocínio automático (que inclui o planeamento, a programação, a representação do conhecimento e o raciocínio, a pesquisa e a otimização) e a robótica (que inclui o controlo, a perceção, os sensores e atuadores, bem como a integração de todas as outras técnicas em sistemas ciberfísicos)*"[44].

A Comissão Europeia avança, na proposta de «Regulamento IA», com um conceito pragmático, que pretende ser tecnologicamente neutro e flexível, de modo a adaptar-se à evolução na tecnologia e no mercado, ao mesmo tempo que abrangente. A definição inscrita no artigo 3º, 1) é funcional: "*um programa informático desenvolvido com uma ou várias das técnicas e abordagens enumeradas no anexo I, capaz de, tendo em vista um determinado conjunto de objetivos definidos por seres humanos, criar resultados, tais como conteúdos, previsões, recomendações ou decisões, que influenciam os ambientes com os quais interage.*" Esta definição é complementada, no Anexo I[45], pela enunciação das técnicas e abordagens, lógicas e estatísticas, utilizadas nos sistemas de IA: a) Abordagens de aprendizagem automática, incluindo aprendizagem supervisionada, não supervisionada e

[42] *Orientações Éticas para uma IA de Confiança*, cit., p. 47.

[43] High-Level Expert Group on Artificial Intelligence, *A Definition of AI: Main Capabilities and Scientific Disciplines. Definition developed for the purpose of the deliverables of the High-Level Expert Group on AI*, Bruxelas, 2018, p. 7, *in* https://ec.europa.eu/futurium/en/system/files/ged/ai_hleg_definition_of_ai_18_december_1.pdf (31.10.2022).

[44] *Orientações Éticas para uma IA de Confiança*, cit., p. 47.

[45] *Anexos da Proposta de Regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial (Regulamento Inteligência Artificial) e altera determinados atos legislativos da União*, p. 1, *in* https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0004.02/DOC_2&format=PDF (31.10.2022)

por reforço, utilizando uma grande variedade de métodos, designadamente aprendizagem profunda; b) Abordagens baseadas na lógica e no conhecimento, nomeadamente representação do conhecimento, programação (lógica) indutiva, bases de conhecimento, motores de inferência e de dedução, sistemas de raciocínio (simbólico) e sistemas periciais; c) Abordagens estatísticas, estimação de Bayes, métodos de pesquisa e otimização. Esta lista pode ser alterada pela Comissão, através de atos delegados, para acompanhar a evolução do mercado e da tecnologia.

O *Documento de Compromisso* alcançado pela Presidência eslovena, em 29 de novembro de 2021[46], introduziu algumas modificações a essa noção, acolhendo algumas críticas que consideravam a noção proposta pela Comissão demasiado extensa, incluindo sistemas de *software* que não são considerados IA. A tónica é colocada na forma autónoma de decisão: "*any such system should be capable of determining how to achieve a given set of human defined objectives by learning, reasoning or modelling.*" Assim, o § 6 do Preâmbulo da Proposta é reformulado para salientar essa dimensão funcional: "*[t]he definition should be based on key functional characteristics of the software of artificial intelligence*". Consequentemente, os sistemas de IA "*should be intended as having the ability, on the basis of machine and/or human based data and inputs, to infer the way to achieve a given set of human-defined objectives through learning, reasoning or modelling and to generate specific outputs in the form of content for generative AI systems (such as text, video or images), as well as predictions, recommendations, or decisions, which influence the environment with which the system interacts, be it in a physical or digital dimension*". Este documento reformula a definição de IA do artigo 3º: *'artificial intelligence system' (AI system) means a system that (i) receives machine and/or human-based data and inputs*, (ii) *infers how to achieve a given set of human-defined objectives using learning, reasoning or modelling implemented with the techniques and approaches listed in Annex I, and* (iii) *generates outputs in the form of content (generative AI systems), predictions, recommendations or decisions, which influence the environments it interacts with*.

O *Draft Report* do PE[47] acrescenta a referência a um outro resultado dos sistemas de IA: a formulação de hipóteses. Por outro lado, não impõe que os objetivos sejam definidos por seres humanos, abordagem que admite a possi-

[46] Council of the European Union, *Presidency compromise text*, 14278/21, *in* https://data.consilium.europa.eu/doc/document/ST-14278-2021-INIT/en/pdf (31.10.2022).

[47] Parlamento Europeu, *Draft Report on the proposal for a regulation of the European Parliament and of the Council on harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts*, 20.04.2022, 2021/0106(COD), *in* https://www.europarl.europa.eu/doceo/document/CJ40-PR-731563_EN.pdf (31.10.2022).

bilidade futura de os sistemas de IA virem a ter capacidade para definir objetivos de modo autónomo. A questão fundamental aqui é se o desenvolvimento desse tipo de sistemas deve ser admitido.

Podem distinguir-se «sistemas de IA de utilização geral», (também designados como «*foundation models*»), tais como sistemas de reconhecimento de imagem/vídeo, geração de áudio/vídeo, deteção de padrões, pergunta resposta, tradução, etc., que podem ser aplicados em diversos domínios (saúde, finanças, educação, etc.) e contextos, que podem desempenhar diferentes tarefas e alcançar múltiplos resultados, necessitando ou não de adaptação, incluindo integrados em outros sistemas de IA[48], e «sistemas de utilização específica» («*fixed purpose*»), criados para uma finalidade determinada e que só podem desempenhar as tarefas para que foram treinados (enquanto os anteriores podem desempenhar tarefas para que não foram originalmente treinados). A abordagem assente no risco coloca a ênfase na utilização do sistema de IA e não nas respetivas funcionalidades, pelo que desde logo se suscitou perplexidade em relação ao regime a aplicar aos «sistemas de IA de utilização geral», dado que nenhuma previsão especial se previu na proposta de Regulamento apresentada pela Comissão. Tal poderia significar a exclusão da aplicação da regulamentação aos criadores destes sistemas que são depois afinados para uma utilização específica num domínio especial. O *Documento de Compromisso da Presidência eslovena*[49] prevê que não é suficiente para ficar sob a aplicação da regulamentação proposta a colocação no mercado, em serviço ou uso de um «sistema de IA de utilização geral». A finalidade[50], sim, continua a ser determinante e é o utilizador que fica sob a aplicação do regime a instituir. A não coincidência entre o criador do sistema de IA e o seu utilizador, que até podem estar muito distantes, implica a não coincidência entre quem assume o risco e quem efetivamente pode perceber o sistema e as possíveis soluções para os problemas que dele resultam[51]. O *Documento de Compromisso*

[48] Sobre estes sistemas e as dificuldades de qualificação, Carlos I. Gutierrez *et al.*, "A Proposal for a Definition of General Purpose Artificial Intelligence Systems", Future of Lige Institute – Working Paper, November 2022, *in* https://futureoflife.org/wp-content/uploads/2022/11/SSRN-id4238951-1.pdf (5.11.2022). Na ausência de um consenso, os autores defendem que o critério dessa classificação seja(m) a(s) «tarefa(s)» desempenhada(s).

[49] Novo considerando 70(a) do Preâmbulo.

[50] Carlos I. Gutierrez et al. criticam esta opção, pois ela não garante devidamente a segurança e a confiança na IA. "A Proposal for a Definition of General Purpose Artificial Intelligence Systems", cit., p. 3.

[51] Carlos I. Gutierrez *et al.*, "A Proposal for a Definition of General Purpose Artificial Intelligence Systems", cit., p. 1.

da presidência francesa[52] introduziu um regime especial aplicável a esses sistemas[53], independentemente de ser um modelo pré-treinado ou de ter de ser objeto de adaptação pelo utilizador.

O desenvolvimento fulgurante dos modelos fundacionais, tal como o ChatGPT, entretanto ocorrido, conduziu a emendas na posição do PE, aprovada em maio de 2023, com o objetivo de incluir no âmbito do regime a instituir esse tipo de modelos.

Os sistemas de IA podem ser sistemas de *software*, tais como os motores de busca, assistentes de voz, programas de análise de imagens ou sistemas de reconhecimento facial e de discurso. Podem estar integrados («*embedded*») num produto ou contribuir para o funcionamento do produto sem estarem integrados nele («*non-embedded*»). Estão frequentemente integrados em outros sistemas mecânicos, digitais, etc. e não são sistemas independentes, como sucede com os robôs, pelo que a regulamentação de produtos em geral é também relevante[54].

3. A Elaboração de uma Política

3.1. Comunicação Inteligência Artificial para a Europa, 25.04.2018, COM(2018) 237 final

Segundo a Comissão, a política europeia para a IA deve assegurar a excelência e a confiança. Este documento propõe três pilares: 1 – Reforçar da capacidade tecnológica e industrial da UE e a aceitação da IA em toda a economia, tanto pelo setor privado como pelo público; 2 – Preparar as mudanças socioeconómicas decorrentes da IA; 3 – Garantir um quadro ético e jurídico apropriado, baseado nos valores da União e em consonância com a Carta dos Direitos Fundamentais da União Europeia (de ora em diante, CDF).

A intenção é promover a liderança da UE na revolução da inteligência artificial, "à sua maneira e com base nos seus valores", de forma sustentável e inclusiva. A prossecução dos objetivos integra o setor privado e o setor público num quadro do tipo «PPP» (Parceria Público-Privada), através da instituição de uma ampla plataforma multilateral – a «Aliança Europeia para a IA»[55].

[52] *In* https://artificialintelligenceact.eu/wp-content/uploads/2022/05/AIA-FRA-Art-34-13-May.pdf (31.10.2022).

[53] A sua definição é acrescentada no artigo 3º da proposta e as novas disposições 4a e 4b definem o regime de aplicação.

[54] Em primeiro lugar, a Diretiva 2006/42/CE relativa às máquinas e que alterou a Diretiva 95/16/CE. No momento em que apresentou a proposta de «Regulamento IA», a Comissão apresentou a proposta de «Regulamento máquinas», [COM (2021) 202 final].

[55] Sobre esta, *vide* https://futurium.ec.europa.eu/en/european-ai-alliance (31.10.2022).

Convoca-se o modelo do Método Aberto de Coordenação, em que sobressaem o intercâmbio de boas práticas, a identificação de sinergias e o alinhamento de ações com o objetivo de criar um «ecossistema europeu». A preocupação em apoiar as PME a integrarem a IA como ferramenta para a competitividade está na base da criação da plataforma «*AI4EU*»[56], uma plataforma de IA a pedido, que disponibiliza algoritmos, dados e ferramentas de IA.

A concretização desta abordagem fez-se através do *Plano Coordenado*, de 2018, e que foi objeto de uma revisão em 2021. Participam no Plano os EM, a Noruega e a Suíça. Lança as bases para a coordenação das políticas no domínio da IA e propõe 70 ações conjuntas. Os Estados devem elaborar estratégias nacionais[57] e a Comissão defende o aumento da capacidade através da promoção de redes restritas de investigação de excelência, a promoção de ações na educação e formação e, especialmente, a criação de espaços comuns de dados, dada a indispensabilidade de acesso a quantidades gigantes de dados para treinar os sistemas de IA. A promoção da cooperação internacional é colocada entre as prioridades e será a via para a UE assumir o seu protagonismo no desenvolvimento da governação internacional do espaço digital, como «potência normativa», através do «*Brussels effect*» [58]. Com esse objetivo, a Comissão lançou a *International alliance for a human-centric approach to Artificial Intelligence* e a *International outreach for human-centric artificial intelligence initiative*[59].

O reforço da confiança aconselha o desenvolvimento de diretrizes éticas e de um quadro regulatório, mas também uma atenção à segurança e à robustez da tecnologia. Impõe-se a consideração de princípios fundamentais na abordagem político-regulatória, «desde a conceção» e «por defeito». Os seus

[56] Sobre esta, *vide* https://www.ai4europe.eu/ (31.10.2022).

[57] A Estratégia portuguesa encontra-se em *AI PORTUGAL 2030. An innovation and growth strategy to foster Artificial Intelligence in Portugal in the European contexto*, *in* https://www.incode2030.gov.pt/sites/default/files/julho_incode_brochura.pdf (31.10.2022).

[58] O PE acompanha essa pretensão na Resolução do Parlamento Europeu, de 3 de maio de 2022, sobre a inteligência artificial na era digital (2020/2266(INI)), *in* https://www.europarl.europa.eu/doceo/document/TA-9-2022-0140_PT.html (31.10.2022). Um estudo encomendado pelo Comité AIDA do PE versa exatamente sobre a dimensão externa. Ulrike Franke, *Artificial Intelligence diplomacy. Artificial Intelligence governance as a new European Union external policy tool. Study for the special committee on Artificial Intelligence in a Digital Age (AIDA)*, Policy Department for Economic, Scientific and Quality of Life Policies, European Parliament, Luxembourg, 2021, *in* https://www.europarl.europa.eu/RegData/etudes/STUD/2021/662926/IPOL_STU(2021)662926_EN.pdf (31.10.2022). O «RGPD» é o *gold standard* em matéria de proteção de dados. Mikal Ipek, "EU Draft Artificial Intelligence Regulation: Extraterritorial application and Effects", *EU Blog*, 17 February 2022, *in* https://europeanlawblog.eu/2022/02/17/eu-draft-artificial-intelligence-regulation-extraterritorial-application-and-effects/ (31.10.2022).

[59] *In* https://digital-strategy.ec.europa.eu/en/policies/international-outreach-ai (31.10.2022).

eixos fundamentais são uma IA responsável, humana e explicável, recusando a opacidade do designado «efeito caixa negra». A questão da responsabilidade é objeto do documento anexo *Liability for emerging digital technologies*[60].

A importância do desenvolvimento de todo o potencial da IA impõe um outro princípio, o princípio da inovação, que tempera a concretização daqueles eixos. Este princípio densifica-se através de um conjunto de orientações e instrumentos que asseguram que as iniciativas da Comissão favorecem a inovação. O seu desenvolvimento é explicitado na Comunicação *Aumentar a Confiança numa Inteligência Artificial centrada no Ser Humano*. Um exemplo de uma regulamentação facilitadora da inovação é a exceção prevista para a prospeção de textos e de dados na iniciativa apresentada para a modernização das regras da UE em matéria de direitos de autor[61]. Uma regulamentação praticamente limitada aos sistemas de IA de elevado risco na proposta de «Regulamento IA» pode também ser vista como uma abordagem pró-inovação (numa perspetiva menos benigna, será antes uma abordagem pró-indústria num setor dominado por gigantes mundiais e que secundariza os interesses das pessoas). O *Documento de compromisso* da Presidência eslovena vai mais longe. Inclui no seu objeto medidas de apoio à inovação (nova al. e) do artigo 1º) e propõe excluir do «Regulamento IA» os sistemas de IA especificamente desenvolvidos e colocados em operação exclusivamente para fins de I&D[62].

3.2. Comunicação Aumentar a Confiança numa Inteligência Artificial centrada no Ser Humano, 8.04.2019, COM(2019) 168 final

Depois daquele documento genérico, em que a Comissão delineou a estratégia europeia, este constitui um instrumento de continuidade para a concreti-

[60] European Commission, *Staff Working Document Liability for emerging digital technologies Accompanying the document Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions Artificial intelligence for Europe*, Brussels, 25.04.2018, SWD(2018) 137 final, *in* https://eur-lex.europa.eu/legal-content/en/ALL/?uri=CELEX%3A52018SC0137 (31.10.2022) . Por sua vez, o PE aprovou uma resolução com recomendações para a Comissão relativas a um regime de responsabilidade civil para a inteligência artificial. *Resolução do Parlamento Europeu, de 20 de outubro de 2020, que contém recomendações à Comissão sobre o regime de responsabilidade civil aplicável à inteligência artificial* (2020/2014(INL) [2020] OJ C404/107, *in* https://www.europarl.europa.eu/doceo/document/TA-9-2020-0276_PT.html (31.10.2022) . A proposta de regulamento para a IA, surpreendentemente, nada previu nesta matéria, o que foi objeto de críticas severas. Em setembro de 2022, a Comissão apresentou uma proposta de diretiva sobre a responsabilidade da IA.

[61] Artigos 3º e 4º da Diretiva (UE) 2019/790 do Parlamento Europeu e do Conselho de 17 de abril de 2019 relativa aos direitos de autor e direitos conexos no mercado único digital e que altera as Diretivas 96/9/CE e 2001/29/CE, JO L 130/92, 17.05.2019.

[62] Novo considerando 12(a).

zação dos dois últimos pilares enunciados na Comunicação do ano anterior. A Comissão recorda os valores europeus, plasmados no artigo 2º TUE e declara que serão estes a nortear a evolução da IA com o objetivo de colocar as pessoas no centro. Elegem-se três componentes essenciais para uma IA de confiança: 1) Conformidade com a legislação; 2) Respeito dos princípios éticos; 3) Robustez. Este documento é efetivamente uma síntese do documento produzido pelo Grupo de Peritos *Orientações Éticas para uma IA de confiança*[63]. O foco de ambos os documentos é colocado nos dois últimos componentes. Assegurar a confiança impõe o respeito pela legislação, mas também por princípios éticos e depende da segurança e fiabilidade, técnica e social, dos sistemas de IA. São quatro os princípios éticos para a IA: 1. Respeito da autonomia humana; 2. Prevenção de danos; 3. Equidade; 4. Explicabilidade.

A densificação dos valores fundamentais e dos componentes indicados, exige o respeito por sete requisitos: a) Iniciativa e controlo por humanos; b) Robustez e segurança; c) Privacidade e governação dos dados; d) Transparência; e) Diversidade, não discriminação e equidade; f) Bem-estar societal e ambiental; g) Responsabilização.

a) Iniciativa e controlo por humanos

A iniciativa e controlo humano exigido nas tecnologias de IA desdobra-se em três formas possíveis de participação de pessoas no seu funcionamento: a intervenção humana no respetivo funcionamento, com uma interação máquina – ser humano («HCI» – *Human-computer interaction*) que assegura um acompanhamento e uma curadoria capaz de dar «valor e significado humano» à ação da IA através de *feedback* de um ser humano ao longo do desempenho da tecnologia de IA (*human-in-the-loop* – «HITL»)[64], a supervisão humana

[63] Este, por sua vez, colhe inspiração no trabalho do Grupo Europeu para a Ética na Ciência e Novas Tecnologias e da Agência Europeia dos Direitos Fundamentais.

[64] Várias experiências podem ser encontradas em https://hai.stanford.edu/news/humans-loop-design-interactive-ai-systems (31.10.2022). A saúde é um dos domínios onde o seu potencial é maior. Esta via promove igualmente a transparência e explicabilidade da IA. Por outro lado, além da eficiência e exatidão, introduz-se entre os objetivos da tecnologia a preferência, gosto ou juízo e a «*agency*» humana no processo de decisão da IA. Saleema Amershi *et al*, "Power to the People: The Role of Humans in Interactive Machine Learning", *A.I. Magazine (35):4*, pp.105–120, *in* https://ojs.aaai.org/index.php/aimagazine/article/view/2513 (31.10.2022). A investigação mais recente revela que os sistemas *human-in-the loop* podem mesmo melhorar a eficiência. As dificuldades de efetivação dessa interação significativa têm sido igualmente salientadas, desde logo pela rapidez de funcionamento dos sistemas de IA e pela imensidão de informação utilizada. Zhe Liu, Yufan Guo, Jalal Mahmud, "When and Why does a Model Fail? A Human-in-the-loop Error Detection Framework for Sentiment Analysis", *Proceedings of NAACL HLT 2021: Industry Track Papers*, June 6-11,

(*human-on-the-loop* – «HOTL») e o controlo por humanos (*human-in-command*). O modo como a Comissão entende a supervisão humana, mesmo quando o sistema de IA é um instrumento para a tomada de decisão, mas esta tem de ter a validação humana (*v.g.* rejeitar um pedido de prestação social), não corresponde necessariamente à primeira forma indicada, pois nesta a interação não é apenas final, mas deve corresponder a uma interação recorrente ao longo das várias fases do processo de funcionamento do sistema de IA. Na verdade, a distinção entre os sistemas *human-on-the-loop* e *human-in-command* não é simples.

A forma adequada de controlo humano depende de cada sistema de IA e da respetiva aplicação[65], mas deve estar sempre submetida a fiscalização. A possibilidade de uma IA *human-out-of-the loop* («HOOTL»), em que não há qualquer intervenção humana, cada vez mais suscita as maiores reticências. O impacto sobre os direitos fundamentais e a transparência na interação dos sistemas de IA com os seres humanos deve estar integrado na conceção dos sistemas[66].

b) Robustez e segurança

A robustez e segurança concretizam-se por diversas vias: o desenvolvimento de algoritmos seguros, fiáveis e suficientemente robustos para lidar com erros ou incoerências durante todas as fases do ciclo de vida do sistema de IA e para lidar adequadamente com resultados errados; a existência de mecanismos de proteção e de segurança desde a conceção tendo em conta a segurança física e mental de todas as partes envolvidas, com minimização e, sempre que possível, a reversibilidade de consequências ou de erros não intencionais na operação do sistema; o estabelecimento de processos destinados a clarificar e

2021, pp. 170–177, *in* https://aclanthology.org/2021.naacl-industry.22.pdf (31.10.2022); Sundar Narayanan, *Human-in-the-loop or on-the-loop is not a silver bullet. Evaluate their effectiveness*, *in* https://medium.com/mlearning-ai/human-in-the-loop-or-on-the-loop-is-not-a-silver-bullet-evaluate-their-effectiveness-82f37835d765 (31.10.2022); Kobi Leins, Anja Kaspersen, *7 Myths of Using the Term "Human on the Loop": "Just What Do You Think You Are Doing, Dave?", Artificial Intelligence & Equality Initiative Nov 9, 2021*, *in* https://www.carnegiecouncil.org/media/article/7-myths-of-using-the-term-human-on-the-loop (31.10.2022); Eduardo Mosqueira-Rey *et al.*, "Human-in-the-loop machine learning: a state of the art", Artif Intell Rev (2022), *in* https://doi.org/10.1007/s10462-022-10246-w (31.10.2022).

[65] O CENELEC elaborou, em outubro de 2022, um *Draft* com orientações de um modelo de interação colaborativa ser humano-máquica. "Design methodology of advanced human–robot collaborative cells in personalized HRC systems", *in* https://www.cencenelec.eu/media/CEN-CENELEC/News/Workshops/2022/2022-10-21%20-%20SHARE/cwa_hrccells_v2_forpubliccommenting.pdf (31.10.2022).

[66] A inovação mais recente leva mais longe a interação máquina-humano com a possibilidade de integrar em *chips* informáticos tecidos humanos («*Organ on chip*»), a fim de mimetizar em laboratório o funcionamento de órgãos humanos em investigação médica, por ex.

avaliar os riscos potenciais associados à utilização de sistemas de IA em diferentes áreas de aplicação.

c) Privacidade e governação dos dados
Preconiza-se a garantia de privacidade, proteção, qualidade e integridade dos dados em todas as fases do ciclo de vida do sistema de IA. As pessoas têm de manter sempre o controlo pleno dos seus dados e tem de assegurar-se a qualidade e integridade dos dados para impedir enviesamentos e inexatidões.

d) Transparência
A transparência e explicabilidade do processo de decisão da IA são essenciais e não podem ficar isentas de efetividade. A transparência é assegurada com a exigência de rastreabilidade e explicabilidade através de registo, documentação e informação, de modo a que a IA seja «explicável», na medida do possível, e não uma «*black-box*».

e) Diversidade, não discriminação e equidade
As exigências de diversidade, não discriminação e equidade impõem, desde a conceção e ao longo de todo o funcionamento, a participação de peritos e consultores com diferentes competências e que representem os potenciais *stakeholders* e eventuais afetados, pois essa é a solução para melhor responder a qualquer falha ou resultado indesejado[67]. Estas exigências também impõem que a recolha e seleção dos dados de treino do sistema de IA não padeça de limitações e enviesamentos, deliberados ou não.

f) Bem estar societal e ambiental
A exigência de sustentabilidade societal e ambiental impõe um cuidado com o impacto na sociedade e no ambiente e outros seres sencientes. Na primeira, salienta-se o especial cuidado com os processos democráticos e com as competências sociais; no segundo, refere-se a sustentabilidade e responsabilidade ecológica. Não se esquece a referência aos «ODS».

[67] Kobi Leins, Anja Kaspersen, *7 Myths of Using the Term "Human on the Loop…"*, cit. As autoras consideram que é sobretudo na fase de conceção que se deve assegurar essa participação diversa, pois no funcionamento será praticamente impossível a intervenção de um ser humano. Nos EUA, o *Blueprint for an AI Bill of rights* considera que essa diversidade de participação é uma exigência de segurança e declara, no princípio «*Safe and Effective Systems*», "*[a]utomated systems should be developed with consultation from diverse communities, stakeholders, and domain experts to identify concerns, risks, and potential impacts of the system*". White House Office of Science and Technology Policy, *Blueprint for an AI Bill of Rights: Making Automated Systems Work for the American People*, Washignton, October 2022, *in* https://www.whitehouse.gov/wp-content/uploads/2022/10/Blueprint-for-an-AI-Bill-of-Rights.pdf (31.10.2022).

g) Responsabilização

O princípio da responsabilidade desdobra-se em duas exigências: auditabilidade dos sistemas de IA, interna e externa, através de relatórios e avaliação de impacto, em particular quando possam ter impacto sobre direitos fundamentais; "reparação dos impactos adversos e injustos". O uso desta expressão mais flexível em vez de "ressarcimento de danos" é abrangente, e inclui preocupações como a minimização e comunicação dos impactos negativos, mediação compromissória e disponibilização de informação sobre vias de recurso para clientes e utilizadores[68]. Porém, a formulação vaga não acautela com certeza condições para uma tutela efetiva através da responsabilidade civil e/ou criminal e na proposta de «Regulamento IA» essa dimensão não foi contemplada, sendo objeto de uma harmonização muito limitada na proposta de diretiva apresentada em setembro de 2022.

O Grupo de Peritos delineou um quadro para a avaliação da confiança dos sistemas de IA[69], destinada, sobretudo, aos criadores e implantadores. Foi conduzida uma avaliação-piloto junto de 350 *stakeholders* no âmbito das plataformas «Aliança para a IA» e «AI4EU».

Finalmente, a Comissão compromete-se nesta *Comunicação* a prosseguir esforços, bilaterais e multilaterais[70], para que a abordagem regulatória europeia se torne o *gold standard* que reúna consenso à escala global, como sucedeu com a proteção de dados pessoais[71]. Por um lado, espera-se que as grandes empresas tecnológicas, mesmo situadas fora da UE, vão seguir os requisitos impostos na UE nos sistemas de elevado risco, pois não é racional ter diversos padrões (*de facto* «*Brussels effect*»). Por outro lado, outros Estados seguirão o modelo regulatório europeu (*de iure* «*Brussels effect*»). *Fora* como a ONU, o G7, o G20, a OCDE e o Conselho da Europa e mesmo novos formatos inclusivos de participação multilateral, tais como a Aliança para o Multilateralismo, o Fórum da Paz de Paris e a Cimeira «*Finance in Common*» são objeto de atenção[72]. Em 2019, A Comissão lançou o projeto *International alliance for a*

[68] Grupo de Peritos, *Orientações...*, cit., pp. 39-40.

[69] Grupo de Peritos, *Orientações...*, pp. 32-41.

[70] O Regulamento (UE) nº 234/2014 do Parlamento Europeu e do Conselho, de 11 de março de 2014, criou um Instrumento de Parceria para a cooperação com países terceiros (JO L 77 de 15.3.2014, p. 77).

[71] Charlotte Siegmann and Markus Anderljung, *The Brussels Effect and Artificial Intelligence: How EU regulation will impact the global AI market*, Centre for the Governance of AI, University of Oxford, 2022, *in* https://uploads-ssl.webflow.com/614b70a71b9f71c9c240c7a7/630534b77182a3513398500f_Brussels_Effect_GovAI.pdf (31.10.2022).

[72] Segundo o Alto Representante da União para os Negócios Estrangeiros e Política Segurança, na *Comunicação conjunta ao Parlamento Europeu e ao Conselho relativa ao reforço da contribuição da UE*

Human-centric Approach to Artificial Intelligence, e, em 2020, a iniciativa diplomática *International Outreach for human-centric Artificial Intelligence*, a fim de promover a visão europeia da IA no mundo. A última iniciativa da Comissão, em agosto de 2022, foi a Recomendação para negociar a convenção-quadro que está em preparação no Conselho da Europa[73]. Até ao momento, apenas foram adotados atos de *soft law* ou orientações éticas.

A orientação europeia e a orientação norte-americana sobre a regulação da IA estão a convergir[74]. Em 2020, o Congresso norte-americano aprovou o *National AI Initiative Act*[75] e com essa base o *National Institute of Standards and Technology* está a elaborar o *AI Risk Management Framework*. Em 2022, o governo norte-americano publicou o *Blueprint for an AI Bill of Rights. Making Automated Systems Work for the American People*. Neste documento identificam-se cinco princípios que devem orientar a conceção, utilização e desempenho de sistemas de IA, e inclui-se um manual (designado de *Technical Companion*) para a respetiva implementação na formulação de políticas e na produção e utilização desses sistemas (*From Principles to Practice*), dimensão em que se tem revelado mais complexo assegurar os princípios em torno dos quais se tem sedimentado o consenso. Recorre à ilustração com exemplos práticos e salienta que só uma cooperação *multistakeholder* entre a indústria, a ciência, os políticos e a sociedade civil poderá assegurar o desiderato pretendido[76].

para um multilateralismo assente em regras, Bruxelas, 17.2.2021 JOIN(2021) 3 final, "[a] ação da UE nas instâncias multilaterais terá de encontrar um equilíbrio delicado entre a necessidade de defender a soberania tecnológica e garantir a abertura da Internet e os direitos fundamentais. (...) a União deve colaborar com todos os parceiros internacionais para dar resposta aos atuais desafios em matéria de governação digital", p. 10, *in* https://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=CELEX:52021JC0003&from=PT (31.10-2022).

[73] Visa-se impedir a criação de um quadro jurídico internacional que vincule os EM à margem das orientações projetadas na UE, em especial quando se encontra em processo de decisão o «Regulamento IA». Esta pretensão encontra cobertura normativa no artigo 3º, nº 2 do TFUE.

[74] Segundo o *Politico.eu*, está em preparação um "*Joint roadmap on AI evaluation and measurement tools for [trustworthy] AI and risk managemen*t", *in* https://www.politico.eu/newsletter/digital-bridge/platforms-on-the-hook-transatlantic-ai-rulebook-lets-talk-data-transfers/ (31.10.2022). Porém, há ainda um caminho a prosseguir. Future of Life Institute, *Lessons from the NIST AI RMF for the EU AI Act Input for the US-EU TTC*, April 2022, *in* https://futureoflife.org/wp-content/uploads/2022/08/Lessons_from_NIST_AI_RMF-v2.pdf (31.10.2022).

[75] *In* https://www.congress.gov/116/crpt/hrpt617/CRPT-116hrpt617.pdf#page=1210 (31.10.2022). O National Institute of Standards and Technology (NIST) está a preparar o *AI Risk Management Framework*. No entanto, ao contrário da abordagem europeia, nos EUA o quadro é voluntário, ainda que salvaguarde a aplicação de normas vinculativas, nacionais ou internacionais.

[76] White House Office of Science and Technology Policy, *Blueprint for an AI Bill of Rights*, cit., p. 14. Além dos 10 princípios enunciados no *Canada's Digital Charter*, *in* https://ised-isde.

Inclui-se um glossário em que se esclarece o significado e alcance de diversos conceitos e definições, como «discriminação algorítmica» ou «sistema automático». Este documento salienta que abordagem normativa e regulatória da IA tem de ser norteada pela garantia dos direitos e liberdades individuais, sob a categoria norte-americana «*civil rights, civil liberties, and privacy*», mas não preconiza uma regulação imperativa da produção ou comercialização dos produtos equivalente à seguida na UE.

3.3. Livro Branco sobre a Inteligência Artificial – Uma Abordagem Europeia virada para a Excelência e a Confiança, 19.2.2020, COM(2020) 65 final

O Livro Branco apresenta as opções políticas da política regulatória para a IA para alcançar um ecossistema digital de confiança e excelência para a sociedade e economia no seu conjunto (cidadãos, empresas e serviços de interesse público).

No *Inception Impact Assessment*[77] resumem-se as diversas opções políticas, entre a não intervenção e cinco alternativas de intervenção, bem como os impactos económicos (benefícios e encargos), sociais, ambientais, administrativos e nos direitos fundamentais do regime a instituir em cada uma das alternativas regulatórias.

O *Livro Branco* apresenta múltiplas ações a desenvolver: 1. Trabalhar com os Estados-Membros e rever o plano coordenado; 2. Centrar os esforços da comunidade de investigação e inovação e criar de centros de excelência e de

canada.ca/site/innovation-better-canada/sites/default/files/attachments/1020_04_19-Website_Placemat_v09.pdf (31.10.2022), o quadro normativo vigente no Canadá tem um alcance muito mais limitado do que o regime proposto pela Comissão Europeia. Law Commission of Ontario and The Chair on Accountable AI in a Global Context, *Comparing European and Canadian AI Regulation*, November 2021, *in* https://www.lco-cdo.org/wp-content/uploads/2021/12/Comparing-European-and-Canadian-AI-Regulation-Final-November-2021.pdf (31.10.2022. Em 16.06.2022, foi proposto o *Digital Charter Implementation Act, 2022*, que vai avançar com regulamentação da IA mais transversal e geral, aproximando-se do modelo europeu. No Reino Unido a iniciativa apresentada em julho de 2022 aponta para uma abordagem setorial e não geral como é a da UE. UK Government, *Policy paper Establishing a pro-innovation approach to regulating AI*, 18.07.2022, *in* https://www.gov.uk/government/publications/establishing-a-pro-innovation-approach-to-regulating-ai/establishing-a-pro-innovation-approach-to-regulating-ai-policy-statement (5.11.2022).

[77] Uma análise de impacto exaustiva consta do *Commission Staff Document Impact Assessment Accompanying the Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts*, Brussels, 21.4.2021, SWD(2021) 84 final (ambos os documentos *in* https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12527-Inteligencia-artificial-Requisitos-eticos-e-legais_pt (31.10.2022).

teste que possam combinar investimentos europeus, nacionais e privados; 3. Definir Competências num Plano de Ação para a Educação Digital; 4. Dar relevo às PME e estabelecer Pólos de Inovação Digital e uma plataforma de IA a pedido; 5. Promover a parceria com o setor privado; 6. Incentivar o setor público a adotar a IA, em especial nos setores dos cuidados de saúde, das administrações rurais e dos serviços públicos; 7. Garantir o acesso aos dados e às infraestruturas de computação, articulando-se com a Estratégia Europeia de Dados; 8. Promover a cooperação e governação internacional multinível e em rede, entre os setores público e privado e no âmbito da ONU, OCDE, Conselho da Europa, G20 ou União Internacional de Telecomunicações[78].

Os principais desafios apontados são a falta de transparência, vulgarmente designada como «opacidade da IA», que, por sua vez, dificulta a identificação e a prova de possíveis violações da legislação, incluindo de disposições legais que protegem os direitos fundamentais, a imputação de responsabilidades e o preenchimento das condições para pedir uma indemnização. Prevalece a incerteza na responsabilidade ao longo da cadeia de valor. Por outro lado, além da permanente natureza evolutiva dos sistemas de IA, as tecnologias digitais configuram-se cada vez mais como «serviços» («DaaS» – «*Device as a Service*»), o que obriga a repensar o próprio conceito de segurança e a abordagem legislativa nesta matéria, pois a legislação europeia sobre segurança incide sobre produtos[79].

Em geral, a abordagem propugnada, além de flexível, deve basear-se no «risco», seja na eleição dos setores prioritários ou das utilizações de tecnolo-

[78] Têm-se multiplicado os documentos com propostas de princípios jurídicos e éticos para a IA. Além da OCDE, já referida, o G20 aprovou, em junho de 2019, a *G20 Ministerial Statement on Trade and Digital Economy*, *in* https://wp.oecd.ai/app/uploads/2021/06/G20-AI-Principles.pdf (31.10.2022). Em outubro de 2021, foi a vez da NATO aprovar *An Artificial Intelligence Strategy for NATO*, *in* https://www.nato.int/docu/review/articles/2021/10/25/an-artificial-intelligence-strategy-for-nato/index.html (31.10.2022). Em novembro de 2021, a Conferência Geral da UNESCO adotou a *Recommendation on the Ethics of Artificial Intelligence*, *in* https://unesdoc.unesco.org/ark:/48223/pf0000381137 (31.10.2022). Entre as instituições jurídicas da sociedade civil destacamos o documento do European Law Institute/Teresa Rodríguez de las Heras, *Ballell, Guiding Principles for Automated Decision-Making in the EU. ELI Innovation Paper*, 2022, *in* https://www.europeanlawinstitute.eu/fileadmin/user_upload/p_eli/Publications/ELI_Innovation_Paper_on_Guiding_Principles_for_ADM_in_the_EU.pdf (31.10.2022) Em 2017, o Future of Life Institute tinha elaborado os designados «*Asilomar AI Principles*», in https://futureoflife.org/2017/08/11/ai-principles/ (31.10.2022).

[79] A proposta sobre a responsabilidade por produtos defeituosos inclui no conceito de produto os serviços digitais. Nesta matéria, há ainda que considerar o «Regulamento Serviços Digitais» e o «Regulamento Mercados Digitais».

gias de IA. Os critérios de classificação do risco são dois e são, em princípio, cumulativos: a utilização nos setores em que "dadas as características das atividades tipicamente realizadas, se pode esperar que ocorram riscos significativos"; a finalidade ou o impacto da utilização. A Comissão admitia que algumas utilizações/potenciais impactos devem classificar-se como de «elevado risco» independentemente do setor, exemplificando com a utilização de sistemas biométricos.

O *Livro Branco* propugna uma regulação dos sistemas de «alto risco» em relação aos elementos seguintes: dados de treino; conservação de registos e de dados; prestação de informações; robustez e exatidão; supervisão humana; requisitos específicos para determinadas aplicações de IA, tais como as utilizadas para fins de identificação biométrica à distância.

Algumas questões controvérsias se levantam, tal como a compatibilização do registo e da conservação de dados com as regras da proteção de dados pessoais ou a proteção de outros direitos e interesses que condicionam essa conservação.

O *Livro Branco* foi acompanhado pelo *Relatório sobre as implicações em matéria de segurança e de responsabilidade decorrentes da inteligência artificial, da Internet das coisas e da robótica*.

3.4. Comunicação da Comissão "*Fomentar uma Abordagem Europeia da Inteligência Artificial*", 21.04.2021, COM(2021) 205 final

Este é o último documento do processo de elaboração da política regulatória da IA na UE e foi publicado na mesma data em que foi apresentada a proposta de «Regulamento IA». A Comissão relembra a importância estratégica das tecnologias de IA para a UE, sumaria a abordagem a seguir e o quadro da regulamentação da proposta. O relevo da experiência da pandemia Covid-19 e do Mecanismo de Recuperação e Resiliência são considerados na análise.

Em anexo, encontra-se a revisão do Plano Coordenado de 2018, na qual se preveem quatro eixos para o desenvolvimento da estratégia europeia de coordenação das políticas e dos investimentos no domínio da IA: a) criar condições propícias ao desenvolvimento e à adoção da IA na UE (reforço da cooperação e desenvolvimento das infraestruturas, em especial nos dados e capacidade de computação); b) fazer da UE um local onde a excelência impera «do laboratório ao mercado» (financiamento de redes de excelência, parcerias e polos digitais); c) assegurar que a IA está ao serviço das pessoas e é uma força positiva para a sociedade (sustentável, segura, inclusiva, acessível e fiável, respeitadora dos valores e direitos fundamentais); d) reforçar a liderança estratégica em setores de impacto elevado (alterações climáticas e

ambiente, saúde, setor público, robótica, mobilidade, segurança e assuntos internos, e agricultura). Para a respetiva implementação, propõe 40 ações a concretizar pela UE e EM.

4. A Regulação da Inteligência Artificial

4.1. Proposta de Regulamento («Regulamento IA»), 21.4.2021, COM(2021) 206 final

A proposta de Regulamento para a IA foi apresentada em 21 de abril de 2021. Tem sido objeto de um significativo debate. O procedimento legislativo ainda decorre e está longe da conclusão[80]. A Presidência eslovena elaborou um documento de compromisso que a reformulou significativamente, com o objetivo de construir um consenso entre os governos dos EM. A Presidência francesa apresentou um segundo documento de compromisso que respondeu a algumas insuficiências relevantes. Até 02.04.2023, apenas seis câmaras legislativas nacionais tinham apresentado os respetivos pareceres à luz do princípio da subsidiariedade, entre eles a Assembleia da República portuguesa. Em 14.06.2023, o PE aprovou a sua posição, em primeira leitura, propondo numerosas emendas que serão objeto de negociação com o Conselho.[81]

Esta análise elege somente alguns pontos essenciais mais controversos e é apenas preliminar.

Como instrumento do tipo «*New Legislative Framework*» («NLF»), a proposta de regulamento pretende instituir um regime global para a produção, disponibilização e utilização de sistemas de IA. A opção foi por um instrumento legislativo horizontal aplicável uniformemente em toda a União, ainda que a Irlanda e a Dinamarca beneficiem de um regime diferenciado. Tem um âmbito de aplicação genérico no âmbito de utilizações não exclusivamente particulares, embora se excluam os sistemas desenvolvidos ou utilizados com

[80] O Comité Económico e Social Europeu e o Comité das Regiões já deram o respetivo Parecer. O BCE, no respetivo Parecer, apresentou várias sugestões de clarificação sobre a supervisão. Ao invés da generalidade das opiniões, propôs uma abordagem menos restritiva, que inclui a exclusão da categoria de sistemas de IA de risco elevado concebidos para serem utilizados para avaliar a capacidade de endividamento de pessoas singulares ou estabelecer a sua pontuação de crédito, e que tirem partido da utilização autónoma da regressão linear ou logística ou de árvores de decisão sob supervisão humana desde que o impacto dessas abordagens na avaliação da solvabilidade ou da pontuação de crédito das pessoas singulares seja reduzido, *in* https://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=CELEX:52021AB0040&from=EN (31.10.2022). Após a aprovação da posição do PE, em maio de 2023, há a expectativa de que possa concluir-se a adoção do Regulamento até ao final de 2023.

[81] Documento disponível em https://www.europarl.europa.eu/doceo/document/TA-9-2023-0236_EN.html

fins exclusivamente militares e abrangidos no âmbito da PESC[82]. Vincula os EM e a UE, suas instituições, órgãos e organismos. Por outro lado, integra um programa legislativo mais vasto e deve ser articulado com a proposta de Regulamento relativo à segurança geral dos produtos [COM(2021) 346] e a proposta de Regulamento relativo às máquinas e seus componentes e acessórios [COM(2021) 202 final]. Tem como base jurídica os artigos 16º (proteção de dados) e 114º TFUE (harmonização legislativa do mercado interno).

A proposta avança uma definição de IA que pretende ser tecnologicamente neutra e funcional. Apresenta-se como "humanamente determinada", pois não considera que os objetivos possam ser definidos de modo autónomo pelo sistema de IA. Já o PE afasta-se desse condicionalismo e admite que os sistemas de IA possam funcionar segundo objetivos "implícitos", não definidos explicitamente por seres humanos. O PE é muito mais exaustivo nas características que inclui.

Os objetivos gerais da proposta são dois: aumentar a transparência e minimizar os riscos dos sistemas de IA. Coloca no centro das preocupações um nível elevado de proteção de interesses públicos, como a saúde, a segurança e a proteção dos direitos fundamentais. Contudo, para assegurar que a IA se desenvolve segundo esta «marca europeia» a abordagem é menos repressiva do que orientadora, pretendendo sobretudo impedir a consolidação de *standards* não europeus[83], embora não se limite a orientações éticas, como sucede em outras jurisdições, antes inclua proibições para sistemas de IA de risco inaceitável e regras imperativas para sistemas de IA de elevado risco. Como sucedeu com o «RGPD», a UE assume o protagonismo internacional.

Os objetivos gerais densificam-se em quatro objetivos específicos: garantir que os sistemas de IA colocados no mercado da União e utilizados são seguros e respeitam a legislação em vigor em matéria de direitos fundamentais e valores da União; garantir a segurança jurídica para facilitar os investimentos e a inovação no domínio da IA; melhorar a governação e a aplicação efetiva da legislação em vigor em matéria de direitos fundamentais e dos requisitos de segurança aplicáveis aos sistemas de IA; facilitar o desenvolvimento de um

[82] Não se aplica igualmente a entidades públicas exteriores à União e a organizações internacionais que utilizem sistemas de IA no âmbito de acordos com a União ou os EM. Ainda que se possa compreender como decorrência do respeito pela soberania dos Estados terceiros e autonomia das organizações internacionais, o Comité Europeu para a Proteção de Dados e a Agência Europeia para a Proteção de Dados criticaram esta exclusão. *Parecer conjunto 5/2021 do CEPD e da AEPD*, cit.

[83] Oreste Pollicino and Giovanni De Gregorio, "Constitutional Law...", cit., p. 13.

mercado único para as aplicações de IA legítimas, seguras e de confiança e evitar a fragmentação do mercado. A conjugação destes objetivos pode resumir-se na expressão «inovação responsável», utilizada no preâmbulo da proposta.

Estabelecem-se regras harmonizadas para a colocação no mercado, a colocação em serviço e a utilização de sistemas de IA na União Europeia, amparadas por regras de fiscalização e vigilância no mercado.

O princípio estruturante da intervenção é o da intervenção limitada e dinâmica, fundada na classificação dos riscos e respetiva eliminação/mitigação baseada no tipo de risco e no princípio da proporcionalidade, completada por códigos de conduta a seguir voluntariamente para os sistemas de IA que não são de risco elevado, seguindo a «opção 3+» entre as alternativas de intervenção avaliadas.

O quadro regulatório é determinado pelos diferentes níveis de risco, à luz dos seguintes princípios fundamentais: a) princípio da universalidade, segundo o qual o regime previsto se aplica a todos os operadores, estabelecidos na UE ou fora, e a todos os utilizadores que se encontram na UE, e em função do local (na UE) onde ocorrem os efeitos ou o resultado da intervenção do sistema de IA; b) princípio da precaução, que, verificados certos requisitos, impõe a testagem e controlo *ex ante*; c) princípio da transparência, que impõe a rastreabilidade/não opacidade e controlo *ex post*; d) princípio da responsabilidade do fornecedor, i.e. a pessoa que coloca no mercado ou em serviço de um sistema de IA de risco elevado, independentemente de ser ou não a pessoa que concebeu ou desenvolveu o sistema (prevê-se obrigações específicas, nomeadamente de cooperação, e salvaguarda da responsabilidade de intermediários, nos termos da diretiva 2000/31 e regulamento dos serviços digitais que a substituiu, e utilizadores profissionais).

A natureza do risco fica dependente de dois critérios: o setor e a finalidade[84]. De acordo com o nível de risco, o regime é variável entre a proibição e a transparência.

O Título II da proposta inclui as práticas de IA proibidas:

– práticas que tenham um potencial significativo para manipular as pessoas por meio de técnicas subliminares que lhes passam despercebidas ou para explorar as vulnerabilidades de grupos específicos, como as crianças ou as pessoas com deficiência, para distorcer substancialmente o seu comportamento de uma forma que seja suscetível de causar danos psicológicos ou físicos a essa ou a outra pessoa;

[84] A OCDE aponta como critérios a gravidade dos efeitos adversos e o respetivo âmbito. "OECD Framework for the classification of AI systems", cit. p. 67.

- sistemas que permitam às autoridades públicas desenvolver atividades de «classificação social» genérica ou para fins diferentes daqueles para que foram legitimamente recolhidos com efeitos prejudiciais ou desfavoráveis.
- as práticas de identificação biométrica à distância (por exemplo, instrumentos de reconhecimento facial para identificar transeuntes em espaços públicos) em tempo real são, em princípio, proibidas, exceto nas situações especiais previstas e sujeitas a autorização prévia concedida por uma autoridade judiciária ou por uma autoridade administrativa independente (este regime é direito especial em relação à Diretiva 2016/680)[85]. A autorização deve passar por um processo de avaliação da conformidade *ex ante*, realizado por um organismo notificado para verificar a conformidade com os requisitos aplicáveis aos sistemas de IA de risco elevado, e ficará sujeita a requisitos mais exigentes em termos de registo e supervisão humana. Para outras finalidades, aplicam-se a proibição e as exceções que constam do «RGPD».

O «Documento de Compromisso da Presidência eslovena» introduz algumas melhorias. Em relação às práticas manipuladoras, não faz depender a proibição da intenção manipuladora (como resulta da expressão "para distorcer"), proibindo quer as práticas com tal objetivo ou simplesmente com tal efeito[86]. Contudo, a proteção é insuficiente. Deverá ir-se mais longe e contemplar o impacto negativo mais vasto sobre a autonomia individual, sobre grupos e sobre a sociedade[87] (nomeadamente, os valores previstos no artigo 2º TUE) e não simplesmente acautelar danos físicos e psicológicos de indivíduos[88]. Deve ir-se além do efeito lesivo físico ou psicológico sobre pessoas

[85] O seu âmbito de aplicação remete para a Decisão-quadro do Mandado de Detenção Europeu e, portanto, abrange os 32 tipos de crime aí previstos, entre os quais crimes económico-financeiros, o que pode ser excessivo.

[86] O requisito da intenção poderia tornar impossível alcançar o objetivo, não apenas porque é difícil de provar, como podem os sistemas ser passíveis de usos distintos a configurar pelo utilizador, sendo a possibilidade de reconfiguração cada vez mais comum, sob o eufemismo de «democratização da tecnologia». Consequentemente, a proibição poderia ser ineficaz. M. Veale, F. Z. Borgesius, "Demystifying the Draft EU Artificial Intelligence Act. Analysing the good, the bad, and the unclear elements of the proposed approach", *Computer Law Review International*, 4/21, 97-112, *in* https://osf.io/preprints/socarxiv/38p5f/ (31.10.2022).

[87] Sobre a manipulação na tecnologia e os respetivos perigos, D. Susser, B. Roessler, H. Nissenbaum, "Technology, autonomy, and manipulation", *Internet Policy Review*, 8(2), 2019, DOI: 10.14763/2019.2.1410, *in* https://policyreview.info/pdf/policyreview-2019-2-1410.pdf (31.10.2022).

[88] O CEPD e a AEPD também apontaram esta crítica.

e grupos vulneráveis. Deverá simplificar-se, não visando apenas as «técnicas subliminares» de manipulação mas qualquer técnica manipulativa[89]. Não corresponde às exigências de uma «IA humana» admitir como lícitas técnicas manipulativas, desde que não sejam subliminares. Mais, deve introduzir-se uma cláusula genérica de proteção de estados de vulnerabilidade de qualquer pessoa, ainda que se tenha alargado as causas de vulnerabilidade à situação económica e social[90]. Sendo um requisito a verificação ou suscetibilidade de dano concreto sobre pessoas concretas, a não comprovação de um resultado danoso identificável para alguém em particular exclui a proibição de sistemas com intento manipulativo, mas que produzem um dano em resultado de efeitos cumulativos, a que são especialmente sensíveis as crianças[91]. Ainda que outros atos[92], como o «RGPD», o «DMA» ou o «DSA» possam cuidar desses riscos, o instrumento geral de regulação da IA deve contemplar essa proteção alargada, em especial quando grande parte dos sistemas de IA não ficarão submetidos a exigências regulatórias imperativas, só aplicáveis aos que cabem na categoria de «risco elevado». Os que não caem nessa categoria, não são isentos de risco e podem atentar contra esses valores, o que, em virtude da sua ação reiterada, pode acarretar um «risco sistémico» para a sociedade, em especial o processo democrático e a igualdade[93].

Este documento acolhe as críticas do CEPD e da AEPD e proíbe as práticas de «classificação social», algo obviamente contrário ao artigo 21º da CDF, independentemente de serem conduzidas por autoridades públicas, tal como

[89] Risto Uuk, *Manipulation and the AI Act,* April 2022, *in* https://futureoflife.org/wp-content/uploads/2022/08/FLI-Manipulation_AI_Act.pdf (31.10.2022). A ambiguidade do conceito de «técnica subliminar» foi salientada, ainda que haja índices para a sua verificação (p. 4).

[90] A vulnerabilidade não é apenas uma condição estrutural de certas categorias ou grupos, mas é uma condição que todos os seres humanos podem sofrer em algum momento, fruto de circunstâncias contextuais e pontuais. Martha Fineman, "The Vulnerable Subject and the Responsive State", Emory Law Journal, 60(2), 251–275, *in* https://scholarlycommons.law.emory.edu/elj/vol60/iss2/1 (31.10.2022). Através da interação com as pessoas, as tecnologias de IA podem detetar esses momentos de vulnerabilidade (estados depressivos ou eufóricos, por ex.) e em vez de operarem para a sua exploração deveriam acautelar essa situação, de acordo com um princípio de precaução ou até de lealdade.

[91] M. Veale, F. Z. Borgesius, "Demystifying the Draft EU Artificial Intelligence Act. Analysing the good, the bad, and the unclear elements of the proposed approach", *Computer Law Review International*, 4/21, 97-112. Os autores fazem uma crítica severa da proposta de regulamento e falam de "retórica" da Comissão (p. 99), *in* https://osf.io/preprints/socarxiv/38p5f/ (31.10.2022).

[92] Federico Galli, "AI and Consumers Manipulation: What the Role of EU Fair Marketing Law?", *Católica Law Review 4*, no. 2 (May 1, 2020), pp. 35–64, *in* https://doi.org/10.34632/catolicalawreview.2020.9320 (31.10.2022).

[93] Risto Uuk, *Manipulation and the AI Act*, cit., p. 5.

preconizado pelo CEPD e AEPD. Tal é de saudar, ainda que possa ter um impacto significativo em práticas já correntes de marketing e não deixe manter ambiguidades, como sucede nos *multi-stage profilling* e nos «*dual use*» *systems*[94].

Não prevê a proibição os sistemas de «policiamento preditivo», que o CEPD e a AEPD consideram contrários à dignidade humana. De igual modo, não bane a utilização de sistemas de IA para inferir as emoções de um indivíduo, cuja proibição aqueles organismos preconizaram.

O PE propõe modificações substanciais, acrescentando diversas proibições: a identificação biométrica remota em tempo real em espaços acessíveis ao público; a identificação biométrica diferida, exceto na investigação criminal com autorização judicial; a categorização biométrica com base em características individuais, como género, raça, etnia, cidadania, religião ou orientação política; sistemas de reconhecimento de emoções no local de trabalho, na gestão fronteiriça, em instituições de ensino, a recolha indiscriminada de dados biométricos das redes sociais e circuitos internos de televisão; práticas de policiamento preditivo, com base em perfis, localização ou práticas passadas, pois, além de estarem comprovados riscos sérios de discriminação, tal põe em causa a presunção da inocência e a dignidade humana. Esclarece que a lista de proibições prevista não afeta as proibições resultantes de outra legislação, seja de proteção de dados pessoais, seja de proteção do consumidor ou de proibição de discriminação.

O documento da Presidência eslovena não limitava as condições impostas à identificação biométrica à sua prática remota, respondendo em parte a críticas feitas, nomeadamente pelo Comité Europeu para a Proteção de Dados (CEPD) e pela Agência Europeia para a Proteção de Dados (AEPD); porém, estende as justificações para o seu uso legítimo (por ex. não exigindo a ameaça iminente ou incluindo a finalidade de proteção da saúde ou de infraestruturas críticas) e admitia que em situações de urgência possa avançar sem a autorização judiciária ou independente. Esta possibilidade suscitava sérias reservas e agravava um regime que já era objeto de críticas na fórmula proposta pela Comissão. Por outro lado, introduzia uma cláusula que exceciona a sujeição a esse regime quando ocorre o consentimento da pessoa, regime que nos suscita as maiores reticências. O CEPD e a AEPD defenderam que

[94] Pense-se nas plataformas sociais em que a manipulação envolve fatores alheios ao operador ou fornecedor do sistema (segundo o considerando 16, nesse caso a intenção de manipulação não se presume), como os sistemas de recomendação e *rating* com *biased results* já comprovados como sucede com anfitriões negros.

se banisse genericamente o reconhecimento pessoal em espaços públicos em qualquer contexto, remoto ou não, incluindo os sistemas de identificação à distância em grande escala em espaços *online* seja o reconhecimento de rostos, marcha, impressões digitais, ADN, voz, toques de teclas e outros sinais biométricos ou comportamentais[95]. Numa IA centrada no ser humano e respeitadora dos valores da UE e dos direitos fundamentais, deveria prevalecer o princípio da precaução[96]. Outros problemas se levantam. O regime proposto permite que se venda e coloque no mercado para utilização fora da UE com essa finalidade proibida na União (poderão ser abrangidos pelo Regulamento (UE) 2021/821 que cria um regime da União de controlo das exportações, corretagem, assistência técnica, trânsito e transferências de produtos de dupla utilização). Mais, não proíbe esse uso *online*, na medida em que este não se integre no conceito de «espaço público», ou de "espaço acessível ao público", na formulação do PE",embora tal se possa incluir nas outras categorias de risco e não deixa de estar sujeito ao «RGPD»[97].

Os títulos III e IV da proposta da Comissão preveem o regime aplicável aos sistemas de IA de elevado risco.

A classificação de um sistema de IA como um «sistema de risco elevado» baseia-se nos seguintes critérios: domínio de utilização e finalidade prevista, gravidade e probabilidade de ocorrência dos possíveis danos.

São qualificados como sistemas de IA de risco elevado o sistema que se destine a ser utilizado como um componente de segurança de um produto ou o sistema que é, eles mesmo, um produto abrangido pela legislação de harmonização da União enumerada no anexo II, e deva ser sujeito a uma avaliação da conformidade por terceiros com vista à colocação no mercado ou à colocação em serviço, mesmo que seja colocado no mercado ou em serviço separado desses produtos. O Anexo II elenca 19 atos. O Anexo III contém um elenco de domínios de utilização de sistemas de IA que lhe conferem a natureza de «elevado risco»: identificação biométrica e categorização de pessoas singulares (em tempo real e diferido); gestão e funcionamento de infraestruturas

[95] *Parecer conjunto* 5/2021, cit., p. 12-14.

[96] Há relatos de pressão da indústria sobre o Grupo de Peritos de Alto Nível e na preparação do *Livro Branco*, fazendo cair orientações defendidas pelo CEPD e AEPD, como a proibição total do reconhecimento facial. M. Veale, F. Z. Borgesius, "Demystifying the Draft EU Artificial Intelligence Act", cit., p. 98.

[97] Para uma perspetiva crítica, M. Veale, F. Z. Borgesius, "Desmystifying the Draft EU Artificial Intelligence Act...", cit., p. 102. Segundo os autores as obrigações em relação a sistemas que façam categorização biométrica ou de reconhecimento de emoções não acrescentam nada significativo ao 13º RGDP (p. 107).

críticas; educação e formação profissional; emprego, gestão de trabalhadores e acesso ao emprego por conta própria; acesso (e usufruto) a serviços privados e a serviços e prestações públicas essenciais; manutenção da ordem pública; gestão da migração, do asilo e do controlo das fronteiras; administração da justiça e processos democráticos. Em cada um deles, são indicadas as utilizações abrangidas.

A adaptabilidade do regime à evolução tecnológica fica assegurado com a Comissão a poder aditar sistemas de IA à listagem do Anexo III, através de atos delegados, quando preenchidas condições respeitantes ao domínio de utilização e ao risco equiparável de impacto adverso sobre os direitos fundamentais. A questão é se a Comissão terá a «agilidade» necessária, considerando a velocidade da inovação neste domínio.

O *Draft Report* do PE acrescentou algumas utilizações no Anexo III, nomeadamente os serviços de internet no âmbito das infraestruturas críticas (surpreendentemente, a proposta da Comissão não colocou as comunicações entre estas). Acompanhou a crítica relativa à proteção das crianças e incluiu entre os sistemas de elevado risco os sistemas a utilizar por crianças e que tenham um significativo impacto no seu desenvolvimento pessoal cognitivo ou emocional. Ainda assim, a sistematização que integra estes sistemas no domínio educativo não é apropriada, dado que esses sistemas frequentemente são utilizados em produtos e serviços de entretenimento. O PE alargou a qualificaçao de risco elevado, incluindo o potencial de dano sobre a saúde, a segurança, os direitos fundamentais e o ambiente. Coloca igualmente entre os sistemas de IA de elevado risco os sistemas utilizados em ações de campanha partidária e nos sistemas eleitorais, incluindo para a contagem de votos, prevendo obrigações para os sistemas de recomendação das grandes plataformas.

Por sua vez, o documento da Presidência eslovena colocou mais um ato no Anexo II, introduziu modificações para precisar as práticas integradas no Anexo III (particularmente, associando à categoria de autoridades públicas outras que ajam em sua representação) e acrescentou outras, entre as quais se salienta, no domínio das infraestruturas críticas, a inclusão da proteção do ambiente[98] e a infraestrutura digital, ou, entre os serviços essenciais, o domínio dos seguros (também incluído no documento do PE). Lamentavelmente, não acompanhou a preocupação com os sistemas utilizados por crianças.

Os sistemas de IA de risco elevado têm de respeitar um conjunto de requisitos específicos, que incluem a utilização de conjuntos de dados de alta quali-

[98] A ausência de preocupação com os riscos ambientais e o respetivo controlo na avaliação do risco dos sistemas de IA da parte da Comissão mereceu crítica, que encontrou uma primeira resposta no Documento da Presidência eslovena e depois por parte do PE.

dade, a elaboração de documentação adequada para melhorar a rastreabilidade, a partilha de informações importantes com o utilizador, a conceção e aplicação de medidas de supervisão humana adequadas, bem como a satisfação dos mais elevados graus de exigência em termos de solidez, proteção, cibersegurança e exatidão («*by concept*» e «*by defect*»). Em geral, seguem as regras gerais dos produtos, mas devem ser avaliados quanto à conformidade com estes requisitos antes de serem colocados no mercado ou colocados em serviço (nos sistemas biométricos a avaliação deve ser feita por terceiros independentes quando o sistema de IA não siga integralmente normas harmonizadas ou especificações comuns), prevendo-se a sujeição a testes[99]. A não sujeição de todos a uma avaliação independente é criticada pelo CEPD e pela AEPD.

Há que assegurar que a aprendizagem autónoma não se traduza numa modificação que se exima a uma reavaliação. Com esse objetivo impõe-se a criação de um sistema de gestão de riscos através de um processo iterativo contínuo, executado ao longo de todo o ciclo de vida do sistema de IA, o que requer atualizações regulares sistemáticas com vista a eliminar ou mitigar os riscos, bem como a obrigação de comunicação de incidentes graves e anomalias que constituam um incumprimento de obrigações impostas pela legislação da União destinada a proteger os direitos fundamentais.

As medidas de gestão de riscos dos sistemas de IA de risco elevado podem visar, quando possível, a eliminação ou redução do risco, a atenuação dos riscos que não possam ser eliminados, bem como a prestação de informações e de formação aos utilizadores. O cuidado com os dados de treino, validação e teste implica um conjunto de exigências quanto à escolha, recolha, preparação e acompanhamento. O tratamento de dados pessoais com o objetivo de controlar, detetar e corrigir enviesamentos, é objeto de um regime especial, ficando a coberto do artigo 9º do RGPD, do artigo 10º da Diretiva (UE) 2016/680 e do artigo 10º, nº 1, do Regulamento (UE) 2018/1725, embora com a proteção dos direitos fundamentais através de pseudonimização e cifragem. O *Draft Report* do PE recusou esse regime especial e admite antes que o utilizador titular ou detentor dos dados possa ser responsável pela violação do regime previsto para os dados.

O objetivo da rastreabilidade impõe que os sistemas de IA de elevado risco assegurem o registo de eventos ao longo do seu ciclo de vida, com especial

[99] Os sistemas de IA «*embedded*» em máquinas com funções de segurança ficam sujeitos a uma certificação independente ao serem incluídos pela proposta de «Regulamento máquinas» no respetivo Anexo I (Máquinas e seus componentes e acessórios de alto risco, n.os 24 e 25), mesmo que sigam standards harmonizados.

exigência para os sistemas biométricos, onde se inclui o registo das pessoas que verificam os resultados. O objetivo de explicabilidade traduz-se na imposição de obrigações de transparência e informação aos utilizadores sobre as características, capacidades e limitações do sistema de IA, bem como as medidas de supervisão humana.

A supervisão humana sobre os sistemas de IA de elevado risco pode ter alcances diferenciados, seja assegurando a explicabilidade, seja impondo a confirmação humana de um resultado, seja revertendo a decisão do sistema ou interrompendo o seu funcionamento. Exclui-se a modalidade «*HOOL*». Porém, permanecem incertezas e ambiguidades.

As obrigações previstas são impostas a fornecedores, importadores, distribuidores, utilizadores profissionais e outras partes. Quando não se identifique um importador na União, os fornecedores exteriores estabelecidos fora da União devem ter um mandatário, estabelecido por escrito, na União, a quem compete manter informação e representar o fornecedor junto das autoridades.

No Título IV, acrescem regras de transparência para sistemas de IA concebidos para interagir com pessoas singulares. Devem informar as pessoas de que estão a interagir com um sistema de IA ou um conteúdo artificial – robôs de conversação, sistemas de reconhecimento de emoções, sistemas de categorização biométrica ou de manipulação de conteúdos de imagem, áudio ou vídeo («falsificações profundas»), salvo se tal se revelar óbvio dadas as circunstâncias[100] e o contexto de utilização (é excluída a esfera penal). A *disclosure obligation* de sistemas que criam conteúdos sintéticos está prevista para os utilizadores, não para os *providers*, o que deixa falhas possíveis, pois aqueles podem não estar plenamente conscientes da natureza desses sistemas[101]. A isenção no âmbito da criação artística ou da utilização pessoal está longe de ser clara, em especial quando a confusão entre a atividade puramente pessoal e uma atividade económica é cada vez maior, como sucede com a atividade de *influencer*. O CEPD e a AEPD consideraram que a utilização deste tipo

[100] Estará em formação um «direito à verdade» no mundo virtual?

[101] Deve aplicar-se aqui a distinção entre «*frontend operator*» e «*backend operator*». Christiane Wendehorst, "Liability for Artificial Intelligence. The Need to Address Both Safety Risks and Fundamental Rights Risks", in Silja Voeneky, Philipp Kellmeyer, Oliver Mueller and Wolfram Burgard (Eds.), *The Cambridge Handbook of Responsible Artificial Intelligence. Interdisciplinary Perspectives*, Cambridge, Cambridge University Press, 2022, pp. 187-209, p. 201. Disponível em https://doi.org/10.1017/9781009207898.016 (27.11.2022). Os produtores e programadores de sistemas de IA não se limitam a definir originariamente o sistema e continuam a atualizá-lo e a controlar sua «alimentação» e funcionamento, correspondem a «backend operators».

de sistemas devia ser objeto de maior restrição[102]. Recomendam a proibição de sistemas de IA que categorizem os indivíduos com base em qualquer dos fatores proibidos pelo artigo 21º da CDF. Proibição semelhante defendem para os sistemas de deteção de emoções, que só deveriam ser possíveis em casos específicos, tal como razões de saúde. O *Draft Report* do PE reforçava esta exigência de transparência com uma obrigação de informação aos seres humanos afetados que estão a ser objeto de uma utilização de sistemas de IA de risco elevado, que se estende até a criações artísticas.

Caso estejam em conformidade com standards harmonizados publicados no Jornal Oficial da UE há uma presunção, os sistemas de IA beneficiam de uma presunção de cumprimento dos requisitos impostos[103].

Prevê-se (Título VII) o registo obrigatório dos sistemas de IA de risco elevado autónomos numa base pública da Comissão, antes da sua colocação no mercado ou em serviço.

Os sistemas de IA que não são de risco elevado não são sujeitos a regras obrigatórias, mas são incentivados (Título IX) a desenvolver quadros de auto-regulação através de códigos de conduta voluntários modelados pelo regime dos sistemas de IA de risco elevado.

A proposta de regulamento propõe, no apoio à inovação (Título V), as «*regulatory sandboxes*», i.e. ambientes controlados de testagem de conformidade com a regulamentação das tecnologias inovadoras durante um período limitado, habitualmente designados como «Zonas Livres Tecnológicas (ZLT)» (Título V), a partir dos quais se poderão elaborar manuais de boas práticas e orientações na UE. Estas possibilitam o acesso a polos de inovação digital e o acesso a instalações para realização de testes e experimentação, o que ajudará as empresas inovadoras, as PME e as empresas em fase de arranque a continuarem a inovar – v.g. as Instalações de Ensaio e Experimentação de Referência ou a Plataforma «IA a pedido» [104]. Estas soluções colocam problemas contro-

[102] Esta crítica é acompanhada por G. Maigieri e M. Ienca, "The EU regulates AI but forgets to protect our mind", *EU Law Blog*, 7 July 2021, *in* https://europeanlawblog.eu/2021/07/07/the-eu-regulates-ai-but-forgets-to-protect-our-mind/ (31.10.2022).

[103] Esse empreendimento está em curso e já há alguns standards concluídos: https://www.iso.org/committee/6794475/x/catalogue/p/0/u/1/w/0/d/0 (31.10.2022). https://standards.cencenelec.eu/dyn/www/f?p=205:105:0::::: (31.10.2022). Uma panorâmica dos standards internacionais para a IA pode encontrar-se em Peter Cihon, *Standards for AI Governance: International Standards to Enable Global Coordination in AI Research & Development*, Future of Humanity Institute, University of Oxford, 2019, *in* https://www.fhi.ox.ac.uk/wp-content/uploads/Standards_-FHI-Technical-Report.pdf (31.10.2022).

[104] Sobre esta abordagem regulatória, https://www.europarl.europa.eu/RegData/etudes/STUD/2020/652752/IPOL_STU(2020)652752_EN.pdf (31.10.2022).

versos e não são abordadas questões relevantes, nomeadamente em relação à utilização e ao tratamento de dados pessoais, que foram objeto de especial atenção pelo CEPD e pela AEPD. Estas têm sido utilizadas em especial no domínio financeiro, junto e sob a supervisão das autoridades de supervisão financeira[105]. A primeira *regulatory sandbox* para a IA, uma «*pilot sandbox*» para testar tecnologia de IA de acordo com a proposta de regulamento, foi apresentada em junho deste ano pela Comissão e o governo de Espanha[106].

No que respeita à governação, o primeiro nível de governação efetiva regulada abrange os próprios criadores/fornecedores de sistemas de IA e o Anexo VI da proposta prevê um procedimento de avaliação de conformidade baseado no controlo interno. Não obstante, a proposta não propõe um quadro de referência para a auto-avaliação do risco, ao invés do que sucede nos EUA[107]. O Grupo de Peritos estruturou um quadro de avaliação a implementar internamente e a aplicar nos diversos níveis de decisão empresarial, do Conselho de Administração até ao nível operacional, passando pelos departamentos jurídico e de responsabilidade social.

A governação do ecossistema de IA proposto pela Comissão envolve organismos europeus e nacionais, num quadro em rede multinível[108]. Entre os primeiros encontra-se a Comissão Europeia, o Comité Europeu para a IA e a Autoridade Europeia de Proteção de Dados. O PE propoe o reforço significativo do Comité para a IA, convertendo-o numa verdadeira agência europeia, com personalidade jurídica, uma estrutura organizativa reforçada, com competências que vão da supervisão e monitorização do respeito pelo regula-

[105] Financial Conduct Authority, *Regulatory Sandbox*, 27.03.2022, *in* https://www.fca.org.uk/firms/innovation/regulatory-sandbox (31.10.2022). J. J. Goo e J.-Y Heo, "The Impact of the Regulatory Sandbox on the Fintech Industry, with a Discussion on the Relation between Regulatory Sandboxes and Open Innovation", *in* https://scholar.google.pt/scholar_url?url=https://www.mdpi.com/2199-8531/6/2/43/pdf&hl=pt-PT&sa=X&ei=6xhEY4v3N4n8mgGB8Z2ADw&scisig=AAGBfm1Y4OXVjEElDMYqTCAVK633jEmpGQ&oi=scholarr (31.10.2022).

[106] European Commission, *First regulatory sandbox on Artificial Intelligence presented*, 27.06.2022, *in* https://digital-strategy.ec.europa.eu/en/news/first-regulatory-sandbox-artificial-intelligence-presented (31.10.2022). Os testes iniciam-se em outubro de 2022 e os primeiros resultados serão apresentados durante a presidência espanhola no segundo semestre de 2023. Em Portugal, o DL nº 67/2021, de 30 de julho, institui o primeiro enquadramento legal das designadas ZLT (zonas livres tecnológicas). Em Portugal, foi criada pela Marinha a primeira Zona Livre Tecnológica em julho de 2022 – a Zona Livre Tecnológica (ZLT) Infante D. Henrique (Portaria nº 189/2022 de 25 de julho, DR 1ª Série, p. 6) e localiza-se nos concelhos de Sesimbra, Setúbal e Grândola.

[107] Nos EUA, o *National Institute of Standards and Technology* está a elaborar o *AI Risk Management Framework*.

[108] Foi considerado demasiado complexo. Law commission of Ontario and the Chair of Accountable AI in Global Context, *Comparing...*, p. 33.

mento, à promoção da literacia financeira, passando pela elaboração de orientações, coordenação das autoridades nacionais, investigações conjuntas e até mediação de divergências entre aquelas. Entre os segundos, prevê-se entre as autoridades competentes nacionais a designação de autoridades de controlo/notificadoras nacionais e organismos de avaliação de conformidade. Trata-se de um modelo de coordenação de tipo setorial que impõe a articulação com as autoridades de fiscalização do mercado do Regulamento (UE) 2019/1020 e que não está isenta de ambiguidade, pois as autoridades nacionais são simultaneamente reguladores de produto e supervisores do respeito das obrigações de utilização, ainda que sem um estatuto ou meios semelhantes às autoridades de dados[109]. Para a fiscalização e controlo, prevê-se o acesso aos conjuntos de dados de treino, validação e teste utilizados pelo fornecedor, incluindo através de interfaces de programação de aplicações ou outros meios e ferramentas técnicas adequadas que possibilitem o acesso remoto e pode mesmo ser exigido o acesso ao código-fonte.

A efetivação do regime imposto ocorre no nível europeu e no nível nacional. Preveem-se coimas para a violação das obrigações instituídas. Algumas são diretamente instituídas e tipificadas no Regulamento: coimas até 30 000 000 EUR ou, se o infrator for uma empresa, até 6% do seu volume de negócios anual, consoante o que for mais elevado, para o incumprimento das proibições ou não conformidade das exigências relativas aos dados de treino nos sistemas de risco elevado; a não conformidade do sistema de IA com os restantes requisitos ou obrigações por força do presente regulamento fica sujeita a coimas até 20 000 000 EUR ou, se o infrator for uma empresa, até 4 % do seu volume de negócios anual; o fornecimento de informações incorretas, incompletas ou enganadoras aos organismos notificados e às autoridades nacionais competentes em resposta a um pedido fica sujeito a coimas até 10 000 000 EUR ou, se o infrator for uma empresa, até 2 % do seu volume de negócios anual a nível mundial no exercício anterior, consoante o que for mais elevado. No restante, os Estados-Membros devem estabelecer o regime de sanções, incluindo coimas, aplicáveis em caso de infração ao presente regulamento e devem tomar todas as medidas necessárias para garantir que o mesmo é aplicado correta e eficazmente. As sanções previstas devem ser efetivas, proporcionadas e dissuasivas. Devem ter especialmente em conta os interesses dos fornecedores de pequena dimensão e das empresas em fase de arranque e a respetiva viabilidade económica. Podem ser aplicadas por

[109] Para uma crítica, M. Veale, F. Z. Borgesius, "Desmystifying the Draft EU Artificial Intelligence Act...", cit., p. 111.

tribunais ou outras autoridades, em conformidade com o direito nacional. Na UE, a sua efetivação é da competência da Autoridade Europeia de Proteção de dados. O PE propôs o agravamento das sanções.

O *Draft Report* do PE, inspirado no modelo do *Digital Services Act*, propôs modificações no quadro de governação, reforçando a independência e o papel do Comité Europeu para a IA, incluindo para dirimir divergências entre autoridades de supervisão nacional através de recomendações. Inclui no Comité Europeu para a IA a Agência Europeia dos Direitos Fundamentais. Propõe, também, o reforço significativo da fiscalização do cumprimento do regime previsto atribuído à Comissão, inspirado no regime do controlo da concorrência no mercado interno, em detrimento do papel dos Estados-Membros. Propõe ainda um regime de tutela efetiva que garanta os direitos e interesses dos particulares através de um sistema de queixas e reclamações junto da autoridade nacional de supervisão e o recurso judicial das decisões daquela. Já o CEPD e a AEPD criticaram a proeminência da Comissão no Comité e consideram que tal põe em causa a indispensável independência desse organismo.

O «Documento de compromisso da Presidência eslovena» acrescentou um título dedicado aos «sistemas de IA de utilização geral», ou seja, sistemas de *software* «inteligente» que podem ser utilizados em qualquer domínio, tais como sistemas de reconhecimento de imagem, tradução ou deteção de padrões. São cada vez mais utilizados na medicina, finanças, química, programação, etc.[110] A possibilidade da sua utilização em numerosos domínios, com diferentes níveis de risco, exponencia o risco de efeitos negativos de qualquer falha ou viés, muitas vezes impossíveis de controlar, rastrear e corrigir pelos utilizadores em cada um dos domínios[111]. Os criadores desses sistemas são quem está melhor colocado para perceber os problemas e não poderão ficar isentos de responsabilização. Tem sido defendido que os criadores desses sistemas devem ser submetidos às obrigações do regime a instituir e a responsabilidade deve ser repartida entre os vários níveis de intervenção de acordo com o nível de controlo, recursos e capacidades[112]. Além da segurança ficar

[110] Um elenco desses sistemas pode ser encontrado em Future of Life, *General Purpose AI and the AI Act*, May 2022, p. 3, *in* https://artificialintelligenceact.eu/wp-content/uploads/2022/05/General-Purpose-AI-and-the-AI-Act.pdf (31.10.2022).

[111] Já se comprovaram esses efeitos negativos, seja na disseminação de conteúdos extremistas, discriminatórios ou na exposição de dados pessoais. Future of Life, *General Purpose AI and the AI Act*, cit., p. 4.

[112] Access Now *et al.*, *Call for better protections of people affected at the source of the AI value chain*, 25.10.2022, *in* https://futureoflife.org/wp-content/uploads/2022/10/Civil-society-letter-GPAIS-October-2022.pdf (31.10.2022).

comprometida, colocar a responsabilidade exclusivamente nos utilizadores acabará por limitar o acesso à inovação por parte de organizações menores, em virtude do receio da responsabilização por efeitos que não estão completamente sob o respetivo controlo[113]. Já foi proposto que sejam sujeitos a uma avaliação de conformidade, incluindo independente, antes da respetiva colocação no mercado ou em serviço[114]. A proposta eslovena afastou-se desta proposta e prevê que é a sua colocação no mercado, ao serviço ou utilização para uma determinada finalidade que implica a qualificação como fornecedor, seja ou não *software* aberto, e não acompanha a preocupação referida. A evolução surpreendente dos sistemas de AI generativa não encontrava resposta na proposta da Comissão. O PE propõe que os fornecedores de modelos generativos – entre os quais o Chat GPT é o mais famoso – fiquem obrigados, não apenas ao respeito do regime de obrigações aplicável, como a especiais obrigações de transparência que informem que o conteúdo é produzido por um sistema de IA e sobre a utilização de dados protegidos por propriedade intelectual no treino do sistema. Prevê ainda o seu registo numa base de dados europeia.

Este Regulamento é pioneiro no estabelecimento de uma regulamentação horizontal da IA[115]. No entanto, não é evidente que a regulamentação acompanhe as exigências colocadas por uma IA centrada no ser humano[116]. O regime parece insuficiente[117], particularmente considerando o imperativo do "nível de proteção elevado" fixado pelo artigo 114º, nº 3 TFUE e as exigências do artigo 16º TFUE. A ausência de um regime de tutela dos danos resultantes da ação dos sistemas de IA não é completamente ultrapassada pela harmonização resultante das propostas de diretiva em matéria de responsabilidade apresentadas em setembro. Algumas pretensões ficaram pelo preâmbulo e diversas soluções são questionáveis. A pretensão de instituir uma harmonização global

[113] Future of Life, *General Purpose AI and the AI Act*, cit.

[114] Future of Life, *General Purpose AI and the AI Act*, cit.

[115] Constitui um exemplo de «*New Legislative Framework*», de acordo com o Regulamento (CE) 765/2008, a Decisão 768/2008 e o Regulamento (EU) 2019/1020.

[116] M. Veale e F. Z. Borgesius acusam a proposta da Comissão de ser uma espécie de *patchwork* datado e confuso das fórmulas de regulação dos produtos e dos direitos fundamentais dos anos 80 e 90 e apresentam uma visão muito crítica do regime previsto, que consideram insuficiente e obscuro. "Desmystifying the Draft EU Artificial Intelligence Act...", cit., p. 112. O CEPD e a AEPD afirmam que "há ainda muito trabalho a fazer até que a proposta possa dar origem a um quadro jurídico efetivo, que complete de forma eficiente o RGPD na proteção dos direitos humanos fundamentais, promovendo simultaneamente a inovação". *Parecer Conjunto 5/2021*..., cit., p. 26.

[117] Na opinião de M. Veale e F. Z. Borgesius, a efetiva regulação da IA vai fazer-se através da standardização, que a Comissão pretende que avance em simultâneo (*Impact Assessment*, COM 2021 (206) final, cit., p. 57). "Desmystifying the Draft EU Artificial Intelligence Act...", cit., p. 105.

pode, em virtude do primado e preempção, excluir regimes nacionais mais rigorosos[118], que imponham, por exemplo, a regulação das emissões de carbono[119] ou que imponham um regime imperativo a sistemas de IA que não sejam de alto risco[120]. A velocidade da inovação pode tornar este regime obsoleto muito rapidamente e, dado que só será aplicável dois anos após a entrada em vigor, há uma elevada probabilidade de já estar ultrapassado nesse momento. O objetivo de garantir uma IA segura e de confiança não se cumprirá.

4.2. Proposta de Diretiva relativa à responsabilidade de produtos defeituosos, 28.09.2022, COM(2022) 495 final, e Proposta de Diretiva relativa à adaptação das regras de responsabilidade civil extracontratual à inteligência artificial (Diretiva Responsabilidade da IA), 28.09.2022, COM(2022) 496 final

A proposta de Regulamento para a IA não trata da reparação dos danos resultantes do funcionamento de sistemas de IA[121]. Obviamente, aplicam-se as regras nacionais de responsabilidade, seja contratual, seja extracontratual[122], bem como diversa legislação europeia[123]. A diversidade de regimes nacionais e a especificidade deste setor recomendam uma intervenção legislativa a nível da União, especialmente no âmbito das condições da prova[124], quer

[118] Não foi incluída uma cláusula que acautelasse essa possibilidade, nos termos do artigo 114º, nº 10 TFUE, pelo que restará a possibilidade de aplicação das cláusulas dos nºs 4, 5 e 8. Cláusulas que preveem expressamente a aplicabilidade de regras nacionais *praeter legem*, como o artigo 29º, nº 2 da proposta, confortam esta interpretação.

[119] É notória uma falta de preocupação com a imposição de padrões ambientais específicos para os sistemas de IA.

[120] Em França, a legislação impõe requisitos de transparência algorítmica superiores (*Loi n° 2016-1321 du 7 octobre 2016 pour une République numérique; décret n° 2017-330 du 14 mars 2017 relatif aux droits des personnes faisant l'objet de décisions individuelles prises sur le fondement d'un traitement Algorithmique*). Lilian Edwards and Michael Veale, "Enslaving the Algorithm: From a "Right to an Explanation" to a "Right to Better Decisions"?", *IEEE Security & Privacy*, 16(3), 2018, pp. 46-58.

[121] A secundarização do tema da responsabilidade foi criticada por Christiane Wendehorst, "Liability for Artificial Intelligence...", cit., p. 188.

[122] A diversidade de regimes nacionais é imensa.

[123] No âmbito contratual, a Diretiva (EU) 2019/770 sobre certos aspetos relativos aos contratos de fornecimento de conteúdos e serviços digitais e a Diretiva (UE) 2019/771 relativa a certos aspetos dos contratos de compra e venda de bens contêm regras sobre a responsabilidade contratual. No âmbito extracontratual, o principal instrumento europeu é a Diretiva 85/374/CEE relativa à responsabilidade por produtos defeituosos. São ainda relevantes o «RGPD» e a Diretiva 2004/35/CE relativa à responsabilidade ambiental.

[124] European Commission, Directorate-General for Justice and Consumers, *Liability for artificial intelligence and other emerging digital technologies*, Publications Office, Luxembourg, 2019, *in* https://data.europa.eu/doi/10.2838/573689 (31.10.2022). Comissão Europeia, *Relatório da Comissão ao*

para assegurar a proteção dos direitos dos lesados por sistemas de IA, quer para garantir a certeza jurídica, fundamental para os operadores deste setor e para incrementar a utilização da IA[125].

Em 28.09.2022, a Comissão apresentou duas propostas de diretiva sobre a responsabilidade extracontratual – a proposta de Diretiva relativa à responsabilidade decorrente dos produtos defeituosos [COM(2022) 495 final][126] e a proposta de Diretiva relativa à adaptação das regras de responsabilidade civil extracontratual à inteligência artificial [COM (2022) 496 final][127].

A primeira altera a Diretiva 85/374/CEE relativa à responsabilidade por produtos defeituosos, adaptando-a aos desenvolvimentos tecnológicos, incluindo a IA. A complexidade dos sistemas de IA, a autonomia e a tendencial opacidade do seu funcionamento dificultam e tornam especialmente onerosas as ações de responsabilidade. Também a segurança jurídica reclama certeza legislativa nesta matéria, em especial num setor eminentemente transnacional. Por isso, em paralelo, foi apresentada a segunda proposta de diretiva, que acrescenta um regime especial para a responsabilidade em relação a produtos de IA. Deste modo, a Comissão responde a algumas das críticas à proposta de «Regulamento IA» e melhora as condições para o exercício da tutela jurisdicional efetiva de direitos fundamentais lesados pela ação de sistemas de IA.

Parlamento Europeu Europeu, ao Conselho e ao Comité Económico e Social Europeu sobre as implicações em matéria de segurança e de responsabilidade decorrentes da inteligência artificial, da Internet das coisas e da robótica, Bruxelas, 19.02.2020, COM/2020/64 final, *in* https://eur-lex.europa.eu/legal-content/PT/TXT/?qid=1593079180383&uri=CELEX:52020DC0064 (31.10.2022). European Commission, Directorate-General for Justice and Consumers, Karner, E., Koch, B., Geistfeld, M., *Comparative law study on civil liability for artificial intelligence*, Publications Office of the European Union, 2021, *in* https://data.europa.eu/doi/10.2838/77360 (31.10.2022). O PE, numa Resolução de 20.20.2020, solicita à Comissão uma proposta de regulamento sobre a responsabilidade pela operação de sistemas de Inteligência Artificial, *in* https://www.europarl.europa.eu/doceo/document/TA-9-2020-0276_PT.html (31.10.2022).

[125] Segundo Herbert Zech, o afastamento da culpa em favor da responsabilidade objetiva pode favorecer a inovação por assegurar a mais elevada certeza jurídica e como internaliza os riscos para os operadores incentiva-os a desenvolverem o conhecimento da IA e a investirem na respetiva segurança. "Liability for AI: public policy considerations", ERA Forum (2021) 22:147–158, pp. 152-153. Disponível em https://link.springer.com/content/pdf/10.1007/s12027-020-00648-0.pdf (27.11.2022).

[126] Proposta de Diretiva do Parlamento Europeu e do Conselho relativa à responsabilidade decorrente dos produtos defeituosos, Bruxelas, XXX, COM(2022) 495 final, *in* https://eur-lex.europa.eu/legal-content/PT/TXT/?uri=COM%3A2022%3A495%3AFIN&qid=1664465004344 (31.10.2022).

[127] Proposta de Diretiva do Parlamento Europeu e do Conselho relativa à adaptação das regras de responsabilidade civil extracontratual à inteligência artificial (Diretiva Responsabilidade da IA), Bruxelas, 28.9.2022, COM(2022) 496 final, *in* https://eur-lex.europa.eu/legal-content/PT/ALL/?uri=celex:52022PC0496 (31.10.2022).

Estas duas iniciativas[128] estão estreitamente ligadas e formam um pacote legislativo, uma vez que as ações abrangidas pelo respetivo âmbito de aplicação dizem respeito a diferentes tipos de responsabilidade. A proposta de «Diretiva Responsabilidade dos Produtos Defeituosos» abrange a responsabilidade objetiva do produtor por produtos defeituosos, conduzindo a uma indemnização por certos tipos de danos, principalmente sofridos por particulares. A proposta de «Diretiva Responsabilidade da IA» abrange as ações nacionais de indemnização baseadas na culpa de qualquer interveniente na cadeia, com vista a indemnizar todo o tipo de danos e todo o tipo de vítimas. O regime da responsabilidade destes dois atos deve ainda articular-se com a Diretiva 2020/1828 em relação às ações coletivas.

O processo ainda se encontra no seu início. Seguem-se apenas alguns tópicos fundamentais.

A proposta de Diretiva [COM(2022) 495 final] estabelece, ao longo dos 20 artigos, o regime-base da responsabilidade extracontratual por danos provocados por produtos defeituosos. Os produtos e serviços digitais, incluindo o *software*, em geral, e os sistemas de IA, em especial, são qualificados como «produtos», embora o *software* livre/aberto desenvolvido ou fornecido fora do âmbito de uma atividade comercial seja objeto de uma exceção. Os serviços conexos são integrados como componentes no conceito de produto, quando se incorporem ou interliguem com um produto e sejam essenciais para o desempenho de uma função daquele, o que pode ser especialmente importante para sistemas de IA, *v.g.* sistemas de navegação.

Este regime não abrange produtos objeto de utilização exclusivamente profissional pelo lesado, ou seja, protege apenas os particulares. De igual modo, não integra entre os sujeitos responsáveis, fornecedores particulares. Entre os danos ressarcíveis, inclui nas lesões corporais os danos à saúde psicológica, o que é muito importante nos produtos digitais, em especial em relação aos consumidores mais vulneráveis, e abrange igualmente a perda de dados.

[128] Sobre a proposta COM(2022) final, vide Samar Abbas Nawaz, "The proposed AI Lisbility Rules: Ease or Burden", 07.11.2022 *in* https://europeanlawblog.eu/2022/11/07/the-proposed-eu-ai-liability-rules-ease-or-burden/#more-8626 (08.11.2022). O Autor é crítico e considera que o regime proposto não contribui efetivamente para facilitar o ressarcimento de prejuízos produzidos por sistemas de IA. Também o Irish Council for Civil Liberties, em 04.10.2022, sublinhou as insuficiências do regime proposto (https://www.iccl.ie/news/new-liability-rules-on-product-and-ai-are-encouraging-but-need-improvement/, 24.11.2022). Do outro lado, a indústria, numa carta aberta aos Comissários Breton e Reynders de 24.08.2022, manifestou as suas reticências a uma intervenção regulatória que vá além do que resulta do regime geral sobre produtos defeituosos. Disponível em https://www.developersalliance.org/statement-on-pld-ai (24.11.2022).

Os sistemas de IA ficam sujeitos ao regime da responsabilidade por produtos defeituosos quando caiam no âmbito da noção de produto defeituoso que aí se propõe. Nesse caso, não é necessária a culpa para a responsabilização de algum dos operadores que a diretiva estabelece como responsáveis, ainda que o demandante tenha de provar a qualidade defeituosa do produto, o dano e o nexo de causalidade. A presunção da qualidade defeituosa quando se comprove que o produto não cumpre os requisitos de segurança obrigatórios estabelecidos no direito da União ou no direito nacional destinados a proteger contra o risco do dano ocorrido pode ser importante para a responsabilização por danos resultantes de sistemas de IA. De igual modo, a não aplicação da isenção de responsabilidade quando a qualidade defeituosa tiver surgido posteriormente à colocação no mercado ou em serviço aos serviços conexos, *software* e respetivas atualizações, será particularmente importante nos sistemas de IA. Os produtores podem ser responsabilizados por danos causados por modificações que possam reconduzir-se a uma «qualidade defeituosa» do produto resultantes de atualizações do produto, mesmo que autónomas, como sucede com o «*machine learning*», sob a fórmula de «controlo do produto», o que se revela especialmente importante na IA. Mas o regime previsto é passível de crítica, pois não se aplicará a sistemas colocados no mercado antes do momento de produção de efeitos da diretiva, i.e. 12 meses após a entrada em vigor da diretiva, mesmo em relação a atualizações posteriores.

A proposta de Diretiva [COM(2022) 496 final] é um documento muito circunscrito, com apenas nove artigos e um âmbito muito limitado[129]. Incide sobre a responsabilidade civil extracontratual resultante de danos provocados por sistemas de IA, independentemente da sua natureza defeituosa. Exclui liminarmente do seu âmbito a responsabilidade criminal, mas integra no âmbito subjetivo de proteção todos os sujeitos, profissionais e não profissionais. Abrange as instituições do Estado. Salvaguarda-se igualmente o regime da responsabilidade por produtos defeituosos e regimes derrogatórios e especiais de responsabilidade, nomeadamente no domínio dos transportes, ou as isenções de responsabilidade e as obrigações de diligência devida previstas no «Regulamento Serviços Digitais».

[129] O Parlamento Europeu defendeu que a regulação desta matéria se fizesse através de Regulamento, juntamente com a regulação dos sistemas de IA (Pontos 5 e 26 da Resolução do Parlamento Europeu, de 20 de outubro de 2020, que contém recomendações à Comissão sobre o regime de responsabilidade civil aplicável à inteligência artificial, disponível em https://www.europarl.europa.eu/doceo/document/TA-9-2020-0276_PT.html (31.10.2022). Em geral, o PE propunha uma abordagem regulatória mais uniformizadora e mais estrita.

Entre as opções possíveis, a Comissão opta por uma abordagem cautelosa, que exclui a responsabilidade objetiva numa primeira fase e, numa segunda fase, avaliará a possibilidade de avançar com normas harmonizadas para um regime de responsabilidade pelo risco para riscos específicos (eventualmente associado a um seguro obrigatório)[130].

Por causa da opacidade, autonomia e complexidade dos sistemas de IA pode ser excessivamente difícil assegurar a responsabilidade quando se exige a culpa (sob a forma de dolo ou de negligência) e a prova de um nexo de causalidade entre o comportamento de um sujeito jurídico e o resultado danoso. A conectividade em rede dos sistemas de IA introduz uma complexidade suplementar, pois a origem múltipla da informação em que se baseia o facto lesivo dificulta a determinação da causa e pode tornar muito mais complexa a determinação do responsável[131]. Esta situação, insustentável num Estado-de-direito, abala a confiança social necessária para a aceitação dos novos desenvolvimentos tecnológicos[132]. Ainda assim, a abordagem da Comissão é cautelosa e seletiva. Segundo o preâmbulo, o objetivo da proposta é "assegurar que as pessoas que pedem uma indemnização por danos que lhes sejam causados por um sistema de IA gozam de um nível de proteção equivalente ao nível de que gozam as pessoas que pedem uma indemnização por danos causados sem o envolvimento de um sistema de IA" (considerando 7). Não pretende harmonizar "os aspetos gerais da responsabilidade civil (...) tais como a definição de «culpa» ou de «causalidade», os diferentes tipos de danos que dão origem a direitos de indemnização, a repartição da responsabilidade por vários infratores, a conduta que concorre para a origem do dano, o cálculo dos danos ou os prazos de prescrição" (considerando 10). Não se afastam os

[130] Comissão Europeia, *Documento de trabalho dos Serviços da Comissão. Relatório do Resumo da avaliação de impacto que acompanha o documento Proposta de diretiva do Parlamento Europeu e do Conselho relativa à adaptação das regras de responsabilidade civil extracontratual à inteligência artificial (Diretiva Responsabilidade da IA)*, Bruxelas, 28.9.2022, SWD(2022) 320 final, *in* https://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=CELEX:52022SC0320&from=PT (31.10.2022). Este documento sumaria a avaliação feita em European Commission, *Commission Staff Working Document. Impact Assessment Report Accopanying the document Proposal for a Directive of the European Parliament and of the Council on adapting non-contractual civil liability rules to artificial intelligence*, Brussels, 28.9.2022, SWD(2022) 319 final, *in* https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52022SC0319&from=PT (31.10.2022).

[131] Este ponto é salientado por Herbert Zech, "Liability for AI...", cit., p. 149.

[132] Segundo Herbert Zech, a responsabilidade objetiva favoreceria a aceitação social da IA, ainda que sujeita a limites. "Liability for AI...", cit., p. 153. Se o risco é muito elevado, a alternativa poderá ser a proibição, o que será mais inibidor da inovação.

regimes nacionais aplicáveis, embora possam ser aplicáveis outros atos da UE e não o direito nacional, *v.g.* relativos à segurança dos produtos[133].

A futura diretiva aplicar-se-á quando se exige a culpa. A proposta prevê a harmonização de algumas regras em matéria de prova, incluindo a introdução de presunções ilidíveis, com vista a facilitar a prova por parte dos lesados em resultado de ações de sistemas de IA para o ressarcimento dos danos. Aplicam-se as regras nacionais que determinam sobre que parte recai o ónus da prova, o grau de certeza necessário no que respeita ao nível de prova ou a definição de «culpa», salvo no que diga respeito às presunções ilidíveis previstas. O facto culposo tem de ser produzido pelo sistema de IA, não sendo aplicável este regime quando é praticado (ação ou omissão) por uma pessoa com o apoio de um sistema de IA. Introduz modificações no ónus da prova, nomeadamente quando haja incumprimento das regras impostas no regulamento. O seu objetivo fundamental é garantir que os lesados por um sistema de IA têm acesso a um «processo equitativo».

Propõe a harmonização de algumas regras aplicáveis no âmbito de ações de indemnização de direito civil relativas a responsabilidade culposa extracontratual por danos causados por um sistema de IA, nomeadamente facilitando o acesso às informações registadas/documentadas de acordo com as regras de segurança dos produtos do «Regulamento Inteligência Artificial» em processos judiciais, relevantes para que a vítima identifique e prove a ação/omissão que causou o dano. Também permite aos tribunais presumir (presunção ilidível), preenchidas diversas condições indicadas, que quando se comprove o incumprimento das obrigações impostas no «Regulamento IA» aos sistemas de IA de risco elevado se verifica o nexo de causalidade entre esse incumprimento e o dano verificado. Os utilizadores não profissionais beneficiam de um regime especial que atenua a presunção de causalidade. No entanto, não se prescinde da culpa e não há uma presunção de culpa[134].

A proposta representa um avanço, em especial para a difícil tarefa de recolha de provas e determinação do responsável num âmbito em que a assimetria de informação entre criadores/fornecedores e eventuais lesados é abissal.

[133] Diretiva 2001/95/CE («Diretiva Segurança Geral dos Produtos»).

[134] O Relatório do *Expert Group on Liability and New Technologies – New Technologies Formation* propôs a opção pela aplicação da presunção na culpa e no nexo de causalidade. Além de uma análise exaustiva das várias modalidades de responsabilidade em vigor nos EM, o documento ilustra essa análise com exemplos. *Liability for Artificial Intelligence and other emerging digital technologies*, Luxembourg, Publications Office of the European Union, 2019, disponível em https://op.europa.eu/en/publication-detail/-/publication/1c5e30be-1197-11ea-8c1f-01aa75ed71a1/language-en/format-PDF (31.10.2022). Este documento serviu de base ao documento [COM(2020) 64 final].

Várias críticas podem fazer-se logo numa primeira análise.

A primeira e mais importante é a incoerência com a «*risk approach*» da proposta de «Regulamento da IA», pois, embora o regime proposto tenha em consideração o regime de obrigações diferenciado previsto neste último, daí não decorre um regime de responsabilidade «paralelo»[135]. Não se prevê um regime de responsabilidade objetiva pelo risco para os sistemas de IA de risco elevado, nem sequer se prevê um regime de presunções (mesmo ilidíveis) de culpa ou sobre o nexo de causalidade quando as obrigações impostas não são respeitadas. O regime não tem um recorte preciso, rigoroso e coerente. Sendo vários os intervenientes ao longo do ciclo do sistema de IA, a informação pode estar disseminada na cadeia, pelo que a operacionalização desse regime enfrenta dificuldades. Como se determinará qual a informação pertinente e necessária a ser divulgada ao requerente? Tal é um problema para além da óbvia e acautelada necessidade de proteção de segredos comerciais e informações confidenciais.

A responsabilidade por violação dos direitos fundamentais, como a discriminação, a violação da integridade dos dados pessoais, da privacidade, a manipulação, etc., apresenta dificuldades particulares, em virtude da menor precisão em relação aos danos materiais[136].

Há uma profusão de conceitos indeterminados, tais como «[elementos de prova] razoavelmente acessíveis ao demandante», «razoavelmente provável» e «conhecimentos especializados suficientes», ou «excessivamente difícil». As modificações no regime da prova ainda colocam sobre o demandante uma exigência excessiva, pois as condições impostas para a inversão do ónus da prova são significativas. Espera-se que ao longo do procedimento legislativo haja um aperfeiçoamento. O Comité Económico e Social Europeu, no seu Parecer de 24.01.2023, ainda que favorável a uma harmonização mínima, salientou precisamente a importância de necessidade de estabelecer definições jurídicas claras.

A Resolução do PE de 20 de outubro de 2020 preconizava uma proteção mais reforçada, através da responsabilidade objetiva para os sistemas de IA de risco elevado e um regime de presunções de culpa ilidíveis para outras

[135] Uma análise crítica encontra-se em Susana Navas Navarro, "The future regulation on non-contractual civil liability for AI systems", *Official blog Union Essays*, 24.11.2022, disponível em https://officialblogofunio.com/2022/11/ (27.11.2022). Christiane Wendehorst, "Liability for Artificial Intelligence...", cit., apresenta uma abordagem assente na categorização dos riscos e faz uma avaliação da adequação das diferentes modalidades de responsabilidade a esses riscos e à natureza da IA.

[136] Christiane Wendehorst, "Liability for Artificial Intelligence...", cit., p. 205.

categorias[137]. Na consulta pública, as opiniões dos cidadãos eram favoráveis a uma intervenção em matéria do regime de prova, mas também no sentido da responsabilidade objetiva e até da imposição de um seguro obrigatório. Parece-nos que esta última solução será a mais adequada, quer para o eventual lesado quer para o fornecedor. Não obstante, será necessário determinar qual dos intervenientes no ciclo de vida dos sistemas de IA (criador/implantador, distribuidor, utilizador, etc.) deverá ficar sujeito a essa obrigação e os seus termos[138]. A solução adequada deverá diferir entre os diferentes tipos de IA[139]. As empresas mostraram posições mais diversas e foi notória uma divisão entre os maiores operadores no mercado e as PME, estas últimas com posições mais próximas dos cidadãos[140]. A opção da Comissão pode ser considerada uma posição pró-inovação, de modo a favorecer o desenvolvimento de novos produtos/serviços na Europa, ou pode ser visto como um favorecimento aos interesses dos principais *players* do setor.

A «Diretiva da responsabilidade da IA» é uma diretiva com um âmbito muito circunscrito e mesmo nesse âmbito limitado é um ato de harmonização mínima[141], que admite regimes mais favoráveis ao demandante na legislação nacional, incluindo regimes de responsabilidade objetiva. Por isso, o objetivo de impedir a fragmentação no mercado interno não é efetivamente atingido.

O futuro seguirá, com certeza, a via da responsabilidade objetiva, pelo risco, associada a um regime de seguros obrigatórios[142] para alguns sistemas de IA de risco elevado tipificados, ou a um fundo social de compensação[143]. As duas diretivas preveem um reexame cinco anos após o prazo de transposi-

[137] Porém, a imposição da responsabilidade aos «frontend» e «backend operators» pode não ser suficiente e deixar de fora outros intervenientes importantes na cadeia de criação e operação dos sistemas de IA. Herbert Zech, "Liability for AI...", cit., pp. 155-156.

[138] Herber Zech defende que nos sistemas de risco elevado a responsabilidade recais sobre os produtores. "Liability for AI...", cit., p. 155.

[139] Os «sistemas de IA geral» são objeto de adaptações para a respetiva utilização num particular setor, o que levanta problemas para a determinação de quem deveria ser o responsável num regime de responsabilidade objetiva. Christiane Wendehorst, "Liability for Artificial Intelligence...", cit., p. 196. A autora não é favorável a um regime de responsabilidade objetiva para os sistemas de IA, em geral, mas apenas para certas categorias. "Liability for Artificial Intelligence...", cit., p. 203. Sugere igualmente cautela na abordagem regulatória dos danos não patrimoniais (p. 202),

[140] Posição fácil de compreender, na medida em que estas são sobretudo utilizadoras e não produtoras de sistemas de IA.

[141] Ao contrário da «Diretiva da responsabilidade por produtos defeituosos».

[142] Para Herbert Zech esta opção distribui o risco de modo mais equitativo, embora reconheça a potencial iniquidade entre os intervenientes na cadeia da IA. "Liability for AI...", cit., p. 154.

[143] Herbert Zech defende um sistema de seguro social financiado pelos diferentes intervenientes na cadeia de criação e funcionamento dos sistemas de IA. "Liability for AI...", cit., pp. 156-157.

ção, em especial para avaliar a adequação de avançar para a responsabilidade objetiva, considerando ainda a evolução de soluções no setor dos seguros. O regime proposto acaba por ser um regime temporário, num momento em que o potencial da IA ainda não é absolutamente certo, desse modo não condicionando a inovação e permitindo que os principais *stakeholders* testem diferentes opções regulatórias nacionais.

5. Conclusão

O processo legislativo ainda está em curso e previsivelmente tem um longo caminho pela frente. Numa conclusão preliminar, as iniciativas da EU marcam a agenda regulatória global e colocam a UE na vanguarda normativa. A abordagem, seguindo a lógica do «risco», pretende assegurar um equilíbrio entre a proteção dos valores e direitos fundamentais, de acordo com o «*European way*», e a salvaguarda da competitividade e a promoção da inovação. Várias insuficiências e incoerências foram descritas, em especial no âmbito da responsabilidade. Espera-se que sejam superadas ao longo do processo de adoção dos atos. Finalmente, o recorte dos regimes da IA é intrinsecamente evolutivo, nomeadamente prevendo a atualização das categorias de IA, para acompanhar a evolução tecnológica, e o agendamento de uma avaliação com vista à redefinição do regime da responsabilidade. Em suma, este pacote de propostas legislativas é positivo, mas pode ser melhorado até à sua adoção.

Bibliografia

AAVV, Joint Letter on the PLD and AI Directive, em https://www.developersalliance.org/statement-on-pld-ai (24.11.2022)

Access Now *et al.*, *Call for better protections of people affected at the source of the AI value chain*, 25.10.2022, *in* https://futureoflife.org/wp-content/uploads/2022/10/Civil-society-letter-GPAIS-October-2022.pdf (31.10.2022)

Agência Europeia para os Direitos Fundamentais, *Getting the future right – Artificial intelligence and fundamental rights*, *in* https://fra.europa.eu/en/publication/2020/artificial-intelligence-and-fundamental-rights (31.10.2022)

Almenzar, Marina Estevez *et al*, *Glossary of human-centred artificial intelligence*, Publications Office of the European Union, Luxembourg, 2022, doi:10.2760/860665, JRC129614, *in* https://publications.jrc.ec.europa.eu/repository/handle/JRC129614 (5.11.2022)

Alto Representante da União para os Negócios Estrangeiros e Política Segurança, na *Comunicação conjunta ao Parlamento Europeu e ao Conselho relativa ao reforço da contribuição da UE para um multilateralismo assente em regras*, Bruxelas, 17.2.2021 JOIN(2021) 3 final, *in* https://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=CELEX:52021JC0003&from=PT (31.10.2022)

Amershi, Saleema *et al*, "Power to the People: The Role of Humans in Interactive Machine Learning", *A.I. Magazine (35):4*, pp.105–120, *in* https://ojs.aaai.org/index.php/aimagazine/article/view/2513 (31.10.2022)

Balakrishnan, Tara / Chui, Michael / Hall, Bryce, *McKinsey Global Survey The State of AI 2020*, de 17.11.2020, *in* https://www.mckinsey.com/capabilities/quantumblack/our-insights/global-survey-the-state-of-ai-in-2020 (31.10.2022)

CENELEC, "Design methodology of advanced human–robot collaborative cells in personalized HRC systems", 2022-10, *in* https://www.cencenelec.eu/media/CEN-CENELEC/News/Workshops/2022/2022-10-21%20-%20SHARE/cwa_hrccells_v2_forpubliccommenting.pdf (5.11.2022)

CEPD e AEPD, *Parecer conjunto 5/2021 do CEPD e da AEPD sobre a proposta de regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial (Regulamento Inteligência Artificial)*, 18.06.2021, *in* https://edpb.europa.eu/system/files/2021-10/edpb-edps_joint_opinion_ai_regulation_pt.pdf (31.10.2022)

Cihon, Peter, *Standards for AI Governance: International Standards to Enable Global Coordination in AI Research & Development*, Future of Humanity Institute, University of Oxford, 2019, *in* https://www.fhi.ox.ac.uk/wp-content/uploads/Standards_-FHI-Technical-Report.pdf (31.10.2022)

Comissão Europeia, *Anexos da Proposta de Regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial (Regulamento Inteligência Artificial) e altera determinados atos legislativos da União*, Bruxelas, 21.4.2021, COM(2021) 206 final ANNEXES 1 to 9, *in* https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0004.02/DOC_2&format=PDF (31.10.2022)

Comissão Europeia, Comunicação da Comissão ao Parlamento Europeu, ao Conselho Europeu, ao Conselho, ao Comité Económico e Social Europeu e ao Comité das Regiões *Inteligência artificial para a Europa*, Bruxelas, 25.4.2018, COM(2018) 237 final, *in* https://eur-lex.europa.eu/legal-content/PT/TXT/HTML/?uri=CELEX:52018DC0237&from=EN (31.10.2022)

Comissão Europeia, *Documento de trabalho dos Serviços da Comissão. Relatório do Resumo da avaliação de impacto que acompanha o documento Proposta de diretiva do Parlamento Europeu e do Conselho relativa à adaptação das regras de responsabilidade civil extracontratual à inteligência artificial (Diretiva Responsabilidade da IA)*, Bruxelas, 28.9.2022, SWD(2022) 320 final, *in* https://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=CELEX:52022SC0320&from=PT (31.10.2022)

Comissão Europeia, Proposta de Diretiva do Parlamento Europeu e do Conselho relativa à responsabilidade decorrente dos produtos defeituosos, Bruxelas, XXX, COM(2022) 495 final, *in* https://eur-lex.europa.eu/legal-content/PT/TXT/?uri=COM%3A2022%3A495%3AFIN&qid=1664465004344 (31.10.2022)

Comissão Europeia, Proposta de Diretiva do Parlamento Europeu e do Conselho relativa à adaptação das regras de responsabilidade civil extracontratual à inteligência artificial (Diretiva Responsabilidade da IA), Bruxelas, 28.9.2022, COM(2022) 496 final, *in* https://eur-lex.europa.eu/legal-content/PT/ALL/?uri=celex:52022PC0496 (31.10.2022)

Comissão Europeia, Proposta de Regulamento do Parlamento Europeu e do Conselho relativo aos requisitos horizontais de cibersegurança dos produtos com elementos digi-

tais e que altera o Regulamento (UE) 2019/1020, Bruxelas, 15.09.2022, COM(2022) 454 final, *in* https://eur-lex.europa.eu/resource.html?uri=cellar:864f472b-34e9-11ed-9c68-01aa75ed71a1.0008.02/DOC_1&format=PDF (31.10.2022)

Comissão Europeia, *Recomendação de Decisão do Conselho que autoriza a abertura de negociações, em nome da União Europeia, tendo em vista uma convenção do Conselho da Europa sobre inteligência artificial, direitos humanos, democracia e Estado de direito*, Bruxelas, 18.8.2022, COM(2022) 414 final, *in* https://ec.europa.eu/transparency/documents-register/detail?ref=COM(2022)414&lang=en (31.10.2022)

Comissão Europeia, *Relatório da Comissão ao Parlamento Europeu Europeu, ao Conselho e ao Comité Económico e Social Europeu sobre as implicações em matéria de segurança e de responsabilidade decorrentes da inteligência artificial, da Internet das coisas e da robótica*, Bruxelas, 19.02.2020, COM(2020) 64 final, *in* https://eur-lex.europa.eu/legal-content/PT/TXT/?qid=1593079180383&uri=CELEX:52020DC0064 (31.10.2022)

Comissão Europeia, *Study to Support an Impact Assessment of Regulatory Requirements for Artificial Intelligence in Europe. Final Report*, 21.04.2021, *in* https://op.europa.eu/en/publication-detail/-/publication/55538b70-a638-11eb-9585-01aa75ed71a1 (31.10.2022)

Conselho da União Europeia, em 21.10.2020, *Conclusões sobre a Carta dos Dreitos Fundamentais no contexto da inteligência artificial e da transformação digital*, 11481/20, 21.10.2020, *in* https://data.consilium.europa.eu/doc/document/ST-11481-2020-INIT/pt/pdf (31.10.2022)

Council of Europe, *Algorithms and Human Rights. Study on the human rights dimensions of automated data processing techniques (in particular algorithms) and possible regulatory implications*, Strasbourg, março 2018, *in* https://rm.coe.int/algorithms-and-human-rights-en-rev/16807956b5 (31.10.2022)

Council of the European Union, *Compromis de la Presidence*, 9029/22, Bruxelles, le 13 mai 2022, *in* https://artificialintelligenceact.eu/wp-content/uploads/2022/05/AIA-FRA-Art-34-13-May.pdf (31.10.2022)

Council of the European Union, *Presidency compromise text*, 14278/21, Brussels, 29.11.2021, *in* https://data.consilium.europa.eu/doc/document/ST-14278-2021-INIT/en/pdf (31.10.2022)

Edwards, Lilian / Veale, Michael, "Enslaving the Algorithm: From a "Right to an Explanation" to a "Right to Better Decisions"?", *IEEE Security & Privacy*, 16(3), 2018, pp. 46-58

European Commission, *Commission Staff Document Impact Assessment Accompanying the Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts*, Brussels, 21.4.2021, SWD(2021) 84 final, *in* https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12527-Inteligencia-artificial-Requisitos-eticos-e-legais_pt (31.10.2022)

European Commission, *Commission Staff Working Document. Impact Assessment Report Accopanying the document Proposal for a Directive of the European Parliament ando f the Council on adapting non-contractual civil liability rules to artificial intelligence*, Brussels, 28.9.2022, SWD(2022) 319 final, *in* https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52022SC0319&from=PT (31.10.2022)

European Commission, Directorate-General for Justice and Consumers, *Liability for artificial intelligence and other emerging digital technologies*, Publications Office, Luxembourg, 2019, *in* https://data.europa.eu/doi/10.2838/573689 (31.10.2022)

European Commission, Directorate-General for Justice and Consumers, Karner, E., Koch, B., Geistfeld, M., *Comparative law study on civil liability for artificial intelligence*, Publications Office of the European Union, 2021, *in* https://data.europa.eu/doi/10.2838/77360 (31.10.2022)

European Commission, *European enterprise survey on the use of technologies based on artificial intelligence*, Luxemburgo, 2020, *in* https://op.europa.eu/en/publication-detail/-/publication/f089bbae-f0b0-11ea-991b-01aa75ed71a1 (31.10.2022)

European Commission, *First regulatory sandbox on Artificial Intelligence presented*, 27.06.2022, *in* https://digital-strategy.ec.europa.eu/en/news/first-regulatory-sandbox-artificial-intelligence-presented (31.10.2022)

European Commission, *Staff Working Document Liability for emerging digital technologies Accompanying the document Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions Artificial intelligence for Europe*, Brussels, 25.04.2018, SWD(2018) 137 final, *in* https://eur-lex.europa.eu/legal-content/en/ALL/?uri=CELEX%3A52018SC0137 (31.10.2022)

European Law Institute/Teresa Rodríguez de las Heras *Ballell*, *Guiding Principles for Automated Decision-Making in the EU. ELI Innovation Paper*, 2022, *in* https://www.europeanlawinstitute.eu/fileadmin/user_upload/p_eli/Publications/ELI_Innovation_Paper_on_Guiding_Principles_for_ADM_in_the_EU.pdf (31.10.2022)

Financial Conduct Authority, *Regulatory Sandbox*, 27.03.2022, *in* https://www.fca.org.uk/firms/innovation/regulatory-sandbox (31.10.2022)

Fineman, Martha, "The Vulnerable Subject and the Responsive State", Emory Law Journal [em linha], 60(2), 251–275, *in* https://scholarlycommons.law.emory.edu/elj/vol60/iss2/1 (31.10.2022)

Franke, Ulrike, *Artificial Intelligence diplomacy. Artificial Intelligence governance as a new European Union external policy tool. Study for the special committee on Artificial Intelligence in a Digital Age (AIDA)*, Policy Department for Economic, Scientific and Quality of Life Policies, European Parliament, Luxembourg, 2021, *in* https://www.europarl.europa.eu/RegData/etudes/STUD/2021/662926/IPOL_STU(2021)662926_EN.pdf (31-10-2022)

Future of Life Institute, «*Asilomar AI Principles*», 2017, *in* https://futureoflife.org/2017/08/11/ai-principles/ (31.10.2022)

Future of Life Institute, *Emerging Non-European Monopolies in the Global AI Market*, November 2022, *in* https://futureoflife.org/wp-content/uploads/2022/11/Emerging_Non-European_Monopolies_in_the_Global_AI_Market.pdf (5.11.2022)

Future of Life Institute, *Lessons from the NIST AI RMF for the EU AI Act Input for the US-EU TTC*, April 2022, *in* https://futureoflife.org/wp-content/uploads/2022/08/Lessons_from_NIST_AI_RMF-v2.pdf (31.10.2022)

Future of Life, *General Purpose AI and the AI Act*, May 2022, p. 3, *in* https://artificialintelligenceact.eu/wp-content/uploads/2022/05/General-Purpose-AI-and-the-AI-Act.pdf (31.10.2022)

G20 Ministerial Statement on Trade and Digital Economy, 2021, *in* https://wp.oecd.ai/app/uploads/2021/06/G20-AI-Principles.pdf (31.10.2022)

Goertzel, Ben, "Should Humanity Build a Global AI Nanny to Delay the Singularity Until It's Better Understood?", *Journal of Consciousness Studies*, January 2012 19(1-2),

pp. 96-111, *in* https://www.researchgate.net/publication/233497378_Should_Humanity_Build_a_Global_AI_Nanny_to_Delay_the_Singularity_Until_It%27s_Better_Understood (31.10.2022)

Goo, J. J. / Heo, J.-Y, "The Impact of the Regulatory Sandbox on the Fintech Industry, with a Discussion on the Relation between Regulatory Sandboxes and Open Innovation", *in* https://scholar.google.pt/scholar_url?url=https://www.mdpi.com/2199-8531/6/2/43/pdf&hl=pt-PT&sa=X&ei=6xhEY4v3N4n8mgGB8Z2ADw&scisig=AAGBfm1Y4OXVjEElDMYqTCAVK633jEmpGQ&oi=scholarr (31.10.2022)

Governo português, *AI PORTUGAL 2030. An innovation and growth strategy to foster Artificial Intelligence in Portugal in the European contexto, in* https://www.incode2030.gov.pt/sites/default/files/julho_incode_brochura.pdf (31.10.2022)

Grupo Independente de Peritos de Alto-Nível sobre a Inteligência Artificial, *Orientações Éticas para uma IA de Confiança*, Bruxelas, 2019, pp. 41-45, *in* https://op.europa.eu/en/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1/language-pt/format-PDF (31.10.2022)

Gutierrez, Carlos I. *et al*., "A Proposal for a Definition of General Purpose Artificial Intelligence Systems", *Future of Lige Institute – Working Paper*, November 2022, *in* https://futureoflife.org/wp-content/uploads/2022/11/SSRN-id4238951-1.pdf (5.11.2022)

Hassani, Hossein *et al*, "Artificial Intelligence (AI) or Intelligence Augmentation (IA): What Is the Future?", *AI 2020, 1*, 143–155; doi:10.3390/ai1020008, *in* https://www.researchgate.net/publication/340594734_Artificial_Intelligence_AI_or_Intelligence_Augmentation_IA_What_Is_the_Future (31.10.2022)

High-Level Expert Group on Artificial Intelligence, *A Definition of AI: Main Capabilities and Scientific Disciplines. Definition developed for the purpose of the deliverables of the High-Level Expert Group on AI*, Bruxelas, 2018, *in* https://ec.europa.eu/futurium/en/system/files/ged/ai_hleg_definition_of_ai_18_december_1.pdf (31.10.2022)

Independent High-Level Expert Group on Artificial Intelligence, 23.07.2020, *in* https://futurium.ec.europa.eu/en/european-ai-alliance/document/ai-hleg-sectoral-considerations-policy-and-investment-recommendations-trustworthy-ai (31.10.2022)

Innovation, Science and Economic Development Canada, *Canada's Digital Charter*, *in* https://ised-isde.canada.ca/site/innovation-better-canada/sites/default/files/attachments/1020_04_19-Website_Placemat_v09.pdf (5.11.2022)

International Organization for Standardization, *ISO/IEC 22989:2022. Information technology – Artificial intelligence – Artificial intelligence concepts and terminology*, *in* https://www.iso.org/obp/ui/#iso:std:iso-iec:22989:ed-1:v1:en (5.11.2022)

Ipek, Mikal, "EU Draft Artificial Intelligence Regulation: Extraterritorial application and Effects", *EU Blog*, 17 February 2022, *in* https://europeanlawblog.eu/2022/02/17/eu-draft--artificial-intelligence-regulation-extraterritorial-application-and-effects/ (31.10.2022)

Irish Council for Civil Liberties, https://www.iccl.ie/news/new-liability-rules--on-product-and-ai-are-encouraging-but-need-improvement/, 24.11.2022

Joint Research Center/European Commission, *The Future of Government 2030+. A Citizen Centric Perspective on New Government Models*, 2017-2019 (o Projeto, o Relatório final e as Recomendações podem ser encontrados *in* https://blogs.ec.europa.eu/eupolicylab/futurgov/ (31.10.2022)

Law Commission of Ontario and The Chair on Accountable AI in a Global Context, *Comparing European and Canadian AI Regulation*, November 2021, *in* https://www.lco-cdo.org/wp-content/uploads/2021/12/Comparing-European-and--Canadian-AI-Regulation-Final-November-2021.pdf (31.10.2022)

Leins, Kobi / Kaspersen, Anja, "7 Myths of Using the Term "Human on the Loop": "Just What Do You Think You Are Doing, Dave?", *Artificial Intelligence & Equality Initiative Nov 9, 2021*, *in* https://www.carnegiecouncil.org/media/article/7-myths-of-using--the-term-human-on-the-loop (31.10.2022)

Lemoine, Black, *Is LaMDA Sentient? – an Interview*, 11.06.2022, *in* https://cajundiscordian.medium.com/is-lamda-sentient-an-interview-ea64d916d917 (31.10.2022)

Liu, Zhe / Guo, Yufan / Mahmud, Jalal, "When and Why does a Model Fail? A Human-in-the-loop Error Detection Framework for Sentiment Analysis", *Proceedings of NAACL HLT 2021: Industry Track Papers*, June 6-11, 2021, pp. 170–177, *in* https://aclanthology.org/2021.naacl-industry.22.pdf (31.10.2022)

Maigieri, G. / Ienca, M., "The EU regulates AI but forgets to protect our mind", *EU Law Blog*, 7 July 2021, *in* https://europeanlawblog.eu/2021/07/07/the-eu-regulates-ai--but-forgets-to-protect-our-mind/ (31.10.2022)

Mosqueira-Rey, Eduardo *et al*., "Human-in-the-loop machine learning: a state of the art", *Artif Intell Rev* (2022), *in* https://doi.org/10.1007/s10462-022-10246-w (31.10.2022)

Narayanan, Sundar, *Human-in-the-loop or on-the-loop is not a silver bullet. Evaluate their effectiveness*, *in* https://medium.com/mlearning-ai/human-in-the-loop-or-on-the-loop--is-not-a-silver-bullet-evaluate-their-effectiveness-82f37835d765 (31.10.2022)

NATO, *An Artificial Intelligence Strategy for NATO*, 2021, *in* https://www.nato.int/docu/review/articles/2021/10/25/an-artificial-intelligence-strategy-for-nato/index.html (31.10.2022)

Navarro, Susana Navas, "The future regulation on non-contractual civil liability for AI systems", *Official blog Union Essays*, *24*.11.2022, disponível em https://officialblogofunio.com/2022/11/ (27.11.2022)

Nawaz, Samar Abbas, "The proposed AI Lisbility Rules: Ease or Burden", 07.11.2022 *in* https://europeanlawblog.eu/2022/11/07/the-proposed-eu-ai-liability-rules-ease--or-burden/#more-8626 (08.11.2022)

Noruega e Estados-membros da UE, *Declaration Cooperation on Artificial Intelligence*, 10.04.2018, in https://ec.europa.eu/jrc/communities/en/node/1286/document/eu--declaration-cooperation-artificial-intelligence (31.10.2022)

OECD, "OECD Framework for the classification of AI systems", OECD Digital Economy Papers, February 2022, No. 323, *in* https://www.oecd-ilibrary.org/science-and-technology/oecd-framework-for-the-classification-of-ai-systems_cb6d9eca-en (31.10.2022)

OECD, "Scoping the OECD AI principles: Deliberations of the Expert Group on Artificial Intelligence at the OECD (AIGO)", *OECD Digital Economy Paper*s, November 2019, Nº 291, *in* https://doi.org/10.1787/d62f618a-en (31.10.2022)

OECD, *Artificial Intelligence in Society*, Paris, OECD Publishing, 2019, *in* https://read.oecd--ilibrary.org/science-and-technology/artificial-intelligence-in-society_eedfee77--en#page13 (31.10.2022)

OECD, *Recommendation of the Council on Artificial Intelligence*, 22.05.2019, *in* https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449 (31.10.2022)

PARLAMENTO EUROPEU, *Draft Report on the proposal for a regulation of the European Parliament and of the Council on harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts*, 20.04.2022, 2021/0106(COD), *in* https://www.europarl.europa.eu/doceo/document/CJ40-PR-731563_EN.pdf (31.10.2022)

PARLAMENTO EUROPEU, *Resolução que contém recomendações à Comissão sobre o regime de responsabilidade civil aplicável à inteligência artificial*, (2020/2014(INL), 20.10.2020, *in* https://www.europarl.europa.eu/doceo/document/TA-9-2020-0276_PT.html (31.10.2022)

PARLAMENTO EUROPEU, *Resolução sobre a inteligência artificial na era digital* (2020/2266(INI)), 3.05.2022, *in* https://www.europarl.europa.eu/doceo/document/TA-9-2022-0140_PT.html (31.10.2022)

POLLICINO, ORESTE / GREGORIO, GIOVANNI DE, "Constitutional Law in the Algorithmic Society", *in* HANS-W. MICKLITZ *et al*, *Constitutional Challenges in the Algorithmic Society*, Cambridge, Cambridge University Press, pp. 3-24, *in* https://www.cambridge.org/core/books/constitutional-challenges-in-the-algorithmic-society/constitutional-law-in-the-algorithmic-society/969E0889109C8092AD0AB57019E507E3 (31.10.2022)

SAMOILI, SOFIA *et al*, *AI Watch. Defining Artificial Intelligence 2.0*, EUR 30873 EN, Publications Office of the European Union, Luxembourg, 2021, doi:10.2760/019901, JRC126426, *in* https://publications.jrc.ec.europa.eu/repository/handle/JRC126426 (31.10.2022)

SARTOR, GIOVANNI, "Artificial intelligence and human rights: Between law and ethics", *Maastricht Journal of European and Comparative Law*, 27 (2020), issue 6, pp. 705-719

SCOTT, MARK, "Digital Bridge: Platforms on the hook – Transatlantic AI rulebook – Let's talk data transfers", *Politico.eu*, 6.10.2022, *in* https://www.politico.eu/newsletter/digital-bridge/platforms-on-the-hook-transatlantic-ai-rulebook-lets-talk-data-transfers/ (31.10.2022)

SIEGMANN, CHARLOTTE / ANDERLJUNG, MARKUS, *The Brussels Effect and Artificial Intelligence: How EU regulation will impact the global AI market*, Centre for the Governance of AI, University of Oxford, 2022, *in* https://uploads-ssl.webflow.com/614b70a71b9f71c9c240c7a7/630534b77182a3513398500f_Brussels_Effect_GovAI.pdf (31.10.2022)

SUSSER, D. / ROESSLER, B., NISSENBAUM / H., "Technology, autonomy, and manipulation", *Internet Policy Review*, 8(2), 2019, DOI: 10.14763/2019.2.1410, *in* https://policyreview.info/pdf/policyreview-2019-2-1410.pdf (31.10.2022)

TEGMARK, MAX, *Life 3.0. Ser-se Humano na Era da Inteligência Artificial*, (trad. port. de J. Van Zeller), Alfragide, 2019

UK GOVERNMENT, *Policy paper Establishing a pro-innovation approach to regulating AI*, 18.07.2022, *in* https://www.gov.uk/government/publications/establishing-a-pro-innovation-approach-to-regulating-ai/establishing-a-pro-innovation-approach-to-regulating-ai-policy-statement (31.10.2022)

UNESCO, *Recommendation on the Ethics of Artificial Intelligence*, 2022, *in* https://unesdoc.unesco.org/ark:/48223/pf0000381137 (5.11.2022)

US CONGRESS, *National Initiative Act 2020*, *in* https://www.congress.gov/116/crpt/hrpt617/CRPT-116hrpt617.pdf#page=1210 (31.10.2022)

UUK, RISTO, *Manipulation and the AI Act*, April 2022, *in* https://futureoflife.org/wp-content/uploads/2022/08/FLI-Manipulation_AI_Act.pdf (31.10.2022)

Veale, M. / Borgesius, F. Z., "Demystifying the Draft EU Artificial Intelligence Act. Analysing the good, the bad, and the unclear elements of the proposed approach", *Computer Law Review International*, 4/21, 97-112, *in* https://osf.io/preprints/socarxiv/38p5f/ (31.10.2022)

Wang, Ge, *Humans in the Loop: The Design of Interactive AI Systems*, Stanford University Human-Centered Artificial Intelligence, 20.10.2019, *in* https://hai.stanford.edu/news/humans-loop-design-interactive-ai-systems (31.10.2022)

Wendehorst, Christiane, "Liability for Artificial Intelligence The Need to Address Both Safety Risks and Fundamental Rights Risks", *in* Silja Voeneky, Philipp Kellmeyer, Oliver Mueller and Wolfram Burgard (Eds.), *The Cambridge Handbook of Responsible Artificial Intelligence. Interdisciplinary Perspectives*, Cambridge, Cambridge University Press, pp. 187-209, disponível em https://doi.org/10.1017/9781009207898.016 (27.11.2022)

White House Office of Science and Technology Policy, *Blueprint for an AI Bill of Rights: Making Automated Systems Work for the American People*, Washignton, October 2022, *in* https://www.whitehouse.gov/wp-content/uploads/2022/10/Blueprint-for--an-AI-Bill-of-Rights.pdf (31.10.2022)

Zech, Herbert, "Liability for AI: public policy considerations", *ERA Forum (2021)* 22:147–158, p. 149. Disponível em https://link.springer.com/content/pdf/10.1007/s12027-020-00648-0.pdf (27.11.2022)

Inteligência artificial e inteligência coletiva

Artificial Intelligence and collective intelligence

Luísa Neto*

RESUMO: Se o enquadramento democrático impõe a transparência da esfera pública, a inteligência artificial tem uma importância fundamental numa sociedade democrática como forma de legitimação da cidadania política e de organização pública. Os novos meios de construir uma esfera pública inclusiva do ponto de vista da argumentação devem ter em conta a necessidade de não minar novas exigências de transparência no que diz respeito à sensibilização política. Se um Estado de Direito Democrático pode de alguma forma ser definido como uma combinação adequada da consagração dos direitos fundamentais e do equilíbrio e separação de poderes, o quadro digital reúne novos desafios e também novas soluções possíveis.

PALAVRAS -CHAVE: Contrato Social; Democracia; Espaço Público; Inteligência Artificial; Inteligência Coletiva

ABSTRACT: If democracy sets its pace along with transparency of the public sphere, artificial intelligence has a key importance in a democratic society as a way of legitimation of political citizenship and of organising the public sphere.

* Presidente do Conselho Diretivo do Instituto Nacional de Administração, I. P. Professora Associada com Agregação da FDUP e CIJ-FDUP. lneto@direito.up.pt. ORCID – 0000-0001-5561-6302

Indeed, new means of building up an inclusive public sphere and reasoning must take into account the need to not undermine new demands of transparency in what concerns political awareness. If a democratic state of law can be somehow defined as an adequate combination of the enshrinement of fundamental rights and of the balance and separation of powers, the digital framework brings together new challenges and also new possible solutions.

KEYWORDS: Democracy; Social Contract; Public sphere; Artificial Intelligence; Collective Intelligence

1. A redefinição do espaço público

Na aldeia planetária do advento da telemática – e mesmo que consigamos com dificuldade permanecer alheios ao terror concentracionista de George Orwell –, um sistema informativo é hoje verdadeiro suporte da ação política e do diálogo entre os órgãos do poder e os cidadãos, da expressão dos acontecimentos e opiniões com impacte público, sendo a difusão de conhecimentos essencial ao desenvolvimento económico-social.

Ora, ainda que "[E]xistam diversas definições de "Sociedade da Informação e do conhecimento", (...) podemos assumir (Livro Verde para a Sociedade da Informação (datado de 1997), que a *expressão 'Sociedade da Informação' se refere a um modo de desenvolvimento social e económico em que a aquisição, armazenamento, processamento, valorização, transmissão, distribuição e disseminação de informação conducente à criação de conhecimento e à satisfação das necessidades dos cidadãos e das empresas, desempenham um papel central na atividade económica, na criação de riqueza, na definição da qualidade de vida dos cidadãos e das suas práticas culturais."* [1] Assim, se a informação sempre funcionou como válvula catártica de escape no espaço público, no contexto cibernético atual, a intersecção dos bens protegidos implica uma reconceptualização dos fundamentos em causa.

[1] Veja-se ainda, em termos históricos, o tratamento feito por JOSÉ DE OLIVEIRA ASCENSÃO em *A sociedade da informação*, Coimbra, Coimbra Editora, 1999, pps. 163-184, Sep. de Direito da Sociedade da Informação, Vol. 1 e *Direito da sociedade da informação*, ed. Faculdade de Direito da Universidade de Lisboa, Associação Portuguesa do Direito Intelectual, Alberto de Sá e Mello *et al*, Coimbra, Coimbra Editora, 1999-2003.

Como afirmou Hillary Clinton na Washington University em 15 de fevereiro de 2011[2]: *"The Internet has become the public space of the 21st century – the world's town square, classroom, marketplace, coffeehouse, and nightclub".*

Sabemos que o ambiente de Estado de Direito Democrático reclama um contexto de transparência tão bem sintetizado por David Brin no título da sua obra de 1998[3]. Neste sentido, as matérias atinentes à sociedade de informação suscitam o interesse da comunidade em geral e integram os "fundamentais do Estado", por serem hoje pedra de toque do discurso democrático e da construção de uma sociedade formal e materialmente inclusiva, colocando desafios importantes quanto ao próprio desenho do que sejam as tradicionalmente entendidas divisão e separação de poderes.

Assim, a atual *ágora* ou o espaço público moderno delineado por Habermas[4] apresenta necessidades e condicionantes diferentes, numa tentativa de sopesagem de tendências de desregulamentação, privatização e liberalização e num duplo sentido, complementar, de mundialização e localização.

Na síntese de Ignacio Ramonet[5]:"três esferas que antes eram autónomas: de um lado, a cultura de massa, com sua lógica comercial, suas criações populares, seus objetivos basicamente mercantis; de outro, a comunicação, no sentido publicitário, o marketing, a propaganda, a retórica da persuasão; e, finalmente, a informação, com suas agências de notícias, boletins de radiodifusão ou de televisão, a imprensa, as redes de informação contínua (...) foram-se misturando pouco a pouco, até constituírem uma única esfera, ciclópica, na qual

[2] Hillary Clinton, "Internet Rights and Wrongs: Choices & Challenges in a Networked World", George Washington University, 15 de Fevereiro 2011, disponível em https://2009-2017.state.gov/secretary/20092013clinton/rm/2011/02/156619.htm

[3] David Brin, *The Transparent Society,* Boston: Addison Wesley, 1998.

[4] J. Habermas, *L'Espace Public, Archéologie de la publicité comme dimension constitutive de la société bourgeoise*, Paris: Payot, 1978 (ed. orig., 1962)/ Jurgen Habermas, *Mudança Estrutural da Esfera Pública*, Rio de Janeiro: Tempo Brasileiro, 1984. Em termos de actualização, vejam-se por todos, Cândido Monzón, *Opinión Pública, Comunicación y Política,* Madrid: Tecnos, 1996, James Curran, "Rethinking mass communication", in James Curran, David Morley e Valerie Walkerdine (eds), *Cultural Studies and Communications,* Londres: Arnold, 1996 e N. Fraser, "Rethinking the public sphere: a contribution to the critique of actually existing democracy", in C. Calhoun (ed), Peter Dahlgren (ed), *Communication and citizenship: Journalism and the public sphere*, London and NY: Routledge, 1997 (reprint). Veja-se ainda Paulo Ferreira da Cunha, *"Sociedade da Informação e Estado de Direito Democrático, Direito à Informação ou deveres de protecção informativa do Estado"*, Revista da FDUP, Edição comemorativa dos dez anos da FDUP, Coimbra: Coimbra Editora, 2006, pp. 623-651, em especial quanto aos conceitos de "comunicação hierarquizada" e "comunicação reticular".

[5] Editorial do *Monde Diplomatique* de Outubro de 2003.

é cada vez mais difícil distinguir as actividades pertencentes à cultura de massa, à comunicação ou à informação". [6]

Numa outra abordagem, Javier Echevarria[7] lembrava que "a nova forma de ser súbdito consiste em dedicar quotidianamente algum tempo a escutar ou a ver a corte". Ora, se a inteligência artificial não é mais do que uma forma de preparação – repito, de preparação – da informação, fundamento da democracia, novos desafios resultam hoje da necessidade de considerar suportes diferenciados que permitam o *animus* sociocultural da nova sociedade da informação [8], em contextos que são hoje também os de uma – ao menos no entender de alguns – ciberdemocracia.[9]

Sucede que esta – *quod erat demonstrandum* – ciberdemocracia, que enlaça a sociedade da informação, impõe exigências diferenciadas ao ramo de Direito que delas se ocupa, dada a necessidade de contemporização com a reserva do "tecnologicamente possível"[10] e a inexistência de claros limites espaciais de aplicação normativa.

Assim, se a proteção da individualidade da pessoa em contexto cibernético deve ser a continuidade da proteção garantida pelo Estado na realidade *offline,* a imaterialidade convoca todavia desafios acrescidos. Trata-se pois de saber se é ainda possível a utilização dos "odres velhos" para conter e disciplinar o "vinho novo" da Internet[11], em especial quanto à suscetibilidade de aplicação da dogmática comum da responsabilidade civil e penal e da proteção dos direitos fundamentais à realidade fáctica da inteligência artificial.

[6] *Idem*, op. e loc. cits. Vejam-se ainda Luísa Neto, "Ciência da Informação e Direito: um novo paradigma de reconstrução do papel do Estado no contexto da sociedade global de informação", *in A informação jurídica na era digital*, 24 e 25 de Fevereiro de 2011, Porto: Afrontamento, CETAC Media, 2012, pp. 41-55 e Luísa Neto, "Um outro tipo de "freios e contrapesos": a comunicação social no contexto do Estado de Direito Democrático", *Estudos em homenagem ao Prof. Doutor Jorge Miranda,* FDUL, Coimbra: Coimbra Editora, Volume II (Direito Constitucional e Justiça constitucional) Coimbra Editora, 2012, pp. 455-509.

[7] *Apud* Francisco Rui Cádima, *Desafios dos novos media: a nova ordem política e comunicacional*, 2ª ed, Lisboa: Editorial Notícias, 1999, p. 26.

[8] J. Ferreira Salgado, *Informação e civilização*, Ordem dos Advogados, Conselho Distrital do Porto, 1973, Conferência proferida pelo autor em 11 de Maio de 1973.

[9] Vejam-se Timothy Fenoulhet, *Democracy and the information society*, European Commission, DG XIII, January 1996, Noam Chomsky, *Necessary Illusions*, London: Pluto, 1989 e Pierre Lévy, *Ciberdemocracia*, Lisboa: Instituto Piaget, 2003, pps. 28, 33 e ss e 171.

[10] Sobre este conceito, leia-se Carla Amado Gomes, "Estado Social e concretização de DF na era tecnológica", RFDUP, Coimbra: Coimbra Editora, A. 7, 2010, pp. 19-34.

[11] Expressão celebrizada por Manuel Carneiro da Frada em "Vinho novo em odres velhos?»/A responsabilidade civil das «operadoras de Internet» e a doutrina comum da imputação de danos", Revista da Ordem dos Advogados, ano 59, II, Abril de 1999), depois objeto de republicação em Direito da *Sociedade de Informação*, II, Coimbra: Coimbra Editora, 2001.

Fundamentalmente, a discussão da responsabilização encontra escolhos na imensidão da informação disponível, não filtrada, agregada, trabalhada, manipulada, constantemente atualizada e virtualmente impossível de apagar[12], contribuindo para a qualificação de Castells de sociedade em rede ou para aquilo a que Domingos Farinho já qualificou como "verdadeiro ecossistema virtual"[13], composto por uma miríade de diferenciados agentes – não raro com interesses últimos contraditórios e de difícil compatibilização, como Estados, operadores de telecomunicações, fornecedores de acesso à Internet, fornecedores de bens e serviços *online*, gestores dos *websites* e, claro, os utilizadores, ou cibernautas – e ancora-se em um carácter não espacial de aplicação da regulação e de escolha e, ainda, no *a priori* lícito anonimato digital, que potencia o ferrão da "letra escarlate digital" de que fala Daniel Solove.[14]

2. A inteligência artificial

É que o principal objetivo dos sistemas de inteligência artificial, é o de executar funções que, caso fossem executadas por um ser humano, seriam consideradas inteligentes. É um conceito amplo, e que recebe tantas definições quantos os significados diferentes que damos à palavra 'inteligência', como a capacidade de raciocínio, de aprendizagem, de reconhecimento de padrões e inferência.

O seu desenvolvimento tem extrapolado os clássicos programas de xadrez ou de conversão e envolvido áreas como visão computacional, análise e síntese da voz, lógica difusa, redes neurais artificiais e muitas outras. Se inicialmente a inteligência artificial visava reproduzir o pensamento humano, hoje abraçou a ideia de reproduzir faculdades humanas como criatividade, auto-aperfeiçoamento e uso da linguagem.

São meramente exemplificativas as aplicações práticas de técnicas e métodos de inteligência artificial no âmbito da governança pública, da contribuição para os objetivos de políticas públicas, ou para as forças armadas, no âmbito da saúde ou do comércio eletrónico.

[12] A 'reserva do tecnologicamente possível' impede a real concretização do *right to be forgotten (droit à l'oubli)* em atual equação acrescida desde a respetiva previsão no artigo 17º do Regulamento Geral de Proteção de Dados – Regulamento (UE) 2016/679 do Parlamento e do Conselho, de 27 de abril de 2016, relativo à proteção das pessoas singulares no que diz respeito ao tratamento de dados pessoais e à livre circulação desses dados -, a que em Portugal acresce a consideração da Lei nº 58/2019, de 8 de Agosto.

[13] Domingos Soares Farinho, *Intimidade da Vida Privada e Media no Ciberespaço*, Coimbra: Almedina, 2007, p. 14.

[14] Veja-se Daniel J. Solove, *The Future of Reputation: Gossip, Rumor, and Privacy on the Internet*, Yale University, 2007, pp. 76 ss.

As vantagens da redução de erros, da possível síntese de esforços não deixam de considerar as desvantagens do custo, da eventual falta de criatividade, da projeção de um crescimento do desemprego.

Se a inteligência artificial surge como objeto de estudo desde os anos 40 do século XX, a inevitável ponderação da ligação aos direitos fundamentais só emerge a partir dos anos 70 do mesmo século[15] muito por influência da discussão sobre os *big data* e segue hoje um périplo pontuado por considerações prospetivas sobre as neurociências e a discussão da pessoalidade em detrimento da personalidade.

De facto, elenquem-se a título de inventário:

- A concessão da cidadania pela Arábia Saudita ao robot Sophia em Outubro de 2017 e discussão da atribuição da personalidade jurídica a entidades computorizadas[16];
- A resolução do Parlamento Europeu, de 16 de fevereiro de 2017, que contém recomendações à Comissão sobre disposições de Direito Civil sobre Robótica (2015/2103(INL));[17]
- O guia ético para a inteligência artificial lançado pela União Europeia desde Dezembro de 2018, da responsabilidade do HighLevel Expert Group on Artificial Intelligence (AI HLEG);
- A possibilidade de técnicas não invasivas de identificação de correlações mentais com especial importância para efeitos de marketing direto, abrangidas pelo termo geral de 'neurotecnologia persuasiva';[18]
- As possibilidades de estimulação magnética transcraniana e de estimulação cerebral: a utilização das neurociências para melhorar a função cerebral normal e as questões relativas à autonomia e privacidade/intimidade[19];

[15] Marcello Ienca, "Towards new human rights in the age of neuroscience and neurotechnology, *Life Sciences, Society and Policy*, 13, 5, 2017, em *https://doi.org/10.1186/s40504-017-0050-1*

[16] Vide Carlos Rogel (coord.), *Los robots y el Derecho*, Madrid, Reus Editorial, Colección Jurídica General, 2018.

[17] Veja-se sobre o tema em especial, Javier Ercilla García, *Normas de derecho civil y robótica*, Pamplona, Aranzadi, 2018.

[18] A. Fernandez, N. Sriraman, B. Gurevitz, O. Ouiller, *Pervasive Neurotechnology: A Groundbreaking Analysis of 10,000+ Patent Filings Transforming Medicine, Health, Entertainment and Business*, SharpBrains, 2015.

[19] I. Persson e J. Savulescu, "The perils of cognitive enhancement and the urgent imperative to enhance the moral character of humanity", *Int J Appl Philos.*, 25(3), 162–77, 2008; W. Sententia, "Neuroethical considerations: cognitive liberty and converging technologies for improving human cognition", *Ann N Y Acad Sci,* 1013(1), pp. 221–8, 2004.

- Em especial, a discussão quanto ao melhoramento cognitivo[20] (*neuro-enhancement): intelligence pill* ou *intelligent chip*[21]*;*
- A autoconsciência e a susceptibilidade de mudança de personalidade[22]*;*
- A emergência da arquitectura da escolha e do *nudging* [23] por oposição à racionalidade deliberativa de John Rawls e como sucedâneo das teorias paternalistas (em especial, libertárias);[24]

[20] Martha Farah *et al* , "Neurocognitive Enhancement: what can we do and what should we do?, *Nature Reviews Neuroscience*, v. 5, n. 5, p. 421-425, 2004, disponível em https://repository.upenn.edu/neuroethics_pubs/9, 2004 e Martha Farah, Emerging Ethical Issues in Neuroscience. Nature neuroscience. 5. 1123-9. 10.1038/nn1102-1123, 2002

[21] Veja-se o artigo "The Ethics of Brain Science: Open Your Mind", *The Economist*, 2002 May 25; 363(8274): 5 p. [Online]. http://web.ebscohost.com/ehost/delivery?vid=8&hid=5&sid=c902ab2a-0616-456...

Veja-se ainda Luísa Neto, em co-autoria com Rui Vieira da Cunha) *Compulsory Neuro-Interventions: Metaphysical and Conceptual Foundations of the Subject of Responsibility and Autonomy of Choice*, Estudos Comemorativos dos 20 anos da FDUP, FDUP, 2017, Vol II., pp. 119-144. John E. Harris, "Moral Enhancement and Freedom", Bioethics, 25(2), 102–111; 2011. Elisabeth Hildt e Andreas G. Franke (eds.), *Cognitive Enhancement: An Interdisciplinary Perspective*, Dordrecht, Springer, 2013. Neil LEVY, "Is Neurolaw Conceptually Confused?", 18(2) J. Ethics, 2014. Kasper LIPPERT-RASMUSSEN, "Neuroprediction, Truth-Sensitivity, and the Law", 18(2) J. Ethics, 2014. Michael Pardo e Dennis Patterson, *Minds, Brains, and Law: The Conceptual Foundations of Law and Neuroscience*, New York, Oxford University Press, 2013. Ingmar Persson e Julian Savulescu, *Unfit for the Future: The Need for Moral Enhancement*. Oxford, Oxford University Press, 2012. Thomas Søbirk Petersen, "(Neuro)prediction, Dangerousness, and Retributivism, 18(2)" J. Ethics, 18(2), 2014. Julian Savulescu e Nick Bostrom (eds.), *Human Enhancement*. Oxford, Oxford University Press, 2010; Stephen J. Morse, *The Promise of Neuroscience for Law: Hope or Hype?, in* Palgrave Handbook of Philosophy and Public Policy 77–96 (David Boonin ed., 2018).

[22] C. Lewis, *et al,* "Subjectively perceived personality and mood changes associated with subthalamic stimulation in patients with Parkinson's disease", *Psychol Med*., 45(01), 73–85, 2015; R. Mackenzie, "Who should hold the remote for the new me? Cognitive, affective, and behavioral side effects of DBS and authentic choices over future personalities", *Ajob Neurosci*, 2(1), pp. 18–20, 2011.

[23] Richard H. Thaler e Cass R. Sunstein, *Nudge: Improving decisions about health, wealth, and happiness*, New Haven, Yale University Press, CT, 2008.

[24] Ayala Arad e Ariel Rubinstein, "The People's Perspective on Libertarian-Paternalistic Policies," July 2017, disponível em https://www.tau.ac.il/~aradayal/LP.pdf. Veja-se ainda Adrien Barton e Till Grüne-Yanoff, "From Libertarian Paternalism to Nudging–and Beyond", Review of Philosophy and Psychology 6 (3), 341-359, 2015; Andreas Kapsner & Barbara Sandfuchs, "Nudging as a Threat to Privacy", *Review of Philosophy and Psychology* 6 (3):455-468, 2015; Robert Lepenies & Magdalena Małecka, "The Institutional Consequences of Nudging – Nudges", Politics, and the Law, *Review of Philosophy and Psychology* 6 (3):427-437, 2015; T. M. Wilkinson, "Nudging and manipulation", *Political Sudies*, Volume: 61 issue: 2, page(s): 341-355, 2013.

- A discussão da "boa manipulação" e a rampa escorregadia para intervenções não consentidas, em especial no caso do *nudging* preventivo ou antecipatório
- A suscetibilidade das implicações de desigualdade económica e social;
- As relações com os tribunais e o interesse especial no caso das relações com o dever de obediência à lei e a responsabilidade para efeitos do direito dos contratos e do direito penal[25];
- A suscetibilidade agravada de violação do direito à privacidade, à liberdade de pensamento, de integridade pessoal, de protecção legal contra a discriminação, de direito a um julgamento justo, de proteção contra a não auto-incriminação, de pode resultar a eventual emergência de novos direitos: o direito à liberdade cognitiva, ou direito à privacidade mental, o direito à integridade mental e o direito à continuidade psicológica, todos ainda sem referência expressa nos instrumentos internacionais dos direitos do homem mas com apoio consistente na *jurisprudence of the mind* defendida por Glen Boire[26] na formulação ampliada da protecção da privacidade resultante da protecção do artigo 12º da DUDH e do artigo 8º da CEDH;

3. Da inteligência artificial à inteligência coletiva e vice-versa

Ora, como resumiria *avant la lettre* o filosofo Pierre Lévy – em *A Inteligência Colectiva, Para uma antropologia do ciberespaço*, obra com primeira edição em 1994[27], "[A] questão é: como usaremos as novas tecnologias de forma significativa para aumentar a inteligência humana coletiva?"

E se computadores pudessem capturar as palavras que digitamos na internet e convertê-las em uma linguagem que descreva seus significados?

[25] Sobre este tópico, vejam-se, Valerie Gray Hardcastle, *Why Brain Images (Probably) Should Not Be Used in US Criminal Trials*, *in* Palgrave Handbook of Philosophy and Public Policy 25–37 (David Boonin ed., 2018); Natalie S. Gordon & Mark R. Fondacaro, *Rethinking the Voluntary Act Requirement: Implications from Neuroscience and Behavioral Science Research*, 36 Behav. Sci. L. 426–36, 2018; William Hirstein, Katrina L. Sifferd & Tyler Fagan, *Responsible Brains: Neuroscience, Law, and Human Culpability*, 2018; A. Zangrossi, S. Agosta, G. Cervesato, F. Tessarotto, & G. Sartori Zangrossi, *"I Didn't Want To Do It!" The Detection of Past Intentions*, 9 Front Hum Neurosci. 608, 2015.

[26] RG Boire, "Mind matters", *Journal of Cognitive Liberties*, 4, (1), 7–10, 2003.

[27] Pierre Lévy, *A Inteligência Colectiva, Para uma antropologia do ciberespaço*, obra com primeira edição em 1994. Lévy baseou-se em tríades inspiradas na conexão tripla entre o signo, a coisa representada e a cognição produzida na mente, em termos previamente definidos pelo semiólogo americano Charles Sanders Peirce.

Os dados analisados revelariam *insights* de questões mais profundas sobre emoções e motivações humanas, em vez de meras estatísticas e dados.

É claro que a espécie humana é um maravilhoso exemplo de inteligência coletiva que cria cultura porque assente em uma inteligência pessoal reflexiva que aumenta a capacidade da inteligência coletiva global.

A questão é a do aumento da inteligência humana coletiva através de sistemas simbólicos, no estágio digital ou, como refere Pierre Lévy, no estágio algorítmico.

Neste sentido, a possível externalização da memória humana coletiva e dos processos intelectuais aumentou a autonomia individual e a auto-organização das comunidades humanas. A disponibilidade de grandes fluxos de dados – *big data* – é apenas uma atualização do potencial da *internet* e conduz a uma esfera pública global e hipermediada.

De facto, o maior potencial do meio algorítmico não é a transmissão de informação: é a transformação automática de dados. A denominada "web semântica" baseia-se em conexões lógicas entre dados e em modelos algébricos de lógica.

Se para Lévy, a inteligência coletiva pode ser dividida em inteligência técnica, conceitual e emocional, no mundo atual a atualização das ideias enquanto capital só pode ser atingido quando se pensa em conjunto.

4. A integração da inteligência artificial no contrato social

Muitos são os que alertam para o facto de a política e de a economia dependerem fortemente da informação – e sobretudo da informação personalizada que permita ao mercado atuar de forma cada vez mais precisa. É aliás provavelmente nesta sede que surge como mais evidente a intersecção com a evidência de que já se falou: a informação como mercadoria, neste caso paredes meia com a ideia da regulação dos serviços da sociedade da informação como incluindo o comércio eletrónico.[28] Aliás, não pode deixar de se assinalar que a Internet opera, ainda, a transformação do direito de autor em ato económico de consumo, sendo curioso que as matérias do "direito de autor tecnológico" estejam hoje a ser tratadas no domínio do regime do comércio eletrónico[29] ou da concorrência desleal.

[28] Neste sentido específico, vejam-se B. D. Loader, "The governance of cyberspace. Politics, technology and global restructuring", *in* B. D. Loader (ed.), *The Governance of Cyberspace*, 1997, Londres, Routledge, pp. 1-19, J. Slevin, *The Internet and Society*, 2000, Oxford: Polity Press, pp. 1-10 e pp. 214-235.

[29] Esta é matéria que está naturalmente apartada do escopo do nosso tema, nos termos em que foi o mesmo delimitado previamente. Ainda assim, para além do Documento Orientador da Iniciativa

O Direito não é alheio à necessidade de regulação *hoc casu*. Há normas, e há sanções. Já questão diferente é a da suscetibilidade ou da facilidade na sua aplicação.

Tal como salienta José de Oliveira Ascensão,[30] está provavelmente hoje em causa a opção entre a regulamentação ou regulação, ou a visão da regulamentação como entrave à regulação. Mas este debate tem variadas perspectivas paralelas e reciprocamente subsidiárias: económica, jurídica, tecnológica.

A proposta de consideração e justificação do contrato social – de um *overlapping consensus* – mais não fez do que pretender conferir lastro normativo ao fundamento do exercício do poder.

No entanto, abstendo-se de ingerências preventivas ilícitas e de excessivas cláusulas de limitação, o Estado pode adotar uma de várias atitudes não exclusivas ou excludentes: a de abstenção e/ou neutralidade, o traçar de linhas de política, a participação na gestão, a total gestão. Trata-se de gerir as opções políticas de base que sustentam um direito à verdade e à transparência, discutindo a preferência por um modelo de auto-regulação (por muitos referenciada como *netiquette*), co-regulação – *v.g.* com aprovação de códigos de conduta *inter pares* – ou hetero-regulação ou, ainda, a da assunção de um modelo cumulativo.

Existem mil maneiras de juntar os homens, não existe senão uma de os unir.[31] *clamava Rousseau*. "É necessário considerar ainda que a deliberação pública que pode obrigar todos os súbditos para com o soberano (...) não pode pela razão contrária obrigar o soberano para consigo próprio, e que, por consequência, é contra a natureza do corpo político que o soberano imponha a si próprio uma lei que ele não possa infringir. Não podendo considerar-se senão sob uma só e mesma relação, ele está então na mesma situação de um particular que tivesse realizado um contrato consigo próprio; por isto se vê que não há nem pode haver nenhuma espécie de Lei Fundamental obrigatória para o corpo do Povo, o que não significa que esse corpo não possa perfeitamente comprometer-se com outrem, pelo menos naquilo que não é contrário à sua natureza (...)", propunha Rousseau.[32]

Nacional para o Comércio Electrónico, veja-se ainda Alexandre Libório Dias Pereira, *Serviços da sociedade da informação: alguns problemas jurídicos do comércio electrónico na Internet,* Lisboa: Faculdade de Direito da Universidade Nova de Lisboa, 2001.

[30] José de Oliveira Ascensão, "E agora? Pesquisa do Futuro Próximo", in AA.VV., *Sociedade da Informação. Estudos Jurídicos*, Coimbra: Livraria Almedina, 1999, p. 62.

[31] Jean Jacques Rousseau, *O Contrato Social*, Capítulo V.

[32] *Idem*, Capítulo III..

No contexto cibernético, se todos realçamos a importância da proteção dos direitos fundamentais na complexidade das relações jurídico-sociais, se têm aqui operância as teses quanto à vinculação do legislador ordinário pelos direitos fundamentais, há também que não ter medo de afirmar que se deve atender à chamada teoria das bagatelas como princípio de interpretação e conformação social: de facto, tal como o direito penal exclui a proteção típica quando entende não existir adequação social, é importante que percebamos que o âmbito de proteção do direito à reserva da vida privada exige hoje, no mundo atual, uma configuração obviamente distinta daquela a que se referia Brandeis no final do século XIX. Mormente no cenário *online*.

É este o nó górdio da racionalidade deliberativa[33] também para Rawls: "A pessoa que decide possui uma certa competência – conhece as características gerais das suas necessidades e objetivos, tanto presentes como futuros, e é capaz de calcular a intensidade relativa dos seus desejos e de decidir, se necessário, o que é que realmente deseja."

5. Conclusão

Em suma, "[O]s novos meios tecnológicos conseguiram, simultânea e paradoxalmente, alargar a liberdade de informação e desresponsabilizar os intervenientes."[34] A ideia de sociedade transparente a que já aludimos mostra bem que a questão hoje em dia se deslocou da discussão sobre "o que cabe no âmbito de protecção" para a necessidade de tutela face a novos meios. Afirmada uma linha de continuidade entre o *online* e o *offline* não conseguiremos, no entanto, e provavelmente, a solução apenas com um mero *update* das regras existentes, ainda que a transparência dos procedimentos de restrição de direitos deva ser sempre centrada num modelo de legitimação procedimental ancorado num teste triplo da proporcionalidade – necessidade, adequação e proibição do excesso -, que parece simplista, mas que é, ao invés, exigente e clarificador.

Quanto mais soubermos as respostas para estas questões, maior será a transparência da fonte e maior nossa confiança nela. E repare-se que este imperativo de fonte de informação transparente é muito próxima do pensamento científico. Porque o conhecimento científico deve ser capaz de res-

[33] John Rawls, *Uma teoria da justiça, Lisboa, Fundamentos,* 1993, p. 321.

[34] Vejam-se ainda Paulo Ferreira da Cunha, "*Sociedade da Informação e Estado de Direito Democrático, Direito à Informação ou deveres de protecção informativa do Estado",* op.cit., p. 341, e Dominique Wolton, *Internet et après ?*, Paris: Flammarion, Coll. "Champs", 2000 (Dominique Wolton, *E depois da Internet? Para uma teoria crítica dos novos media*, Lisboa: Difel, 2000).

ponder questões tais quais: de onde vem os dados? De onde vem a teoria? De onde vêm os financiamentos? A transparência é a nova objetividade. Rawls falaria hoje da necessidade de contemporizar – ou de rasgar – um outro tipo de véu da ignorância.

Lista Bibliográfica

Fernandez, A., Sriraman, N., Gurevitz, B., Ouiller, O., *Pervasive Neurotechnology: A Groundbreaking Analysis of 10,000+ Patent Filings Transforming Medicine, Health, Entertainment and Business*, SharpBrains, 2015.

Arad, Ayala e Rubinstein, Ariel, "The People's Perspective on Libertarian-Paternalistic Policies," July 2017, disponível em https://www.tau.ac.il/~aradayal/LP.pdf

Ascensão, José De Oliveira, "E agora? Pesquisa do Futuro Próximo", in AA.VV., *Sociedade da Informação. Estudos Jurídicos*, Coimbra, Livraria Almedina, 1999.

Ascensão, José De Oliveira, *A sociedade da informação,* Coimbra, Coimbra Editora, 1999 Ascensão, José De Oliveira, *Direito da sociedade da informação*, ed. Faculdade de Direito da Universidade de Lisboa, Associação Portuguesa do Direito Intelectual, Alberto de Sá e Mello *et al*, Coimbra, Coimbra Editora, 1999-2003.

Barton, Adrien e Grüne-Yanoff, Till, "From Libertarian Paternalism to Nudging–and Beyond", Review of Philosophy and Psychology 6 (3), 341-359, 2015.

Boire, RG, "Mind matters", *Journal of Cognitive Liberties*, 4, (1), 7–10, 2003.

Brin, David, *The Transparent Society,* Boston: Addison Wesley, 1998.

Cádima, Francisco Rui, *Desafios dos novos media: a nova ordem política e comunicacional,* 2ª ed, Lisboa, Editorial Notícias, 1999.

Chomsky, Noam, *Necessary Illusions*, London, Pluto, 1989.

Cunha, Paulo Ferreira, *"Sociedade da Informação e Estado de Direito Democrático, Direito à Informação ou deveres de protecção informativa do Estado"*, Revista da FDUP, Edição comemorativa dos dez anos da FDUP, Coimbra, Coimbra Editora, 2006.

Curran, James, "Rethinking mass communication", in James Curran, David Morley e Valerie Walkerdine (eds), *Cultural Studies and Communications,* Londres: Arnold, 1996.

E. Harris, John, "Moral Enhancement and Freedom", Bioethics, 25(2), 102–111 ; 2011.

Farah, Martha *et al* , "Neurocognitive Enhancement: what can we do and what should we do?, *Nature Reviews Neuroscience*, v. 5, n. 5, p. 421-425, 2004, disponível em https://repository.upenn.edu/neuroethics_pubs/9, 2004.

Farah, Martha, Emerging Ethical Issues in Neuroscience. Nature neuroscience. 5. 1123-9. 10.1038/nn1102-1123, 2002.

Farinho, Domingos Soares, *Intimidade da Vida Privada e Media no Ciberespaço*, Coimbra: Almedina, 2007.

Fenoulhet, Timothy, *Democracy and the information society*, European Commission, DG XIII, January 1996.

Frada, Manuel Carneiro, "Vinho novo em odres velhos?»/A responsabilidade civil das «operadoras de Internet» e a doutrina comum da imputação de danos*"*, Revista da Ordem dos Advogados, ano 59, II, Abril de 1999).

Fraser, N., "Rethinking the public sphere: a contribution to the critique of actually existing democracy", in C. Calhoun (ed), Peter Dahlgren (ed), *Communication and citizenship: Journalism and the public sphere*, London and NY: Routledge, 1997 (reprint).

García, Javier Ercilla, *Normas de derecho civil y robótica*, Pamplona, Aranzadi, 2018.

Gomes, Carla Amado, "Estado Social e concretização de DF na era tecnológica", RFDUP, Coimbra, Coimbra Editora, A. 7, 2010, pp. 19-34.

Gordon, Natalie S. & Mark R. Fondacaro, *Rethinking the Voluntary Act Requirement: Implications from Neuroscience and Behavioral Science Research*, 36 Behav. Sci. L. 426–36, 2018; HIRSTEIN, William, Katrina L. Sifferd, & Tyler Fagan, *Responsible Brains: Neuroscience, Law, and Human Culpability*, 2018

Habermas, J., *L'Espace Public, Archéologie de la publicité comme dimension constitutive de la société bourgeoise*, Paris, Payot, 1978 (ed. orig., 1962)/ Jurgen Habermas, *Mudança Estrutural da Esfera Pública*, Rio de Janeiro: Tempo Brasileiro, 1984.

Hardcastle, Valerie Gray, *Why Brain Images (Probably) Should Not Be Used in US Criminal Trials*, *in* Palgrave Handbook of Philosophy and Public Policy 25–37 (David Boonin ed., 2018).

Hildt, Elisabeth & Franke, Andreas G. (eds.), *Cognitive Enhancement: An Interdisciplinary Perspective*, Dordrecht, Springer, 2013.

Levy, Neil, "Is Neurolaw Conceptually Confused?", 18(2) J. Ethics, 2014.

Ienca, Marcello, "Towards new human rights in the age of neuroscience and neurotechnology, *Life Sciences, Society and Policy*, 13, 5, 2017, em *https://doi.org/10.1186/s40504-017-0050-1*

Kapsner, Andreas & Barbara Sandfuchs, "Nudging as a Threat to Privacy", *Review of Philosophy and Psychology* 6 (3):455-468, 2015.

Lepenies, Robert & Małecka, Magdalena, "The Institutional Consequences of Nudging – Nudges", Politics, and the Law, *Review of Philosophy and Psychology* 6 (3):427-437, 2015.

Wilkinson, T.M., "Nudging and manipulation", *Political Sudies*, Volume: 61 issue: 2, page(s): 341-355, 2013.

Lévy, Pierre, *A Inteligência Colectiva, Para uma antropologia do ciberespaço*, (primeira edição em 1994).

Lévy, Pierre, *Ciberdemocracia*, Lisboa, Instituto Piaget, 2003.

Lewis, C. *et al,* "Subjectively perceived personality and mood changes associated with subthalamic stimulation in patients with Parkinson's disease", *Psychol Med.*, 45(01), 73–85, 2015.

Lippert-Rasmussen, Kasper, "Neuroprediction, Truth-Sensitivity, and the Law", 18(2) J. Ethics, 2014

Loader, B.D., "The governance of cyberspace. Politics, technology and global restructuring", *in* B. D. Loader (ed.), *The Governance of Cyberspace*, 1997, Londres, Routledge, pp. 1-19

Mackenzie, R., "Who should hold the remote for the new me? Cognitive, affective, and behavioral side effects of DBS and authentic choices over future personalities", *Ajob Neurosci,* 2(1), pp. 18–20, 2011.

Monzón, Cândido, *Opinión Pública, Comunicación y Política,* Madrid: Tecnos, 1996.

Neto, Luísa & Rui Vieira da Cunha, *Compulsory Neuro-Interventions: Metaphysical and Conceptual Foundations of the Subject of Responsibility and Autonomy of Choice*, Estudos Comemorativos dos 20 anos da FDUP, FDUP, 2017, Vol II., pp. 119-144.

Neto, Luísa, "Um outro tipo de "freios e contrapesos": a comunicação social no contexto do Estado de Direito Democrático", *Estudos em homenagem ao Prof. Doutor Jorge Miranda,* FDUL, Coimbra: Coimbra Editora, Volume II (Direito Constitucional e Justiça constitucional) Coimbra Editora, 2012, pp. 455-509.

Neto, Luísa, "Ciência da Informação e Direito: um novo paradigma de reconstrução do papel do Estado no contexto da sociedade global de informação*", in A informação jurídica na era digital*, 24 e 25 de Fevereiro de 2011, Porto, Afrontamento, CETAC Media, 2012

Pardo, Michael & Patterson, Dennis, *Minds, Brains, and Law: The Conceptual Foundations of Law and Neuroscience*, New York, Oxford University Press, 2013.

Pereira, Alexandre Libório Dias, *Serviços da sociedade da informação: alguns problemas jurídicos do comércio electrónico na Internet,* Lisboa, Faculdade de Direito da Universidade Nova de Lisboa, 2001.

Persson, I. e Savulescu, J., "The perils of cognitive enhancement and the urgent imperative to enhance the moral character of humanity", *Int J Appl Philos.*, 25(3), 162–77, 2008

Persson, Ingmar & Savulescu, Julian, *Unfit for the Future: The Need for Moral Enhancement.* Oxford, Oxford University Press, 2012.

Petersen, Thomas Søbirk, "(Neuro)prediction, Dangerousness, and Retributivism, 18(2)" J. Ethics, 18(2), 2014.

Rawls, John, *Uma teoria da justiça,* Lisboa, Fundamentos, 1993.

Rogel, Carlos (coord.), *Los robots y el Derecho*, Madrid, Reus Editorial, Colección Jurídica General, 2018.

Salgado, J. Ferreira, *Informação e civilização*, Ordem dos Advogados, Conselho Distrital do Porto, 1973, Conferência proferida pelo autor em 11 de Maio de 1973.

Savulescu, Julian & Bostrom, Nick (eds.), *Human Enhancement.* Oxford, Oxford University Press, 2010; MORSE, Stephen J., *The Promise of Neuroscience for Law: Hope or Hype?*, *in* Palgrave Handbook of Philosophy and Public Policy 77–96 (David Boonin ed., 2018).

Sententia, W., "Neuroethical considerations: cognitive liberty and converging technologies for improving human cognition", *Ann N Y Acad Sci,* 1013(1), pp. 221–8, 2004.

Slevin, J., *The Internet and Society*, 2000, Oxford: Polity Press, pp. 1-10 e pp. 214-235.

Solove, Daniel J., *The **Future** of **Reputation**: Gossip, Rumor, and Privacy on the Internet*, Yale University, 2007

Thaler, Richard H. e Sunstein, Cass R., *Nudge: Improving decisions about health, wealth, and happiness*, New Haven, Yale University Press, CT, 2008.

Wolton, Dominique, Internet et après?, Paris: Flammarion, Coll. "Champs", 2000 (Dominique Wolton, *E depois da Internet? Para uma teoria crítica dos novos media*, Lisboa: Difel, 2000).

Zangrossi, A., Agosta, S., Cervesato, G., Tessarotto, F. & Sartori Zangrossi, G., *"I Didn't Want To Do It!" The Detection of Past Intentions*, 9 Front Hum Neurosci. 608, 2015.

La necesaria modificación de la Directiva 85/374/CEE a propósito de la inteligencia artificial

The necessary modification of Directive 85/374/EEC on intelligence matters artificial

Nerea Díaz Ortiz*

RESUMO: A inteligência artificial está começando a tomar conta da sociedade, o que em teoria facilitará o dia a dia das pessoas, mas também pode ser fonte de certos danos. Para garantir a confiança da sociedade na inteligência artificial, é necessário um regulamento claro sobre responsabilidade civil, razão pela qual é essencial saber quais os regulamentos aplicáveis, uma vez que o Parlamento Europeu e o Conselho apresentaram uma proposta de regulamento sobre responsabilidade civil pelo funcionamento de sistemas de inteligência artificial, que teriam de coexistir com a Diretiva sobre produtos defeituosos. Por este motivo, nesta comunicação analisamos a proposta de Regulamento e, em particular, analisamos quais as alterações à Directiva que são necessárias para a adequar à realidade actual e futura e garantir a segurança jurídica.

PALAVRAS-CHAVE: consumidores, inteligência artificial, produto defeituoso, responsabilidade civil, regulamento, futuro.

ABSTRACT: Artificial intelligence is beginning to take over society, which in theory will make people's daily lives easier, but it could also be the source of cer-

* Universidad de Cantabria. Departamento de Derecho Privado. Investigadora predoctoral. diazortizn@unican.es

tain damages. In order to ensure that society has confidence in artificial intelligence, a clear regulation on civil liability is necessary, which is why it is essential to know which regulations would be applicable, since the European Parliament and the Council have presented a proposal for a Regulation on civil liability for the operation of artificial intelligence systems, which would have to coexist with the Directive on defective products. For this reason, in this communication we analyze the proposed Regulation and, in particular, we analyze which amendments in the Directive are necessary to bring it into line with the current and future reality, and to ensure legal certainty.

KEYWORDS: consumers, artificial intelligence, defective product, civil liability, regulation, future.

1. Introducción

Se puede afirmar rotundamente que la vida diaria en el futuro se verá totalmente modificada, y esto es así, como consecuencia del desarrollo de la inteligencia artificial (en adelante IA). Actualmente, la IA comienza a apoderarse de la sociedad, pero, se espera, que en un futuro cercano, estemos rodeados de aquello que se denomina coloquialmente "robots"[1]. Aparentemente, desempeñará un papel destacado, ya que permitirá paliar la escasez de los recursos en una sociedad futura más envejecida que la actual, pero, además, permitirá paliar las desigualdades existentes en los países del tercer mundo.

Quizá, uno de los aparatos que se esperan en mayor medida son los vehículos automatizados[2], y sí que es cierto, que uno de los sectores que se verán afectados en mayor proporción es el sector automovilístico. Pero, en el presente trabajo,

[1] Aunque hay que saber que los conceptos de "inteligencia artificial" y "robots" no son sinónimos, ya que en el caso de la inteligencia artificial es inmaterial, mientras que el robot es material, pudiendo incorporar inteligencia artificial. Es decir, no siempre un robot incorpora inteligencia artificial, y no siempre la inteligencia artificial está incorporada a un robot.
PEDRO PORTELLANO, "Inteligencia artificial y responsabilidad por productos", *Revista de Derecho Mercantil,* núm. 316, Editorial Civitas, SA, 2020.

[2] MÓNICA NAVARRO-MICHEL, "Vehículo automatizados y responsabilidad por producto defectuoso", *Revista de Derecho Civil*, vol. VII, núm. 5 (octubre-diciembre, 2020), Estudios.

se procederá al análisis de la responsabilidad de los daños generados por la IA, en otro ámbito en el que actualmente ya se utiliza la inteligencia artificial, y en el que es fundamental la protección y seguridad, como es el ámbito sanitario.

Las ventajas que genera la inteligencia artificial en este ámbito son muy variadas, comenzado en primer lugar, con la posibilidad de cruzar los datos de forma precisa y rápida, lo que permite proporcionar mayor facilidad y satisfacción a nivel interno en el tratamiento de datos, pero también esto beneficia a los pacientes. Esto es así, como consecuencia de que en ocasiones los datos se encuentran fragmentados y en formatos diferentes, de forma que permite que se unifiquen. De forma que, la información contenida en los historiales de los pacientes podrá ser analizada con mayor rapidez.

Pero, además, también permitirá analizar e interpretar las imágenes médicas, y en consecuencia, se podrá diagnosticar y tratar a los paciente con mayor rapidez y precisión.

Quizá donde más impresión práctica genera, es en el caso de los procedimientos quirúrgicos, en los que en muchas ocasiones, se hace uso de robots, aunque sí que es cierto que el robot autónomo parece todavía algo imposible, existe hoy en día ya una clasificación en base a su autonomía, siendo el nivel cero aquel que depende por completo del cirujano, y el nivel cinco, sería aquel totalmente autónomo[3].

Por tanto, en la actualidad, el médico es el que realiza la operación con la ayuda de los robots, no obstante, se pretende avanzar y conseguir, mas pronto que tarde, que los robots sean programados para realizar directamente las intervenciones quirúrgicas.

Ya que, por ejemplo, un robot quirúrgico, en una operación en la que está reproduciendo las técnicas utilizadas por el médico, si en un momento determinado cambian las circunstancias en la operación, cabría plantearse que ocurriría. Si es el médico el que estuviese operando, con sus conocimientos y con su experiencia previa, que está resultará del todo fundamental, podrá o por lo menos, intentará superar el problema que ha surgido de forma totalmente inesperada. En cambio, un robot que carece de consciencia, que tiene los conocimientos, quizá en ese momento no actúa y se paraliza, o actúa en un sentido inadecuado. Quién sabe.

También resulta posible que exista un defecto de fabricación en el robot, como por ejemplo, que exista un error en su memoria, que en un momento determinado, le impide recordar aquello que tiene que realizar.

[3] Jesús Esteban Cárcar Benito, "La inteligencia artificial (IA) como aplicación jurídica y razonable: la cuestión sanitaria", *IUS ET SCIENTIA,* Vol. 7, nª1, 2021.

En un primer momento, la IA va a facilitar la vida de la población, pero, tal y como se ha enunciado es posible que sea el origen de determinados daños. Por ello, resulta fundamental que la regulación sobre responsabilidad sea clara y fácilmente comprensible.

2. Regulación actual

Para conseguir que la sociedad muestre confianza en la IA es necesaria la regulación desde una doble perspectiva, en primer lugar, en materia de seguridad, y en segundo lugar, para el caso de que se produzcan daños, es decir, en materia de responsabilidad civil. Esta doble regulación permite proteger en mayor medida a los consumidores, pero, a su vez, también facilita que se genere confianza en estas nuevas tecnologías, lo que repercute en que la industria se sienta segura para realizar inversiones en esta materia. Tal y como se ha enunciado con anterioridad, es el futuro, y hay que conseguir que la regulación sea adecuada, y en consecuencia, facilite su desarrollo. Ya que, por ejemplo, uno de los grandes problemas en materia de responsabilidad civil, es identificar a la persona legitimada pasivamente, todo ello, como consecuencia, de los distintos operadores económicos que participan en la cadena de suministro del producto, además, dicha problemática aumenta en el supuesto de que un sujeto, en un momento posterior a la comercialización, sea quien incorpore al producto un sistema de IA.

Por tanto, la Unión Europea ha procedido al estudio de estas dos materias, y en el presente trabajo, como consecuencia de la gran cantidad de resoluciones de los últimos años, únicamente se procede al estudio de los distintos trabajos que son más recientes en el tiempo.

Todo el estudio previo se ha plasmado en el Informe que acompaña al Libro Blanco. De hecho, la Comisión, el 21 de abril de 2021, ha presentado una propuesta de reglamento en materia de inteligencia artificial[4], pero este Reglamento tiene como finalidad que los productos que posean IA funcionen de manera correcta, fiable y coherente, es decir, lo que regula es la seguridad.

Pero, en cambio, surge el problema de la regulación de la responsabilidad civil, y para ello, en el presente trabajo se va a proceder al análisis de esta materia.

[4] Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en materia de inteligencia artificial (Ley de inteligencia artificial) y se modifican determinados actos legislativos de la Unión {SEC(2021) 167 final} – {SWD(2021) 84 final} – {SWD(2021) 85 final}.

Actualmente, esta materia se encuentra regulada a nivel europeo por la Directiva 85/374/CEE[5] relativa a la responsabilidad por los daños causados por productos defectuosos[6], y además, existen las normas nacionales (y por tanto, que no se encuentran armonizadas) sobre responsabilidad civil.

Por tanto, ¿son suficientes las normas nacionales sobre responsabilidad civil y la directiva de responsabilidad por producto defectuoso, o, en cambio, son necesarias nuevas normas europeas que armonicen la materia relativa a la responsabilidad civil?

Lo que se establece en el informe que acompaña al Libro Blanco[7], es que se ha de garantizar la reparación de los daños que ocasionen aquellos productos y servicios que incorporen inteligencia artificial, de forma que la primera función en la sociedad de las normas de responsabilidad civil es garantizar a las víctimas una indemnización, y en segundo lugar, proporcionar incentivos económicos al responsable de forma que así se consiga evitar que cause dicho perjuicio.

La existencia de la doble regulación genera ciertas diferencias, ya que sí que es cierto, que a nivel europeo existe un nivel de protección uniforme basado en la responsabilidad objetiva del productor por los daños causados por los defectos de los productos. En cambio, en cada país, están en vigor diversos sistemas, algunos de ellos, son sistemas de responsabilidad subjetiva (la víctima ha de probar la culpa de la persona responsable, el daño y la causalidad entre la culpa y el daño), y, en cambio, en otros países rigen sistemas de responsabilidad objetiva (el legislador nacional ha atribuido la responsabilidad a una determinada persona, y no es necesario, que la víctima pruebe la culpa, la relación de causalidad o el daño).

[5] Directiva 85/374/CEE del Consejo, de 25 de julio de 1985, relativa a la aproximación de las disposiciones legales, reglamentarias y administrativas de los Estados Miembros en materia de responsabilidad por los daños causados por productos defectuosos (DO L 210 de 7.8.1985, p.29). Esta Directiva fue traspuesta al ordenamiento jurídico español por la Ley 22/1994, de 6 de julio, de responsabilidad civil por los daños causados por los productos defectuosos. Esta a su vez, fue derogada por el Real Decreto Legislativo 1/2007, de 16 de noviembre, por el que se aprueba el texto refundido de la Ley General para la Defensa de los Consumidores y Usuarios y otras leyes complementarias.

[6] Actualmente, regulado por los artículos 128 y siguientes de Ley General para la Defensa de los Consumidores y Usuarios y otras leyes complementarias.

[7] Informe de la Comisión al Parlamento Europeo, al Consejo y al Comité Económico y Social Europeo. Informe sobre las repercusiones en materia de seguridad y responsabilidad civil de la inteligencia artificial, el internet de las cosas y la robótica, fecha 19 de febrero de 2020. Se procede al análisis de la sección relativa a la responsabilidad civil, la cual, se basa en la evaluación de la Directiva sobre responsabilidad por los daños causados por productos defectuosos.

En el informe se pone como ejemplo las distintas responsabilidades que surgen por un accidente de tráfico, en el que están legitimados pasivamente, tanto el propietario del vehículo[8] por su responsabilidad subjetiva, pero, también, en el supuesto de ausencia de culpa. Esto es así, en la medida en que en la mayoría de los Estados Miembros, la responsabilidad civil objetiva se aplica a la persona a cuyo nombre está matriculado el vehículo, pero, incluso, se podrá dirigir la acción contra el asegurador, ya que en aplicación de la Directiva 2009/103/CE, existe la obligación de asegurar el vehículo. Por tanto, en la práctica, es la entidad aseguradora la que compensa a la víctima y protege al asegurado ya que no ha de satisfacer la indemnización.

Además, se podrá dirigir la acción contra el fabricante en aplicación de la Directiva sobre responsabilidad por los daños causados por productos defectuosos.

Para que esta legitimación pasiva tan amplia funcione, es necesario que, existan normas de responsabilidad civil que sean claras, de forma, que se pueda reclamar en última instancia al responsable de los daños. Un ejemplo muy claro de está cuestión se produce cuando se reclama al asegurador, y este una vez que ha indemnizado a la víctima, solicita el reintegro de la indemnización al productor, como consecuencia de que es este sujeto el responsable.

Por tanto, la existencia, de esta doble regulación, plantea la cuestión de si son suficientes las normas en vigor, o en cambio, es necesario, modificar las existentes, o, incluso, llegar a dictar nuevas normas.

Esto es así, en la medida en que, pueden identificarse ciertos problemas que en el futuro podrían surgir. Uno de dichos problemas, es que en los países donde rigen la responsabilidad subjetiva, puede ocurrir que la prueba sea gravosa o excesivamente onerosa, de forma que, el consumidor no pueda demostrar, la relación de causalidad. También, podrían surgir ciertos problemas en relación a la forma de aplicación de la norma en el caso en concreto como puede ser con el concepto de "culpa". Esto podría suponer una desincentivación de las inversiones, además, que aumentaría los costes de la información y de los seguros, e incluso, podría aumentar la fragmentación si es que cada país es el que regula esta materia, lo que podría reducir el comercio transfronterizo.

[8] *op.cit.* Mónica Navarro-Michel, p. 1.

2.1. Propuesta de Reglamento

Una vez analizadas las distintas propuestas, el 20 de octubre de 2020, se presenta la propuesta de Reglamento[9] relativo a la responsabilidad civil por el funcionamiento de los sistemas de inteligencia artificial.

Para establecer el ámbito de aplicación del Reglamento, en el artículo 3, se prevé la definición de operador inicial como "toda persona física o jurídica que ejerce un grado de control sobre un riesgo asociado a la operación y el funcionamiento del sistema de IA y se beneficia de su funcionamiento", y en cambio, se define al operador final como "toda persona física o jurídica que define, de forma continuada, las características de la tecnología y proporciona datos y un servicio de apoyo final de base esencial y, por tanto, ejerce también grado de control sobre un riesgo asociado a la operación y el funcionamiento del sistema de IA".

Por tanto, la propuesta de reglamento se centra en las reclamaciones contra el operador de un sistema de IA, entendiéndose por tal, tanto al operador final como inicial. Prevalece el Reglamento si el operador inicial es también el productor, o cuando solo hay un operador del sistema. En cambio, prevalecerá la Directiva, si el operador final se puede definir, también, como productor.

Se prevén dos sistemas diferentes de responsabilidad, y para establecer el sistema de responsabilidad aplicable, se han de diferenciar los sistemas de IA en alto riesgo y bajo riesgo.

De forma que se calificará de alto riesgo[10] aquel que funciona de forma autónoma para causar daños o perjuicios a una o más personas de manera aleatoria y que excede lo que cabe esperar razonablemente; para ello, se tendrá en cuenta que la magnitud del potencial depende de la relación entre la gravedad del posible daño o perjuicio (sobre la base de factores determinantes como la magnitud del daño potencial resultante del funcionamiento en las personas afectadas, incluida la posible afectación a derechos fundamentales), el grado de autonomía de la toma de decisiones, la probabilidad de que el riesgo se materialice (se han de tener en cuenta factores tales como el papel de los cálculos algorítmicos en el proceso de toma de decisiones, la complejidad de la decisión y la reversibilidad de los efectos), y el modo y el contexto en que se utiliza el sistema de IA (si puede tener efectos jurídicos o fácticos sobre derechos importantes de la persona afectada jurídicamente protegidos y de si los efectos pueden evitarse razonablemente). Y en sentido negativo,

[9] Propuesta de Reglamento del Parlamento Europeo y del Consejo relativo a la responsabilidad civil por el funcionamiento de los sistemas de inteligencia artificial.

[10] Artículo 3.1.c) de la Propuesta de Reglamento.

se definirá a la IA de bajo riesgo, y para entender, la diferencia entre los dos conceptos, se puede utilizar un símil, que es que una IA débil sabe jugar al ajedrez, y en cambio, una IA fuerte se plantea que entre el ajedrez y la damas hay cierto parecido[11].

Por tanto, el operador de un sistema de IA de alto riesgo será objetivamente responsable (artículo 3) de cualquier daño o perjuicio que causa, de forma que solo se exoneraran en el caso, de que concurra fuerza mayor. Además, es obligatorio un seguro de responsabilidad civil, y el importe de la indemnización será hasta un máximo de dos millones de euros (fallecimiento o daños causados a la salud o la integridad física) o hasta un máximo de un millón de euros (daños morales significativos que resulten de una pérdida económica comprobable o en daños a bienes). Si se ha de abonar a varias personas, los importes no superaran el límite máximo establecido (artículo 5).

El plazo de prescripción varía, será de 30 años (daños a la vida, la salud o integridad física) o de diez años a partir de la fecha en la que se produjo el menoscabo a los bienes o existencia de daño moral, o de treinta años a partir de la fecha en que tuvo lugar la operación del sistema de IA de alto riesgo que causó posteriormente el menoscabo a los bienes o el daño moral (el plazo que venza antes) (artículo 7).

En cambio, el operador de un sistema de IA que no constituye un sistema de alto riesgo, está sujeto a responsabilidad subjetiva, exonerándose cuando el sistema se activó sin su consentimiento (al tiempo que se tomaron todas las medidas razonables y necesarias para evitar dicha activación fuera del control del operador), o se observó la diligencia debida a través de la realización de una serie de acciones, que son las siguientes: la selección de un sistema de IA adecuado para las tareas y las capacidades pertinentes, la correcta puesta en funcionamiento del sistema de IA, el control de las actividades, y el mantenimiento de la fiabilidad operativa mediante la instalación periódica de todas las actualizaciones disponibles (artículo 8). Y los plazos de prescripción aplicables son los que se contemplan a nivel nacional (artículo 9).

Este doble sistema de responsabilidad, en mi opinión, no es adecuado, ya que el sistema en todo caso, tendría que ser de responsabilidad objetiva. El exigir la culpa, en muchas ocasiones, resultaría que no se podría probar la misma, ya que, se espera, que, en un futuro, el robot pueda seguir desarrollándose, y consiga aprender de forma autodidacta nuevas actividades, y esto

[11] Alejandro Zornoza Somolinos, *"Breves antes a la propuesta de Reglamento del Parlamento Europeo sobre Responsabilidad Civil en materia de inteligencia artificial"*, *R.E.D.S* núm. 17, Julio-Diciembre 2020.

puede generar problemas en relación con el régimen de responsabilidad por culpa. Porque, al fin y al cabo, hoy en día, no nos encontramos en dicho escenario, pero, aparentemente, en un futuro cercano, sí que puede que estemos en dicho escenario.

Y de hecho, se comparte la opinión de Portellano[12], que no llega a comprender porque se debería de beneficiar de un régimen de responsabilidad más benigno (responsabilidad por culpa con inversión de la carga de la prueba) el productor de un robot con inteligencia artificial que toma la decisión de utilizar un producto para limpiar el sofá que causa un incendio que, a su vez, produce la muerte por asfixia de un habitante de la casa, que el productor de una simple batidora que, al soltarse la cuchilla por defectuosa unión de los componentes, causa la lesión de un tendón de la mano (sometido al régimen de responsabilidad objetiva).

Todo ello, porque, al contrario de otros autores[13][14], se considera que la inteligencia artificial conlleva desconocimiento, y por ello, conlleva, riesgo y peligro. Por tanto, hay que partir de esa premisa para poder regular la materia de forma adecuada.

3. Modificación de la Directiva de productos defectuosos

En el considerando noveno de la propuesta de Reglamento de Responsabilidad Civil, se hace referencia a la Directiva 85/374/CEE[15], y establece que será de aplicación en las reclamaciones por responsabilidad civil presentadas contra el productor de un sistema de IA defectuoso por una parte que sufra

[12] *loc. cit*. Pedro Portellano p.1.

[13] Alejandro Platero Alcón, "Breves notas sobre el régimen de responsabilidad civil derivado de los sistemas de inteligencia artificial: especial referencia al algoritmo de recomendaciones de Netflix", *IUS ET SCIENTIA*, Nº1, Vol.7, 2021, *apud* Manuel Ortiz Fernández, *"Reflexiones acerca de la Responsabilidad Civil derivada del uso de la inteligencia artificial: los "principios" de la Unión Europea"*, *Revista de Direito da ULP Law Review/Revista de Direito da ULP*, Vol. 14, nº1, 2021.

[14] "No se puede contemplar los avances científicos en materia de inteligencia artificial como sinónimos de dificultad probatoria y, por consecuencia, de inconveniencia o desasosiego para la víctima. Más bien al contrario, cuanto mayor se la técnica del robot, más fácilmente se podrá detectar el origen del defecto, en tanto más perfeccionada se encontrará la trazabilidad ordenada por la Unión Europea en sus recomendaciones, de modo que podría llegarse al punto en que fuera perfectamente detectable la causa del fallo, y por tanto, el sujeto responsable."
Isabel Zurita Martin, "Las propuestas de reforma legislativa del Libro Blanco Europeo sobre inteligencia artificial en materia de seguridad y responsabilidad civil", *Actualidad Jurídica Iberoamericana*, Nº 14, febrero 2021.

[15] Directiva 85/374/CEE del Consejo, de 25 de julio de 1985, relativa a la aproximación de las disposiciones legales, reglamentarias y administrativas de los Estados Miembros en materia de responsabilidad por los daños causados por productos defectuosos (DO L 210 de 7.8.1985, p.29).

un daño o perjuicio, e incluso establece que "*En consonancia con los principios de mejora de la legislación de la Unión, todo ajuste legislativo necesario debe debatirse durante la revisión necesaria de dicha Directiva*".

Por tanto, se procede a analizar las posibles cuestiones que requieren de una actualización en la Directiva, poniendo en relación dichas cuestiones con el ámbito sanitario.

3.1. Concepto de producto

En la Directiva se define al producto en el artículo 2: "*A los efectos de la presente Directiva, se entiende por "producto" cualquier tiene mueble, excepto las materias primas agrícolas y los productos de la caza, aun cuando está incorporado a otro bien mueble o a un inmueble. (...)*".

Se hace referencia a la necesidad de aclarar el ámbito de aplicación de la Directiva, en relación con la consideración de "producto" al contenido digital. En mi opinión, la actual definición, permite de forma clara, identificar como producto al contenido digital[16]. Es decir, que en el concepto de producto no solo se han de incluir los robots cirujanos, sino que también se han de incluir los programas informáticos que permiten la recolección de datos, o el caso de los programas que se encargan de analizar las imágenes, que aunque no se hayan implementado en ningún dispositivo. Todo ello, en aplicación del artículo 335 del Código Civil[17], ya que son susceptibles de apropiación, es decir, pueden recaer derechos subjetivos sobre ellos[18]. Por tanto, hoy en día, se ha de calificar a los robots inteligentes como productos con características propias, pero como productos[19].

Por tanto, no es necesaria la modificación del concepto de producto, ya que se puede deducir que se incluye dentro del concepto a la IA. Pero, también es cierto, que ya que se va a proceder a la modificación de la Directiva, no resulta complicado, introducir el concepto de inteligencia artificial en dicho

[16] Aún así, si que es cierto, que existe una falta de claridad entre los conceptos de producto y servicio, sobretodo con los programas informáticos, que a veces funcionan de forma autónoma, y en cambio, en otras se puede considerar una parte de un producto físico. Isabel Zurita Martin, *op. cit.*, p. 6.

[17] *"Se reputan bienes muebles los susceptibles de apropiación no comprendidos en el capítulo anterior, y en general todos los que se pueden transportar de un punto a otro sin menoscabo de la cosa inmueble a que estuvieren unidos*".

[18] Joaquín Ataz López, "Daños causados por las cosas. Una nueva visión a raíz de la robótica y de la inteligencia artificial", *en* Mariano Herrador Guardia (Ed.), *Derecho de daños,* Madrid: Lefebvre, 2020.

[19] Silvia Díaz Alabart, *Robots y responsabilidad civil,* Reus, Madrid, 2018.

artículo 2 de la Directiva, simplemente para promocionar seguridad plena sobre la aplicación de la Directiva.

3.2. Concepto de servicios

Aparentemente, la Directiva como su propio nombre indica, no resulta de aplicación a los servicios, por tanto, ¿Qué ocurre cuando en la prestación de servicios sanitarios se utiliza un producto defectuoso?

Esto, claramente genera problemas en la práctica actualmente, pero es cierto, que como se ha dicho con anterioridad, parece, que cada vez en mayor medida se va a hacer uso de la IA en el seno de las operaciones quirúrgicas. Y de hecho, se podría llegar a elucubrar sobre la existencia de robots cirujanos totalmente autónomos en un futuro. Por tanto, resulta necesario que se contemple el concepto de servicios en la Directiva[20].

Ya que la inclusión de este concepto, no solo supone un avance en materia de IA, sino que supondrá en la práctica un gran avance en materia de responsabilidad en el ámbito sanitario. Para poder entender mejor, la importancia de esta cuestión, se ha de traer a colación la STS 1806/2020, de 21 de diciembre de 2020, en la misma se rechaza la responsabilidad de la administración sanitaria por la aplicación de un producto defectuoso en el seno de una operación quirúrgica. El supuesto de hecho de la sentencia, se corresponde con la aplicación del *gas perfluoroctano Ala Octa* en operaciones quirúrgicas de desprendimiento de retina, ocho lotes del producto sanitario, fueron defectuosos, y por ello, a consecuencia de su aplicación, provocó la pérdida de visión de los pacientes.

El Tribunal Supremo ha rechazado la responsabilidad de la administración sanitaria, aunque el artículo 148 del Texto Refundido de la Ley General para la Defensa de los Consumidores y Usuarios, contempla la responsabilidad objetiva en los "servicios sanitarios". Pero, la interpretación jurisprudencial de este concepto, es inadecuada, así la STS 669/2010, de 4 de noviembre, establece que "*dada la específica naturaleza, este tipo de responsabilidad no afecta a los actos propiamente dichos, dado que es inherente a los mismos la aplicación de criterios de responsabilidad fundados en la negligencia por incumplimiento de la lex artis ad hoc. Por consiguiente, la responsabilidad establecida por la legislación de consumidores únicamente es aplicable en relación con los aspectos organizativos o de prestación de servicio s sanitarios, ajenos a la actividad médica propiamente dicha*".

Por tanto, solamente, se aplica a los aspectos organizativos o funcionales, y el Tribunal Supremo, considera que la aplicación de un producto defectuoso

[20] Véase *infra* p. 9.

en el seno de una operación quirúrgica, no es un aspecto organizativo o funcional, y por tanto, se aplica, el sistema general de responsabilidad contemplado en la normativa administrativa.

Por ello, resulta necesario, que no solo se regule a nivel comunitario la responsabilidad objetiva por producto defectuoso, sino que también se ha de regular la responsabilidad por servicios. Porque, hay que tener en cuenta, que las divergencias existentes en las normativas internas, generan en la práctica la desprotección al consumidor. Y claro ejemplo de esta cuestión, es la STJUE de 21 de diciembre de 2011, Asunto C-495/10, en la que un paciente en una intervención quirúrgica sufre unas lesiones como consecuencia de que el colchón térmico utilizado (era un producto defectuoso). Esta situación, ocurrió, en un hospital francés, y se plantea si es compatible con la Directiva la existencia de una línea jurisprudencial nacional que reconoce la responsabilidad también de la administración sanitaria, en cuyo seno se ha hecho uso del producto defectuoso. El Tribunal de Justicia de la Unión Europea estableció que es compatible. Y, esto a su vez, demuestra, las divergencias existentes a nivel nacional entre los distintos países miembros. Ya que si esto hubiese ocurrido en España, no se hubiese reconocido la responsabilidad de la administración sanitaria.

Con este ejemplo, se pretende demostrar, la necesidad de la armonización de esta cuestión a nivel comunitario, ya que la utilización de los robots, va a generar daños, y se ha de proteger al paciente.

3.3. Concepto de defecto

Se entiende por producto defectuoso aquel que no ofrece la seguridad a la que una persona tiene legítimamente derecho, teniendo en cuenta todas las circunstancias, incluso: la presentación del producto; el uso razonablemente pudiera esperarse del producto; y el momento en el que el producto se puso en circulación. Y no se considerará defectuoso por la única razón de que, posteriormente, se haya puesto en circulación un producto más perfeccionado[21].

Claramente, en ocasiones, se podrá identificar la existencia de un defecto, entendido como un fallo en la fabricación del producto, pero, en otras ocaciones, no será así, por tanto, más que hablar de producto defectuoso, se tendría que valorar la creación de una nueva categoría de productos: productos peligrosos o inseguros. Porque, si en el transcurso de una operación quirúr-

[21] Artículo 6 de la Directiva 85/374/CEE del Consejo, de 25 de julio de 1985, relativa a la aproximación de las disposiciones legales, reglamentarias y administrativas de los Estados Miembros en materia de responsabilidad por los daños causados por productos defectuosos.

gica, surge un contratiempo, y el robot actúa de forma inadecuada o incluso, se paraliza, el robot cirujano no presenta un defecto de fabricación al uso. Aunque, se podría llegar a plantear que sí que existe tal defecto, en tanto en cuanto en su fabricación, no se tuvo en cuenta la situación remota a la que se tendría que enfrentar el robot en la operación quirúrgica.

Como consecuencia de que el robot está diseñado para realizar determinadas actividades o realizar ciertas funciones con la seguridad que cabría legítimamente esperar en el usuario según su naturaleza, siempre que el robot no responda a estas expectativas, se ha de calificar al producto como defectuoso, al ser inseguro. Por ello, el productor ha de delimitar claramente cuales son las actividades para las que está programado el robot[22].

Por tanto, cuando el robot no actúa tal y como se espera de él, cuando su decisión se desvía de lo esperado, se ha de considerar que se está ante un defecto del mismo. Si un robot en una operación quirúrgica, en vez de extraer el riñón, extrae el pulmón, claramente se está ante un producto defectuoso. Pero, también, se estará ante un producto defectuoso, cuando ante un fallo cometido a la hora de extraer el pulmón, se bloquee o realice una actuación contraria a la adecuada.

Todo ello, en base a que existen ciertas expectativas que son razonables o legítimas sobre el producto. Como en el caso de un robot que se encargue de extraer órganos, pues lo que se espera es que no genere daños, sino, que todo lo contrario, que de forma adecuada extraiga el órgano afectado.

3.4. Concepto de productor

En el artículo tercero de la Directiva se define al productor como la persona que fabrica un producto acabado, que produce una materia prima o que fabrica una parte integrante, y toda aquella persona que se presente como productor poniendo su nombre, marca o cualquier otro signo distintivo en el producto. Además, toda persona que importe un producto con vistas a su venta, alquiler, arrendamiento financiero o cualquier otra forma de distribución en el marco de una actividad comercial, también es considerado productor. Y en el caso de que el productor no pudiera ser identificado, cada suministrador del producto será considerado como productor, salvo que informe al perjudicado de la identidad del productor o de la persona que le suministró el producto dentro de un plazo de tiempo razonable. Y lo mismo ocurre con los productos importados.

[22] Isabel Zurita Martin, *op. cit.*, p. 6.

De hecho, en el Considerando 18 de la propuesta de Reglamento, se establece que se han de incluir en esta definición a otros sujetos, y no solo delimitar este concepto a la significación actual, y por ello, se han de incluir, también, a los fabricantes, desarrolladores, programadores, prestadores de servicios y operadores finales (Considerando 18).

Se tendría que considerar como productor a toda persona que intervenga en el desarrollo, diseño y fabricación del robot, además de los programadores y diseñadores del software. Al fin y al cabo, el consumidor se dirigirá frente al productor del producto finalizado, ya que será la persona que "conozca", pero nada debería impedir que se dirija frente al realmente responsable. No solo beneficia al consumidor, sino que, también beneficiaría al productor del producto completo. Aunque también es cierto, que en ocasiones esto no será posible, por el desconocimiento de la población sobre esta materia, y por tanto, el desconocimiento de cual es el verdadero problema. De forma que el productor del robot debe responder por el inadecuado funcionamiento del mismo, con independencia del tipo de defecto causante de los daños, pero en cambio, el programador solo responderá de los defectos del programa. Ya que hay que tener en cuenta que cuando se comercializa el robot en su conjunto no surgen problemas, pero cuando se han adquirido los componentes de forma separada, la carga de la prueba puede resultar muy gravosa[23].

En el ámbito sanitario, y en base a la definición actual, se ha negado la calificación de productor al hospital que en el seno de una operación quirúrgica, hace uso de un producto defectuoso. Así, en el asunto C-495/10, 21 de diciembre de 2011, en el que en una operación quirúrgica el paciente sufre unas quemaduras como consecuencia de que el colchón térmico utilizado en la misma era defectuoso, se establece que no resulta de aplicación la Directiva. Ya que se establece que la responsabilidad en que puede incurrir un usuario, como es el hospital, que utiliza en una prestación de servicios médicos el producto adquirido o aparato, no se encuentra dentro del ámbito de aplicación de la Directiva. Ya que a este usuario no se le puede considerar como participante de la cadena de fabricación y de comercialización del producto, y por ello, no se puede calificar ni como productor ni como suministrador del bien.

Por tanto, con la correspondiente modificación del concepto de productor, se ha de considerar, que en un caso análogo al enunciado, sí que sería de apli-

[23] La doctrina distingue entre la comercialización del robot como un todo (" a bundle of card-and software" o "closed systems") cerrado a las interferencias e los usuarios, y la adquisición de los componentes por separado ("open systems") que provienen de distintos suministradores. Isabel Zurita Martin, *op. cit.*, p. 6.

cación la Directiva, y por tanto, no se acudiría a la normativa nacional. Esto resultaría beneficioso para el consumidor, porque existiría uniformidad en todo el territorio de la Unión Europea, y así se garantiza la igualdad. Y así, si un paciente es víctima de un daño provocado por un mal funcionamiento de un programa informático, el prestador del servicio, ya sea una entidad pública o privada, deberá responder de igual modo que el que procura una prótesis defectuosa en una operación quirúrgica[24].

3.5. Causas de exoneración

Alguna de las causas de exoneración de responsabilidad, pueden generar problemas en este ámbito, así es, con la que se contempla en el apartado b) del artículo 7 de la Directiva:

> "o que, teniendo en cuenta las circunstancias, sea probable que el defecto que causó el daño no existiera en el momento en el que él puso el producto en circulación o que este defecto apareciera más tarde"

La inteligencia artificial sufre constantes modificaciones o mejoras, que es lo que se conoce comúnmente como actualizaciones, por ello, surge el problema de que si el defecto surge con posterioridad a una actualización, se pueda llegar a alegar esta causa de exoneración, como consecuencia de que en el momento de la puesta en circulación del producto, el mismo no existía.

Por ello, resulta necesario que se contemple la relación existente entre la puesta en circulación del producto con la continua actualización o mejora del software. Así, de esta forma, el productor no podrá exonerarse de responsabilidad si no facilita el "update" o "upgrade" necesaria para la correcta funcionalidad del producto. Una vez, que se han garantizado los medios de actualización, recaerá, sobre el usuario la responsabilidad por su falta de diligencia al no incorporarlos a la máquina inteligente[25].

Se está haciendo referencia a dos cuestiones que son distintas. Por un lado, la obligación del productor de facilitar las actualizaciones del producto, y que de esta forma, no pueda acogerse a esta causa de exoneración. Pero, es que, yo considero, que también, habría que tener en cuenta que una vez, que el productor facilita la actualización al usuario, y esté la lleva a cabo, puede ser el origen del defecto que genera un daño con posterioridad. Y por tanto, si ese daño deriva de la actualización, no se podrá por el productor alegar que en el momento de la puesta en circulación, el producto carecía de dicho defecto.

[24] Isabel Zurita Martin, *op. cit.*, p. 6.
[25] Isabel Zurita Martin, *op. cit.*, p. 6.

En relación con esta cuestión, hay que traer a colación, el artículo 11 de la Directiva, que establece que los derechos conferidos por la misma se extinguen transcurrido el plazo de diez años a partir de la fecha en que el productor hubiera puesto en circulación el producto que causó el daño, salvo que el perjudicado hubiera ejercicio una acción judicial contra el productor.

Este plazo de 10 años, resulta del todo inadecuado, todo ello, porque se obtienen "robots" con la finalidad de que duren muchos años, y lo que se espera es que en un futuro, duren cada vez más tiempo. Además, en el ámbito sanitario, los robots que actualmente existen, poseen un elevado coste de adquisición, y por ello, se busca que sea una inversión a largo plazo. Si la responsabilidad cesa a los 10 años de adquirir el producto, quizá en un ámbito como el sanitario, en el que se pone en juego la vida de los pacientes, suponga que se tenga que deshacer del "robot" transcurrido dicho periodo de tiempo. Todo ello, para proteger al paciente. Además, como se ha dicho con anterioridad, se pueden generar daños, a partir de las continuas actualizaciones del producto[26][27].

4. Conclusiones

En primer lugar, se ha de llegar a la conclusión, de que lo que se tendría que haber hecho es una modificación de la Directiva, en vez de proponer un nuevo Reglamento, que presenta una regulación completamente divergente a la existente actualmente.

Resulta del todo contradictorio e inadecuado la regulación de un doble sistema, siendo uno de ellos de responsabilidad objetiva, y el otro de responsabilidad subjetiva. Ya que este Reglamento, ha de convivir con la Directiva, y como bien se conoce, la Directiva establece un sistema de responsabilidad objetiva del productor, sin diferenciar entre el mayor o menor riesgo. Por tanto, esto en la práctica, puede generar grandes desigualdades. Este nuevo doble sistema, no se debería introducir en la Directiva, porque, supondría un retroceso en los derechos de los consumidores, que verían como la responsabilidad objetiva contemplada, se transforma en ocasiones en responsabilidad subjetiva.

En mi opinión, el sistema adecuado es el de responsabilidad objetiva, porque el exigir la culpa, lo que generará en la práctica son problemas relativos a la prueba de la culpa. Ya que hay que tener en cuenta que la IA, se ha de relacionar con desconocimiento, y esto en la práctica lo que genera es riesgo

[26] Pedro Portellano, *op. cit.*, p. 1.
[27] Joaquín Ataz López, *op. cit.*, p.8.

y peligro. Además, esto no desincentivará el progreso, en la medida en que la Directiva, ya contempla dicho sistema de responsabilidad objetiva, y no ha tenido dicha consecuencia.

Tal y como se ha analizado, desde la perspectiva del ámbito sanitario, se ha de modificar la Directiva, y se puede sintetizar en las siguientes ideas:

Es aconsejable la concreción del concepto de "producto", y a su vez, la inclusión del concepto de "servicios", ya que en el ámbito sanitario, puede tener una gran transcendencia dicha inclusión. En la medida en que, podrá garantizar que los daños generados por la IA, sean indemnizados al paciente, ya que no sólo estará legitimado pasivamente el productor, sino también el prestador de servicios, que en este caso, sería la administración sanitaria. Todo ello, conllevará una mayor protección al paciente.

El concepto de defecto ha de ser modificado, para poder incluir, aquellos supuestos en los que la inteligencia artificial genera un daño, sin que exista como tal un defecto de fabricación al uso, sino que cuando un robot actúa de forma diferente a la que se espera de él, se ha de considerar, que es un producto defectuoso. Es decir, que cuando no cumpla con las expectativas que legítimamente se pueden esperar de él, se ha de calificar como defectuoso. Por ello, se aboga por llamarlos productos peligrosos o inseguros, más que defectuosos. Como, cuando por ejemplo, se queda paralizado ante una situación no contemplada, o cuando, mediante el autoaprendizaje realiza alguna actuación inadecuada.

También, es necesaria la modificación del concepto de productor, para de esta forma, incluir en dicha definición, a toda persona que participa en todo el proceso productivo.

Además, se ha de impedir, que se pueda alegar la causa de exoneración relativa a la puesta en circulación, en la medida en que, en este ámbito, las actualizaciones tienen una gran importancia. Por tanto, además de obligar al productor a poner a disposición del consumidor dichas actualizaciones, hay que tener en cuenta, que el origen del defecto puede ser dicha actualización. Por todo ello, se ha de modificar el concepto de puesta en circulación.

En relación con esta última problemática, se encuentra el plazo de prescripción de 10 años, porque tal y como se ha dicho, es posible que el defecto surja en el momento de la actualización del producto, y no en el momento de la puesta en circulación como tal. Además, de que se espera que la vida útil de estos productos sea elevada en el tiempo.

Bibliografía

Álvarez Olalla, Pilar, "Propuesta de Reglamento en materia de responsabilidad civil por el uso de inteligencia artificial, del Parlamento Europeo, de 20 de octubre de 2020", *Revista CESCO*, Nº 38/2021, ISSN 2254-2582.

Ataz López, Joaquín, "Daños causados por las cosas. Una nueva visión a raíz de la robótica y de la inteligencia artificial". En M. Herrador Guardia (Ed.), *Derecho de daños 2020* (317-375). Madrid: Lefebvre.

Cárcar Benito, Jesús Estebán, "La inteligencia artificial (IA) como aplicación jurídica y razonable: la cuestión sanitaria", *IUS ET SCIENTIA*, 2021, Vol. 7, nª1, ISSN 2444-8478.

Díaz Alabart, Silvia, *Robots y responsabilidad civil*, Reus, Madrid, 2018, pp-100-101.

García Teruel, Rosa María, "El derecho de daños ante la inteligencia artificial y el machine learning: una aproximación desde las recomendaciones del Parlamento Europeo y del Grupo de Expertos de la Comisión Europea", en COBACHO GÓMEZ, JOSÉ ANTONIO (Ed.), *Cuestiones clásicas y actuales del derecho de daños. Estudios en Homenaje al Profesor DR. Roca Guillamón (1009-1055)*. Pamplona: Aranzadi.

Gómez Ligüerre, Carlos, Y García Micó, Gabriel, "Responsabilidad por daños causados por la Inteligencia Artificial y otras tecnologías emergentes. (Liability for Artificial Intelligence and Other Emeging Technologies). Revista para el análisis del Derecho, (1), 501-511.

Minero Alejandre, Gemma, "Robots y derecho civil. Algunas cuestiones a tener en cuenta desde la perspectiva europea", en GONZÁLEZ PULIDO, IRENE y BUENO DE MATA, FEDERICO (Eds), FOERTICS 8.0: estudios sobre tecnologías disruptivas y justicia (pp.55-66). Granada: Comares.

Navarro-Michel, Mónica, "Vehículo automatizados y responsabilidad por producto defectuoso", *Revista de Derecho Civil*, vol. VII, núm. 5 (octubre-diciembre, 2020), Estudios, pp. 175-223, ISSN 2341-2216.

Ortiz Fernández, MANUEL, "Reflexiones acerca de la Responsabilidad Civil derivada del uso de la inteligencia artificial: los "principios" de la Unión Europea", *Revista de Direito da ULP Law Review/Revista de Direito da ULP*, Vol. 14, nº1.

Platero Alcón, Alejandro, "Breves notas sobre el régimen de responsabilidad civil derivado de los sistemas de inteligencia artificial: especial referencia al algoritmo de recomendaciones de Netflix", *IUS ET SCIENTIA*, Nº1, Vol.7, 2021, ISSN 2444-8478, pp. 135-154.

Portellano, Pedro, "Inteligencia artificial y responsabilidad por productos", *Revista de Derecho Mercantil*, núm. 316/2020, Editorial Civitas, SA.

Ramón Fernández, Francisca, "Robótica, inteligencia artificial y seguridad: ¿Cómo encajar la responsabilidad civil?, *Diario de la Ley*, Nº9365, Sección Doctrina, 25 de febrero de 2019, Editorial Wolters Kluwer,

Zapata Sevilla, José, *Inteligencia artificial y responsabilidad civil: el caso de las organizaciones descentralizadas autónomas*, Congreso sobre la Regulación de los Algoritmos, Oviedo, 2019.

Zornoza Somolinos, Alejandro, "Breves antes a la propuesta de Reglamento del Parlamento Europeo sobre Responsabilidad Civil en materia de inteligencia artificial", *R.E.D.S* núm. 17, Julio-Diciembre 2020, ISSN 2340-4647.

Zurita Martin, Isabel, "Las propuestas de reforma legislativa del Libro Blanco Europeo sobre inteligencia artificial en materia de seguridad y responsabilidad civil", *Actualidad Jurídica Iberoamericana,* Nº 14, febrero 2021, ISSN 2386-4567, pp. 438-487.

Legislación

– Directiva 85/374/CEE del Consejo, de 25 de julio de 1985, relativa a la aproximación de las disposiciones legales, reglamentarias y administrativas de los Estados Miembros en materia de responsabilidad por los daños causados por productos defectuosos (DO L 210 de 7.8.1985, p.29).
– Ley 22/1994, de 6 de julio, de responsabilidad civil por los daños causados por los productos defectuosos.
– Real Decreto Legislativo 1/2007, de 16 de noviembre, por el que se aprueba el texto refundido de la Ley General para la Defensa de los Consumidores y Usuarios y otras leyes complementarias.
– Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en materia de inteligencia artificial (Ley de inteligencia artificial) y se modifican determinados actos legislativos de la Unión {SEC(2021) 167 final} – {SWD(2021) 84 final} – {SWD(2021) 85 final}.

Jurisprudência

Acórdão do STS de 4.11.2010 (669/2010).
Acórdão do STJUE de 21.12.2011 (asunto C-495/10).
Acórdão do STS de 21.12.2020 (1806/2020).

Desordem informativa, algoritmos e inteligência artificial: uma cruzada europeia

Information disorder, algorithms, and artificial intelligence: a European crusade

Tiago Morais Rocha*

RESUMO: A desinformação é um dos principais anátemas dos regimes políticos contemporâneos. A necessidade de combatê-la tornou-se omnipresente no discurso político do "mundo ocidental". A União Europeia encabeça essa ambição, procurado liderar a busca por soluções jurídicas adequadas. Conforme a história recente revela, este é um dos domínios onde os *policy makers* atuam como funâmbulos, tentando alcançar um equilíbrio impossível entre a ordem informativa e os inadmissíveis constrangimentos às liberdades e aos valores fundamentais das sociedades democráticas.

Será a desinformação um produto da inteligência artificial (IA), dos algoritmos e da datificação? Recorremos a um conceito amplo de IA, que combina dados, algoritmos e capacidade computacional, para concluir que as tecnologias associadas ou baseadas na IA agravam significativamente o problema da desinformação, tornando virtualmente impossível distinguir a *verdade* da *manipulação*, ou aumentado exponencialmente o potencial de eficácia das campanhas de desinformação, através da automação e da personalização de conteúdos.

PALAVRAS-CHAVE: Desinformação; algoritmos; inteligência artificial; personalização de conteúdos; democracia.

* Assistente Convidado na FDUP. Investigador Colaborador do CIJ. trocha@direito.up.pt. ORCID: 0000-0003-1817-997X

ABSTRACT: Disinformation is one of the biggest anathemas of existing political regimes. The fight against disinformation has become ubiquitous in the political discourse of Western civilizations. The European Union leads that ambition, striving to seek appropriate legal frameworks. As recent history shows, this is one of the domains where policymakers act like funambulists, trying to reach an impossible balance between the informational order and the inadmissible constraints to fundamental freedoms and values of free and democratic societies.

Is disinformation a product of artificial intelligence (AI), algorithms, and datafication? We depart from a broad concept of AI, which combines data, algorithms, and computational capacity, to conclude that the technologies associated or based on AI may significantly exacerbate the problem of disinformation, making it virtually impossible to distinguish the truth from manipulation, or exponentially increase the effectiveness potential of disinformation campaigns, through automation and personalization of content.

KEYWORDS: Disinformation; algorithms; artificial intelligence; personalization of content; democracy.

> *«It is an old, very old story told all over again: one can use axes to hew wood or to cut heads. The choice does not belong to the axes, but to those who hold them. Whatever the holders' choices, the axes don't mind. And however sharp the edges which it may be currently cutting, technology would not 'advance democracy and human rights' for (and instead of) you»*[1]

1. Considerações prévias

A desinformação é um dos principais anátemas dos regimes políticos contemporâneos. Para os regimes democráticos, a desinformação, nos moldes em que atualmente é criada e difundida, corresponde a um dos subprodutos da era digital e, ainda, da erosão do papel dos órgãos de comunicação social enquanto *gatekeepers*, tendo em conta a predominância das redes, das plataformas e, de um modo mais genérico, do aparato digital na definição,

[1] Zygmunt Bauman, *A Chronicle of Crisis: 2011-2016*, [s.l.], Social Europe Edition, 2017, p. 81.

distribuição e difusão (mas, mais do que isso, na própria construção e amplificação) das narrativas sociais e políticas hoje dominantes. Para os regimes de feição não democrática, a desinformação é já um instrumento ao serviço das suas idiossincrasias e dos seus projetos de manutenção do *poder*, quando não mesmo uma arma híbrida usada em ataques contra outros Estados e pessoas (coletivas ou singulares)[2].

A necessidade de combater a desinformação tornou-se omnipresente no discurso político do "mundo ocidental". A União Europeia procura, pelo menos desde 2018, encabeçar essa ambição, liderando a busca por soluções políticas, jurídicas e tecnológicas adequadas. Sem embargo, e conforme a história recente vem revelando, este é um dos domínios onde os *policy makers* atuam como funâmbulos, tentando alcançar um equilíbrio por vezes impossível entre a ordem informativa e os inadmissíveis constrangimentos às liberdades e aos valores fundamentais das sociedades democráticas[3]. O próprio legislador europeu reconhece esta circunstância, referindo que, por via do seu alcance e impacto, os conteúdos disponíveis nas plataformas digitais podem «*afetar um número considerável de destinatários dos serviços em diferentes Estados-Membros e causar grandes danos à sociedade*», porém, a identificação, o combate e a correção desses desvios pode ser «*particularmente complexa*»[4].

Um número considerável de textos sobre a temática da desinformação associa o fenómeno das "*fake news*" à emergência de uma "era da pós-verdade".

[2] O conceito de "guerra híbrida" tem sido desenvolvido pelos estudos militares, especialmente pela Organização do Tratado Atlântico Norte (NATO), tendo ganho nova tração depois da anexação da Crimeia pela Federação da Rússia, em 2014. A "guerra híbrida" traduz a complexidade dos conflitos modernos, os quais combinam métodos convencionais com métodos não-convencionais, militares e não militares, empregues por atores estatais e não estatais. O propósito da "guerra híbrida" é criar «*confusão e ambiguidade quanto à natureza, à origem e aos objetivos*» das ações empregues. Uma componente chave deste tipo de conflito é a utilização da desinformação, da mentira e do logro de modo a influenciar determinados públicos-alvo, recorrendo-se quer aos meios de comunicação tradicionais, quer, cada vez com mais frequência, às plataformas digitais. Cfr., *inter alia*, Jan Joel Andersson, *Hybrid Operations: Lessons from the past, 2015, EU Institute for Security Studies*, disponível em: https://bityli.com/dclroIZkY, consult. em: 26/11/2022. NATO Strategic Communications Centre of Excellence, *Social Media as Tool of Hybrid Warfare,* 2016, NATO Strategic Communications, disponível em: https://bityli.com/ctxuDvfOeU, consult. em: 26/11/2022. Comunicação Conjunta ao Parlamento Europeu, ao Conselho Europeu e ao Conselho, *Aumentar a resiliência e reforçar a capacidade de enfrentar ameaças híbridas*, Bruxelas, 13.6.2018, JOIN(2018) 16 final.

[3] Tiago Morais Rocha, *A Era Digital e o Estado de Direito Democrático na União Europeia*, Porto, Faculdade de Direito da Universidade do Porto, 2020, p. 23-4 (Tese de Mestrado).

[4] Regulamento (UE) 2022/2065 do Parlamento Europeu e do Conselho, de 19 de outubro de 2022, relativo a um mercado único para os serviços digitais e que altera a Diretiva 2000/31/CE (doravante, apenas Regulamento dos Serviços Digitais). JO L 277 de 27.10.2022, p. 36, considerando 137.

O termo "*post-truth*" acabou mesmo por merecer a honra de ser escolhido, em 2016, como palavra do ano pelos Dicionários Oxford. Embora não se negue a existência de uma relação direta entre a desinformação e a "pós-verdade", facto não escamoteável é que a primeira é apenas um dos elementos da segunda, que a extravasa largamente. Como explica D'Allonnes, o prefixo «pós-» em "pós-verdade" não nos remete para um tempo que sucedeu a um que o precedeu; antes, o prefixo «pós-» indica «*claramente que a noção à qual o prefixo é acrescentado – a verdade – se tornou secundária, desprovida, até, de pertinência*», concluindo que a «*"pós-verdade" pôs em causa o carácter essencial da verdade*»[5]. Assim, quando falamos numa era em que domina a política da pós-verdade, queremos menos referirmo-nos a um tempo que está para lá dos factos, mas a um momento histórico em que domina «*uma política que torna a própria verdade irrelevante ou não significativa, uma política que deslegitima o valor da verdade*»[6].

É possível que o aparente triunfo de uma política da pós-verdade se fique a dever, pelo menos em parte, às características das sociedades modernas e pós-modernas. Esse é, parece-nos, o sentido fundamental da posição de Hannah Arendt quando explica que uma das características principais das massas modernas[7] se prende com a circunstância de estas não «*acreditarem na realidade da sua própria experiência*»[8], mas apenas nas ficções que constroem. Segundo a Autora, o que «*convence as massas não são os factos, mesmo que sejam inventados, mas apenas a coerência com o sistema do qual esses factos fazem parte*»[9]. Tal como no totalitarismo, a pós-verdade, que já não a propaganda, reduz e simplifica o complexo sistema social através de explicações lineares, independentemente do seu absurdo, de narrativas bem urdidas, de bodes expiatórios e de teorias da conspiração que permitem a crença numa "explicação universal" que justifica o acaso de que a realidade é feita.

Numa entrevista de 2017, Evan Williams, um dos cofundadores do Twitter, passando em revista a ascensão das plataformas digitais, dizia ter acreditado que «*assim que toda a gente pudesse falar livremente e trocar informação e ideias, o mundo seria, automaticamente, um lugar melhor*», para logo acrescentar

[5] Myriam Revault d'Allonnes, *A Verdade Frágil: o que a Pós-verdade Faz ao Nosso Mundo Comum*, Lisboa, Edições 70, 2020, p. 22.

[6] *Idem*.

[7] Sobre o conceito de massa, *vide*, por todos, Ortega y Gasset em a *A Rebelião das Massas*. Segundo o Autor, «*massa é o "homem médio"*», aquele que repete em si «*um tipo genérico*». José Ortega y Gasset, *A Rebelião das Massas*, Lisboa, Relógio D'Água, 2019, p. 41.

[8] Hannah Arendt, *As Origens do Totalitarismo,* Lisboa, Dom Quixote, 2017, p. 465.

[9] *Idem*.

que «*estava enganado*»[10]. Por certo que Williams, quando fez aquele comentário, estava longe de imaginar que o Twitter havia de, em apenas cinco anos, ser vendido a um multimilionário que se autoproclama um "absolutista da liberdade de expressão"[11] e que pretende transformar radicalmente a "*town square*" em que aquela rede social se tornou, dando nova força ao debate em torno da publicização de aspectos críticos das plataformas digitais de grande dimensão[12/13].

Se no que à inteligência artificial diz respeito persiste uma certa cautela pública e técnica, até por força das muitas influências culturais[14], o alvor da internet foi visto não só como um marco histórico, mas como uma via democrática promissora. A era digital, na sua ubiquidade e conectividade, pode efetivamente contribuir de forma determinante para o aperfeiçoamento dos sistemas e processos democráticos, podendo, até, aumentar a promoção e respeito pelos direitos fundamentais e humanos. Todavia, e como já tivemos

[10] David Streitfeld, *'The Internet Is Broken': @ev Is Trying to Salvage It,* The New York Times, 20/05/2017, disponível em: https://www.nytimes.com/2017/05/20/technology/evan-williams-medium-twitter-internet.html, consult. em: 26/11/2022.

[11] *Vide* Francisco de Abreu Duarte, *From Platforms To Musk To... Protocols?*, The Digital Constitutionalist – The Future of Constitutionalism, 2022, disponível em: https://digi-con.org/from-platforms-to-musk-to-protocols/, consult. em: 03/12/2022.

[12] A este propósito, defendendo que as plataformas digitais são também empresas de comunicação social, pelo que devem estar sujeitas ao mesmo quadro regulamentar que decorre da existência de um interesse público prevalente, *vide* Philip M. Napoli, *Social Media and the Public Interest: Media Regulation in the Disinformation Age*. New York: Columbia University Press, 2019.

[13] Refere Eli Pariser que a aquisição do Twitter por Elon Musk «*é a inevitável consequência de uma escolha coletiva que fizemos de ceder a esfera pública a empresas centralizadas, movidas pela publicidade e controladas por um pequeno punhado de homens. O resultado tem sido um ambiente digital funcionalmente autocrático, no qual se pode twittar o que se quiser – mas em que se precisa de 44 biliões de dólares para alterar a dinâmica da própria plataforma. Esse ambiente tem sido desastroso para as democracias, para as comunidades e para muitas pessoas que sofrem na pele o ódio, a opressão política e o pior que resulta de sermos uma preocupação posterior na economia da atenção*». Eli Pariser, *Musk's Twitter Will Not Be the Town Square the World Needs*, Wired, 28/10/2022, disponível em: https://www.wired.com/story/elon-musk-twitter-town-square/, consult. em: 26/11/2022. *Vide*, ainda, Oreste Pollicino, *L'uomo più ricco al mondo non è il miglior custode della libertà di espressione*, 16/11/2022, Il Sole 24 Ore, disponível em: https://www.ilsole24ore.com/art/l-uomo-piu-ricco-mondo-non-e-miglior-custode-liberta-espressione-AEfpOGHC?refresh_ce=1, consult. em: 03/12/2022 e Massimo Sideri, *Se su Twitter a parlare è «la voce del padrone»*, 30/11/2022, Corriere Della Sera, disponível em: https://www.corriere.it/editoriali/22_novembre_30/se-twitter-parla-la-voce-padrone-76e235d2-70df-11ed-9572-e4b947a0ebd2.shtml, consult. em: 03/12/2022.

[14] Basta pensarmos no distópico filme "*Eu, Robot*", estreado em 2004, e largamente inspirado na série de contos que Isaac Asimov publicou entre os anos 50 e os anos 80 do passado século.

oportunidade de defender[15], se a expansão em linha, virtualmente ilimitada, dos horizontes das liberdades resulta em mais (liberdade de) expressão (e) de opinião, o aumento da participação pode não ser necessariamente benéfico para a democracia e para o Estado de Direito, principalmente num quadro em que o discurso genuíno, factual e racional é preterido, desclassificado e sufocado por vozes e narrativas inautênticas, falsas e altamente emocionais, destinadas a polarizar as sociedades, a destruir a confiança nas instituições públicas e privadas e a criar um tal estado de confusão em que «*os factos deixam de importar ou sequer se reconhecem existir* [...], *empoderando-se cada pessoa a escolher a sua própria realidade*»[16].

Neste quadro, a inteligência artificial e os algoritmos são um passo adicional, uma *nova tecnologia* a acrescentar às *velhas técnicas*, na criação, produção, difusão e amplificação da desinformação. Porém, estas novas tecnologias agravam significativamente o problema, tornando virtualmente impossível distinguir a *verdade* da *manipulação*, e aumentado exponencialmente a eficácia das campanhas de desinformação, através da automação e da personalização de conteúdos. A luta contra a desinformação transforma-se, pois, numa verdadeira cruzada, em prol de «*um debate público inclusivo e pluralista*» na União Europeia[17], sem o qual uma sociedade democrática, e a própria União, não pode existir[18].

[15] Tiago Morais Rocha, *A Era Digital e o Estado de Direito Democrático na União Europeia*, Porto, cit. p. 23-4.

[16] Stephan Lewandowsky, *et. al.*, *Beyond Misinformation: Understanding and Coping with the 'Post-Truth' Era*, Journal of Applied Research in Memory and Cognition, Vol. 6., nº 4, 2017, p. 353-369 (p. 361).

[17] Comunicação da Comissão ao Parlamento Europeu, ao Conselho, ao Comité Económico e Social Europeu e ao Comité das Regiões, *Combater a desinformação em linha: uma estratégia europeia*, Bruxelas, 26.4.2018, COM(2018) 236 final, p. 1.

[18] Esta tem sido a posição constante do Tribunal Europeu dos Direitos Humanos (TEDH). Cfr., *inter alia*, o Acórdão do TEDH no caso nº 64569/09, *DELFI AS v. ESTONIA*, 16/06/2015. No plano europeu, o pluralismo é uma das características da sociedade europeia, conforme afirmado pelo artº 2º do Tratado da União Europeia (TUE). O Tribunal de Justiça da União Europeia afirmou recentemente que os «*valores que constam do artigo 2º TUE foram identificados e são partilhados pelos Estados-Membros. Definem a própria identidade da União enquanto ordem jurídica comum.* [...] *O artigo 2.° TUE não constitui uma simples enunciação de orientações ou intenções de natureza política, mas contém valores que são abrangidos* [...] *pela própria identidade da União enquanto ordem jurídica comum, valores que são concretizados em princípios que comportam obrigações juridicamente vinculativas para os Estados-Membros*». Acórdão do TJUE de 16.02.2022, *Polónia v. Parlamento Europeu e Conselho*, C-157/21, § 145 e 264.

2. Desinformação e a *vexata quaestio* da sua definição

A desinformação não é uma novidade do séc. XXI. Antes, a desinformação é tão antiga quanto a existência dos meios de difusão da informação, sendo que cada avanço tecnológico, desde o telégrafo no séc. XIX aos modernos algoritmos das redes sociais, veio criar novas possibilidades de artifício e fabricação[19]. Apesar de essencial para que se possam elaborar políticas públicas eficazes e até para combater o fenómeno, ainda não foi possível alcançar um consenso terminológico e concetual em torno do ecossistema desinformativo. Efetivamente, são cada vez mais frequentes as soluções tecnológicas baseadas em algoritmos e na inteligência artificial, designadamente no *machine learning*, que procuram identificar, classificar e indexar a desinformação, o discurso de ódio e outros conteúdos ilegais em linha. Projetos financiados pela União Europeia no âmbito do programa Horizonte 2020, de que são exemplos os projetos INVID e FANDANGO, apoiam-se na IA para proceder à verificação de conteúdos online[20]. Todas estas ferramentas tecnológicas, e bem assim as medidas regulamentares que possam ser adotadas, têm como pressuposto lógico que se defina com rigor e precisão que conteúdos constituem desinformação[21].

A partir de 2016, e na sequência das campanhas de desinformação em massa e em linha que ocorreram durante o processo eleitoral norte-americano de 2016[22] e, no mesmo ano, o referendo sobre a permanência do Reino Unido na União Europeia[23], tornou-se popular a expressão "*fake news*", que rapidamente se elevou a termo *catch-all* com variados significados[24]. As "*fake news*"

[19] Axel Gelfert, *Fake News: A Definition,* Informal Logic, vol. 38, nº 1, 2018, p.84-117 (p. 90)

[20] Relatório da Comissão ao Parlamento Europeu, ao Conselho Europeu, ao Conselho, ao Comité Económico Social Europeu e ao Comité das Regiões sobre a aplicação da Comunicação «Combater a desinformação em linha: uma estratégia europeia», Bruxelas, 5.12.2018, COM(2018) 794 final, p. 8.

[21] Maria D. Molina; S. Shyam Sundar; Thai Le; Dongwon Lee, *"Fake News" Is Not Simply False Information: A Concept Explication and Taxonomy of Online Content*, American Behavioral Scientist, Vol. 65, nº 2, 2021, p. 180-212 (p. 181)

[22] Para uma descrição detalhada dos acontecimentos *vide* o "*Report of the Select Committee on Intelligence of the United States Senate on Russian Active Measures Campaigns and Interference in the 2016 U.S. Election*", em 5 volumes, disponível em https://bityli.com/QXAHzUJjZ, consult. em: 27/11/2022.

[23] House Of Commons – Digital, Culture, Media and Sport Committee, *Disinformation and 'fake news': Final Report*, London, 2019, disponível em: https://bityli.com/yzpqubQqF, consult. em: 27/11/2022.

[24] Darren Lilleker, *Evidence to the Culture, Media and Sport Committee 'Fake news' inquiry presented by the Faculty for Media & Communication,* Bournemouth University, 2017, disponível em: https://eprints.bournemouth.ac.uk/28610/, consul. em: 27/11/2022.

podem ser definidas como as notícias que «*pretendem descrever acontecimentos do mundo real, geralmente mimetizando os meios de comunicação tradicional, embora os seus criadores saibam que o conteúdo é significativamente falso, transmitindo-o com dois objetivos: ser largamente difundido e enganar pelo menos parte da audiência*»[25], embora o termo seja igualmente utilizado para descrever notícias satíricas, paródias, rumores, conspirações, lendas urbanas, estratégias comunicacionais e publicitárias tendenciosas destinadas a exagerar ou ocultar certos factos, e, ainda, factos verdeiros que atentam «*contra as crenças e as emoções de uma pessoa*»[26] ou as suas posições políticas[27].

Precisamente por se tratar de um termo vago, ambíguo e «*contencioso, graças à sua natureza altamente politizada*»[28], a doutrina foi paulatinamente abandonando a sua utilização, optando por uma terminologia menos politicamente conotada. É esta, por exemplo, a opção adotada desde 2018 pelo Parlamento do Reino Unido, que recomendou ao governo a rejeição do termo *"fake news"* em favor das expressões 'desinformação' e 'informação incorreta ou errada', uma vez que o primeiro é usado para descrever qualquer declaração que o destinatário da mensagem não gosta ou não concorda[29].

Em 2017, Claire Wardle, partindo da apontada inadequação do termo "notícias falsas", introduziu o conceito de "desordem informativa" para descrever um ecossistema mediático complexo onde convivem diferentes tipos de conteúdos, produzidos e disseminados com diferentes motivações e através de diversos meios[30]. Para Wardle é imperativo distinguir entre três categorias de conteúdos[31]: *dis-information* (desinformação, correspondendo aos conteúdos falsos e criados intencionalmente para causar dano a uma pessoa, grupo

[25] Regina Rini, *Fake News and Partisian Epistemology*, Kennedy Institute of Ethics Journal, Vol. 2, nº 2, 2017, p. 43-64.

[26] Nick Rochlin, *Fake news: belief in post-truth*, Library Hi Tech, Vol. 35, nº 3, 2017, p. 386-392 (p. 388)

[27] Soroush Vosoughi; Deb Roy; Sinan Aral, *The spread of true and false news online*, Science, Vol. 359, nº 6380, 2018, p. 1146-1151.

[28] Maria D. Molina; S. Shyam Sundar; Thai Le; Dongwon Lee, *"Fake News" Is Not Simply False Information: A Concept Explication and Taxonomy of Online Content*, cit., p. 184.

[29] Cfr., *inter alia*, House Of Commons – Digital, Culture, Media and Sport Committee, *Disinformation and 'fake news': Final Report*, London, 2019, p. 10.

[30] Claire Wardle, *Fake news. It's complicated*, First Draft News, 2017, disponível em: https://firstdraftnews.org/latest/fake-news-complicated, consult. em: 12/03/2020.

[31] Claire Wardle; Hossein Derakhshan, *Information Disorder: Toward an interdisciplinary framework for research and policy making*, Strasbourg, Council of Europe, 2017, disponível em: https://rm.coe.int/information-disorder-toward-an-interdisciplinary-framework-for-researc/168076277c, consult. em: 14/04/2022, p. 20.

social, organização ou país), *mis-information* (informação incorreta, correspondendo a conteúdos falsos mas sem o elemento intencional) e *mal-information* (informação danosa, correspondendo a conteúdos com base verídica, usados para infligir intencionalmente dano numa pessoa, organização ou país). Esta classificação, como veremos, foi aquela que ganhou tração, acabando por ser adotada, no essencial, pelas instituições europeias[32].

A concetualização adotada por WARDLE não é, todavia, consensual, sendo criticada pelo *Global Disinformation Index* (GDI), uma organização sem fins lucrativos que avalia o risco da presença de desinformação nos *websites* dos *media* tradicionais. Para o GDI, uma distinção entre desinformação e informação incorreta baseada no elemento intencional é insuficiente, não descrevendo o fenómeno da desordem informativa em toda a sua dimensão[33]. Um exemplo ilustrativo das insuficiências de um conceito de desinformação baseado exclusivamente na associação da veracidade do conteúdo ao elemento intencional pode ser diariamente encontrado no *website* "BREITBART". Uma das páginas daquele *website* agrega um conjunto de artigos sob a manchete "crise na fronteira", onde se podia ler, a 27/11/2022, entre outros, o seguinte conjunto de títulos: "*mais dois abusadores sexuais de menores detidos em apenas um dia depois de cruzarem a fronteira e entrarem no Texas*", "*10 mil migrantes apreendidos na fronteira do Texas em apenas uma semana*", "*dois abusadores sexuais apanhados a passar a fronteira para o Texas*". É possível que cada uma destas notícias seja verdadeira, mas a escolha de curadoria que é feita de apresentar estes conteúdos agregados e carregados de efeitos emocionais serve apenas para alimentar a narrativa de que os imigrantes são, na sua maioria, criminosos, o que é desmentido pelas estatísticas[34]. Segundo o GDI, a noção de desinformação deve, em vez da notícia (ou para além dela), focar-se no conceito de "narrativa adversarial",

[32] Segundo a Comissão, numa Comunicação de 2018, «*a desinformação é entendida como informação comprovadamente falsa ou enganadora que é criada, apresentada e divulgada para obter vantagens económicas ou para enganar deliberadamente o público, e que é suscetível de causar um prejuízo público. O prejuízo público abrange ameaças aos processos políticos democráticos e aos processos de elaboração de políticas, bem como a bens públicos, tais como a proteção da saúde dos cidadãos da UE, o ambiente ou a segurança*». Comunicação da Comissão ao Parlamento Europeu, ao Conselho, ao Comité Económico e Social Europeu e ao Comité das Regiões, *Combater a desinformação em linha: uma estratégia europeia*, Bruxelas, 26.4.2018, COM(2018) 236 final, p. 4.

[33] DANNY ROGERS, *Disinformation as Adversarial Narrative,* Global Disinformation Index, 2022, disponível em: https://www.disinformationindex.org/blog/2022-06-22-disinformation-as-adversarial-narrative-conflict/, consult. em 27/11/2022.

[34] MICHAEL T LIGHT JINGYING HE; JASON P. ROBEY, *Comparing crime rates between undocumented immigrants, legal immigrants, and native-born US citizens in Texas*, PNAS, Vol. 117, nº 51, 2020, p. 32340-32347.

havendo desinformação sempre que é produzida e disseminada uma narrativa intencionalmente enganadora e (este é que constitui a pedra de toque do conceito) adversarial face a um grupo ou elemento de risco (uma minoria, uma instituição, a democracia, a ciência, a saúde pública, etc.)[35]. A valia desta concetualização prende-se, na nossa ótica, por destacar que a transmissão da desinformação ocorre não exclusivamente através de conteúdos estritamente noticiosos, mas por via daquilo que o GDI define como os "artefactos da *web*" alojados em plataformas digitais. Esses artefactos podem ser diversos tipos, desde *memes*[36], vídeos, artigos, sondagens, petições, eventos, conteúdo de *infuencers* e *merchandise*[37]. Nestas condições, as "narrativas adversariais" são «narrativas distribuídas intencionalmente sem uma rígida cronologia ou sequência de conteúdos (*"artefactos"), que procuram enfurecer e polarizar os utilizadores da* internet, *sendo caracterizadas pelo conflito ou oposição entre diferentes atores e interesses*»[38]. Curiosamente, o legislador português parece ter-se aproximado da noção de desinformação enquanto narrativa, definindo-a na Carta Portuguesa de Direitos Humanos na Era Digital (Lei nº 27/2021, de 17 de maio) como «*toda a narrativa comprovadamente falsa ou enganadora criada, apresentada e divulgada para obter vantagens económicas ou para enganar deliberadamente o público, e que seja suscetível de causar um prejuízo público, nomeadamente ameaça aos processos políticos democráticos, aos processos de elaboração de políticas públicas e a bens públicos*» (artº 6º, nº 2, entretanto revogado pela Lei nº 15/2022, de 11 de agosto). Sem prejuízo da revogação daquele número, o mais relevante, parece-nos, é a manutenção na ordem jurídica portuguesa de um "direito à proteção contra a desinformação", cuja concetualização é ainda insípida, mas que julgamos ser decorrência de um direito fundamental à integridade informacional, como corolário da liberdade de expressão e de informação reconhecida no plano europeu (artº 11º da Carta dos Direitos Fundamentais da União Europeia). De facto, toda a produção legislativa recente, quer nos Estados-Membro, quer na própria União, aponta para a emergência, enquanto consequência da liberdade de expressão e de informação, de um direito dos cidadãos a não somente

[35] Danny Rogers, *Disinformation as Adversarial Narrative,* cit.

[36] *Memes* são conteúdos virais sob a forma de imagem, texto ou vídeo (ou a combinação de alguns destes elementos) que são partilhados, muitas vezes sob vestes humorísticas, para comentar, parodiar ou referenciar algum símbolo cultural, alguma ideia social ou política ou algum evento da atualidade.

[37] Ben Decker, *Adversarial Narratives: A New Model for Disinformation*, Global Disinformation Index, 2019, disponível em: https://www.disinformationindex.org/blog/2019-8-1-adversarial-narratives-are-the-new-model-for-disinformation/, consult. em: 27/11/2022, p. 6.

[38] Ben Decker, *Adversarial Narratives: A New Model for Disinformation*, cit., p. 7.

serem informados, mas também a receberem informações verdadeiras e não manipuladas. Este tema, todavia, extravasa o âmbito do presente trabalho.

Nesta encruzilhada terminológica, a União Europeia foi chamada a tomar posição. Em 2018, o Grupo Independente de Alto Nível sobre Notícias Falsas e Desinformação (HLEG) propôs à Comissão um conceito amplo de desinformação, reconhecendo o seu caráter multifacetado, abrangendo não só as informações falsas, mas também aquelas que sejam «*imprecisas ou enganadoras, criadas, apresentadas e promovidas de modo a causar, intencionalmente, um dano público ou para a obtenção de lucro*»[39]. Os danos públicos então identificados prendiam-se com a integridade dos processos democráticos, os valores europeus, a educação, a saúde, a ciência, a economia[40], entre outros. Desde esse momento, e como traço comum a todo o percurso posterior, a noção de desinformação foi entendida como excluindo a sátira e a paródia e outras formas de discurso e conteúdos ilegais em linha de acordo com as leis nacionais e com a própria legislação da União (aqui se insere o discurso de ódio, a difamação, o incitamento à violência, a violação dos direitos da propriedade intelectual, a pornografia infantil, etc.).

A Comissão, na sua Comunicação «*Combater a desinformação em linha: uma estratégia europeia*», de 2018, restringe o âmbito do conceito de desinformação, entendo-a como a «*informação comprovadamente falsa ou enganadora que é criada, apresentada e divulgada para obter vantagens económicas ou para enganar deliberadamente o público, e que é susceptível de causar um prejuízo público*»[41], prejuízo esse que tanto se pode reportar a ameaças aos processos políticos democráticos e aos processos de elaboração de políticas, como a bens públicos, tais como a proteção da saúde dos cidadãos da UE, o ambiente ou a segurança. Segundo a Comissão, a desinformação não abrange erros na comunicação de informa-

[39] High Level Group On Fake News And Online Disinformation, *A multi-dimensional approach to disinformation*. Comissão Europeia, 2018, disponível em: https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=50271, consult. em 19/01/2023, p. 10.

[40] O impacto da desinformação na economia pode ser atestado pelo recente caso da farmacêutica "Eli Lilly and Co.". Depois de ter sido adquirido por Elon Musk, o Twitter anunciou que as suas contas verificadas (a que a plataforma apõe um rótulo azul) passariam a ter um custo mensal de 8 dólares. Nessa sequência, um utilizador criou uma conta falsa da farmacêutica, que acabou por ser verificada pelo Twitter como oficial, tendo aí publicado que a insulina passaria a ser gratuita, o que levou a que valor das ações da empresa caísse 4,37%. Cfr. Bruce U. Lee, *Fake Eli Lilly Twitter Account Claims Insulin Is Free, Stock Falls 4.37%*, Forbes, 28/11/2022, disponível em: https://www.forbes.com/sites/brucelee/2022/11/12/fake-eli-lilly-twitter-account-claims-insulin-is-free-stock-falls-43/?sh=6b5f8f741a3d, consult. em: 28/11/2022.

[41] Comunicação da Comissão ao Parlamento Europeu, ao Conselho, ao Comité Económico e Social Europeu e ao Comité das Regiões, *Combater a desinformação em linha: uma estratégia europeia*, cit., p. 4.

ções por parte dos *media* (informações incorretas ou *misinformation*), sátiras, paródias ou notícias e comentários identificados como partidários. Assim, o conceito de desinformação da Comissão faz depender a classificação de três elementos[42]: [i.] o caráter "comprovadamente" falso ou enganador do conteúdo, o que pressupõe uma qualquer verificação objetiva através, por exemplo, do *fact-checking*; [ii.] a intenção subjacente à divulgação dos conteúdos, que só pode ser a obtenção de vantagens económicas e/ou o logro deliberado da opinião pública; e [iii.] a consequência, que se traduz necessariamente num prejuízo público, que tanto pode ser reportado ao processo democrático (por exemplo, a interferência em eleições) como a certos bens públicos. Esta mesma definição foi aceite pelos signatários do primeiro Código de Conduta da UE sobre Desinformação.

Não obstante, à medida que foram sendo adoptadas medidas tendentes a combater a desinformação, ficou clara a existência de dois problemas. Em primeiro lugar, o ecossistema digital complexifica-se e adapta-se a cada nova medida destinada a impedir e travar a criação, a divulgação e a amplificação da desinformação. Isto é, o fenómeno da desinformação tinha, em 2016, determinadas características (por exemplo, uma relativa prevalência dos esforços de "*clickbaiting*" para *websites* que procuravam mimetizar órgãos de comunicação social genuínos para obtenção de proveitos económicos com a publicidade aí colocada) que foram evoluindo, exigindo uma maleabilidade ou elasticidade dos conceitos (a desinformação está hoje presente noutros formatos (imagem, som e vídeo) e não se limita à propagação de conteúdos claramente falsos, antes se orientando para a construção de narrativas manipuladoras que até podem ter por base factos verdadeiros ou só parcialmente falsos, para a deturpação da informação, para a indução do público em erro ou para a utilização de técnicas de interação digital artificial[43]) se se quiser repor a ordem informativa.

Em segundo lugar, e conforme destaca James Pamment na sua série de artigos de investigação comissionados pelo Serviço Europeu de Ação Externa (EEAS), as instituições, as agências e todos os demais órgãos e organismos da UE, em ordem a adotar um plano de defesa efetivo contra as várias ameaças que enfrentam no espaço informacional, têm necessariamente de adotar

[42] Conforme já apontámos em Tiago Morais Rocha, *A Era Digital e o Estado de Direito Democrático na União Europeia*, Porto, cit. p. 30.

[43] Comunicação da Comissão ao Parlamento Europeu, ao Conselho, ao Comité Económico e Social Europeu e ao Comité das Regiões sobre o *plano de ação para a democracia europeia*, Bruxelas, 3.12.2020 COM(2020) 790 final, p. 23.

um quadro terminológico comum que distinga entre os vários aspetos da luta contra a desinformação[44]. A proposta do Autor, que acabou por ser aceite no Plano de Ação para a Democracia Europeia de 2020, giza uma linha de distinção entre a desinformação em sentido amplo, usada como conceito-chapéu que abarca todas as realidades ligadas ao fenómeno da desinformação, nomeadamente a desinformação em sentido estrito (abrangendo um conjunto de técnicas de comunicação intencionalmente deceptivas que se afastam das práticas aceitáveis do discurso público *online*), a informação incorreta ou *misinformation* (definida como a distribuição de informação comprovadamente falsa sem a intenção de enganar ou causar danos), as operações de influência (reportando-se às ações coordenadas e adversariais por natureza, geralmente associadas a campanhas híbridas, que traduzem um esforço para influenciar a sociedade, como é o caso da interferência em processos eleitorais através de campanhas de manipulação em massa ou "*leaks*" estratégicos) e a interferência estrangeira (definida como as tentativas de impedir ou dificultar a livre expressão da vontade política, particularmente no caso de intervenção nos processos políticos soberanos de outro Estado)[45].

No Plano de Ação para a Democracia Europeia de 2020, a Comissão seguiu de perto o trabalho de Pamment, referindo expressamente a importância de distinguir fenómenos diversos que geralmente são agrupados sob o rótulo da desinformação. Assim, definiu a informação incorreta (ou *misinformation*) como o «*conteúdo falso ou enganador, partilhado sem intenção de prejudicar mas cujos efeitos podem causar danos, por exemplo quando as pessoas partilham informações falsas com amigos e familiares de boa-fé*»[46]. Quanto à desinformação, manteve a definição anterior, retirando apenas a necessidade de que a informação fosse "comprovadamente" falsa ou enganadora. De seguida, a Comissão apresentou a noção de "tentativas de exercer influência sobre informação", como os «*esforços coordenados desenvolvidos por intervenientes internos ou externos com o objetivo de influenciar uma audiência-alvo através de vários meios enganadores, incluindo a supressão de fontes de informação independentes em combinação com a desinformação*»[47]. Está-se aqui a reconhecer, implicitamente, a relevância daqueles conteúdos

[44] James Pamment, *The EU's Role in Fighting Disinformation: Crafting A Disinformation Framework – Working Paper nº 2*, Washington, Carnegie Endowment for International Peace, 2020, disponível em: https://carnegieendowment.org/files/Pamment_-_Crafting_Disinformation_1.pdf, consult. em: 28/11/2022, p. 1.

[45] *Idem*, p. 2 e ss.

[46] Comunicação da Comissão ao Parlamento Europeu, ao Conselho, ao Comité Económico e Social Europeu e ao Comité das Regiões sobre o *plano de ação para a democracia europeia*, cit., p. 21.

[47] *Idem*.

não posicionados na dicotomia entre o facto e o não-facto, o falso e o verdadeiro, antes se reportando a uma "narrativa adversarial" ou a um conjunto de conteúdos insuscetíveis de prova, inverificáveis ou carregados de efeitos emocionais, destinados a causar o medo ou a exacerbar divisões e polarizações políticas e sociais, naquilo que se traduz em «*manipulação através de práticas informacionais agressivas*»[48]. Estas práticas manipulativas, podem consistir, por exemplo, na criação e difusão de narrativas distintas e contraditórias entre si[49] ou com a publicação de mensagens contendo juízos de valor impossíveis de verificar objetivamente.

Finalmente, a Comissão define a "interferência estrangeira no espaço de informação" como os «*esforços desenvolvidos por um interveniente estatal estrangeiro ou pelos seus agentes para corromper a livre formação e expressão da vontade política dos indivíduos, utilizando meios coercivos e enganadores*»[50], os quais ocorrem geralmente no âmbito de operações híbridas mais vastas.

Apesar deste depuramento concetual, por certo útil, a verdade é que a própria Comissão continua, no mesmo documento, a recorrer à utilização do termo desinformação como um *"catch-all"*. De facto, algumas linhas depois de proceder às distinções concetuais enunciadas, refere que a «*desinformação*» pode ser combatida através de uma estratégia de comunicação proativa quando «*não haja intenção de enganar, causar danos públicos ou obter um ganho económico*»[51]. Acontece, porém, que segundo a própria terminologia da Comissão, já não estaremos aí perante desinformação, mas perante "informação incorreta".

Em todo o caso, esta distinção de conceitos, sendo um trabalho em permanente evolução, é francamente positivo[52]: por um lado, permite que as insti-

[48] Judit Bayer; Natalija Bitiukoava; Petra Bárd; Judit Szakács; Alberto Alemanno; Erik Uszkiewicz, *Disinformation and propaganda – impact on the functioning of the rule of law in the EU and its Member States*, 2019, European Parliament, disponível em: https://www.europarl.europa.eu/RegData/etudes/STUD/2019/608864/IPOL_STU(2019)608864_EN.pdf, consult. em: 12/04/2022, p. 27-28.

[49] A 21 de maio de 2016, no Texas, agentes associados à *IRA*, uma organização russa encarregue de várias ciberoperações de manipulação e influência interna e externa, convocaram, através de duas páginas do Facebook (ambas geridas pela *IRA*), dois protestos em Houston, o primeiro com o título "*Stop Islamization of Texas*" e o segundo com o título "*Save Islamic Knowledge*". Em ambos os casos recorreu-se a anúncios pagos micro-direcionados. Os protestos acabaram em confrontos entre os manifestantes. Cfr. o "*Report of the Select Committee on Intelligence of the United States Senate on Russian Active Measures Campaigns and Interference in the 2016 U.S. Election*", Vol. 2, cit., p. 47.

[50] Comunicação da Comissão ao Parlamento Europeu, ao Conselho, ao Comité Económico e Social Europeu e ao Comité das Regiões sobre o *plano de ação para a democracia europeia*, cit., p. 21.

[51] *Idem*.

[52] Cfr. James Pamment, *The EU's Role in Fighting Disinformation: Crafting A Disinformation Framework – Working Paper nº 2*, cit., p. 1.

tuições e organismos da UE, ao adotarem um quadro comum de significados, reforcem a consistência e a coordenação das suas ações. Depois, possibilita que os Estados-Membro tenham *standards* comuns quando adotam medidas nacionais sobre esta problemática. Finalmente, contribui para que em relação a cada um dos fenómenos identificados se desenvolva um conjunto de ferramentas e medidas específicas e diferenciadas.

Não só nas relações interinstitucionais é útil a adoção de um quadro terminológico comum. As plataformas digitais, no âmbito das suas obrigações voluntárias ou regulamentares, beneficiam da harmonização e delimitação de conceitos[53]. Malogrado objetivo, no Código de Conduta Reforçado da UE sobre Desinformação, adotado em junho de 2022, ainda que partindo das destrinças terminológicas realizadas pela Comissão no Plano de Ação para a Democracia Europeia, a elas fazendo expressas referências, as plataformas digitais preferiram adotar o conceito amplo de desinformação sempre que a ela se referem, o qual abrange a totalidade das realidades convocadas[54].

Apesar de todos os esforços desenvolvidos pela doutrina[55], a taxonomia relacionada com a "desordem informativa" continua esquiva e pouco rigorosa, numa área em que, justamente, é o contrário que se impõe. Na luta contra a desinformação podem – e são – adotadas medidas que contendem diretamente com os direitos, liberdades e garantias dos cidadãos em sociedades democráticas e pluralistas, pelo que essa limitação, se se pretender legitimar nos riscos e danos causados pela desinformação *lato sensu*, tem de partir de uma clara e rigorosa identificação e delimitação do que constitui, afinal, a desinformação. Ora, a multiplicação e sobreposição de conceitos quase *ad infinitum* torna inviável esse processo, criando incerteza e insegurança numa área – a dos direitos fundamentais – que a tal é avessa, além de contribuir para o descrédito e relativismo dos conceitos. Precisamente, o Alto Comis-

[53] Desde logo, convém destacar que a ponderação a efectuar entre o combate à desinformação (em sentido amplo) e a liberdade de expressão e outros direitos fundamentais, sendo sempre delicado, é especialmente difícil – e restritivo – em relação a algumas categorias, nomeadamente no que respeita à informação incorrecta.

[54] Cfr. o Considerando (A) do Preâmbulo Código de Conduta Reforçado da UE sobre Desinformação, disponível em https://digital-strategy.ec.europa.eu/en/library/2022-strengthened-code-practice-disinformation, ainda não disponível em português.

[55] Veja-se, *inter alia*, o contributo de Maria D. Molina; S. Shyam Sundar; Thai Le; Dongwon Lee, *"Fake News" Is Not Simply False Information: A Concept Explication and Taxonomy of Online Content*, cit., oferecendo um conjunto de critérios que podem ser "ensinados" a máquinas e traduzidos em algoritmos para auxiliar na tarefa de deteção da desinformação. Os Autores adotam a seguinte taxonomia: notícias falsas, comentário e opinião, erros de jornalismo, conteúdo polarizado e sensacionalista, "jornalismo-cidadão", sátira e informação persuasiva.

sariado das Nações Unidas para os Direitos Humanos, tendo em consideração medidas legislativas adotadas (algumas delas entretanto revogadas) em diversos países, como é o caso da França[56], da Alemanha[57], de Itália e da Malásia, veio esclarecer que definições ambíguas, vagas e subjetivas sobre o que constitui as "*fake news*", de modo a estabelecer proibições genéricas da sua difusão, violam os direitos humanos conforme reconhecidos pelo Pacto Internacional sobre os Direitos Civis e Políticos (artº 19º e 25º), devendo ser abolidas[58]. De facto, qualquer limitação a direitos, liberdades e garantias deve basear-se em critérios claros, racionais, objetivos e proporcionais, ora para negar restrições ilegítimas à liberdade de expressão e de informação, ora para evitar o arbítrio, a insegurança e a incerteza jurídica ou a ineficácia dos regimes legais.

3. Velhos e novos métodos: o papel da inteligência artificial e dos algoritmos na criação e disseminação da desinformação

Com os avanços tecnológicos, as campanhas de desinformação são cada vez mais sofisticadas e micro-direcionadas[59]. A imensidão e variedade de dados fornecidos pelos utilizadores dos serviços digitais, conscientemente ou não, tem permitido que a desinformação seja produzida e difundida como se de um "fato à medida" se tratasse, contribuindo para o afunilamento do discurso público nas plataformas digitais e para a criação de "*echo chambers*" e "*filter bubles*"[60]. Por outro lado, e conforme já referido, a desinformação hoje disseminada escapa frequentemente às tradicionais categorias da "verdade" e da "falsidade". Antes, o ecossistema desinformativo preencheu-se de "narrativas adversariais" e técnicas manipulativas agressivas, as quais exploram

[56] Para uma descrição detalhada das medidas legislativas adotadas em França *vide* Oreste Pollicino; Giovanni De Gregorio, *Constitutional Democracy, Platform Powers and Digital Populism*, Constitutional Studies, Vol. 8, 2022, p. 11-34 (pp. 27 e ss.)

[57] *Idem*.

[58] Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, *Freedom of Expression and Elections in the Digital Age*, Research Paper 1/2019, 2019, disponível em: https://www.ohchr.org/sites/default/files/Documents/Issues/Opinion/ElectionsReportDigitalAge.pdf, consult. em: 29/11/2022, p. 9 e ss.

[59] Carme Colomina; Héctor Sánchez Margalef; Richard Youngs, *The impact of disinformation on democratic processes and human rights in the world,* Brussels, European Parliament, 2021, disponível em: https://www.europarl.europa.eu/RegData/etudes/STUD/2021/653635/EXPO_STU(2021)653635_EN.pdf, consult em: 12/06/2022, p. 6.

[60] Sobre estes fenómenos, cfr. Tiago Morais Rocha, *A Era Digital e o Estado de Direito Democrático na União Europeia*, cit., p. 40 e ss.

medos e tensões sociais pré-existentes[61], aumentando a polarização da sociedade, a desconfiança face às instituições democráticas e à sua legitimidade e reduzindo a participação cívica nos processos políticos.

Um dos principais objetivos das campanhas de desinformação que tiveram lugar na segunda metade da passada década foi a interferência em processos eleitorais, por agentes externos ou internos, com motivações políticas ou económicas. De facto, «*a saúde das democracias, quaisquer que sejam o seu tipo e grau, depende de um mísero pormenor técnico: o processo eleitoral. Tudo o resto é secundário. Se o regime eleitoral é acertado, se se ajusta à realidade, tudo vai bem; se não, mesmo que o resto ande otimamente, tudo vai mal.* [...] *Um regime eleitoral é estúpido quando é falso.* [...] *Sem o apoio de sufrágios autênticos, as instituições democráticas ficam no vento. No vento estão as palavras. "A República não era mais do que uma palavra". A expressão é de César.*»[62] A discussão em torno da capacidade da desinformação, dos algoritmos e das arquiteturas digitais influenciarem as atitudes *offline* dos cidadãos, nomeadamente o comportamento eleitoral, não é pacífica na comunidade científica, até porque, neste domínio, é difícil estabelecerem-se relações de causa-efeito[63]. Todavia, alguns estudos sugerem a existência de uma correlação positiva entre a utilização da internet e das redes sociais enquanto fonte de informação política e determinadas opções eleitorais[64].

A montante e a jusante dos processos eleitorais é consensual a importância de uma esfera pública composta por cidadãos *bem* informados, equipados com as ferramentas necessárias à tomada de decisões relevantes sobre o bem comum. Conforme explica John Rawls, «*sem um público informado acerca dos problemas prementes, as decisões políticas e sociais cruciais não podem, simples-*

[61] Judit Bayer; Bernd Holznagel; Katarzyna Lubianiec, et al., *Disinformation and propaganda: impact on the functioning of the rule of law and democratic processes in the EU and its Member States – 2021 update,* European Parliament, 2021, disponível em: https://www.europarl.europa.eu/thinktank/en/document/EXPO_STU(2021)653633, consult. em: 26/09/2022. p. 21. Segundo os Autores, uma parte significativa da base de dados da plataforma EUvsDisinfo é preenchida por entradas cujo conteúdo não pode ser verificado como falso ou enganador. É o caso de uma notícia de 2020 que tinha como alvo a minoria russófona da Letónia e que dava conta que, em cinco anos, a UE iria desagregar e a Letónia quereria juntar-se à Federação da Rússia.

[62] José Ortega y Gasset, *A Rebelião das Massas*, cit. p. 151.

[63] Stephan Lewandowsky; Laura Smillie; David Garcia, et al., *Technology and Democracy: Understanding the influence of online technologies on political behaviour and decision-making*, Luxembourg, Publications Office of the European Union, 2020, p. 19.

[64] Cfr., *inter alia*, Max Schaub; Davide Morisi, *Voter mobilisation in the echo chamber: Broadband internet and the rise of populism in Europe*, European Journal of Political Research, Vol. 59, 2020, p. 752-773 e Michele Cantarella; Nicolò Fraccaroli; Roberto Volpe, *Does fake news affect voting behaviour?*, Research Policy, Vol. 52, nº 1, 2023.

mente, ser tomadas»[65]. Assim, no mundo político ideal, espera-se que o "cidadão democrático" esteja «*bem informado sobre os assuntos políticos*», conhecendo, designadamente, os «*factos relevantes, as alternativas propostas e quais as expectáveis consequências*»[66], de tal modo que a «*informação política é para a democracia o que o dinheiro é para a economia; é a moeda da cidadania*»[67]. Conceber a informação enquanto a "moeda da cidadania democrática" implica, essencialmente, reconhecer que para que possa desempenhar a sua função, a comunidade politicamente organizada tenha de respeitar duas condições: primeiro, os cidadãos devem ter acesso a informação factual que permita a avaliação das políticas públicas; segundo, os cidadãos devem usar esses factos para formar as suas preferências. Numa palavra, a existência de uma democracia funcional «*depende da capacidade dos seus cidadãos tomarem decisões informadas*»[68].

Esta necessidade da *razão pública* enquanto condição de funcionamento de uma democracia é válida tanto para os Estados-Membro como para a União Europeia. Realmente, a UE é um projeto de paz, liberdade e democracia como ideais de governança política, social e económica. O âmago axiológico da União gira justamente em torno da dignidade humana, da liberdade, da democracia, da igualdade e do Estado de Direito (artº 2º TUE), donde decorre, entre outros, uma organização política e institucional baseada na democracia representativa (artº 10º, nº 1 TUE), onde todos os cidadãos são chamados a participar na vida democrática da UE, num quadro de pluralismo ideológico (nº 3), cuja expressão numa esfera pública europeia deve ser promovida pelas instituições (artº 11º, nº 1 TUE).

Nesta senda, quaisquer processos, técnicas ou campanhas que, de forma artificial ou maliciosa, interfiram na formação e expressão da opinião pública, controlando-a ou manipulando-a, constituem perigos para as sociedades democráticas. Este perigo existe tanto no mundo físico como no mundo virtual. Porém, é neste último que ele é mais grave, tendo em conta a facilidade, rapidez, eficácia e escala com se que se podem urdir narrativas e semear ou exacerbar divisionismos. A esfera digital oferece, hoje, ferramentas que tornam a manipulação e a interferência à escala global muito simples, sem que

[65] John Rawls, *The Idea of Public Reason Revisited,* The University Of Chicago Law Review, Vol. 64, nº 3, 1997, p. 765-807 (p. 773)

[66] James H. Kuklinski; Paul J. Quirk; Jennifer Jerit, *et al.*, *Misinformation and the Currency of Democratic Citizenship*, The Journal of Politics, Vol. 62., nº 3, 2000, p. 790-816 (p.790)

[67] *Idem*, p. 791.

[68] Stephan Lewandowsky; Laura Smillie; David Garcia, *et al.*, *Technology and Democracy: Understanding the influence of online technologies on political behaviour and decision-making*, cit., p. 11.

existam ainda adequadas salvaguardas de transparência e regulação[69]. E isto num contexto em que o «*comando*» permanece sendo o «*exercício normal da autoridade. Este baseia-se sempre na opinião pública – sempre, hoje como há dez mil anos, entre os Ingleses como entre os Botocudos. Nunca ninguém mandou na terra alimentando o seu mando essencialmente com outra coisa que não fosse a opinião pública*»[70], ou, dito de outra forma, num quadro em que a «*soberania do Estado sempre começou com o controlo da informação*»[71]. Controlar os elos de difusão e produção dos conteúdos culturais e de informação é, então, uma forma especialmente relevante de poder.

As campanhas de desinformação servem-se actualmente de velhas técnicas, crescentemente aperfeiçoadas, e, ainda, de novos mecanismos, alguns em desenvolvimento.

Quanto aos velhos métodos, destaca-se a contínua utilização dos meios de comunicação social controlados por agentes de desinformação, muitos deles estatais, como forma de distribuição de desinformação, sob a forma de *"fake news"* ou narrativas manipuladas dirigidas a determinadas audiências-alvo[72]. Além dos órgãos de comunicação social ao serviço de determinados projetos de poder iliberais, os meios de comunicação social tradicional continuam a desempenhar um papel relevante na amplificação da desinformação, amiúde destacando-a nos seus ciclos noticiosos enquanto "conteúdo viral", ainda que procedendo à verificação de factos, atraindo público que, de outro modo, não contactaria com esses elementos.

Outra velha técnica utilizada no âmbito das campanhas de desinformação são os "*bots*" e os "*trolls*", em conjunto designadas como as "*cyber troops*". Os "*bots*" (diminuto de "*robot*") operam nas plataformas digitais através de contas falsas ou inautênticas (sem correspondência a um utilizador real) que, através de mecanismos automatizados ou semi-automatizados, interagem entre si e com outras contas (publicando, partilhando, gostando, comentando, etc.), de forma a difundirem e amplificarem rapidamente um conteúdo *online* e, por via da atenção que lhe é dada, aumentando o seu *ranking* nos algoritmos de personalização das plataformas digitais. Os "*trolls*", por sua vez,

[69] Stephan Lewandowsky; Laura Smillie; David Garcia, *et al.*, *Technology and Democracy: Understanding the influence of online technologies on political behaviour and decision-making*, cit., p. 11.

[70] José Ortega y Gasset, *A Rebelião das Massas*, cit., p. 126.

[71] Manuel Castells, *A Galáxia Internet – Reflexões sobre Internet, Negócios e Sociedade*, Lisboa, Fundação Calouste Gulbenkian, 2004, p. 146.

[72] É o caso, por exemplo, do canal televisivo "RT" e da agência noticia "Sputnik", ambas quase exclusivamente financiadas pela Federação Russa, ou, ainda, de vários *websites* noticiosos, como o BREITBART.

correspondem, via de regra, a utilizadores que se mascaram como autênticos e que, recorrendo ao anonimato, espalham desinformação e interagem com outros utilizadores de modo a artificialmente «*saturarem*» as redes «*com comentários, importunado e assediando*»[73] ou simplesmente manifestando um apoio sem correspondência com a realidade ("*astroturfing*"). Estes exércitos ao serviço da desinformação organizavam-se frequentemente em "quintas" ou "fábricas" de "*trolls*" com ou sem objetivos financeiros subjacentes, de que são exemplo a "*Internet Research Agency*" russa, o "*50 Cent Army*" chinês, os "*AK Trolls*" turcos, as "*Trolls farms*" na Macedónia, ou, mais recentemente, a "Cat@Net" na Polónia.

À medida que estes agentes e técnicas se sofisticam e complexificam, a sua deteção torna-se cada vez mais difícil por parte de investigadores, sinalizadores e plataformas digitais[74], tanto mais que se multiplicam no ambiente *online* "*Potemkin personas*", "*Potemkin pages*", redações de notícias e publicações falsas, pseudo "*think thanks*" e organizações inautênticas[75]. "*Potemkin personas*"[76] correspondem, neste contexto, a «*trolls estrangeiros, geralmente russos, que constroem uma presença* online *credível em várias plataformas, misturando as suas mensagens políticas com publicações banais sobre a sua suposta vida quotidiana, tal como um utilizador autêntico faria*»[77]. Estes utilizadores fazem-se passar, frequente-

[73] Jean-Baptiste Jeangène Vilmer; Alexandre Escorcia; Marine Guillaume, *et al.*, *Information Manipulation: A Challenge for Our Democracies,* Paris, Ministry for Europe and Foreign Affairs and Ministry for the Armed Forces, 2018, disponível em: https://www.diplomatie.gouv.fr/IMG/pdf/information_manipulation_rvb_cle838736.pdf, consult. em: 14/04/2022, p. 84.

[74] Judit Bayer; Bernd Holznagel; Katarzyna Lubianiec, *et al.*, *Disinformation and propaganda: impact on the functioning of the rule of law and democratic processes in the EU and its Member States – 2021 update,* cit., p. 21.

[75] Para um reportório completo da utilização destas técnicas no contexto norte-americano, cfr. Renée Diresta; Shelby Grossman, *Pontemkin Pages & Personas: Assessing GRU Online Operations, 2014-2019,* Stanford Internet Observatory, 2019, disponível em: https://fsi-live.s3.us-west-1.amazonaws.com/s3fs-public/potemkin-pages-personas-sio-wp.pdf, consult. em: 03/12/2022.

[76] A utilização da expressão "*Potemkin*" parece dever-se a Gregório Alexandrovich Potemkin, uma alta patente militar russa que, no séc. XVIII, governou o sul da Ucrânia e, para agradar a Catarina, a Grande, terá ordenado a construção de um conjunto de cidades falsas por onde a Czarina passaria em visita, sendo recebida por aldeões em festa, de forma a exponenciar o seu sucesso enquanto governante e a esconder a pobreza. Desde então, o termo "*Potemkin villages*" passou a ser sinónimo de mistificação e fraude. Cfr. Małgorzata Zawadzka, *Today's Potemkin Village: Kremlin Disinformation and Propaganda in Poland*, 2018, disponível em: https://warsawinstitute.org/todays-potemkin-village-kremlin-disinformation-propaganda-poland/, consult. em: 03/12/2022.

[77] Judit Bayer; Bernd Holznagel; Katarzyna Lubianiec, *et al.*, *Disinformation and propaganda: impact on the functioning of the rule of law and democratic processes in the EU and its Member States – 2021 update,* cit., p. 21.

mente, por especialistas, académicos, jornalistas ou escritores, para, recorrendo à falácia do apelo à autoridade, criarem artigos e "artefactos da *web*" que publicam em meios de comunicação social falsos ou genuínos, além de integrarem pseudo-organizações independentes que criam uma aparência de legitimidade no espaço mediático[78].

A este propósito, a Organização do Tratado do Atlântico Norte (NATO) tem alertado para as operações de "branqueamento" ou "lavagem" de informação desenvolvidas na União Europeia por alguns atores hostis, especialmente a Federação Russa. Estas operações destinam-se, no essencial, a legitimar informação falsa ou enganadora através de uma rede de intermediários que recorrem a um conjunto de técnicas para distorcer e obscurecer a fonte original da informação[79]. Identificando as técnicas e as fases do "branqueamento de informação", a NATO alerta para a circunstância de os recursos digitais facilitarem estas operações de influência da opinião pública, além de aumentarem exponencialmente a sua eficácia[80].

Não obstante o impacto e a frequência das operações de interferência estrangeira, tem-se sedimentado, desde o início desta década, uma tendência para que as campanhas de desinformação e de influência da opinião pública sejam maioritariamente resultado de interferências internas/domésticas e não externas, conforme vem testemunhando a investigação em torno da desinformação durante as eleições de 2019 para o Parlamento Europeu[81] e durante as eleições presidenciais dos EUA de 2020[82]. Mais recentemente, a empresa

[78] Refiram-se apenas dois exemplos destas pessoas, páginas e organizações "*Potemkin*": a "Peace Data" e a "NAEBC", montadas e financiadas pela *Internet Research Agency* (IRA) russa, com *webpages* compostas, entre outros, por artigos comprados a jornalistas americanos para criar a aparência de credibilidade. Cfr. Judit Bayer; Bernd Holznagel; Katarzyna Lubianiec, et al., *Disinformation and propaganda: impact on the functioning of the rule of law and democratic processes in the EU and its Member States – 2021 update,* cit., p. 22.

[79] Belén Carrasco Rodríguez, *Information Laundering in Germany,* Nato Strategic Communications Centre of Excellence, 2020, disponível em: https://stratcomcoe.org/publications/information-laundering-in-germany/23, consult. em: 03/01/2023.

[80] Belén Carrasco Rodríguez, *Information Laundering in the Nordic-Baltic region*, Nato Strategic Communications Centre of Excellence, 2020, disponível em: https://stratcomcoe.org/publications/information-laundering-in-the-nordic-baltic-region/26, consult. em: 03/01/2023, p. 10.

[81] Destacando o papel dos grupos extremistas e dos partidos políticos populistas, *vide* Chloe Colliver, *Click Here For Outrage: Disinformation in the European Parliamentary Elections 2019*, Institute for Strategic Dialogue, 2020, disponível em: https://www.isdglobal.org/wp-content/uploads/2020/06/isd_Click-for-Outrage.pdf, consult. em: 03/01/2023, p. 9.

[82] Center for an Informed Public, Digital Forensic Research Lab, Graphika, & Stanford Internet Observatory, *The Long Fuse: Misinformation and the 2020 Election*, 2021, disponível em: https://purl.stanford.edu/tr171zs0069, consult. em: 03/01/2023, p. 187 e ss.

META, dona do Facebook, veio a público comunicar que, em 2022, interrompeu mais de 200 operações de "influência encoberta" em 68 países e 42 línguas, a maioria das quais tendo como alvo os públicos nacionais e apenas 1/3 audiências estrangeiras[83]. Não se exclui, contudo, que os agentes domésticos sejam, em alguns casos, intermediários de agentes externos, tal como reconhece a Comissão Europeia: «*as táticas e discursos utilizados pelas fontes russas para atacar a UE e os seus valores foram frequentemente adotadas por agentes políticos internos, com a participação de outros agentes externos*»[84]. A emergência desta tendência não se estende, todavia, à "infodemia" relacionada com a pandemia de Covid-19, onde claramente se identificou, no espaço da União, campanhas massivas de desinformação rastreáveis até à Federação Russa, à China[85] e ao Irão[86].

Além dos velhos métodos, alguns dos quais aperfeiçoados, novas tendências começam a emergir no ecossistema da desinformação. É o caso da "*deslocação da desinformação das plataformas públicas e abertas para os serviços de mensagem encriptados*"[87], como o WhatsApp ou o Telegram. Esta deslocação, prevalente, por exemplo, no ecossistema informativo do Brasil[88], coloca novos e complexos desafios às plataformas digitais, levantando a questão de como combater a desinformação sem afetar a encriptação dos dados transmitidos[89]. A principal característica destes serviços de mensagens encriptados é a ausência

[83] Informação disponível em https://about.fb.com/news/2022/12/metas-2022-coordinated-inauthentic-behavior-enforcements/?utm_source=POLITICO.EU&utm_campaign=61f7777e66-EMAIL_CAMPAIGN_2023_01_05_12_30&utm_medium=email&utm_term=0_10959edeb5-61f7777e66-%5BLIST_EMAIL_ID%5D e consultada em 05/01/2023.

[84] Comunicação Conjunta ao Parlamento Europeu, ao Conselho Europeu, ao Conselho, ao Comité Económico e Social Europeu e ao Comité das Regiões, *Relatório sobre a execução do plano de ação contra a desinformação*, Bruxelas, 14.6.2019, JOIN(2019) 12 final, p. 3.

[85] Focando-se apenas nos conteúdos produzidos e difundidos pelos órgãos de comunicação social estatais chineses e concluindo que o Partido Comunista Chinês procurou controlar a narrativa em torno da origem e gestão da pandemia, primeiro com a difusão de narrativas positivas, depois com a adesão a teorias da conspiração, cfr. Vanessa Molter; Renee Diresta, *Pandemics & Propaganda: How Chinese State Media Creates and Propagates CCP Coronavirus Narratives*, The Harvard Kennedy School Misinformation Review, Vol. 1., 2020, p. 1-24.

[86] Judit Bayer; Bernd Holznagel; Katarzyna Lubianiec, *et al.*, *Disinformation and propaganda: impact on the functioning of the rule of law and democratic processes in the EU and its Member States – 2021 update,* cit., p. 28.

[87] Carme Colomina; Héctor Sánchez Margalef; Richard Youngs, *The impact of disinformation on democratic processes and human rights in the world*, cit., p. 3.

[88] Gustavo Teixeira de Faria Pereira; Iluska Maria da Silva Coutinho, *WhatsApp, desinformação e infodemia: o "inimigo" criptografado*, *Liinc Em Revista*, Vol. 18., nº 1, 2022, p. 1-22.

[89] Comunicação da Comissão ao Parlamento Europeu, ao Conselho, ao Comité Económico e Social Europeu e ao Comité das Regiões, *Orientações da Comissão Europeia relativas ao reforço do Código de Conduta sobre Desinformação*, Bruxelas, 26.5.2021, COM(2021) 262 final, p. 19.

de moderação de conteúdos (e, por consequência, da verificação de factos) e de sistemas algorítmicos de recomendação e personalização, o que, contudo, não impede a difusão da desinformação em rede e a sua maior popularidade face a conteúdos fidedignos, ainda que a investigação sugira que a escala e o alcance da desinformação nestas plataformas não seja tão «*dramática como é frequentemente retratado por especialistas e comentadores*»[90].

Outra fonte de preocupação prende-se com a difusão de desinformação em formatos multimédia, como imagens e áudios, os quais, segundo a investigação, têm maior poder persuasivo do que o texto[91]. Estes formatos dificultam sobremaneira a actuação dos algoritmos ao serviço da deteção da desinformação, sendo muitas vezes partilhados em plataformas de mensagens encriptados. Em especial, os formatos áudio, em plataformas como o "Clubhouse", que funciona como um serviço de rádio que permite aos utilizadores falarem para cinco mil ouvintes ao mesmo tempo, levantam especiais problemas de monitorização, uma vez que não são guardados registo dos áudios, tornando difícil a sua verificação[92]. Apesar destes serviços não serem ainda populares na Europa, podemos estar perante o próximo passo nas campanhas de desinformação.

Alguns dos novos métodos da desinformação beneficiam dos algoritmos e da inteligência artificial. No documento *«Uma definição de IA: principais capacidades e disciplinas científicas»,* o Grupo Independente de Peritos de Alto Nível sobre a Inteligência Artificial criado pela Comissão Europeia em 2018, advogou o alargamento do conceito de inteligência artificial proposto pela Comissão[93/94]. Para aquele Grupo, a inteligência artificial deve ser defi-

[90] Aliaksandr Herasimenka; Jonathan Bright; Aleksi Knuutila; Philip N. Howard, *Misinformation and professional news on largely unmoderated platforms: the case of telegram*, Journal of Information Technology & Politics, 2022, p. 1-15 (p. 8 e ss.).

[91] Cristian Vaccari; Andrew Chadwick, *Deepfakes and Disinformation: Exploring the Impact of Synthetic Political Video and Deception, Uncertainty, and Trust in News*, Social Media + Society, 2020, p. 1-13 (p. 2)

[92] Judit Bayer; Bernd Holznagel; Katarzyna Lubianiec, *et al.*, *Disinformation and propaganda: impact on the functioning of the rule of law and democratic processes in the EU and its Member States – 2021 update,* cit., p. 24.

[93] Segundo a Comissão, o «*conceito de inteligência artificial (IA) aplica-se a sistemas que apresentam um comportamento inteligente, analisando o seu ambiente e tomando medidas – com um determinado nível de autonomia – para atingir objetivos específicos*». Comunicação da Comissão ao Parlamento Europeu, ao Conselho Europeu, ao Comité Económico e Social Europeu e ao Comité das Regiões, *Inteligência artificial para a Europa*, Bruxelas, 25.4.2018, COM(2018) 237 final, p. 1.

[94] Grupo Independente de Peritos de Alto Nível sobre a Inteligência Artificial, *Uma Definição de IA: Principais capacidades e disciplinas científicas*, Bruxelas, 2019, disponível em: https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60669, consult. em 04/01/2023, p. 1.

nida como o *software* (e, eventualmente, o *hardware*) concebido pelo Homem, que, tendo recebido um objetivo complexo, atua na dimensão física ou digital percecionando o seu ambiente mediante a aquisição de dados, interpretando os dados recolhidos, raciocinando sobre o conhecimento ou processando as informações resultantes desses dados e decidindo as melhores ações a adotar para atingir o objetivo estabelecido[95]. Nestes termos, o conceito de inteligência artificial, e designadamente a disciplina científica que lhe subjaz, abrange diversos modelos e técnicas, tais como a aprendizagem automática (incluindo a aprendizagem profunda ou "*deep learning*"), o raciocínio automático e a robótica. Resulta claro da definição agora apresentada que a inteligência artificial depende de dados, algoritmos e de capacidade computacional. Os algoritmos, na sua forma mais simples, são apenas «*procedimentos informatizados destinados a resolver certos problemas ou a atingir determinados objetivos*»[96], ou seja, conjuntos de instruções ou regras (conjuntos esses que podem ser extremamente complexos) que um computador segue para resolver um problema ou executar uma tarefa, tais como classificar, priorizar, associar e filtrar dados[97]. A grande inovação introduzida pela aprendizagem automática (ou "*machine learning*") prende-se com a inversão dos mecanismos subjacentes aos algoritmos: em vez de as instruções *entrarem* no computador (*input*) para que *saia* um determinado resultado (*output*), o computador é alimentado simultaneamente com os dados e o resultado desejado, *saindo* «*o algoritmo que transforma os primeiros no segundo. Com a aprendizagem automática, os computadores escrevem os seus próprios programas*»[98]. Sem embargo, é preciso notar que qualquer algoritmo, enquanto «*objeto político*», não é uma «*idealização, uma ferramenta linguística ou dispositivo emergente*», antes, ele «*anima materialidades concretas*» e interfere com os «*processos sociais* [...], *mobilizando concretas relações sociais* [...], *reduzindo a complexidade do mundo social em termos de cálculos* [...] *que só podem ser entendidos por máquinas*»[99]. Cathy O'Neil chega ao ponto de afirmar que os modelos algorítmicos não passam de «*opiniões incorporadas em matemática*»[100].

[95] *Idem*, p. 6.

[96] Inês da Silva Costa, *A proteção da pessoa na era dos big data: a opacidade do algoritmo e as decisões automatizadas*, Revista Electrónica de Direito, Vol. 24, nº 1, 2021, p. 34-82 (p. 40)

[97] Samantha Shorey; Philip N. Howard, *Automation, Big Data, and Politics: A Research Review*, International Journal of Communication. Vol. 10, 2016, p. 5032-5055 (p. 5033)

[98] Pedro Domingos, *A Revolução do Algoritmo Mestre*, 17ª ed., Lisboa, Manuscrito, 2017, p. 30.

[99] Josh Scannell, *What Can an Algorithm Do?*, DIS MAGAZINE, disponível em: http://dismagazine.com/discussion/72975/josh-scannell-what-can-an-algorithm-do/, consult. em: 04/06/2020.

[100] Cathi O'Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*, London, Penguin Books, 2016, p. 21.

No domínio da desinformação, os algoritmos podem ser usados ou para combater a desinformação, detetando-a, classificando-a, sinalizando-a, removendo-a ou desclassificando-a[101], ou para a difusão e amplificação do fenómeno, quer através dos *bots*, quer através dos algoritmos de personalização ou recomendação[102], e, ainda, dos algoritmos de micro-direcionamento destinados, com base na recolha de dados pessoais e elaboração de perfis, a garantir que determinado conteúdo personalizado é dirigido a apenas uma pessoa ou a um grupo de pessoas, sendo adaptado em função das características percecionadas dos destinatários[103].

A inteligência artificial, designadamente a aprendizagem automática, vem abrir novos horizontes no que toca ao próprio processo de criação da desinformação, como resulta evidente do potencial de criação de textos do modelo de linguagem natural "ChatGPT"[104]. É o que pode também vir a suceder também com os designados "*deepfakes*" (ou "falsificações profundas"). Os "*deepfakes*" correspondem a qualquer conteúdo digital, seja ele áudio, vídeo ou imagem, criado artificialmente com recurso a algoritmos de aprendizagem automática[105]. A produção destes conteúdos é possível graças à inteligência

[101] Teresa Quintel; Carsten Ullrich, *Self-Regulation of Fundamental Rights? The EU Code of Conduct on Hate Speech, Related Initiatives and Beyond*, 2018, disponível em: https://ssrn.com/abstract=3298719, consult. em: 10/01/2023.

[102] Segundo o Regulamento dos Serviços Digitais, os sistemas de recomendação correspondem a «*um sistema total ou parcialmente automatizado utilizado por uma plataforma em linha para sugerir na sua interface* [...] *informações específicas aos destinatários do serviço ou conferir prioridade a essa informação,* [...], *ou que determine de outra forma a ordem relativa ou a proeminência das informações apresentadas*» (artº 3º, al. s). Preferimos, não obstante, a designação "sistemas" ou "algoritmos" de "personalização", na medida em que os sistemas de recomendação ficam aquém dos sistemas de personalização: estes algoritmos, tendo em conta a sua utilização, não se limitam a recomendar, mas definem verdadeiramente o conteúdo visualizado.

[103] Cfr., *inter alia*, a Proposta de Regulamento do Parlamento Europeu e do Conselho sobre a transparência e o direcionamento da propaganda política, Bruxelas, 25.11.2021, COM(2021) 731 final, p. 3.

[104] *Vide* Nathan E. Sanders; Bruce Schneier, *How ChatGPT Hijacks Democracy*, The New York Times, 15/01/2023, disponível em: https://www.nytimes.com/2023/01/15/opinion/ai-chatgpt-lobbying-democracy.html, consult. em: 16/03/2023. Cfr., ainda, Josh A. Goldstein, *et al.*, *Generative Language Models and Automated Influence Operations: Emerging Threats and Potential Mitigations*, 2023, disponível em: https://arxiv.org/abs/2301.04246, consult. em: 15/01/2023. Defendendo uma suspensão de seis meses no desenvolvimento dos sistemas de inteligência artificial com capacidades superiores às do "GPT-4", cfr. a carta aberta do *Future of Life Institute* de 29/03/2023, disponível em https://futureoflife.org/open-letter/pause-giant-ai-experiments/, consult. em: 04/04/2023.

[105] Keir Giles; Kim Hartmann; Munira Mustaffa, *The Role of Deepfakes in Malign Influence Campaigns*, NATO Strategic Communications Centre of Excellence, 2019, disponível em: https://

artificial, nomeadamente às "*Generative Adversarial Networks*", uma tecnologia que utiliza elementos audiovisuais e sonoros autênticos (*inputs*) como ferramenta de treino das expressões características de uma dada pessoa, fabricando tantas versões quantas as necessárias até alcançar um resultado que considera semelhante ou igual ao real[106]. O *output* desta tecnologia traduz-se em conteúdos digitais que mimetizam na perfeição o que um indivíduo, por exemplo um líder político, diz ou faz[107], tornando potencialmente impossível, mesmo para os próprios sistemas de inteligência artificial, distinguir o artificial do original.

Até à data, ainda não foram encontradas evidências da utilização maligna de "*deepfakes*" em campanhas de desinformação, pelo menos em formato vídeo[108]. Vários fatores podem ser apontados para tal[109]: o conteúdo criado pelos sistemas de inteligência artificial ainda padece de imperfeições que o tornam detetável; a facilidade com que se pode colocar lado a lado o conteúdo fabricado e o conteúdo original no âmbito da verificação de factos pode tornar menos apelativa a utilização destes métodos nas campanhas de desinformação; o facto de as audiências estarem cada vez mais alerta para este fenómeno retirou-lhe parte do seu apelo inicial (por exemplo, são agora comuns os *pivots* de televisão virtuais); o grau de especialidade e capacidade informática que ainda é necessário para produzir os "*deepfakes*", principalmente se comparados com os velhos métodos da desinformação. Em todo o caso, é inegável o potencial dos "*deepfakes*", principalmente se utilizados para fins de manipulação política, na produção de desinformação: as falsificações profun-

stratcomcoe.org/publications/the-role-of-deepfakes-in-malign-influence-campaigns/72, consult. em: 20/09/2022, p. 4. Os Autores defendem, ainda, a inclusão do texto quando utilizado para gerar conteúdos falsos, mas aparentemente genuínos, distribuídos em grande escala.

[106] Cristian Vaccari; Andrew Chadwick, *Deepfakes and Disinformation: Exploring the Impact of Synthetic Political Video and Deception, Uncertainty, and Trust in News*, cit., p. 2-3.

[107] Tom Dobber; Nadia Metoui; Damian Trilling, et al., *Do (Microtargeted) Deepfakes Have Real Effects on Political Attitudes?*, The International Journal of Press/Politics, Vol. 26, nº 1, 2021, p. 69-91 (p. 69)

[108] Judit Bayer; Bernd Holznagel; Katarzyna Lubianiec, et al., *Disinformation and propaganda...*, cit., p. 24. Sem embargo, já ocorreram casos em que os "*deepfakes*" foram utilizados com sucesso para ataques financeiros contra empresas, simulando-se mensagens de voz em que os dirigentes de topo solicitavam aos departamentos financeiros a transferência urgente de avultadas quantias financeiras. Cfr. Keir Giles; Kim Hartmann; Munira Mustaffa, *The Role of Deepfakes in Malign Influence Campaigns,* cit., p. 15. *Vide*, ainda, a recente notícia de que a Microsoft desenvolveu uma tecnologia ("VALL-E") capaz de recriar qualquer voz com base numa amostra de apenas três segundos, disponível em: https://tinyurl.com/2p9dz49s, consult. em: 23/02/2023.

[109] Parte destes fatores são referidos por Keir Giles; Kim Hartmann; Munira Mustaffa, *The Role of Deepfakes in Malign Influence Campaigns,* cit., p. 8 e ss.

das exacerbam o problema da desinformação, tornando-a mais eficaz e crível. Aliás, mesmo quando se conclui que o conteúdo fabricado não é (ainda) capaz de enganar os cidadãos, a investigação avança que a incerteza que ele cria mina a confiança do público nas instituições e nos órgãos de comunicação social[110]. Por outro lado, as técnicas falsificação profunda podem ser colocadas ao serviço dos velhos métodos da desinformação, aperfeiçoando-os. É o que acontece, por exemplo, com a geração artificial de imagens de perfis que replicam na perfeição seres humanos e que depois são usados nas contas falsas dos *bots*, dos *trolls* ou das "*Potemkin personas*".

O potencial dos "*deepfakes*" e, de um modo geral, de todos os métodos de desinformação, pode ser exponenciado pelo recurso ao micro-direcionamento[111] e à personalização de conteúdos. Quando falamos em personalização, referimo-nos a sistemas algorítmicos capazes de prever «*o que é relevante para o utilizador, filtrando e removendo a informação irrelevante e aumentando a relevância daquela informação considerada importante e pertinente para um utilizador individual*»[112]. Estas técnicas têm hoje inegáveis consequências no nosso ambiente digital, tornando possível que dois utilizadores, usando a mesma plataforma digital, sejam expostos a conteúdos e mensagens totalmente distintas, especialmente curadas para si ou para o grupo que partilha determinadas características e onde foram inseridos por um algoritmo[113], além de contribuírem para que os cidadãos se encontrem potencialmente isolados dentro das plataformas digitais, quais ilhas, lendo, vendo e ouvindo apenas ecos da sua própria mundividência. O ecossistema assim criado e explorado põe em perigo o Estado de Direito, a democracia, e vários direitos e liberdades fundamentais, designadamente a liberdade de expressão e de informação[114]:

[110] Cristian Vaccari; Andrew Chadwick, *Deepfakes and Disinformation: Exploring the Impact of Synthetic Political Video and Deception, Uncertainty, and Trust in News*, cit., p. 9.
A este propósito, os Autores citam, apropriadamente, Hannah Arendt: «*A people that no longer can believe anything cannot make up its mind. It is deprived not only of its capacity to act but also of its capacity to think and to judge. And with such a people you can then do what you please*».

[111] Com igual conclusão, *vide* Tom Dobber; Nadia Metoui; Damian Trilling, et al., *Do (Microtargeted) Deepfakes Have Real Effects on Political Attitudes?*, The International Journal of Press/Politics, Vol. 26, nº 1, 2021, p. 69-91 (p. 70)

[112] Engin Bozdag, *Bias in algorithmic filtering and personalization*, Ethics and Information Technology, Vol. 15, 2013, p. 209-227 (p. 211)

[113] Tiago Morais Rocha, *A Era Digital e o Estado de Direito Democrático na União Europeia*, Porto, cit., p. 40.

[114] Maja Brkan defende que a liberdade de expressão e de informação no ambiente digital pode ser colocada em perigo por quatro fatores associados à inteligência artificial: (1) a personalização de conteúdos, (2) a utilização de algoritmos e aprendizagem automática na moderação, bloqueio e remoção de conteúdos, (3) a despromoção e "diluição" de conteúdos legais mas

quando a esfera pública é não só ocupada, mas dominada, pela manipulação, pela distorção e pela desinformação é o próprio propósito primário da liberdade (de expressão e de informação) que se queda frustrado[115].

4. A resposta política e jurídica da União Europeia

Refletindo sobre as alterações impostas à liberdade de expressão no atual ecossistema digital, Jack M. Balkin defende o imperativo de se superar um modelo dualista de regulação do discurso e da liberdade de expressão em prol de um modelo pluralista que reconheça a variedade de atores que hoje dominam o espaço público[116]. Para o Autor, a conceptualização da liberdade de expressão enquanto relação diádica Estado-pessoa que caracterizou o séc. XIX e XX encontra-se ultrapassada por uma conceção triangular, que integra, agora, um novo *player*: um conjunto de infraestruturas digitais de comunicação privadas geridas por empresas que governam a esfera pública digital onde as pessoas se encontram e comunicam[117]. Esta realidade implica novos modelos de regulação, baseados na cooperação entre público e o privado, na co e na autorregulação. Na sua cruzada contra a desinformação, a União Europeia aceita este novo paradigma regulatório.

Efetivamente, reconhecemos na resposta da União Europeia aos desafios colocados pela desordem informativa uma primeira fase, que vai sensivelmente de 2015 a 2022, essencialmente assente em instrumentos de *soft law*, os quais se vieram a relevar relativamente ineficazes atenta a sua natureza não-vincu-

danosos (desinformação) e (4) a implementação, através de processos automáticos, do direito ao esquecimento. Reconhecendo a importância da liberdade de expressão no contexto europeu (cfr. o artº 11º da Carta dos Direitos Fundamentais da União Europeia (CDFUE) e o artº 10º da Convenção Europeia dos Direitos Humanos (CDEH), a Autora problematiza a ausência de eficácia horizontal dos direitos fundamentais, nomeadamente do artº 11º da CDFUE, quando combinado com os poderes *de facto* exercidos pelas plataformas digitais e a forma como eles afetam o ambiente discursivo público e privado de milhões de cidadãos. Uma das vias de superação proposta para a ausência de eficácia horizontal da Carta passa pela conversão da relação horizontal entre as plataformas e os seus utilizadores numa relação "quasi-vertical", «*permitindo a aplicação dos direitos fundamentais contra indivíduos ou empresas que possuam poderes capazes de impactar os interesses de um público mais vasto*». Assim, as plataformas digitais seriam um *tertium genus* ou uma categoria especial de destinatários de direitos fundamentais, entre os indivíduos e as autoridades públicas, partindo-se do reconhecimento do seu (cada vez mais) importante papel na esfera pública europeia, enquanto *gatekeepers*, curadores e moderadores de conteúdos. *Vide* Maja Brkan, *Freedom of Expression and Artificial Intelligence: On Personalisation, Disinformation and (Lack Of) Horizontal Effect of the Charter*, 2019, disponível em: https://ssrn.com/abstract=3354180, consult. em: 10/01/2023.

[115] Tiago Morais Rocha, *A Era Digital e o Estado de Direito Democrático na União Europeia*, Porto, cit., p. 64.

[116] Jack M. Balkin, *Free Speech is a Triangle*, Columbia Law Review, Vol. 118, nº 7, 2018, p. 2011-2056.

[117] *Idem*, p. 2012.

lativa[118], tendo merecido críticas, quer no seio das instituições, quer na doutrina especializada[119]. Um bom exemplo dos problemas então apontados na doutrina é a descrição feita em 2020 por James Pamment quanto à política pública europeia em torno da desinformação: «*caracteriza-se pela falta de clareza terminológica, por bases jurídicas pouco claras e não testadas, por uma fraca base de evidência, por um mandato político pouco fiável e por uma variedade de instrumentos que se desenvolveram de uma forma orgânica em vez de uma forma sistemática*»[120]. A partir de 2022, com a publicação do Regulamento dos Serviços Digitais, inicia-se em definitivo uma nova fase já baseada na regulação através da *hard law*, fase essa que despontou anteriormente, embora em âmbitos distintos, mas igualmente tendo como pano de fundo a moderação de conteúdos *online* pelas plataformas digitais, com a Diretiva relativa aos direitos de autor e direitos conexos no mercado único digital[121] e com o Regulamento relativo ao combate à difusão de conteúdos terroristas em linha[122].

Um pouco por toda a parte têm-se multiplicado as tentativas de encontrar quadros regulatórios satisfatórios para governar (e constitucionalizar) a esfera pública digital. Debelar os problemas causados pela desordem informativa

[118] Teresa Quintel; Carsten Ullrich, *Self-Regulation of Fundamental Rights? The EU Code of Conduct on Hate Speech, Related Initiatives and Beyond*, cit.

[119] Por exemplo, em relação ao Código de Conduta da UE sobre Desinformação, a Comissão, em 2020, avaliou o primeiro ano da sua implementação pelas plataformas digitais, identificando quatro categorias de insuficiências de instrumento de *soft law*: (1) aplicação do Código pelas plataformas signatárias e pelos Estados-Membro de forma inconsistente e incompleta, (2) ausência de definições uniformes, (3) existências de várias falhas na cobertura dos compromissos assumidos, e (4) limitações associadas à própria natureza autorregulatória do Código. Do mesmo modo, o ERGA (Grupo de Reguladores Europeus dos Serviços de Media Audiovisuais) produziu igualmente um relatório de avaliação onde conclui que «*o Código tem fraquezas significativas que necessitam de ser resolvidas para que se possam alcançar os objetivos propostos*». *Vide* Commission Staff Working Document, *Assessment of the Code of Practice on Disinformation – Achievements and areas for further improvement*, Bruxelas, 10.09.2020, SWD(2020) 180 final; European Regulators Group for Audiovisual Media Services, *ERGA Report on Disinformation: Assessment of the Implementation of the Code of Practice*, 2020, disponível em: https://erga-online.eu/wp-content/uploads/2020/05/ERGA-2019-report-published-2020-LQ.pdf, consult. em 11/01/2023.

[120] James Pamment, *The EU's Role in Fighting Disinformation: Taking Back Initiative – Working Paper nº 1*, Washington, Carnegie Endowment for International Peace, 2020, disponível em: https://carnegieendowment.org/2020/07/15/eu-s-role-in-fighting-disinformation-taking-back-initiative-pub-82286, consult. em: 28/11/2022, p. 5.

[121] Diretiva (UE) 2019/790 do Parlamento Europeu e do Conselho de 17 de abril de 2019 relativa aos direitos de autor e direitos conexos no mercado único digital e que altera as Diretivas 96/9/CE e 2001/29/CE, JO L 130, 17.5.2019, p. 92-195.

[122] Regulamento (UE) 2021/784 do Parlamento Europeu e do Conselho de 29 de abril de 2021 relativo ao combate e difusão de conteúdos terroristas em linha, JO L 172, 17.5.2021, p. 79-109.

é tarefa especialmente complexa, na medida em que os tradicionais poderes públicos se vêm confrontados com escolhas de fronteira (ou verdadeiros dilemas), entre a regulação eficaz das plataformas e o pleno respeito pelas liberdades e direitos fundamentais dos cidadãos, nomeadamente a liberdade de expressão e de opinião. Já não se trata tanto de aceitar que as democracias devem, em face dos perigos, ser militantes[123], mas de definir até onde deve ir essa militância de modo a não aniquilar, coartar ou violar, em nome da verdade, da democracia ou do pluralismo, os direitos, liberdades e garantias e, por arrasto, os valores e ideais em que as próprias democracias se fundam. Assim, qualquer solução que possa ser dada ao fenómeno desinformativo implica uma prévia determinação dos limites da liberdade de expressão, de opinião e de informação e a sua ponderação no confronto com outros direitos, liberdades e interesses legítimos[124].

Na busca por uma regulação adequada, um dos principais desafios prende-se com o caráter global ou transnacional destes fenómenos: eles não mais estão confinados às fronteiras de um Estado, pelo que qualquer medida adotada a esse nível padece, à partida e em princípio, de uma anunciada ineficácia[125]. Por outro lado, e regressando ao triângulo da liberdade de expressão de Balkin, a entrada na esfera pública de novos atores-âncora-plataforma no discurso coletivo, torna imperativo o seu envolvimento nos esquemas de regulação, muitas vezes como destinatários das medidas e normas legais. Ou seja, qualquer regulação eficaz do discurso público na esfera digital torna os poderes públicos dependentes de empresas privadas globais na execução das políticas que soberanamente definam. A União Europeia está particularmente empenhada e capacitada, no âmbito de uma regulação subsidiária e multinível, para responder a fenómenos à escola europeia, intervindo diretamente na esfera das plataformas digitais. É a esse percurso a que nos dedicaremos doravante.

Como antedito, a abordagem da União Europeia ao fenómeno da desinformação pode ser decomposta, essencialmente, em duas fases. Na primeira fase, o papel da UE girou sobretudo em torno do estudo do fenómeno desinformativo e da emissão de orientações de política pública, apostando essencialmente

[123] Karl Loewenstein, *Militant Democracy and Fundamental Rights* I, The American Political Science Review, Vol. 31, nº 3, 1937, p. 417-432.

[124] Oreste Pollicino; Giovanni De Gregorio; Laura Somaini, *The European Regulatory Conundrum to Face the Rise and Amplification of False Content Online*, *In* G. Z. Capaldo, (ed.), *The Global Community: Yearbook of International Law and Jurisprudence 2019*, Oxford, Oxford University Press, 2020, p. 319-356.

[125] Ingolf Pernice, *Risk management in the digital constellation – a constitutional perspective (part I)*, Revista D'Internet, Dret i Política, nº 26, 2018, p. 83-94.

em soluções co e autorregulatórias[126]. Os grandes objetivos prosseguidos prendiam-se, então, com a melhoria das capacidades das instituições para detetarem, analisarem e denunciarem a desinformação, com a criação de respostas comuns de combate à desinformação, mobilizando-se o setor privado, promovendo-se a transparência e a autorresponsabilização, e reforçando-se a resiliência das sociedades[127].

Efetivamente, as preocupações das instituições da União com esta temática são expressas no primeiro semestre de 2015, quando o Conselho Europeu solicitou ao Alto Representante para os Negócios Estrangeiros e a Política de Segurança que, em cooperação com os Estados-Membro e as instituições, tomasse medidas contra as campanhas de desinformação levadas a cabo pela Federação da Rússia no contexto do conflito com a Ucrânia[128]. Em resposta, foi criado o grupo de trabalho «East StratCom» com o objetivo de desenhar uma estratégia de comunicação e promoção das políticas da UE junto dos países do Leste europeu vizinhos da UE, no âmbito da política de ação externa cuja finalidade é, entre outras, a consolidação e o apoio à democracia, ao Estado de Direito e aos direitos humanos (cfr. al. d) do nº 2 do artº 21º TUE). O grupo de trabalho, integrado na Divisão de Comunicação Estratégica e Análise de Informação do Serviço Europeu de Ação Externa, desenha produtos e campanhas de comunicação para melhor explicar os valores, os interesses e as políticas da UE junto dos países da "Parceria Oriental" (Arménia, Azerbaijão, Bielorrússia, Geórgia, Moldávia e Ucrânia), procurando reforçar o ambiente mediático nesses países, além de elaborar relatórios periódicos expondo campanhas e narrativas de desinformação com origem russa. O principal produto da «East StratCom» é o «EUvsDisinfo», um *website* que contém um repositório de conteúdos classificados como desinformação, além de produzir um relatório semanal sobre este tema[129].

Só mais tarde, já em junho de 2017, é que o Parlamento Europeu manifestou a sua preocupação sobre esta matéria, solicitando à Comissão que analisasse «*o quadro jurídico em matéria de notícias falsas*» verificando «*a possibilidade de uma intervenção legislativa para limitar a divulgação e difusão de conteúdos falsos*», sempre em reconhecimento de que a livre troca de opiniões é fundamental para a

[126] Oreste Pollicino; Giovanni De Gregorio; Laura Somaini, *The European Regulatory Conundrum to Face the Rise and Amplification of False Content Online*, cit.

[127] Comunicação Conjunta ao Parlamento Europeu, ao Conselho Europeu, ao Conselho, ao Comité Económico e Social Europeu e ao Comité das Regiões, *Plano de Ação contra a Desinformação*, Bruxelas, 5.12.2018, JOIN(2018) 36 final, p. 6 e ss.

[128] Conselho Europeu, *Reunião do Conselho Europeu (19 a 20 de março de 2015) – Conclusões*. EUCO 11/15. Bruxelas, 20.03.2015, p. 5.

[129] Para mais informações, cfr. http://www.euvsdisinfo.eu/.

democracia, algo igualmente válido no ambiente digital[130]. Curiosamente, se até aí a desinformação vinha sendo tratada como um problema de segurança externa, o Parlamento Europeu dá-lhe um novo (e adicional) enquadramento jurídico, ao tratar o tema no âmbito do Mercado Único Digital, designadamente por referência às responsabilidades dos prestadores intermediários, na aceção da Diretiva sobre comércio eletrónico[131].

Na sequência dos apelos, quer do Conselho Europeu, quer do Parlamento, a Comissão, no seu Programa de Trabalho para 2018, inclui, também no âmbito do Mercado Único Digital[132], a intenção de apresentar «*uma iniciativa em matéria de luta contra as notícias falsas*»[133]. Efetivamente, a desinformação deixa a partir daqui de ser percecionada somente como uma ameaça externa, mas como um problema interno[134], com capacidade para afetar a integridade dos processos eleitorais, nomeadamente das eleições para o Parlamento Europeu de 2019. Neste sentido, no início de 2018 foi selecionado o Grupo Independente de Alto Nível sobre Notícias Falsas e Desinformação (HLEG), responsável por aconselhar a Comissão sobre a «*delimitação do fenómeno das notícias falsas, a definição dos papéis e responsabilidades dos intervenientes relevantes,* [...] *balancear as posições em jogo, e formular recomendações*»[135]. O HLEG apresentou o seu relatório final em março de 2018, propondo uma abordagem multidimensional para o fenómeno desinformativo, que inicia, como visto, pela proposta de uma definição de desinformação. A abordagem multidimensional proposta pelo HLEG baseia-se em cinco pilares (aumentar a transparência do ecossistema mediático digital; promover a literacia digital; desenvolver ferramentas que permitam aos utilizadores das plataformas e aos jornalistas reagirem face à

[130] Resolução do Parlamento Europeu, de 15 de junho de 2017, sobre as plataformas em linha e o Mercado Único Digital (2016/2276 (INI)) JO C 331 de 18.9.2018, p. 141.

[131] Diretiva 2000/31/CE do Parlamento Europeu e do Conselho de 08 de Junho de 2000 relativa a certos aspetos legais dos serviços da sociedade de informação, em especial do comércio eletrónico, no mercado interno. JO L 178 de 17.7.2000.

[132] Para uma análise crítica, *vide* Joana Covelo Abreu, *O Mercado Único Digital como o novo mundo para a União Europeia: repercussões na estrutura regulatória social e institucional – a redefinição do serviço universal e do Organismo de Reguladores Europeus das Comunicações Eletrónicas (ORECE),* UNIO – EU Law Journal, Vol. 4, nº 2, 2018, p. 59-72.

[133] Comunicação da Comissão ao Parlamento Europeu, ao Conselho, ao Comité Económico e Social Europeu e ao Comité das Regiões, *Programa de Trabalho da Comissão para 2018*, Estrasburgo, 24.10.2017, COM(2017) 650 final, p. 5.

[134] Tendo em conta, *inter alia*, o referendo sobre a permanência do Reino Unido na UE, as eleições francesas e alemãs, a atuação do governo húngaro, a deriva política em Itália e o referendo pela independência na Catalunha.

[135] *Vide* https://digital-strategy.ec.europa.eu/en/news/experts-appointed-high-level-group-fake-news-and-online-disinformation, consult. em 19/01/2023.

desinformação; proteger a diversidade e sustentabilidade dos meios de comunicação social europeus; e promover a investigação sobre o impacto da desinformação na Europa) em torno dos quais são recomendadas um conjunto de ações e politicas[136].

Pouco tempo depois, em abril de 2018, a Comissão explicitou os princípios orientadores e os objetivos que devem orientar as ações para combater o fenómeno da desinformação de forma eficaz[137], tendo em conta a incapacidade de as plataformas digitais atuarem face ao desafio da desinformação e ao uso manipulativo das suas infraestruturas digitais, a que se somavam dúvidas quanto à suficiência do nível de proteção prestado pelas plataformas aos seus utilizadores[138]. A estratégia da Comissão para combater a desinformação, bebendo do relatório do HLEG, baseou-se em quatro princípios[139]: 1) melhorar a transparência atinente à origem da informação e à forma como esta é produzida, patrocinada, divulgada e direcionada, a fim de permitir que os cidadãos avaliem os conteúdos a que acedem em linha; 2) promover a diversidade de informação, por forma a permitir aos cidadãos tomarem decisões informadas com base no pensamento crítico, mediante o apoio ao jornalismo de alta qualidade, à literacia mediática e ao reequilíbrio da relação entre os criadores e os distribuidores de informação; 3) promover a credibilidade da informação, fornecendo uma indicação da sua fiabilidade, nomeadamente com a ajuda de sinalizadores de confiança, bem como melhorando a rastreabilidade da informação e a autenticação de fornecedores influentes de informação; 4) oferecer soluções inclusivas que envolvam todos os atores relevantes no ecossistema digital. A estratégia da Comissão é clara ao reconhecer que no domínio da desinformação as plataformas têm deveres que extravasam as obrigações legais a que estão vinculadas, os quais decorrem do seu «*papel central*» no ecossistema digital[140]. Não obstante o reconhecimento

[136] High Level Group On Fake News And Online Disinformation, *A multi-dimensional approach to disinformation*. Comissão Europeia, 2018, disponível em: https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=50271, consult. em 19/01/2023, p. 7.

[137] Comunicação da Comissão ao Parlamento Europeu, ao Conselho, ao Comité Económico e Social Europeu e ao Comité das Regiões, *Combater a desinformação em linha: uma estratégia europeia*, cit., p. 5.

[138] *Idem*, p. 2.

[139] É preciso notar que a estratégia da Comissão segue, em traços largos, os princípios adotados pelas organizações internacionais sobre a matéria. *Vide*, em especial, a Declaração Conjunta sobre a Liberdade de Expressão e "Notícias Falsas", Desinformação e Propaganda da Organização das Nações Unidas, da Organização para a Segurança e Cooperação na Europa, da Organização dos Estados Americanos e da Comissão Africana sobre Direitos Humanos e dos Povos de 03.03.2017, disponível em: https://www.osce.org/files/f/documents/6/8/302796.pdf, consult. em 19/01/2023.

[140] Comunicação da Comissão ao Parlamento Europeu, ao Conselho, ao Comité Económico e Social Europeu e ao Comité das Regiões, *Combater a desinformação em linha: uma estratégia europeia*, cit., p. 8.

implícito que esta constatação encerra – o da ausência de mecanismos legais capazes de dar resposta aos desafios da desinformação –, a aposta da Comissão foi no sentido da autorregulação das plataformas, com a convocação de um fórum multilateral sobre a desinformação do qual deveria resultar, até julho de 2018, um código de conduta sobre a desinformação à escala da UE[141]. Além da autorregulação, a Comissão procurou atuar no eixo da fiabilidade da informação, através da criação de uma rede europeia independente de verificadores de facto tendo em vista, entre outros, o estabelecimento de métodos de trabalho comuns e o intercâmbio de boas práticas.

O Conselho Europeu, em junho de 2018, voltou a convidar o Alto Representante e a Comissão para, em conjunto com Estados-Membros, apresentarem um plano de ação com propostas específicas para uma resposta coordenada da União ao desafio da desinformação[142], enquadrando novamente o tema no âmbito da segurança e defesa, em lugar do mercado interno, designadamente o Mercado Único Digital, como vinha fazendo a Comissão, não obstante o facto de, na sua comunicação de 2018, ter a Comissão expressado que a desinformação é, antes de mais, um problema que se relaciona com a qualidade das democracias, com a expressão da razão pública e com os direitos fundamentais, nomeadamente a liberdade de expressão[143].

O fórum multilateral sobre a desinformação, que se realizou no primeiro semestre de 2018, reuniu os principais *stakeholders*, entre representantes das plataformas digitais, dos meios tradicionais de imprensa, de verificadores de factos, de membros da academia e da sociedade civil. Tal como previsto, foi aí concluído o Código de Conduta da UE sobre a Desinformação, o qual foi subscrito por várias empresas e plataformas digitais, nomeadamente a Google, o Facebook, o Twitter, a Mozilla, a Microsoft (em maio de 2019) e o TikTok (em junho de 2020). O Código de Conduta, na sua versão de 2018, é um instrumento de autorregulação no quadro de uma gestão do risco digital multinível, assente num compromisso entre um nível superior e transnacional de legitimação política – a União -, que atua em representação das suas instituições, dos Estados-Membros e dos cidadãos da União, e o setor privado, responsável primário pela gestão das infraestruturas digitais onde a desinformação é difundida, procurando introduzir um sistema de regulação e escrutínio onde cada signatário, reconhecendo o seu específico papel a

[141] *Idem*, p. 9.

[142] Conselho Europeu, *Reunião do Conselho Europeu (28 de junho de 2018) – Conclusões*. EUCO 9/18. Bruxelas, 28.06.2018, p. 6.

[143] Comunicação da Comissão ao Parlamento Europeu, ao Conselho, ao Comité Económico e Social Europeu e ao Comité das Regiões, *Combater a desinformação em linha: uma estratégia europeia*, cit., p. 1.

desempenhar na luta contra a desinformação, assume uma série de compromissos voluntários. Esses compromissos visam, essencialmente, [i.] controlar e limitar a colocação de anúncios em contas e *websites* associadas à desinformação para reduzir as suas receitas (a "desmonetização" dos conteúdos), [ii.] aumentar a transparência e o reconhecimento da publicidade em linha, tornando-a claramente distinguível de outros conteúdos, e, em especial, da "publicidade temática", nomeadamente a política, [iii.] assegurar a integridade dos serviços das plataformas digitais, designadamente contra sistemas automatizados (*bots*), [iv.] dar prioridade a informações «*pertinentes, autênticas e fidedignas nas pesquisas, nos* feeds *de notícias ou noutros canais de distribuição com classificação automática*»[144], muito embora os signatários do Código recusem liminarmente qualquer medida, de caráter voluntário ou não, para «*suprimir ou impedir o acesso a mensagens ou conteúdos ilícitos apenas por se julgar que são "falsos"*»[145], o que, para evitar tornar as plataformas em "árbitros da verdade *online*", as transforma em verdadeiras árbitros de visibilidade[146], [v.] permitir e promover a investigação sobre desinformação e propaganda política, [vi.] e avaliar anualmente as medidas adotadas no combate à desinformação[147].

Não obstante a importância simbólica e a circunstância de abordar vários aspetos pertinentes, o Código de Conduta estava impregnado de "*wishful thinking*" e de linguagem ambígua ou pouco ambiciosa[148] ("compromisso", "envidar esforços comercialmente razoáveis", etc.). Além do caráter voluntário e de ausência de compromissos objetivos ou mensuráveis, o documento não estabelecia uma abordagem comum para todos os subscritores, que podiam optar livremente pelos compromissos que pretendiam subscrever. Sem embargo, teve a virtualidade de reunir as principais plataformas digitais em torno do compromisso de, anualmente, produzirem um relatório de autoavaliação contendo um balanço do seu trabalho no combate à desinformação, relatórios esses analisados por uma organização terceira e independente escolhida pelos próprios signatários[149].

[144] Código de Conduta da UE sobre Desinformação, 26 de setembro de 2018, disponível em: https://combatefakenews.lusa.pt/wp-content/uploads/2020/08/CNECT-2019-20022-00-00-PT-TRA-00pdf-1.pdf. consult. em 20/01/2023, p. 7.

[145] *Idem*, p. 6.

[146] Tiago Morais Rocha, *A Era Digital e o Estado de Direito Democrático na União Europeia*, Porto, cit. p. 73.

[147] Código de Conduta da UE sobre Desinformação, cit.

[148] Oreste Pollicino; Giovanni De Gregorio; Laura Somaini, *The European Regulatory Conundrum to Face the Rise and Amplification of False Content Online*, cit.

[149] Código de Conduta da UE sobre Desinformação, cit., p. 8.

O fórum multilateral donde emergiu o Código de Conduta da UE sobre Desinformação integrava também um "Painel de Consulta", responsável por emitir um parecer sobre o Código. Aquando da sua publicação, o "Painel de Consulta", recorrendo à definição de autorregulação da própria UE[150], concluiu que o Código «*não é, de modo algum, autorregulatório, pelo que as Plataformas, apesar dos seus esforços, não criaram um Código de Conduta*», uma vez que lhe faltam pressupostos essenciais, como mecanismos de auto-monitorização, cumprimento e imposição. Sem prejuízo da avaliação do "Painel de Consulta", a Comissão considerava, no final de 2018, que o Código constituía «*um enquadramento adequado para atingir os objetivos pretendidos*»[151]. Esse entendimento haveria de mudar, porém, aquando da apresentação, pelas plataformas, dos primeiros relatórios de implementação do Código, em janeiro de 2019, o que levou a que a Comissão pedisse aos signatários que redobrassem os seus esforços no âmbito da preparação para as eleições para o Parlamento Europeu de 2019[152].

Ao mesmo tempo, e dando eco às preocupações do Conselho Europeu expressas em outubro de 2018[153], no contexto da preparação das eleições de 2019 para o PE, a Comissão, na sequência do discurso sobre o estado da União proferido pelo Presidente Juncker, preparou um pacote de medidas relativas ao processo eleitoral do ano seguinte. Assim, em setembro de 2018[154], a Comissão

[150] «*A autorregulação constitui um tipo de iniciativa voluntária que oferece aos operadores económicos, aos parceiros sociais, às associações e às organizações não governamentais a possibilidade de adotarem orientações comuns entre si e para si. Cabe-lhes a responsabilidade de elaborar, acompanhar e fazer cumprir as referidas orientações*». Cfr. considerando (14) da Diretiva (UE) 2018/1808 do Parlamento Europeu e do Conselho de 14.11.2018 que altera a Diretiva 2010/13/UE relativa à coordenação de certas disposições legislativas, regulamentares e administrativas dos Estados-Membros respeitantes à oferta de serviços de comunicação social audiovisual (Diretiva Serviços de Comunicação Social Audiovisual), para a adaptar à evolução das realidades do mercado. JO L 303 de 28.11.2018.

[151] Relatório da Comissão ao Parlamento Europeu, ao Conselho Europeu, ao Conselho, ao Comité Económico e Social Europeu e ao Comité das Regiões sobre a aplicação da Comunicação «Combater a desinformação em linha: uma estratégia europeia». Bruxelas, 5.12.2018, COM(2018) 794 final.

[152] Cfr. Comunicado de imprensa da Comissão de 29.01.2019, disponível em: https://ec.europa.eu/commission/presscorner/detail/en/IP_19_746, consult. em 20/01/2023.

[153] Conselho Europeu, *Reunião do Conselho Europeu (18 de outubro de 2018) – Conclusões*, Bruxelas, 18.10.2018, EUCO 13/18, p. 3.

[154] Note-se que já em junho de 2018, a Comissão e o Alto Representante haviam apresentado uma comunicação com o objetivo de aumentar a resiliência da União face a ameaças híbridas, donde se destacava o papel da desinformação e da contrainformação no contexto das campanhas híbridas em curso na União. Cfr. Comunicação Conjunta ao Parlamento Europeu, ao Conselho Europeu e ao Conselho, *Aumentar a resiliência e reforçar a capacidade de enfrentar ameaças híbridas*, Bruxelas, 13.6.2018, JOIN(2018) 16 final.

apresenta uma comunicação tendente à garantia de eleições europeias livres e justas, reconhecendo que as eleições de maio de 2019 para o PE decorreriam num «*contexto muito diferente do contexto de todas as eleições anteriores*»[155], tendo em conta, entre outros, o espetro das campanhas de desinformação direcionada que haviam afetado os processos democráticos de vários países dentro e fora da União. Segundo a Comissão, os períodos eleitorais são momentos propícios à desinformação e à manipulação por países terceiros ou interesses privados, podendo afetar a integridade e a equidade do processo eleitoral e a confiança dos cidadãos, o que tem potencial para colocar em causa a própria democracia e o processo que lhe subjaz[156]. Não obstante o facto de a organização dos processos eleitorais ser uma competência que pertence essencialmente aos Estados-Membro, a Comissão propôs, então, um conjunto de medidas dirigidas aos Estados, donde se destaca a criação de redes de cooperação eleitoral[157], o reforço da transparência das campanhas políticas, designadamente nas plataformas digitais, de modo a que os cidadãos possam saber quem está por detrás das comunicações que recebem, e o reforço das regras em matéria de financiamento dos partidos políticos europeus e da proteção dos dados pessoais dos cidadãos para que não sejam usados abusivamente no âmbito das campanhas eleitorais[158].

Ainda na sequência do mandato do Conselho Europeu, a Comissão e o Alto Representante apresentam, em dezembro de 2018, um "Plano de Ação contra a Desinformação" que teve especialmente em conta o período eleito-

[155] Comunicação da Comissão ao Parlamento Europeu, ao Conselho, ao Comité Económico e Social Europeu e ao Comité das Regiões, *Garantir eleições europeias livres e justas*, Bruxelas, 12.9.2018, COM(2018) 637 final, p. 1.

[156] *Idem*.

[157] Cfr. Commission Recommendation of 12.9.2018 on election cooperation networks, online transparency, protection against cybersecurity incidents and fighting disinformation campaigns in the context of elections to the European Parliament, Brussels, 12.9.2018, C(2018) 5949 final.

[158] Proposta de Regulamento do Parlamento Europeu e do Conselho que altera o Regulamento (UE, Euratom) nº 1141/2014 no que diz respeito a um procedimento de verificação de violações das regras em matéria de proteção de dados pessoais no âmbito das eleições para o Parlamento Europeu, Bruxelas, 12.9.2018, COM(2017) 636 final. Esta proposta acabou por ser adotada pelo Parlamento Europeu e pelo Conselho, dando origem ao Regulamento (UE, Euratom) 2019/493 do Parlamento Europeu e do Conselho, de 25 de março de 2019, que altera o Regulamento (UE, Euratom) n.º 1141/2014 no que diz respeito a um procedimento de verificação de violações das regras em matéria de proteção de dados pessoais no âmbito das eleições para o Parlamento Europeu, JO LI 85 de 27.3.2019, p. 7-10. Nos termos do aditado artº 10º-A, nº 1, um «*partido político europeu ou uma fundação política europeia não pode influenciar ou tentar influenciar de forma deliberada os resultados das eleições para o Parlamento Europeu, aproveitando-se de uma violação cometida por uma pessoa singular ou coletiva das regras aplicáveis em matéria de proteção de dados pessoais*».

ral que se avizinhava em maio do ano seguinte, e, bem assim, «*as mais de 50 eleições presidenciais, legislativas e autárquicas/regionais que irão ter lugar nos Estados-Membros até 2020*»[159]. Este plano começa por reconhecer que «*combater a desinformação exigirá vontade política e uma intervenção comum*»[160], salientando o imprescindível caráter multinível da resposta aos desafios da desordem informativa. O plano de ação, apresentando uma «*estratégia estruturada para dar resposta a questões que exigem tanto esforços reativos (desacreditar e reduzir a visibilidade dos conteúdos de desinformação) como pró-ativos a mais longo prazo (literacia mediática e medidas para melhorar a resiliência da sociedade)*»[161], propõe um conjunto de medidas coordenadas assente em quatro pilares: [i.] melhorar a capacidade das instituições da União para detetar, analisar e denunciar a desinformação, [ii.] reforçar a coordenação e as respostas comuns, [iii.] mobilizar o setor privado e [iv.] sensibilizar os cidadãos e reforçar a resiliência da sociedade. No âmbito do primeiro pilar, prometeu-se aumentar os meios humanos, tecnológicos (incluindo a mobilização da inteligência artificial) e financeiros dedicados à deteção e investigação científica da desinformação. No domínio do segundo pilar, a principal proposta passou pela criação de um "sistema de alerta rápido" com o objetivo de, depois de se identificar num Estado-Membro uma campanha de desinformação, se partilhar essa informação com os demais Estados-Membro e as instituições, tudo em tempo real. O Sistema de Alerta Rápido (SAR) haveria de ser estabelecido em março de 2019, tendo sido utilizado pela primeira vez no âmbito da pandemia de Covid-19, em março de 2020[162]. Por referência ao terceiro e quarto pilares, previu-se apenas a aplicação do Código de Conduta e um conjunto de ações e campanhas de âmbito europeu e local para sensibilizar a sociedade civil para os perigos da desinformação.

Sem embargo, o "Plano de Ação contra a Desinformação" acabou por ser duramente criticado pelo Tribunal de Contas Europeu, que salienta que o plano não dispõe de «*mecanismos abrangentes destinados a garantir que qualquer*

[159] Comunicação Conjunta ao Parlamento Europeu, ao Conselho Europeu, ao Conselho, ao Comité Económico e Social Europeu e ao Comité das Regiões, *Plano de Ação contra a Desinformação*, Bruxelas, 5.12.2018, JOIN(2018) 36 final, p. 3.

[160] *Idem*, p. 5.

[161] Relatório Especial do Tribunal de Contas Europeu 09/2021, *Desinformação na UE: fenómeno combatido, mas não controlado*, 03/06/2021, disponível em: https://www.eca.europa.eu/pt/Pages/DocItem.aspx?did=58682, consult. em: 26/02/2023, p. 14.

[162] *Vide* "EU Rapid Alert System used amid coronavirus disinformation campaign", disponível em: https://www.euractiv.com/section/media/news/eu-alert-triggered-after-coronavirus-disinformation-campaign/, consult. em 21/01/2023.

resposta da UE à desinformação seja bem coordenada, eficaz e proporcional», inexistindo um «*quadro de acompanhamento, avaliação e comunicação de informações conexo ao plano de ação da UE, o que prejudica a responsabilização*»[163].

No rescaldo das eleições de maio para o PE, a Comissão e o Alto Representante, num relatório de junho de 2019 relativo à execução do "Plano de Ação contra a Desinformação", concluíram não ter existido uma campanha de desinformação sistemática e de grande escala com origem externa tendo por alvo as eleições europeias[164] [165]. Sem embargo, a investigação confirma que nos 28 processos nacionais simultâneos que compõem a eleição para o Parlamento Europeu foram detetadas atividades desinformativas com o objetivo de reduzir a taxa de participação e influenciar as preferências dos eleitores. Estas atividades abrangeram um vasto leque de temas, desde a contestação da legitimidade democrática da União à promoção de debates públicos fraturantes sobre as migrações e a soberania[166]. Todavia, a maioria destas ações foi associada a atores domésticos em lugar de agentes externos, não obstante as possíveis ligações existentes[167], o que, em todo o caso, não permitia às instituições e aos Estados-Membro baixar os braços na luta contra a desinformação, qualificada como «*um desafio a longo prazo que diz respeito a todos os segmentos das*

[163] Relatório Especial do Tribunal de Contas Europeu 09/2021, *Desinformação na UE: fenómeno combatido, mas não controlado*, 03/06/2021, disponível em: https://www.eca.europa.eu/pt/Pages/DocItem.aspx?did=58682, consult. em: 26/02/2023, p. 4.

[164] Comunicação conjunta ao Parlamento Europeu, ao Conselho Europeu, ao Conselho, ao Comité Económico e Social Europeu e ao Comité das Regiões, *Relatório sobre a execução do plano de ação contra a desinformação*, Bruxelas, 14.6.2019, JOIN(2019) 12 final, p. 3.

[165] A investigação efetuada é, todavia, menos animadora. O grupo de ativistas Avaaz encontrou mais de 500 páginas e grupos suspeitos no Facebook, os quais eram seguidos por 32 milhões de pessoas e geraram mais de 67 milhões de interações. Em conjunto, estas páginas e grupos tinham quase três vezes mais seguidores do que as páginas dos principais partidos europeus de extrema-direita ou antieuropeus. Segundo a investigação, as redes de desinformação removidas pelo Facebook em antecipação das eleições para o PE produziram conteúdos que foram vistos, estima-se, 533 milhões de vezes, numa média de 6 milhões de visualizações diárias. *Vide* Avaaz, *Far Right Networks of Deception, Avaaz Investigation Uncovers Flood of Disinformation, Triggering Shutdowns of Facebook Pages with over 500 Million Views Ahead of EU Elections*, 22.05.2019, disponível em: https://s3.amazonaws.com/avaazimages.avaaz.org/Networks_Report_Update_Page_July_2019.pdf, consult. em: 21/01/2023, p. 6 e ss.

[166] Oferecendo dezoito exemplos em Estados-Membro como a Polónia, a Alemanha ou Espanha, cfr. Chloe Colliver, *Click Here For Outrage: Disinformation in the European Parliamentary Elections 2019*, cit., p. 7 e ss.

[167] Judit Bayer; Bernd Holznagel; Katarzyna Lubianiec, *et al.*, *Disinformation and propaganda: impact on the functioning of the rule of law and democratic processes in the EU and its Member States – 2021 update*, cit., p. 31.

nossas sociedades e que requer um compromisso e esforços contínuos»[168]. A este respeito, e salientando o esforço das plataformas digitais, a Comissão e o Alto Representante adjetivaram como insuficientes as ações até aí adotadas, logo sinalizando que, após uma avaliação da eficácia do Código de Conduta nos primeiros 12 meses de aplicação, a Comissão poderia, caso os resultados não fossem satisfatórios, «*propor iniciativas suplementares, nomeadamente de caráter normativo*»[169], em sintonia, diga-se, com a posição do Parlamento Europeu, que, em outubro de 2019, solicitou à Comissão que estudasse «*possíveis medidas legislativas e não legislativas suscetíveis de levar as plataformas de redes sociais a intervir com o objetivo de rotular sistematicamente os conteúdos partilhados por robôs digitais, rever os algoritmos, para os tornar tão imparciais quanto possível, e encerrar as contas das pessoas implicadas em atividades ilegais destinadas a perturbar os processos democráticos ou a promover o discurso de ódio, sem que a liberdade de expressão seja comprometida*»[170].

A pandemia de Covid-19 veio, ao longo do ano de 2020, redobrar a necessidade de uma intervenção decisiva europeia no combate à desinformação. Não estava já primacialmente em causa a integridade dos processos democráticos, mas a preservação de um outro bem público: a saúde humana. Efetivamente, a desordem informativa, aproveitando o desconhecimento e a confusão que se gerou em torno de vírus SARS-CoV-2, levou a OMS a declarar, em março, a existência de uma "*infodemia*", caracterizada pela excessiva quantidade de informação disponível acerca da pandemia, muita dela nas redes sociais, misturando-se informação correta, incorreta, teorias da conspiração, rumores, boatos e desinformação, de tal modo que se tornava difícil ao cidadão comum encontrar dados fidedignos sobre o vírus[171]. Este estado de coisas colocava em causa a resposta da UE e dos Estados-Membro em matéria de saúde pública, situação que se haveria de agravar posteriormente com a disponibilização das vacinas[172], tema que já era fértil para a desinformação.

[168] Comunicação conjunta ao Parlamento Europeu, ao Conselho Europeu, ao Conselho, ao Comité Económico e Social Europeu e ao Comité das Regiões, *Relatório sobre a execução do plano de ação contra a desinformação*, cit., p. 1.

[169] *Idem*, p. 6.

[170] Resolução do Parlamento Europeu, de 10 de outubro de 2019, sobre a interferência eleitoral e a desinformação nos processos democráticos nacionais e europeus (2019/2810 (RSP)).

[171] Organização Mundial da Saúde, *Coronavirus disease 2019 (COVID-19) Situation Report – 45*, 05.03.2020, disponível em: https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200305-sitrep-45-covid-19.pdf?sfvrsn=ed2ba78b_4, consult. em: 21/01/2023, p. 2.

[172] Cfr., *inter alia*, Sun Kyong Lee; Juhyung Sun; Seulki Jang, Et Al., *Misinformation of COVID-19 vaccines and vaccine hesitancy*, Scientific Reports, 12, nº 13681, 2022. Os Autores deste estudo concluem existir evidência que suporta que a exposição a informação incorreta associada

Neste contexto, os chefes de Estado e de governo dos Estados-Membro, reunidos informalmente em março, comprometeram-se a combater «*de forma resoluta a desinformação mediante uma comunicação transparente, atempada e baseada em factos*»[173], instando a Comissão e o Alto Representante a adotarem as medidas necessárias a esse fim[174]. Volvidos três meses, a Comissão e o Alto Representante apresentaram a comunicação «Combater a desinformação sobre a COVID-19: repor a verdade dos factos», da qual resulta dois aspetos relevantes: em primeiro lugar, a comunicação aponta claramente para a necessidade de se reforçar o Código de Conduta da UE sobre a Desinformação[175], atentas as suas percecionadas insuficiências, pedindo-se às plataformas digitais signatárias que apresentassem pormenorizados relatórios mensais sobre as suas estratégias e medidas para combater a desinformação sobre a COVID-19; em segundo lugar, e com maior relevância, a comunicação inaugura um novo quadro de compreensão sobre o fenómeno da desinformação, entendendo-o de forma verdadeiramente multidimensional e transversal, na esteira do que havia já sido proposta pelo HLEG dois anos antes. Assim, além do repetidíssimo e mais ou menos inconsequente apelo ao reforço da cooperação entre Estados-Membro e instituições da União (a que se somam, agora, os países terceiros e os parceiros internacionais), a comunicação vem associar o tema de desinformação a outros problemas, políticas e medidas adotadas ou a adotar pela UE, designadamente em torno do reforço da comunicação estratégica, da crise do Estado de Direito e dos *media* tradicionais, da educação digital, da literacia mediática e digital, e da defesa do consumidor[176]. Em todo o caso, é

à crença de que é verdadeira pode aumentar a hesitação no toma das vacinas, reduzindo a intenção comportamental das pessoas se vacinarem.

[173] Declaração comum dos membros do Conselho Europeu, Bruxelas, 26.03.2020, disponível em: https://www.consilium.europa.eu/media/43094/26-vc-euco-statement-pt.pdf, consult. em: 21/01/2023, p. 2.

[174] Apelo repetido pelas demais instituições da União, designadamente o Conselho e o Parlamento Europeu, em abril de 2020. Cfr. https://www.consilium.europa.eu/pt/meetings/fac/2020/04/03/ e Resolução do Parlamento Europeu, de 17 de abril de 2020, sobre a ação coordenada da UE para combater a pandemia de COVID-19 e as suas consequências (2020/2616 (RSP) JO C 316).

[175] Comunicação Conjunta do Parlamento Europeu, do Conselho Europeu, do Conselho, do Comité Económico e Social e do Comité das Regiões, *Combater a desinformação sobre a COVID-19: repor a verdade dos factos*, Bruxelas, 10.6.2020, JOIN(2020) 8 final, p. 9.

[176] Efetivamente, durante a crise pandémica multiplicaram-se nas plataformas digitais os anúncios publicitários enganosos a produtos ilegais ou perigosos, especificamente produtos com alegadas e falsas propriedades curativas ou preventivas contra a COVID-19. A pedido da Comissão, e na sequência de uma inspeção pela Rede de Cooperação no domínio da Defesa do Consumidor, várias plataformas digitais removeram milhões de anúncios publicitários e ofertas comerciais. Cfr. Comunicação Conjunta do Parlamento Europeu, do Conselho Europeu, do Conselho, do Comité

preciso notar que as plataformas digitais, com maior ou menor rapidez, adotaram novos instrumentos para promover e destacar informações fidedignas e oficiais sobre a Covid-19, tais como painéis de informação, cartões, janelas instantâneas, mapas ou mensagens que remetiam os utilizadores para fontes de informação documentadas nas diferentes línguas[177].

Em especial no que tange ao Estado de Direito, a preocupação das instituições justificava-se com a circunstância de as medidas de emergência e de exceção constitucional adotadas um pouco por toda a Europa, pudessem servir como pretexto para restringir abusiva e ilegitimamente os direitos, liberdades e garantias dos cidadãos e das organizações, especialmente a liberdade de expressão, de opinião e de informação[178]. Esses receios vieram a provar-se exatos[179] pelo menos no caso da Hungria, que, durante o estado de exceção constitucional, aprovou um conjunto de alterações legislativas relevantes (incluindo à Constituição), de entre as quais figurava a criminalização da disseminação de desinformação (através da utilização de conceitos vagos e ambíguos como informação "distorcida") durante o estado de emergência, o que causou «*incerteza e autocensura*» à comunicação social e outros agentes[180].

No fim de junho de 2020 foi criado o Observatório Europeu dos Meios de Comunicação Digital (EDMO), uma organização independente financiada pela UE que junta verificadores de factos, investigadores, plataformas digitais, jornalistas e outros atores relevantes, tudo com o objetivo de colaborar na identificação e deteção da desinformação e das suas fontes, a diluir o seu impacto, a apoiar a verificação de factos e a informação de qualidade e a congregar na mesma plataforma as comunidades de peritos e especialistas em desinformação[181].

Económico e Social e do Comité das Regiões, *Combater a desinformação sobre a COVID-19: repor a verdade dos factos*, cit., p. 16.

[177] *Idem*, p. 9.

[178] *Idem*, p. 12.

[179] Joelle Grogan, *Impact of COVID-19 measures on democracy and fundamental rights,* European Parliament, 2022, disponível em: https://tinyurl.com/impactCovid19Rights, consult. em: 22/01/2023. *Vide*, ainda, Comunicação da Comissão ao Parlamento Europeu, ao Conselho, ao Comité Económico e Social Europeu e ao Comité das Regiões, *Relatório de 2021 sobre o Estado de direito. Situação na União Europeia*. Bruxelas, 20.7.2021, COM(2021) 700 final.

[180] Resolução do Parlamento Europeu, de 15 de setembro de 2022, sobre o projeto de decisão do Conselho relativa à verificação, nos termos do artigo 7º, nº 1, do Tratado da União Europeia, da existência de um risco manifesto de violação grave, pela Hungria, dos valores em que a União se funda (2018/0902R(NLE).

[181] Cfr. https://edmo.eu/

Em outubro de 2019, as plataformas digitais entregaram à Comissão os seus primeiros relatórios anuais relativos à aplicação do Código de Conduta da UE sobre a Desinformação. Esses relatórios foram analisados pela ERGA e por uma consultora independente contratada pela Comissão, tendo ambas colocado em evidência as insuficiências do Código decorrentes da sua natureza autorregulatória[182]. A consultora independente destaca que a mera existência do Código é já uma conquista, tendo conduzido as plataformas digitais a uma maior proatividade, o que, em todo o caso, não implica que o Código não tivesse de ser reforçado, nomeadamente no que toca à precisão das definições, à harmonização das práticas (por exemplo, os relatórios de autoavaliação variam de plataforma para plataforma, tornando difícil estabelecer comparações), à abrangência dos compromissos e do número de signatários e à criação de um mecanismo de ação em caso de não-cumprimento[183]. A avaliação da Comissão foi no mesmo sentido, determinando que o Código deveria ser melhorado em vários aspetos de modo a ser aplicado de forma completa e consistente pelas plataformas digitais, através de compromissos claros sujeitos a mecanismos de supervisão adequados[184].

No contexto da pandemia, a partir de agosto de 2020 e até julho de 2021, na sequência da comunicação «*Combater a desinformação sobre a COVID-19: repor a verdade dos factos*», as plataformas digitais submetem mensalmente à Comissão relatórios pormenorizados explicando as ações adotadas para limitar a difusão de desinformação relacionada com a crise pandémica. A submissão desses relatórios evidenciou, do ponto de vista da Comissão, duas questões cuja resolução se impunha: por um lado, faltava um modelo uniforme de comunicação de resultados pelas plataformas digitais, tornando difícil a comparação entre períodos temporais e entre plataformas; por outro lado, os dados disponibilizados, nomeadamente os quantitativos, não eram detalhados o suficiente, nomeadamente através de uma segregação por Estado-Membro[185].

[182] *Vide* nota de rodapé [119].

[183] Iva Plasilova; Jordan Hill; Malin Carlberg , et al., *Study for the "Assessment of the implementation of the Code of Practice on Disinformation"*. European Commission, 2020, disponível em: https://digital-strategy.ec.europa.eu/en/library/study-assessment-implementation-code-practice-disinformation, consult em: 15/01/2023, p. 32 e ss.

[184] Commission Staff Working Document, *Assessment of the Code of Practice on Disinformation – Achievements and areas for further improvement*, Bruxelas, 10.09.2020, SWD(2020) 180 final, p. 19.

[185] Cfr., *inter alia*, o comunicado de imprensa da Comissão de 06/11/2020, disponível em https://digital-strategy.ec.europa.eu/en/library/third-set-reports-fighting-covid-19-disinformation-monitoring-programme, consult. em: 21/01/2023.

Em face das críticas e insuficiências identificadas relativamente às ações em curso, a Comissão procurou, consistentemente, desenvolver novos instrumentos de resposta política e jurídica, sempre com enfoque numa estratégia multinível, tendente à coordenação entre Estados-Membro (primacialmente competentes nestas matérias) e a União Europeia. Em dezembro de 2020, a Comissão publica dois planos de ação com medidas destinadas a evitar a proliferação descontrolada da desinformação: a comunicação «*Os meios de comunicação social da Europa na Década Digital: plano de ação para a apoiar a recuperação e a transformação*»[186] e a comunicação «*sobre o plano de ação para a democracia europeia*»[187]. A primeira comunicação, centrada no ecossistema dos meios de informação e audiovisuais, destacava o "sentido de urgência" numa resposta política forte para evitar a tendência do enfraquecimento da diversidade cultural e do pluralismo dos meios de comunicação social na Europa, tendência essa agravada pela situação pandémica, mas que já resultava de circunstâncias anteriores. Neste contexto, e no que respeita ao combate à desinformação, o plano de ação para os meios de comunicação social foi elaborado numa perspetiva de articulação com o plano de ação para a democracia europeia, designadamente na sua vertente de capacitação dos cidadãos através da literacia mediática[188], prevendo-se, no primeiro, a efetiva aplicação da Diretiva Serviços de Comunicação Social Audiovisual[189] no que tange à promoção de medidas para desenvolver as competências de literacia mediática a cargo dos Estados-Membro (artº 33º-A) e das plataformas de partilha de vídeos (artº 28º-B)[190].

[186] Comunicação da Comissão ao Parlamento Europeu, ao Conselho, ao Comité Económico e Social Europeu e ao Comité das Regiões, *Os meios de comunicação social da Europa na Década Digital: plano de ação para apoiar a recuperação e a transformação*. Bruxelas, 3.12.2020, COM(2020) 784 final.

[187] Comunicação da Comissão ao Parlamento Europeu, ao Conselho, ao Comité Económico e Social Europeu e ao Comité das Regiões, *sobre o plano de ação para a democracia europeia*. Bruxelas, 3.12.2020, COM(2020) 790 final.

[188] Nos termos do plano de ação, a «*literacia mediática inclui todas as capacidades técnicas, cognitivas, sociais, cívicas e criativas que permitem aos cidadãos aceder aos meios de comunicação social, ter uma compreensão crítica dos mesmos e interagir com eles*». Comunicação da Comissão ao Parlamento Europeu, ao Conselho, ao Comité Económico e Social Europeu e ao Comité das Regiões, *Os meios de comunicação social da Europa na Década Digital: plano de ação para apoiar a recuperação e a transformação*, cit., p. 21.

[189] Diretiva 2010/12/UE do Parlamento Europeu e do Conselho de 10 de março de 2010 relativa à coordenação de certas disposições legislativas, regulamentares e administrativas dos Estados-Membros respeitantes à oferta de serviços de comunicação social audiovisual (Diretiva «Serviços de Comunicação Social Audiovisual»), conforme alterada pela Diretiva (UE) 2018/1808 do Parlamento Europeu e do Conselho de 14 de novembro de 2018.

[190] A importância de um projeto efetivo e eficaz de literacia digital havia já sido evidenciado na Comunicação da Comissão ao Parlamento Europeu, ao Conselho, ao Comité Económico e Social

Por sua vez, a Comunicação «*sobre o plano de ação para a democracia europeia*» surge no contexto das Orientações Políticas da Comissão Europeia 2019-2014[191], onde Ursula von der Leyen se comprometeu a apresentar aquele plano e a «*tomar mais medidas*» para proteger a UE de «*interferências externas*», incluindo medidas para evitar que as plataformas digitais sejam utilizadas para «*desestabilizar as* [...] *democracias*», nomeadamente através da desinformação e do discurso de ódio em linha[192]. O Plano de Ação para a Democracia Europeia começa, justamente, por reconhecer que a União se funda na democracia, no Estado de Direito e nos direitos fundamentais, sendo a primeira um valor fulcral que não deve ser dado como garantido, antes dependendo de promoção e defesa[193]. Depois, o Plano de Ação expõe em termos paradoxais as dificuldades com que se confronta a UE, ao salientar que «*as próprias liberdades que procuramos defender, como a liberdade de expressão, foram usadas, em alguns casos, para enganar e manipular*»[194]. Procurando reforçar a resiliência das democracias europeias perante os desafios atuais, o Plano coloca particular enfoque na "transformação digital" das próprias democracias. Conforme sintetiza a Comissão, «*a revolução digital transformou a política democrática*»[195], desde as campanhas eleitorais em linha às novas formas de participação política e cívica, o que, se por um lado veio facilitar o acesso à vida pública e ao debate democrático a alguns grupos (especialmente os mais jovens), trouxe, por outro lado, um conjunto de novas vulnerabilidades e preocupações em torno da integridade das eleições, da liberdade e pluralidade das opiniões e dos meios de comunicação social, e a própria integridade da informação que circula online.

Europeu e ao Comité das Regiões, *Plano de Ação para a Educação Digital 2021-2027. Reconfigurar a educação e a formação para a era digital*. Bruxelas, 30.09.2020, COM(2020) 624 final. Nesta comunicação, a Comissão propôs-se a constituir um grupo pluridisciplinar para a elaboração de orientações comuns para os professores e o pessoal educativo no sentido de promover a literacia digital e combater a desinformação através da educação e da formação. No obstante a previsão de que as orientações estivessem concluídas para serem apresentadas conjuntamente com o plano de ação para os meios de comunicação social, a verdade é que só em 2022 foram publicadas. As orientações estão disponíveis em: https://op.europa.eu/pt/publication-detail/-/publication/a224c235-4843-11ed-92ed-01aa75ed71a1, consult. em: 27/02/2023.

[191] Orientações Políticas para a Próxima Comissão Europeia 2019-2024, *Uma União mais ambiciosa. O meu programa para a Europa pela candidata à função de Presidente da Comissão Europeia Ursula von der Leyen,* disponível em: https://commission.europa.eu/system/files/2020-04/political-guidelines-next-commission_pt.pdf, consult em: 05/03/2023.

[192] *Idem*, p. 25.

[193] Comunicação da Comissão ao Parlamento Europeu, ao Conselho, ao Comité Económico e Social Europeu e ao Comité das Regiões, *sobre o plano de ação para a democracia europeia*, cit., p. 1.

[194] *Idem*.

[195] *Idem*.

Tendo em vista esse objetivo de tornar mais resilientes o espaço democrático europeu, a Comissão propôs-se adotar medidas tendentes a três grandes objetivos de política pública: (1) proteger a integridade das eleições e promover a participação democrática, (2) reforçar a liberdade e o pluralismo dos meios de comunicação social e (3) combater a desinformação. No que tange ao primeiro objetivo, destaca-se a promessa da Comissão, entretanto cumprida, de apresentar nova legislação para garantir maior transparência da publicidade temática de cariz político[196] e de propor a revisão do Regulamento relativo ao estatuto e ao financiamento dos partidos políticos europeus e das funções políticas europeias[197]. O segundo objetivo, relativo à liberdade e pluralismo dos meios de comunicação social na UE, tem de ser lido e interpretado conjuntamente com o outro plano de ação destinado a recuperar os *media* europeus em face da pandemia, passando por melhorar a segurança dos jornalistas, combater a utilização abusiva de ações judiciais estratégicas contra a participação pública (as "SLAPP") e reforçar a cooperação para implementar normas deontológicas.

Finalmente, no âmbito da desinformação, o Plano de Ação para a Democracia Europeia forneceu um novo esquema concetual para a "desordem informativa", como antes já explicitado. Distinguindo entre informação incorreta, desinformação, tentativas de exercer influência sobre a informação e interferência estrangeira no espaço de informação, o Plano conclui pela necessidade de respostas específicas para cada um destes fenómenos. Sem prejuízo, as várias ações propostas pela Comissão neste Plano visavam «*evitar a amplificação manipuladora de conteúdo nocivo, aumentando a transparência, restringindo técnicas manipuladoras e reduzindo os incentivos económicos à propagação da desinformação, bem como estabelecer desincentivos, impondo custos aos responsáveis por tentativas de exercer influência e aos agentes de interferência estrangeira*»[198]. Para tal, a Comissão propôs ações em três eixos: (1) melhorar a capacidade da UE e dos Estados-Membro para combater a desinformação, (2) criar mais obrigações para as plataformas digitais e responsabilizá-las, e (3) capacitar os cidadãos para tomarem decisões informadas.

[196] Proposta de Regulamento do Parlamento Europeu e do Conselho sobre a transparência e o direcionamento da propaganda política. Bruxelas, 25.11.2021, COM(2021) 731 final.

[197] Proposta de regulamento do Parlamento Europeu e do Conselho relativo ao estatuto e ao financiamento dos partidos políticos europeus e das fundações políticas europeias (reformulação). COM(2021) 734 final.

[198] Comunicação da Comissão ao Parlamento Europeu, ao Conselho, ao Comité Económico e Social Europeu e ao Comité das Regiões, *sobre o plano de ação para a democracia europeia*, cit., p. 22.

Quanto ao primeiro eixo, e identificando os riscos da "desordem informativa" para a democracia europeia e para o Estado de Direito, o que não passa apenas pela desinformação em sentido estrito, mas também por outras técnicas manipuladoras destinadas a induzir as audiências em erro, a Comissão propõe um conjunto de ações destinadas a tornar mais estreita a cooperação internacional e na UE (entre as instituições[199] e entre os Estados-Membro) no combate ao fenómeno desinformativo, aproveitando-se as estruturas já existentes (Sistema de Alerta Rápido, Rede Europeia de Cooperação para as Eleições, Agência da UE para a Cibersegurança, Observatório Europeu dos Meios de Comunicação Digital, etc.). Além disso, propôs-se explorar novos instrumentos (não concretamente identificados) que visem neutralizar as técnicas de desinformação frequentemente utilizadas, além de impor custos aos seus responsáveis.

No que toca ao segundo eixo, a Comissão aponta já aí um conjunto de questões e propostas de regulamentação que haveriam de ser vertidas para o Regulamento dos Serviços Digitais. Efetivamente, sobressai do Plano de Ação o objetivo de aumentar a transparência da atuação das plataformas digitais, principalmente as de grande dimensão, designadamente no que respeita à moderação de conteúdos, à utilização de algoritmos e sistemas de recomendação e personalização e, em geral, aos processos algorítmicos e à publicidade em linha. Ainda relevante, a Comissão sinaliza a necessidade de revisão do Código de Conduta da UE sobre a Desinformação, através de uma «*abordagem mais robusta baseada em compromissos claros e sujeita a mecanismos de supervisão adequados*»[200], comprometendo-se a emitir orientações para o reforço desse Código na primavera de 2021.

Por último, quanto ao último eixo, a Comissão propôs-se a intensificar os seus esforços para reforçar a literacia mediática, numa perspetiva multidisciplinar, designadamente apoiando novos projetos ao abrigo de vários programas da UE[201], em especial projetos de organização da sociedade civil e de estabelecimentos de ensino superior, com a participação de jornalistas.

[199] No plano interno, propôs-se a criação de um protocolo de ação interinstitucional que combine conhecimentos e recursos em resposta a situações específicas, por exemplo na preparação das eleições para o Parlamento Europeu.

[200] Comunicação da Comissão ao Parlamento Europeu, ao Conselho, ao Comité Económico e Social Europeu e ao Comité das Regiões, *sobre o plano de ação para a democracia europeia*, cit., p. 25.

[201] É o caso do programa "Media Literacy for All", que financiou vários projetos no âmbito da literacia mediática, tais como o "SMarT-EU", o "Influencers Trust Laber" e o "FREEYOU". Foi também o caso da Semana Europeia para a Literacia Mediática, que se realizou entre 30 de março e 5 de abril de 2020. Cfr. https://digital-strategy.ec.europa.eu/pt/node/1503, consult. em: 07/03/2023.

Em abril de 2021, a Comissão organizou uma reunião entre as diversas partes interessadas para discutir as formas de reforçar o Código de Conduta da UE sobre a Desinformação, em jeito de preparação para as orientações que haveria de apresentar pouco depois[202]. Efetivamente, a 26 de maio de 2021, a Comissão apresenta as suas orientações relativas ao reforço do Código de Conduta[203], aliás na esteira dos objetivos a que já se havia proposta no âmbito do Plano de Ação para a Democracia Europeia. As orientações fornecidas pela Comissão eram abrangentes e completas, procurando não somente corrigir as insuficiências do Código, mas transformá-lo num verdadeiro mecanismo de autorregulação «*forte, estável e flexível, que torne as plataformas em linha mais transparentes e responsáveis desde a conceção*»[204]. Aproveitando a experiência obtida com as medidas adotadas no quadro da pandemia de Covid-19[205], a Comissão abordou estruturalmente as lacunas do Código de Conduta, nomeadamente a ausência de uniformidade e qualidade dos relatórios submetidos pelas plataformas digitais, a ausência de indicadores-chave de desempenho (ICD), a falta de avaliação independente da eficácia do Código e das medidas adotadas pelas plataformas, a falta de suficiente cobertura da verificação de factos, detetando-se que conteúdos que já haviam sido sinalizados como desinformativos voltam a surgir noutras plataformas, e a incapacidade de desmonetizar os conteúdos associados à desinformação, nomeadamente pela colocação de anúncios publicitários[206].

Numa palavra, a Comissão procurou promover um consenso entre as partes interessadas e já signatárias do Código de Conduta de modo garantir a sua eficácia e robustez. Para tanto, recorreu às iniciativas legislativas que já tinha em calha, principalmente o Regulamento dos Serviços Digitais[207].

[202] Cfr. *Summary of the Multi-stakeholder discussions in preparation of the Guidance to strengthen the Code of Practice on Disinformation*, disponível em: https://digital-strategy.ec.europa.eu/en/library/summary-multi-stakeholder-discussions-preparation-guidance-strengthen-code-practice-disinformation, consult. em: 07/03/2023.

[203] Comunicação da Comissão ao Parlamento Europeu, ao Conselho, ao Comité Económico e Social Europeu e ao Comité das Regiões, *Orientações da Comissão Europeia relativas ao reforço do Código de Conduta sobre Desinformação*, Bruxelas, 26.5.2021, COM(2021) 262 final.

[204] *Idem*, p. 3.

[205] Comunicação Conjunta do Parlamento Europeu, do Conselho Europeu, do Conselho, do Comité Económico e Social e do Comité das Regiões, *Combater a desinformação sobre a COVID-19: repor a verdade dos factos*, cit.

[206] Comunicação da Comissão ao Parlamento Europeu, ao Conselho, ao Comité Económico e Social Europeu e ao Comité das Regiões, *Orientações da Comissão Europeia relativas ao reforço do Código de Conduta sobre Desinformação*, cit., p. 4-5.

[207] Proposta de Regulamento do Parlamento Europeu e do Conselho relativo a um mercado único de serviços digitais (Regulamento Serviços Digitais) e que altera a Diretiva 2000/31/CE, Bruxelas, 15.12.2020, COM(2020) 825 final.

Realmente, e como se verá, o Regulamento dos Serviços Digitais (RSD) cria obrigações a cargo das plataformas digitais de muito grande dimensão para que identifiquem os riscos sistémicos associados aos seus serviços e adotem medidas de mitigação, podendo tais medidas passar, *inter alia*, pela elaboração ou adesão a Códigos de Conduta de natureza autorregulatória. Ora, em face da discussão desta proposta de Regulamento, era do interesse das plataformas digitais rever o Código de Conduta sobre a Desinformação de modo a conformá-lo às futuras obrigações resultantes do RSD (*vide*, em especial, o artº 35º e 45º do RSD). Assim, as orientações da Comissão para o reforço do Código de Conduta da UE sobre a Desinformação surgem já num contexto de clara articulação com a proposta de RSD, o que porventura justifique a ambição das alterações propostas e a disponibilidade das plataformas digitais para aderirem à nova versão "forte" do Código.

Em termos substanciais, o destaque das orientações da Comissão vai para, do ponto de vista transversal, o alargamento do âmbito da Código de Conduta, i.e., para a adoção de um conceito amplo de desinformação, em linha com o adotado no Plano de Ação para a Democracia Europeia, e para o alargamento subjetivo do Código a mais plataformas, nomeadamente às emergentes (ainda que de pequena dimensão), aos serviços de mensagens encriptados e às entidades ligadas ao ecossistema da publicidade digital. Depois, a Comissão solicitou a substituição dos compromissos vagos e genéricos que marcaram a primeira versão do Código de Conduta por compromissos específicos e, sobretudo, mensuráveis de modo uniforme entre plataformas, através da implementação de IDC de duas ordens: indicadores do nível de serviço[208], tanto quantitativos como qualitativos, e indicadores estruturais[209]. Além disso, propôs um quadro permanente de acompanhamento da aplicação do Código e da sua adaptação à evolução do fenómeno desinformativo.

Apresentadas as orientações, a Comissão instou os signatários do Código de Conduta da UE sobre a Desinformação a, até ao outono de 2021, apresen-

[208] Estes indicadores destinam-se a medir a aplicação dos compromissos do código e o impacto concreto dessa aplicação nas plataformas. Alguns exemplos avançados pelas orientações: número de verificações de factos, número de recursos relativos a medidas adotadas pelas plataformas digitais em relação a conteúdos, número de páginas, contas ou perfis que partilham desinformação, etc. Cfr. Comunicação da Comissão ao Parlamento Europeu, ao Conselho, ao Comité Económico e Social Europeu e ao Comité das Regiões, *Orientações da Comissão Europeia relativas ao reforço do Código de Conduta sobre Desinformação*, cit., p. 24-5.

[209] Os indicadores estruturais, nos termos das orientações, visam medir o impacto global do Código no fenómeno da desinformação, podendo, por exemplo, basear-se em amostras representativas dos utilizadores de vários Estados-Membro para aferir a exposição dos cidadãos à desinformação. *Idem*.

tarem uma proposta de revisão do Código para posterior debate. O processo de elaboração da proposta de revisão acabou por se prolongar no tempo, na medida em que a discussão foi alargada a vinte e seis possíveis novos signatários[210].

Entretanto, em dezembro de 2021, a Comissão Europeia publicou o seu relatório anual sobre a aplicação da Carta dos Direitos Fundamentais da União Europeia, relatório que naquele ano se focou na proteção dos direitos fundamentais na era digital[211]. Aí se destacou o papel das plataformas digitais na moderação de conteúdos que não sendo ilegais e estando, por isso, protegidos pela liberdade expressão, podem causar significativos danos ou disrupções aos processos democráticos, à autodeterminação informacional, ao debate público e a outros bens e interesses públicos, como a saúde e a segurança. Efetivamente, um espaço público onde predomine a manipulação e a desinformação tende para a hiperpolarização e para a desconfiança generalizada nas instituições e nos processos democráticos[212]. Destacando, na esteira da jurisprudência do TEDH, que, em matéria de liberdade de expressão, os Estados têm não só obrigações negativas (de não interferência e censura), mas também obrigações positivas de promoção de uma esfera pública pluralista, inclusiva e favorável ao debate de ideias opostas sobre o bem comum, o relatório conclui que as plataformas digitais adotam medidas com potencial de restrição dos direitos fundamentais dos cidadãos (nomeadamente a sua liberdade de expressão e de opinião, mas também a liberdade de informação) sem terem de ter em conta o interesse público e sem as correspondentes salvaguardas procedimentais ou mecanismos de responsabilização que, em geral, impendem sobre os atores públicos[213]. Neste enquadramento, o relatório salienta que a luta contra a desinformação não pode, em todo o caso, sacrificar os direitos e as liberdades fundamentais protegidas pela Carta, nomeadamente através de fenómenos de censura.

[210] Cfr. https://digital-strategy.ec.europa.eu/en/news/revision-code-practice-strengthened-code-expected-march-2022, consult. em: 07/03/2023.

[211] Report from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, *Protecting Fundamental Rights in the Digital Age – 2021 Annual Report on the Application of the EU Charter of Fundamental Rights*, Brussels, 10.12.2021, COM(2021) 819 final.

[212] *Idem*, p. 9.

[213] O exercício "delegado" ou "autónomo" em linha de "poderes quasi públicos" por parte das plataformas digitais está, aliás, na base do "constitucionalismo digital" como reação e estratégia normativa à consolidação dos novos poderes digitais privados. Para mais, cfr. Giovanni de Gregorio, *Digital Constitutionalism in Europe: Reframing Rights and Powers in the Algorithmic Society*, Cambridge, Cambridge University Press, 2022, p. 95 e ss.

Em contraciclo com esta posição de princípio, e na sequência da invasão da Ucrânia pela Rússia e do subsequente conflito, o Conselho da União Europeia decidiu, em março de 2022, proibir temporariamente a difusão e transmissão por quaisquer meios (cabo, satélite ou internet) dos órgãos de comunicação social associados ao grupo "RT" e "Sputnik", financiados quase integralmente pela Federação da Rússia[214]. Nos considerandos dos atos jurídicos adotados, o Conselho justifica a sua decisão apelando não só à agressão e conflito em curso, como também às sucessivas e sistemáticas campanhas internacionais «*de manipulação dos meios de comunicação social e de distorção dos factos*»[215], enquadrando essas campanhas no âmbito das ameaças híbridas, as quais visam, na perspetiva do Conselho, os partidos políticos europeus, em especial nos períodos eleitorais, os requerentes de asilo, as minorias étnicas da Rússia, as minorias de género e o funcionamento das instituições democráticas da União e dos Estados-Membro. No quadro da guerra na Ucrânia, o Conselho acusa a Rússia de promover ações coordenadas de "propaganda" dirigidas à sociedade civil da União e aos seus países vizinhos, de modo a «*justificar e apoiar a agressão*»[216] em curso, através da distorção e grave manipulação dos factos.

Não obstante o Conselho referir que a proibição da radiodifusão dos órgãos de comunicação social associados ou financiados pelo Estado russo, quando associada à luta contra a desinformação e em defesa dos valores europeus, é compatível com os valores da União e com a CDFUE, nomeadamente com a liberdade de expressão de informação (artº 11º), tal asserção não é isenta de dúvidas, revelando, aliás, uma contradição insanável nas políticas da UE em torno deste tema. Com efeito, tem sido posição constante da UE o reconhecimento de que a desinformação não corresponde, em si mesma, a um conteúdo ilegal, estando antes protegida pela liberdade de expressão[217]. Note-se

[214] Regulamento (UE) 2022/350 do Conselho, de 1 de março de 2022, que altera o Regulamento (UE) nº 833/2014 que impõe medidas restritivas tendo em conta as ações da Rússia que desestabilizam a situação na Ucrânia. Cfr., ainda, Decisão (PESC) 2022/351 do Conselho, de 1 de março de 2022, que altera a Decisão 2014/512/PESC que impõe medidas restritivas tendo em conta as ações da Rússia que desestabilizam a situação na Ucrânia. JO L 65 de 2.3.2022.

[215] *Idem*, considerando 6.

[216] *Idem*, considerando 7.

[217] Como a jurisprudência tem reconhecido repetidamente, a liberdade de expressão constitui um dos «*fundamentos essenciais de uma sociedade democrática e uma das condições primordiais do seu progresso e do desenvolvimento individual e que, em princípio, protege não só as "informações" ou "ideias" favoravelmente acolhidas ou vistas como inofensivas ou indiferentes mas também as que ofendem, choquem ou inquietem, e isto a fim de garantir o pluralismo, a tolerância e o espírito de abertura sem os quais não existe sociedade democrática*» (Acórdão do Tribunal Geral (9ª Secção) de 15.06.2017, *Kiselev v. Conselho*, T-262/15, § 90),

que nos seus considerandos, o Conselho não se refere diretamente a campanhas de desinformação (em sentido estrito) promovidas pela federação russa, colocando antes o enfoque na "propaganda" e na "manipulação" e "distorção" dos factos, ambos, em princípio, protegidos pela liberdade de expressão e de informação. Aliás, é difícil conciliar, numa perspetiva estrita de combate à desinformação, a decisão do Conselho com as posições das demais instituições, nomeadamente da Comissão. Desde o início que a «*abordagem da UE em matéria de combate à desinformação baseou-se na proteção da liberdade de expressão e de outros direitos e liberdades garantidos pela Carta dos Direitos Fundamentais da UE. Em consonância com esses direitos e liberdades, em vez de criminalizar ou de proibir a desinformação como tal, a estratégia da UE visa tornar o ambiente em linha e os respetivos intervenientes mais transparentes e responsáveis,* [...] *promovendo um debate democrático aberto*»[218]. Poder-se-ia julgar que, no contexto destas proibições, a que em junho e dezembro de 2022 se vieram juntar outros órgãos de comunicação social (Rossiya RTR/RTR Planeta, Rossiya 24/Russia 24, TV Centre International, NTV/NTV Mir, Rossiya 1, REN TV e Pervyi Kanal), estaríamos perante uma inflexão na estratégia da UE, desta vez orientada para o proibicionismo e a restrição da liberdade de expressão e de informação. Não cremos, porém, que assim seja. Por um lado, o Código de Conduta Reforçado da UE sobre a Desinformação e o Regulamento dos Serviços Digitais não confirmam essa alteração estratégica. Por outro lado, parece-nos que a decisão do Conselho, apesar de também surgir justificada em abstrato no quadro da luta contra a desinformação, tem sobretudo (ou em exclusivo) subjacente a aplicação de sanções à Rússia pela invasão da Ucrânia, o que é corroborado pela base jurídica com que foi adotada, i.e., o artº 29º do TUE, referente à Política Externa e de Segurança Comum (PESC), sendo certo que a possibilidade de estabelecer derrogações à liberdade de receção e retransmissão dos serviços de comunicação social audiovisual está já regulada na Diretiva Serviços de Comunicação Social Audiovisual, pelo que bem podia ter sido ter sido essa a base jurídica escolhida, como já o fora no passado (cfr. Ac. *Kiselev v. Conselho*, T-262/15), o que neste caso não sucedeu.

sendo que, no âmbito do discurso político e de questões de interesse público, as restrições à liberdade de expressão devem ser ainda mais exigentes, estando esse discurso sujeito a uma «*forte proteção*» (*idem*, § 91). O TEDH vem também consistentemente referindo que a especial proteção de que goza o debate político e de assuntos de interesse geral só cede perante formas de expressão que promovam ou tentem justificar a violência, o ódio, a xenofobia ou outra forma de intolerância (Acórdão do TEDH no caso nº 27510/08, *PERINÇEK v. SWITZERLAND*, 15/09/2015, § 230).

[218] Comunicação da Comissão ao Parlamento Europeu, ao Conselho, ao Comité Económico e Social Europeu e ao Comité das Regiões, Orientações da Comissão Europeia relativas ao reforço do Código de Conduta sobre Desinformação, cit., p. 1.

Deste modo, a justificação e a legalidade da ação da União não se encontra tanto na luta contra a desinformação, mas, isso sim, na interdição da propaganda em favor da guerra, prevista no nº 1 do artº 20º do Pacto Internacional sobre os Direitos Civis e Políticos, o qual constitui uma norma de direito internacional geral ou comum, integrando, ainda, a tradição constitucional comum dos Estados-Membro, podendo, por isso, considerar-se um princípio geral de direito da União[219]. O Tribunal Geral (TG) teve já oportunidade de se debruçar sobre a legalidade dos atos jurídicos adotados pelo Conselho, os quais passaram no crivo dos juízes do Planalto de Kirchberg[220]. Na sua apreciação, o TG conclui que o «*direito dos meios de comunicação social e, mais especificamente, dos jornalistas de comunicar informações sobre questões de interesse geral é protegido desde que atuem de boa-fé, com base em factos exatos, e forneçam informações "fiáveis e precisas" no respeito da ética jornalística ou, por outras palavras, no respeito pelos princípios de um jornalismo responsável*»[221], pelo que a liberdade de informação pode ser restringida, cumprindo-se os demais requisitos, quando esteja em causa a prossecução de um objetivo de interesse geral, reconhecido como tal pela União. *In casu*, está em causa, segundo o Tribunal, a prossecução de um duplo objetivo: por um lado, os atos adotados visam «*proteger a ordem e a segurança públicas da União, ameaçadas pela campanha internacional sistemática de propaganda desenvolvida pela Federação da Rússia* [...] *a fim de* [...] *apoiar a agressão militar da Ucrânia, o que corresponde a um dos objetivos de política externa e de segurança comum*» (artº 21º, nº 2, al. a) TUE)[222]; por outro lado, e uma vez que a desinformação integra o «*arsenal de guerra moderno, as medidas restritivas em causa inscrevem-se igualmente no âmbito da prossecução pela União dos objetivos, nomeadamente pacíficos, que lhe foram atribuídos pelo artigo 3º, n.os 1 e 5, TUE*»[223].

Resulta, pois, evidente, que o Tribunal Geral, para as considerar legítimas, integra e contextualiza estas medidas não no quadro geral da luta contra a desinformação, mas num «*contexto extraordinário e de extrema urgência*»[224], determinado pela intensificação da agressão militar na Ucrânia, em violação do direito internacional, e como parte de um pacote de medidas de envergadura inédita destinadas a frustrar, com os instrumentos pacíficos de que a União

[219] Com argumentação idêntica, *vide* Björnstjern Baade, *The EU's "Ban" of RT and Sputnik: A Lawful Measure Against Propaganda for War,* VerfBlog, 08/03/2022, disponível em: https://verfassungsblog.de/the-eus-ban-of-rt-and-sputnik/, consult. em: 13/03/2023.

[220] *Vide* Acórdão do Tribunal Geral (Grande Secção) de 27.07.2022, *RT France v. Conselho*, T-125/22.

[221] *Idem*, § 136.

[222] *Idem*, § 161.

[223] *Idem*, § 162.

[224] *Idem*, § 198.

dispunha, a agressão em curso, dissuadindo-a e protegendo as fronteiras da União. Conclui-se, portanto, que proibição de radiodifusão de determinados meios de comunicação social no espaço da União só lateralmente integra o quadro global respostas da UE no combate à desinformação, inserindo-se primacialmente no âmbito das medidas sancionatórias e dissuasoras aplicadas à Federação da Rússia pela intervenção militar na Ucrânia.

Finalmente, em 16 de junho de 2022, e após um prolongado processo de revisão, é assinado, por 34 subscritores, o Código de Conduta Reforçado da UE sobre a Desinformação[225]. O Código Reforçado procura dar resposta às orientações da Comissão Europeia e, ao mesmo tempo, alinhar-se, aquando da sua aplicação, com o Regulamento dos Serviços Digitais, nomeadamente com os seus objetivos em termos de co e autorregulação no domínio da gestão e mitigação dos riscos sistémicos das "plataformas em linha de muito grande dimensão". Efetivamente, os signatários do Código sinalizaram no Preâmbulo que aquele instrumento pretendia tornar-se um "código de conduta" na aceção do artº 45º do RSD, de modo a poder ser considerado uma medida de atenuação de riscos, de acordo com o artº 35º do RSD. Denota-se que o Código de Conduta Reforçado foi, no essencial, desenhado para as grandes plataformas digitais, embora sobrem incentivos para a sua aplicação às de menor dimensão, tanto mais que estas poderão optar por compromissos menos exigentes. Em termos gerais, e como evolução positiva em relação à primeira versão do Código de Conduta, os subscritores comprometem-se a aceitar e implementar todas os compromissos e medidas que digam respeito aos seus serviços, produtos e atividades, devendo apresentar as razões da não implementação ou subscrição de determinadas medidas e compromissos. Por outro lado, os aspetos quantitativos e qualitativos que permitem medir a aplicação e a eficácia do Código foram largamente reforçados, estando os subscritores obrigados a fornecer "relatórios de elementos qualitativos" (QRE) e "indicadores do nível de serviço" (SLI) específicos por referência aos vários compromissos assumidos, de modo a permitir uma comparação efetiva entre diferentes plataformas e períodos temporais.

Estruturalmente, o Código de Conduta Reforçado divide-se em nove capítulos temáticos (além do preâmbulo), contendo 44 compromissos e um total de 128 medidas destinadas a atingir os objetivos inerentes aos compromissos.

[225] Código de Conduta Reforçado da UE sobre Desinformação, disponível em https://digital-strategy.ec.europa.eu/en/library/2022-strengthened-code-practice-disinformation, consult. em: 14/03/2023.

Substancialmente, o primeiro capítulo, reportando-se ao escrutínio da colocação de publicidade em linha, assume como compromisso macro a desmonetização da desinformação através de várias medidas, tais como evitar a colocação de publicidade junto de conteúdos desinformativos, a proibição de publicação de anúncios publicitários por agentes que sistematicamente violem as políticas de desmonetização da desinformação, e a implementação de políticas publicitárias que não permitam a disseminação de desinformação sob a forma de anúncios, nomeadamente através da sua revisão prévia.

O segundo capítulo refere-se especificamente aos anúncios temáticos de natureza política, procurando alinhar-se com a proposta de Regulamento sobre a transparência e o direcionamento da propaganda política[226]. Neste âmbito, as plataformas salientam que em matéria de anúncios de natureza política a prioridade é assegurar a neutralidade da sua colocação online, independentemente das diversas orientações políticas e dos assuntos que abordem. Sem embargo, os signatários assumem o compromisso de identificar e "etiquetar" claramente os anúncios de cariz político (o que abrange igualmente, quando tecnicamente possível, as plataformas de serviços de mensagens encriptadas), tornando disponível aos utilizadores um conjunto de informações que visam garantir a transparência da publicidade desta natureza, tais como a identidade do patrocinador do anúncio político ou da entidade que em última instância o controla o patrocinador, os montantes gastos e o período em que os anúncios correram online, informações essas facilmente acessíveis em repositórios públicos que agregam as informações durante, pelo menos, cinco anos. Ademais, e no caso do direcionamento de propaganda política, as plataformas comprometem-se a fornecer aos utilizadores informações claras, compreensivas e abrangentes sobre as razões (i.e., os critérios) que levam a que determinado utilizador seja exposto a certo anúncio.

No que respeita à integridade dos serviços, matéria integrada no terceiro capítulo, os subscritores comprometem-se a colaborar entre si para continuar a implementar salvaguardas contra a desinformação e a informação incorreta, incluindo políticas contra comportamentos manipulativos empregues nas plataformas digitais, designadamente através da adoção de uma terminologia comum regularmente revista que limite a criação e a utilização de contas falsas e *bots*, operações de *hack-and-leak*, personificação, *deep fakes* maliciosas, compra de "gostos", promoção paga de conteúdos por "*influencers*" de forma não transparente, utilização de contas para amplificar artificialmente o

[226] Proposta de Regulamento do Parlamento Europeu e do Conselho sobre a transparência e o direcionamento da propaganda política. Bruxelas, 25.11.2021, COM(2021) 731 final.

alcance ou a perceção de apoio público para a desinformação, etc. Além disso, os subscritores, designadamente aqueles que utilizem sistemas de inteligência artificial que permitam a criação ou disseminação conteúdos gerados ou manipulados através destes sistemas, comprometem-se igualmente a respeitar as obrigações de transparência em matéria de inteligência artificial previstas no futuro Regulamento[227].

Já no âmbito do capítulo seguinte, dedicado ao "empoderamento" dos utilizadores, os signatários do Código comprometem-se a apoiar e melhorar a literacia mediática, a desenhar de forma transparente e segura os seus sistemas e políticas, nomeadamente os sistemas/algoritmos de recomendação/personalização de conteúdos, de modo a diminuir o risco da disseminação e amplificação da desinformação. No que toca aos sistemas de recomendação, os signatários comprometem-se a divulgar informações claras, acessíveis e facilmente compreensíveis sobre os principais parâmetros e critérios de que os algoritmos se servem para aumentar ou reduzir a visibilidade dos conteúdos, oferecendo aos utilizadores opções para modificar esses parâmetros de acordo com as suas preferências. Neste contexto, as plataformas digitais publicarão essas informações num designado "centro de transparência", que se comprometem a criar. Além destes compromissos, os subscritores do Código de Conduta Reforçado comprometem-se também a disponibilizar aos utilizadores ferramentas que permitam o acesso à proveniência, ao histórico de edições e à autenticidade ou exatidão dos conteúdos digitais. Finalmente, e ainda no âmbito deste capítulo, os signatários, alinhados com o RSD, estabelecem uma série de compromissos e medidas de natureza procedimental referentes à moderação de conteúdos pelas plataformas digitais, nomeadamente a possibilidade de os utilizadores notificarem a presença de desinformação ou de outros conteúdos que violem os termos de serviço (TdS) das plataformas e a manutenção de mecanismos de reclamação/recurso quanto às decisões tomadas pelas plataformas digitais no âmbito das atividades de moderação de conteúdos.

O quinto capítulo (sexto, contando com o preâmbulo) dedica-se às medidas destinadas a permitir o acesso da comunidade científica aos dados armazenados pelas plataformas digitais e aos deveres de colaboração com que estas se comprometem, tudo com o objetivo de estudar e melhor compreender o alcance e as consequências do fenómeno da desinformação. O capítulo sub-

[227] Proposta de Regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial (Regulamento Inteligência Artificial) e altera determinados atos legislativos da União. Bruxelas, 21.4.2021, COM(2021) 206 final.

sequente é exclusivamente dedicado à comunidade de verificadores de factos, com os signatários a assumirem compromissos alargados de cooperação, inclusive financiamento. No sétimo capítulo os subscritores do Código comprometem-se a criar e manter o já referido "centro de transparência" público, através de um *website* dedicado[228], que serve de repositório para toda a informação relevante relacionada com a implementação das medidas e compromissos assumidos no Código de Conduta Reforçado.

Por fim, nos últimos dois capítulos, os signatários estabelecem compromissos gerais em torno da implementação, atualização e monitorização do Código de Conduta, nomeadamente com a criação de uma "Task-Force", presidida pela Comissão e composta pelos signatários do Código e por representantes do Observatório Europeu dos Meios de Comunicação Digital (EDMO), do Grupo de Reguladores Europeus dos Serviços de Media Audiovisuais (ERGA) e do Serviço Europeu de Ação Externa (EEAS), destinada a acompanhar e adaptar o Código de Conduta Reforçado aos desenvolvimentos tecnológicos, societários, económicos e legislativos. Um mês após o período de implementação de seis meses contados a partir da assinatura do Código de Conduta, os signatários comprometeram-se a submeter relatórios harmonizados (desenvolvidos no quadro da "Task-Force") contendo os SLI e QRE relevantes resultantes da implementação das medidas previstas. Após essa primeira série de relatórios ("baseline reports"), as plataformas em linha de muito grande dimensão obrigaram-se a submeter relatórios semestrais e os demais subscritores relatórios anuais.

Referimos no início deste capítulo que, sensivelmente a partir de 2022, é possível identificar uma nova fase nas estratégias de combate à desinformação por parte da União, desta vez orientada para mecanismos de *hard law*, especialmente com a adoção do Regulamento dos Serviços Digitais, em outubro de 2022, após dois anos de negociações entre os co-legisladores europeus. Efetivamente, Oreste Pollicino enquadra o RSD no âmbito das respostas possíveis às preocupações em torno da proteção da liberdade de expressão online face à crescente utilização de sistemas automáticos de moderação de conteúdos, os quais concedem amplas margens de poder aos atores privados, nomeadamente às gigantes tecnológicas que operam os motores de busca e as redes sociais[229]. Segundo o Autor, a União Europeia, «*através da codificação*

[228] Esse *website* encontra-se já disponível em https://disinfocode.eu/, tendo sido lançado em fevereiro de 2023.

[229] Oreste Pollicino, *Judicial protection of fundamental rights on the internet: a road towards digital constitutionalism?*, Oxford, Hart Publishing, 2021, p. 193 e ss.

de algumas das salvaguardas que o TJUE tem identificado nos últimos anos em casos relacionados com a liberdade de expressão online», tem procurado dar uma resposta aos riscos para a liberdade de expressão criados pela gestão de conteúdos na era digital, aí incluindo o recente Regulamento dos Serviços Digitais[230]. Giovanni De Gregorio vai mais longe, considerando o RSD um «*marco*» inserido num "caminho" da UE em direção ao constitucionalismo digital como resposta normativa aos crescentes poderes digitais privados[231].

Não cabendo aqui uma análise detalhada do novo Regulamento, convém referir, em leves pinceladas gerais, que ele visa estabelecer um regime supranacional, horizontal e multinível para mitigar os novos riscos e desafios colocados pela utilização em grande escala dos serviços da sociedade da informação. Embora adotado no âmbito do Mercado Único Digital, tendo, por isso, como base jurídica o artº 114º TFUE, o RSD visa igualmente salvaguardar os direitos fundamentais dos cidadãos da União, tais como previstos na CDFUE, em particular a liberdade de expressão e de informação, a liberdade de empresa, o direito à não-discriminação e a concretização de um elevado grau de defesa do consumidor (considerando (3) e artº 1º, nº 1 RSD). Além disso, o RSD actualiza, mais de vinte anos depois, a Diretiva e-Comércio, todavia mantendo intocado, no essencial, o regime condicional de isenção de responsabilidade dos prestadores de serviços intermediários (embora o atualize à luz da jurisprudência do TJUE e acrescente, no seu artº 7º, uma "cláusula de bom samaritano" que permite às plataformas digitais não perder a isenção de responsabilidade caso, voluntariamente, adotem medidas ativas de investigação destinadas a detetar, identificar, suprimir ou bloquear o acesso a conteúdos ilegais, desde que o façam de boa-fé e de forma diligente[232]).

Em termos estruturais, o objetivo do RSD é aumentar o grau de transparência e a responsabilização das plataformas digitais na gestão e moderação de conteúdos online, visando «*assegurar um ambiente em linha seguro, previsível e fiável, combatendo a difusão de conteúdos ilegais em linha e os riscos sociais que a difusão de desinformação ou de outros conteúdos pode gerar, e no qual os direitos fundamentais consagrados na Carta sejam eficazmente protegidos e a inovação seja facilitada*», como

[230] *Idem*.

[231] Giovanni de Gregorio, *Digital Constitutionalism in Europe: Reframing Rights and Powers in the Algorithmic Society*, cit., p. 211.

[232] Para mais desenvolvimentos cfr., por todos, Folkert Wilman, *Between Preservation and Clarification: The Evolution of the DSA's Liability Rules in Light of the CJEU's Case Law*, *In* Joris Van Hoboken, J. P. Q., Naomi Appelman, Ronan Fahy, Ilaria Buri, Marlene Straub (ed.), *Putting the Digital Services Act into Practice: Enforcement, Access to Justice, and Global Implications*, Berlin, Verfassungsblog gGmbH, 2023, p. 35-49.

decorre do considerando (9). Para esse efeito, é estabelecido um conjunto abrangente de "obrigações de devida diligência" a cargo aos prestadores dos serviços intermediários. Essas medidas procuram ser proporcionais aos riscos e desafios colocados pelos diversos tipos de prestadores, assistindo-se a um crescendo regulatório em função da natureza dos serviços e do tamanho da plataforma digital. Em especial, as plataformas em linha e os motores de busca de muito grande dimensão são onerados com obrigações mais intensas, atenta a sua importância, «*devido ao seu alcance, expresso nomeadamente em número de destinatários do serviço, na facilitação do debate público, das transações económicas e da difusão ao público de informações, opiniões e ideias e na influência que podem exercer sobre a forma como os destinatários obtêm e comunicam informações em linha*» (considerando (75).

No essencial, e sem regular o conteúdo, o RSD introduz um quadro de obrigações e salvaguardas, substantivas e processuais, que exigem às plataformas digitais divulgar informações (transparência), avaliar riscos sistémicos e prever mecanismos de tutela dos direitos fundamentais dos cidadãos.

No quadro do RSD, a desinformação é uma preocupação evidente. Embora não se trate de um conteúdo ilegal, e, por isso, não fique abrangida nem pelo regime de responsabilidade dos prestadores de serviços intermediários (artº 4º a 9º do RSD), nem pelos mecanismos de notificação e ação previstos para os utilizadores no artº 16º do RSD (embora o Código de Conduta Reforçado estenda, no essencial, aquele regime à desinformação), será, pelo menos e na maior parte dos casos, um conteúdo que viola os TdS das plataformas digitais, pelo que sempre que as plataformas em linha atuarem no sentido da restrição desses conteúdos (seja na sua visibilidade, na possibilidade de monetização, na suspensão de prestação do serviço ou da conta do utilizador), devem fundamentar a sua decisão (artº 17º do RSD), criar sistemas internos de gestão de reclamações (artº 20º do RSD) e disponibilizar mecanismos de resolução extrajudicial de litígios (artº 21º do RSD).

Nos considerandos do RSD, são treze as referências à desinformação, quase todas concentradas no âmbito das obrigações adicionais aplicáveis às plataformas em linha e aos motores de busca de muito grande dimensão no que se refere à gestão de riscos sistémicos. Em todo o caso, são várias as disposições do RSD com potencial impacto no ecossistema da desinformação. Desde logo, o artº 14º, aplicável de forma geral a todos os prestadores de serviços intermediários, obriga estes prestadores a, de uma forma "clara, simples, inteligível, facilmente compreensível e inequívoca", incluir nos seus TdS informações sobre quaisquer restrições que imponham em relação à utilização dos seus

serviços, nomeadamente políticas, procedimentos, medidas e instrumentos utilizados para efeitos de moderação de conteúdos, incluindo a tomada de decisões algorítmicas e a análise humana, bem como as regras processuais do respetivo sistema interno de gestão de reclamações. Parece-nos que no âmbito desta obrigação, ficam os prestadores de serviços intermediários obrigados a prever nos seus TdS com suficiente clareza e precisão os conteúdos que consideram contrários às suas políticas, melhorando-se, por essa via, as definições de desinformação vagas que muitas vezes são ainda empregues. Adicionalmente, o artº 15º prevê a obrigação de todos os prestadores de serviços intermediários apresentarem relatórios, pelo menos uma vez por ano, em que disponibilizam ao público informação detalhada sobre as atividades de moderação de conteúdos que tenham desenvolvido, o que se vem articular com os compromissos de reporte já previstos no Código de Conduta Reforçado da UE sobre a Desinformação.

Depois, o conjunto de regras aplicáveis às plataformas em linha no domínio da publicidade (artº 26º) e da transparência dos sistemas de recomendação (artº 27º) tem, também, potenciais impactos. Por um lado, no que respeita à publicidade em linha, as plataformas ficam obrigadas a identificar claramente que aqueles conteúdos correspondem a anúncios pagos, a identificar a pessoa singular ou coletiva que paga o anúncio e a informar os principais parâmetros utilizados para determinar o destinatário da exibição do anúncio publicitário e, se for caso disso, como alterar esses parâmetros. Além disso, e de acordo com o nº 3 do artº 26º do RSD, as plataformas em linha ficam proibidas de direcionar anúncios publicitários com base na definição de perfis que recorra a categorias especiais de dados (origem racial ou étnica, opiniões políticas, convicções religiosas ou filosóficas, dados genéticos, dados biométricos, dados relativos à saúde ou dados relativos à vida sexual ou orientação sexual de uma pessoa). Sendo o micro-direcionamento uma das técnicas utilizadas no âmbito das campanhas de desinformação, são positivas todas as medidas destinadas a conferir transparência e a restringir o recurso a esta técnica. Na mesma senda, e no domínio dos sistemas de recomendação, o artº 27º do RSD prevê que as plataformas em linha que os utilizem ficam obrigadas a, com linguagem clara e inteligível, revelar nos seus TdS os principais parâmetros utilizados pelos algoritmos de personalização, bem como as opções disponíveis para os utilizadores alterarem ou influenciarem esses critérios.

De entre as várias componentes que compõem o RSD, a doutrina vem destacando que, na essência, estamos perante uma «*regulação de due process digital*

integrada por ferramentas de gestão do risco» digital[233]. Justamente, para as plataformas em linha e motores de pesquisa em linha de muito grande dimensão[234] foram reservadas obrigações adicionais, de modo a dar resposta a relevantes preocupações de política pública, nomeadamente tendo em conta a gestão dos riscos sociais que a atividade destas gigantes tecnológicas acarreta. O RSD elenca, no seu artigo 34º, nº 1, quatro categorias de riscos sistémicos que devem ser avaliados de forma aprofundada pelas grandes plataformas: (1) os riscos associados à difusão de conteúdos ilegais (material pedopornográfico, discurso de ódio, incitação à violência, etc.), (2) os riscos reais ou previsíveis no exercício dos direitos fundamentais conforme protegidos pela CDFUE, incluindo, entre outros, a dignidade do ser humano, o direito à liberdade de expressão e de informação, incluindo a liberdade e o pluralismo dos meios de comunicação social, o direito à reserva da intimidade da vida privada, o direito à proteção de dados, o direito à não discriminação, entre outros (deve aqui ter-se em conta, designadamente, a conceção dos sistemas algorítmicos, os mecanismos de moderação de conteúdos, a utilização abusiva ou não autêntica das plataformas ou outros mecanismos usados para silenciar o discurso ou dificultar a sua ocorrência), (3) os riscos negativos, reais ou previsíveis, nos processos democráticos, no discurso cívico e nos processos eleitorais, bem como na segurança pública, e (4) os riscos para a proteção da saúde pública, dos menores e do bem-estar físico e mental das pessoa que decorram da conceção, funcionamento ou utilização das plataformas, nomeadamente através de manipulação. A disseminação de desinformação deve, pois, ser avaliada no âmbito dos riscos associados à violação dos direitos fundamentais, aos processos democráticos, discurso cívico, processos eleitorais e segurança pública e, ainda, no âmbito dos riscos para a saúde pública.

Da avaliação de riscos anual, que deve ter em conta, nomeadamente, a conceção dos sistemas algorítmicos (especialmente, os de recomendação), os sistemas de moderação de conteúdos, os TdS, os sistemas de seleção e exibição de anúncios publicitários e as práticas relacionadas com a gestão de dados (artº 34º, nº 2 RSD), resulta, de acordo com o artº 35º RSD, a obrigação a cargo das grandes plataformas de adotarem medidas "razoáveis, proporcionais e efica-

[233] Martin Husovec, *Will the DSA Work? On Money and Effort*, *In* Joris Van Hoboken, J. P. Q., Naomi Appelman, Ronan Fahy, Ilaria Buri, Marlene Straub (ed.), *Putting the Digital Services Act into Practice: Enforcement, Access to Justice, and Global Implications*, Berlin, Verfassungsblog gGmbH, 2023, p. 19-34 (p. 21)

[234] Por regra, aquelas plataformas e motores de busca que têm um número médio mensal de utilizadores ativos do serviço na União igual ou superior a 45 milhões, nos termos do nº 1 do artº 33º do RSD.

zes" de atenuação dos riscos sistémicos que tiverem identificado, sempre em respeito pelos direitos fundamentais previstos na CDFUE.

Especial enfoque do RDS, no que às grandes plataformas diz respeito, é dado à elaboração de *standards* comuns às plataformas (artº 44º RSD), de códigos de conduta (artº 45º a 47º RSD) e de protocolos de crise (artº 48º RSD), todos facultativos. Esta ênfase na co e autorregulação, torna o Regulamento dos Serviços Digitais num mecanismo híbrido, que se posiciona entre a regulação efetiva (essencialmente no âmbito do domínio da moderação de conteúdos e de certos aspetos dos sistemas de recomendação e da publicidade em linha) e a *soft law*, encarada como mecanismo de atenuação dos riscos sistémicos identificados pelas plataformas em linha e motores de pesquisa de muito grande dimensão.

Só após o início da aplicação da totalidade do Regulamento, em 17 de fevereiro de 2024, se compreenderá se o mesmo cumprirá as elevadas expetativas que sobre ele recaem na efetiva limitação dos poderes privados que se consolidaram na esfera digital, promovendo, como pretende, um ambiente em linha seguro, previsível e fiável, que respeite os direitos fundamentais dos utilizadores e dos cidadãos e promova de forma positiva os processos democráticos e o discurso cívico na esfera digital.

Por fim, em jeito de remate, referência para o mais recente ato, este atípico, adotado pela União Europeia, em 15 de dezembro de 2022: a Declaração Europeia sobre os direitos e princípios digitais para a década digital[235], na esteira da comunicação de 2021 sobre as "orientações para a digitalização até 2030", estabelecendo a "via europeia para a Década Digital"[236], e de um conjunto de outras declarações do Conselho[237]. A Declaração Europeia sobre os direitos e princípios digitais trata-se de uma declaração interinstitucional (Parlamento Europeu, Conselho e Comissão), aprovada à margem do Conselho Europeu, sendo a primeira declaração interinstitucional da UE dedicada exclusivamente aos direitos e princípios fundamentais no contexto

[235] Cfr. Declaração Europeia sobre os direitos e princípios digitais para a década digital, disponível em: https://direito.up.pt/digeucit/2022/12/19/direitos-e-principios-digitais-presidentes-da-comissao-do-parlamento-europeu-e-do-conselho-assinam-declaracao-europeia/, consult. em: 16/03/2023.

[236] Comunicação da Comissão ao Parlamento Europeu, ao Conselho, ao Comité Económico e Social Europeu e ao Comité das Regiões, *Orientações para a Digitalização até 2030: a via europeia para a Década Digital*, Bruxelas, 9.3.2021, COM(2021) 118 final.

[237] Declaração de Tallin de 2017 sobre a administração pública em linha, Declaração de Berlim de 2020 sobre a sociedade digital e a governação digital baseada em valores e Declaração de Lisboa de 2021 – *Democracia Digital com propósito*.

digital. A Declaração integra-se numa tendência para a elaboração de uma "carta de direitos" para a internet[238], tendo como objetivos expressos modelar a transformação digital de acordo com os valores e o direito europeu, colocar no centro da transformação digital os valores e direitos e clarificar a sua aplicação no ambiente digital (no âmbito do designado "paradigma da equivalência normativa", nos termos do qual a garantias offline valem também online[239]), guiar e servir de referência para os intervenientes públicos e privados no desenvolvimento e implementação de novas tecnologias, orientar os decisores na elaboração de políticas públicas e exportar a visão digital da União internacionalmente.

Trata-se de um documento com natureza declarativa e proclamatória, sem efeitos vinculativos, pelo que não afeta o conteúdo das normas jurídicas existentes nem corresponde a uma proposta regulatória. Corresponde, essencialmente, a um compromisso político das instituições europeias e dos Estados-Membro no âmbito das respetivas competências. Em termos de conteúdo, a Declaração reúne num único documento o *acquis* da UE em matéria digital, congregando os direitos e princípios definidos pelo direito da UE, adaptando-os à Era Digital, sem prejuízo de ser possível de encontrar algum caráter inovatório em certos tópicos. A desinformação merece, na Declaração, duas referências no âmbito do capítulo dedicado à participação no espaço público digital, destacando-se aí que as plataformas em linha, em especial as de grande dimensão, devem apoiar o debate democrático livre e proteger a liberdade de expressão, atenuando os riscos do funcionamento dos seus serviços, nomeadamente no que diz respeito às campanhas de informação incorreta e desinformação. Os declarantes comprometem-se, neste domínio, a criar um ambiente digital em que as pessoas estejam protegidas contra a desinformação e a manipulação de informações e outras formas de conteúdos nocivos, incluindo o assédio e a violência baseada no género.

5. Consideração final

Assiste-se, no início desta década, e face às experiências dos últimos anos, a uma intensificação dos esforços da União Europeia, quer em termos quantitativos, quer em termos qualitativos, no combate ao fenómeno da desin-

[238] Dennis Redeker; Lex Gill; Urs Gasser, *Towards digital constitutionalism? Mapping attempts to craft an Internet Bill of Rights*, International Communication Gazette, Vol. 80, nº 4, 2018, p. 302-319.
[239] Cristina Cocito; Paul De Hert, *The Transformative Nature of the EU Declaration on Digital Rights and Principles: Replacing the Old Paradigm (Normative Equivalency of Rights),* 2023, disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4341816, consult. em: 16/03/2023.

formação. A multiplicidade de respostas e instrumentos que procurámos sintetizar nestas páginas evidenciam que a cruzada europeia está longe de estar finalizada.

O Regulamento dos Serviços Digitais e o Código de Conduta Reforçado da UE sobre a Desinformação são, hoje, o repositório das expetativas da União na luta contra a desinformação, embora seja ainda cedo para avaliar se estes instrumentos terão o sucesso pretendido, tornando o debate democrático e a incontornável presença online menos poluídos e mais consentâneos com o espírito da esfera pública.

A aplicação destes e de outros instrumentos não é apenas uma questão de obstinação ideológica em torno de uma sempre impossível purificação do espaço público em linha. Na verdade, está em causa, antes de tudo, a defesa dos valores democráticos europeus, do Estado de Direito e dos direitos fundamentais dos cidadãos. O sucesso destas políticas europeias equivale, por isso, ao próprio sucesso dos sistemas democráticos e do "modo de vida" europeu num momento de transição para o digital.

Bibliografia

Abreu, Joana Covelo – O Mercado Único Digital como o novo mundo para a União Europeia: repercussões na estrutura regulatória social e institucional – a redefinição do serviço universal e do Organismo de Reguladores Europeus das Comunicações Eletrónicas (ORECE) *UNIO – EU Law Journal*. Vol. 4. nº 2 (2018). p. 59-72.

Andersson, Jan Joel – Hybrid Operations: Lessons from the past. 2015. EU Institute for Security Studies. [Consult. em: 26/11/2022]. Disponível em: https://bityli.com/dclroIZkY.

Arendt, Hannah - *As Origens do Totalitarismo*. Lisboa: Dom Quixote, 2017.

Baade, Björnstjern – The EU's "Ban" of RT and Sputnik: A Lawful Measure Against Propaganda for War. 2022. VerfBlog. [Consult. em: 13/03/2023]. Disponível em: https://verfassungsblog.de/the-eus-ban-of-rt-and-sputnik/.

Balkin, Jack M. – Free Speech is a Triangle. *Columbia Law Review*. Vol. 118. nº 7 (2018). p. 2011-2056.

Bauman, Zygmunt - *A Chronicle of Crisis: 2011-2016*. [s.l.]: Social Europe Edition, 2017.

Bayer, Judit; Bitiukoava, Natalija; Bárd, Petra, et al. – Disinformation and propaganda – impact on the functioning of the rule of law in the EU and its Member States. 2019. European Parliament. [Consult. em: 12/04/2022]. Disponível em: https://www.europarl.europa.eu/RegData/etudes/STUD/2019/608864/IPOL_STU(2019)608864_EN.pdf.

Bayer, Judit; Holznagel, Bernd; Lubianiec, Katarzyna, et al. - *Disinformation and propaganda: impact on the functioning of the rule of law and democratic processes in the EU and its Member States – 2021 update*. European Parliament, 2021. [Consult. em: 26/09/2022]. Disponível em: https://www.europarl.europa.eu/thinktank/en/document/EXPO_STU(2021)653633.

___. - *Disinformation and propagand: impact on the functioning of the rule of law and democratic processes in the EU and its Member States (2021 update)*. Brussels: European Parliament 2021. [Consult. em: 21/05/2022]. Disponível em: https://www.europarl.europa.eu/thinktank/en/document/EXPO_STU(2021)653633.

Bozdag, Engin – Bias in algorithmic filtering and personalization. *Ethics and Information Technology*. Vol. 15. (2013). p. 209-227.

Brkan, Maja – Freedom of Expression and Artificial Intelligence: On Personalisation, Disinformation and (Lack Of) Horizontal Effect of the Charter. 2019. [Consult. em: Disponível em: https://ssrn.com/abstract=3354180.

Cantarella, Michele; Fraccaroli, Nicolò; Volpe, Roberto – Does fake news affect voting behaviour? *Research Policy*. Vol. 52. nº 1 (2023).

Castells, Manuel - *A Galáxia Internet – Reflexões sobre Internet, Negócios e Sociedade*. Lisboa: Fundação Calouste Gulbenkian, 2004.

Center For An Informed Public, Digital Forensic Research Lab, Graphika, & Stanford Internet Observatory - *The Long Fuse: Misinformation and the 2020 Election*. 2021. [Consult. em: 03/01/2023]. Disponível em: https://purl.stanford.edu/tr171zs0069.

Cocito, Cristina; Hert, Paul de – The Transformative Nature of the EU Declaration on Digital Rights and Principles: Replacing the Old Paradigm (Normative Equivalency of Rights). 2023. [Consult. em: 16/03/2023]. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4341816.

Colliver, Chloe - *Click Here For Outrage: Disinformation in the European Parliamentary Elections 2019*. Institute for Strategic Dialogue, 2020. [Consult. em: 03/01/2023]. Disponível em: https://www.isdglobal.org/wp-content/uploads/2020/06/isd_Click-for--Outrage.pdf.

Colomina, Carme; Margalef, Héctor Sánchez; Youngs, Richard - *The impact of disinformation on democratic processes and human rights in the world*. Brussels: European Parliament, 2021. [Consult. em: 12/06/2022]. Disponível em: https://www.europarl.europa.eu/RegData/etudes/STUD/2021/653635/EXPO_STU(2021)653635_EN.pdf.

Costa, Inês Da Silva – A proteção da pessoa na era dos big data: a opacidade do algoritmo e as decisões automatizadas. *Revista Electrónica de Direito*. Vol. 24. nº 1 (2021). p. 34-82.

D'Allonnes, Myriam Revault - *A Verdade Frágil: o que a Pós-verdade faz ao Nosso Mundo Comum*. Lisboa: Edições 70, 2020.

Decker, Ben - *Adversarial Narratives: A New Model for Disinformation*. Global Disinformation Index, 2019. [Consult. em: 27/11/2022]. Disponível em: https://www.disinformationindex.org/blog/2019-8-1-adversarial-narratives-are-the-new-model--for-disinformation/.

Diresta, Renée; Grossman, Shelby - *Pontemkin Pages & Personas: Assessing GRU Online Operations, 2014-2019*. Stanford Internet Observatory, 2019. [Consult. em: 03/12/2022]. Disponível em: https://fsi-live.s3.us-west-1.amazonaws.com/s3fs-public/potemkin--pages-personas-sio-wp.pdf.

Dobber, Tom; Metoui, Nadia; Trilling, Damian, et al. – Do (Microtargeted) Deepfakes Have Real Effects on Political Attitudes? *The International Journal of Press/Politics*. Vol. 26. nº 1 (2021). p. 69-91.

Domingos, Pedro - *A Revolução do Algoritmo Mestre*. 17ª ed. Lisboa: Manuscrito, 2017.
Duarte, Francisco De Abreu – From Platforms To Musk To... Protocols? 2022. The Digital Constitutionalist – The Future of Constitutionalism. [Consult. em: Disponível em: https://digi-con.org/from-platforms-to-musk-to-protocols/.
Nato Strategic Communications Centre Of Excellence – Social Media as Tool of Hybrid Warfare. 2016. NATO Strategic Communications. [Consult. em: 26/11/2022]. Disponível em: https://bityli.com/ctxuDvfOeU.
Gasset, José Ortega y - *A Rebelião das Massas*. Lisboa: Relógio D'Água, 2019.
Gelfert, Axel – Fake News: A Definition. *Informal Logic*. Vol. 38. nº 1 (2018). p. 84-117.
Giles, Keir; Hartmann, Kim; Mustaffa, Munira - *The Role of Deepfakes in Malign Influence Campaigns*. NATO Strategic Communications Centre of Excellence, 2019. [Consult. em: 20/09/2022]. Disponível em: https://stratcomcoe.org/publications/the-role-of-deepfakes-in-malign-influence-campaigns/72.
Gregorio, Giovanni De - *Digital Constitutionalism in Europe: Reframing Rights and Powers in the Algorithmic Society*. Cambridge: Cambridge University Press, 2022.
Grogan, Joelle - *Impact of COVID-19 measures on democracy and fundamental rights*. European Parliament, 2022. [Consult. em: 22/01/2023]. Disponível em: https://www.europarl.europa.eu/RegData/etudes/STUD/2022/734010/IPOL_STU(2022)734010_EN.pdf.
Herasimenka, Aliaksandr; Bright, Jonathan; Knuutila, Aleksi, et al. – Misinformation and professional news on largely unmoderated platforms: the case of telegram. *Journal of Information Technology & Politics*. (2022). p. 1-15.
Husovec, Martin - *Will the DSA Work? On Money and Effort*. In JORIS VAN HOBOKEN, J. P. Q., Naomi Appelman, Ronan Fahy, Ilaria Buri, Marlene Straub (ed.) - *Putting the Digital Services Act into Practice: Enforcement, Access to Justice, and Global Implications*. Berlin: Verfassungsblog gGmbH, 2023. p. 19-34.
Kuklinski, James H.; Quirk, Paul J.; Jerit, Jennifer, et al. – Misinformation and the Currency of Democratic Citizenship. *The Journal of Politics*. Vol. 62. nº 3 (2000). p. 790-816.
Lee, Bruce U. - *Fake Eli Lilly Twitter Account Claims Insulin Is Free, Stock Falls 4.37%*. 2022. 28/11/2022. https://www.forbes.com/sites/brucelee/2022/11/12/fake-eli-lilly-twitter-account-claims-insulin-is-free-stock-falls-43/?sh=6b5f8f741a3d
Lee, Sun Kyong; Sun, Juhyung; Jang, Seulki, et al. – Misinformation of COVID-19 vaccines and vaccine hesitancy. *Scientific Reports*. Vol. 12. nº 13681 (2022).
Lewandowsky, Stephan; Ecker, Ullrich; Cook, John – Beyond Misinformation: Understanding and Coping with the 'Post-Truth' Era. *Journal of Applied Research in Memory and Cognition*. Vol. 6. nº 4 (2017). p. 353-369.
Lewandowsky, Stephan; Smillie, Laura; Garcia, David, et al. - *Technology and Democracy: Understanding the influence of online technologies on political behaviour and decision-making*. Luxembourg: Publications Office of the European Union, 2020. [Consult. em:
Light, Michael T.; He, Jingying; Robey, Jason P. – Comparing crime rates between undocumented immigrants, legal immigrants, and native-born US citizens in Texas. *PNAS*. Vol. 117. nº 51 (2020). p. 32340-32347.
Lilleker, Darren – Evidence to the Culture, Media and Sport Committee 'Fake news' inquiry presented by the Faculty for Media & Communication, Bournemouth Univer-

sity. 2017. Bournemouth University. [Consult. em: 27/11/2022]. Disponível em: https://eprints.bournemouth.ac.uk/28610/.

Loewenstein, Karl – Militant Democracy and Fundamental Rights, I. *The American Political Science Review*. Vol. 31. nº 3 (1937). p. 417-432.

Molina, Maria D.; Sundar, S. Shyam; Le, Thai, et al. – "Fake News" Is Not Simply False Information: A Concept Explication and Taxonomy of Online Content. *American Behavioral Scientist*. Vol. 62. nº 2 (2021). p. 180-212.

Molter, Vanessa; Diresta, Renee – Pandemics & Propaganda: How Chinese State Media Creates and Propagates CCP Coronavirus Narratives. *The Harvard Kennedy School Misinformation Review*. Vol. 1. (2020). p. 1-24.

Napoli, Philip M. - *Social Media and the Public Interest: Media Regulation in the Disinformation Age*. New York: Columbia University Press, 2019.

O'neil, Cathi - *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. London: Penguin Books, 2016.

Pamment, James - *The EU's Role in Fighting Disinformation: Crafting A Disinformation Framework – Working Paper nº 2*. Washington: Carnegie Endowment for International Peace, 2020. [Consult. em: 28/11/2022]. Disponível em: https://carnegieendowment.org/files/Pamment_-_Crafting_Disinformation_1.pdf.

___. - *The EU's Role in Fighting Disinformation: Takib Back Initiative – Working Paper nº 1*. Washington: Carnegie Endowment for International Peace, 2020. [Consult. em: 28/11/2022]. Disponível em: https://carnegieendowment.org/2020/07/15/eu-s-role-in-fighting-disinformation-taking-back-initiative-pub-82286.

Pariser, Eli – Musk's Twitter Will Not Be the Town Square the World Needs. *WIRED*. (2022).

Pereira, Gustavo Teixeira De Faria; Coutinho, Iluska Maria Da Silva – WhatsApp, desinformação e infodemia: o "inimigo" criptografado. *Liinc Em Revista*. Vol. 18. nº 1 (2022). p. 1-22.

Pernice, Ingolf – Risk management in the digital constellation – a constitutional perspective (part I). *Revista D'Internet, Dret i Política*. nº 26 (2018). p. 83-94.

Plasilova, Iva; Hill, Jordan; Carlberg, Malin, et al. - *Study for the "Assessment of the implementation of the Code of Practice on Disinformation"*. European Commission, 2020. [Consult. em: 15/01/2023]. Disponível em: https://digital-strategy.ec.europa.eu/en/library/study-assessment-implementation-code-practice-disinformation.

Pollicino, Oreste - *Judicial protection of fundamental rights on the internet: a road towards digital constitutionalism?* Oxford: Hart Publishing, 2021.

___ - *L'uomo più ricco al mondo non è il miglior custode della libertà di espressione*. 2022. 03/12/2022. https://www.ilsole24ore.com/art/l-uomo-piu-ricco-mondo-non-e-miglior-custode-liberta-espressione-AEfpOGHC?refresh_ce=1

Pollicino, Oreste; Giovanni De Gregorio; Somaini, Laura - *The European Regulatory Conundrum to Face the Rise and Amplification of False Content Online*. In Capaldo, G. Z. (ed.) - *The Global Community: Yearbook of International Law and Jurisprudence 2019*. Oxford: Oxford University Press, 2020. p. 319-356.

Pollicino, Oreste; Gregorio, Giovanni De – Constitutional Democracy, Platform Powers and Digital Populism. *Constitutional Studies*. Vol. 8. (2022). p. 11-34.

Quintel, Teresa; Ullrich, Carsten – Self-Regulation of Fundamental Rights? The EU Code of Conduct on Hate Speech, Related Initiatives and Beyond. 2018. [Consult. em: 10/01/2023]. Disponível em: https://ssrn.com/abstract=3298719.

Rawls, John – The Idea of Public Reason Revisited. *The University Of Chicago Law Review*. Vol. 64. nº 3 (1997). p. 765-807.

Redeker, Dennis; Gill, Lex; Gasser, Urs – Towards digital constitutionalism? Mapping attempts to craft an Internet Bill of Rights. *International Communication Gazette*. Vol. 80. nº 4 (2018). p. 302-319.

Rini, Regina – Fake news and partisan epistemology. *Kennedy Institute of Ethics Journal*. Vol. 2. nº 2 (2017). p. 43-64.

Rocha, Tiago Morais - *A Era Digital e o Estado de Direito Democrático na União Europeia*. Porto: Faculdade de Direito da Universidade do Porto, 2020. Tese de mestrado.

Rochlin, Nick – Fake news: belief in post-truth. *Library Hi Tech*. Vol. 35. nº 3 (2017). p. 386-392.

Rodríguez, Belén Carrasco - *Information Laundering in Germany*. Riga: Nato Strategic Communications Centre of Excellence, 2020. [Consult. em: 03/01/2023]. Disponível em: https://stratcomcoe.org/publications/information-laundering-in-germany/23.

___. - *Information Laundering in the Nordic-Baltic region.* Nato Strategic Communications Centre of Excellence, 2020. [Consult. em: 03/01/2023]. Disponível em: https://stratcomcoe.org/publications/information-laundering-in-the-nordic-baltic-region/26.

Rogers, Danny – Disinformation as Adversarial Narrative. 2022. Global Disinformation Index. [Consult. em: 27/11/2022]. Disponível em: https://www.disinformationindex.org/blog/2022-06-22-disinformation-as-adversarial-narrative-conflict/.

Scannell, Josh - *What Can an Algorithm Do?* New York: DIS Magazine, 2015. http://dismagazine.com/discussion/72975/josh-scannell-what-can-an-algorithm-do/

Schaub, Max; Morisi, Davide – Voter mobilization in the echo chamber: Broadband internet and the rise of populism in Europe. *European Journal of Political Research*. Vol. 59. (2020). p. 752-773.

Shorey, Samantha; Howard, Philip N. – Automation, Big Data, and Politics: A Research Review. *International Journal of Communication*. Vol. 10. (2016). p. 5032-5055.

Sideri, Massimo - *Se su Twitter a parlare è «la voce del padrone»*. 2022. 03/12/2022. https://www.corriere.it/editoriali/22_novembre_30/se-twitter-parla-la-voce-padrone-76e235d2-70df-11ed-9572-e4b947a0ebd2.shtml

Streitfeld, David - *'The Internet Is Broken': @ev Is Trying to Salvage It*. New York Times, 2017. https://www.nytimes.com/2017/05/20/technology/evan-williams-medium-twitter-internet.html

Vilmer, Jean-Baptiste Jeangène; Escorcia, Alexandre; Guillaume, Marine, et al. - *Information Manipulation: A Challenge for Our Democracies.* Paris: Ministry for Europe and Foreign Affairs and Ministry for the Armed Forces, 2018. [Consult. em: 14/04/2022]. Disponível em: https://www.diplomatie.gouv.fr/IMG/pdf/information_manipulation_rvb_cle838736.pdf.

Vosoughi, Soroush; Roy, Deb; Aral, Sinan – The spread of true and false news online. *Science*. Vol. 359. nº 6380 (2018). p. 1146-1151.

Wardle, Claire – Fake news. It's complicated. 2017. *First Draft News*. [Consult. em: 12/03/2020]. Disponível em: https://firstdraftnews.org/latest/fake-news-complicated/.

Wardle, Claire; Derakhshan, Hossein - *Information Disorder: Toward an interdisciplinary framework for research and policy making.* Strasbourg: Council of Europe, 2017. [Consult. em: 14/04/2022]. Disponível em: https://rm.coe.int/information-disorder--toward-an-interdisciplinary-framework-for-researc/168076277c.

Wilman, Folkert - *Between Preservation and Clarification: The Evolution of the DSA's Liability Rules in Light of the CJEU's Case Law*. In Joris Van Hoboken, J. P. Q., Naomi Appelman, Ronan Fahy, Ilaria Buri, Marlene Straub (ed.) - *Putting the Digital Services Act into Practice: Enforcement, Access to Justice, and Global Implications*. Berlin: Verfassungsblog gGmbH, 2023. p. 35-49.

Zawadzka, Małgorzata - *Today's Potemkin Village: Kremlin Disinformation and Propaganda in Poland*. 2018. 03/12/2022. https://warsawinstitute.org/todays-potemkin--village-kremlin-disinformation-propaganda-poland/

II
IA: UMA REVOLUÇÃO PROCESSUAL EM CURSO?

La Inteligencia Artificial en el sistema de justicia penal español: algunos proyectos de interés

Artificial intelligence in the Spanish criminal justice system: some interesting projects

Cristina Alonso Salgado*

RESUMO: En la actualidad del siglo veintiuno, la inteligencia artificial está presente en todos los ámbitos de la vida. La justicia, obviamente, no es una excepción. Y en ese aterrizaje sobre lo jurídico, la inteligencia artificial ha irrumpido también el sistema de justicia penal. Hoy día, son ya muchas las manifestaciones de *artificial intelligence* en el ámbito penal español: policía predictiva, puntos calientes, etc.

El presente trabajo abordará de manera breve algunas de esas manifestaciones –fundamentalmente vinculadas a la actividad de las Fuerzas y Cuerpos de Seguridad–, para analizar sus principales potencialidades, sus objeciones y reparos más relevantes, así como algunos de sus aspectos más polémicos o controvertidos.

Únicamente así, examinando pros y contras desde una perspectiva crítica, será posible lograr la incorporación que todos deseamos, la única incorporación posible, esto es, la incorporación garantista de la inteligencia artificial al sistema de Justicia penal español.

PALAVRAS-CHAVE: Inteligencia artificial; algoritmo; policía predictiva; puntos calientes; investigación penal; sistema de justicia penal.

* Profesora Contratada Doctora de Derecho Procesal. Departamento de Derecho Público Especial y de la Empresa. Facultad de Derecho. Universidad de Santiago de Compostela. cristina.alonso@usc.es. https://orcid.org/0000-0002-0383-3169

ABSTRACT: In today's twenty-first century, artificial intelligence is present in all areas of life. And in this landing on the legal field, artificial intelligence has also burst into the criminal justice system. Today, there are already many manifestations of artificial intelligence in the Spanish criminal field: predictive policing, hot spots, etc.

This paper will briefly address some of these manifestations –mainly linked to the activity of the Security Forces and Corps–, in order to analyze their main potential, their most relevant objections and objections, as well as some of their most controversial or controversial aspects.

Only in this way, examining pros and cons from a critical perspective, will it be possible to achieve the incorporation that we all desire, the only possible incorporation, that is, the incorporation of artificial intelligence into the Spanish criminal justice system.

KEYWORDS: Artificial intelligence; algorithm; predictive policing; hot spots; criminal investigation; criminal justice system.

1. A modo de inicio[1]

Con la mejor tradición garantista bajo del brazo, al referirnos a la inteligencia artificial en su interacción con el sistema de Justicia penal, es preciso dar inicio a esta modesta aportación con una posición de cautela. El inciso preliminar deviene exigido por las tentaciones eficientistas que, desde hace algún tiempo, planean sobre lo que ahora constituye nuestro objeto de estudio.

No es nuestro deseo cuestionar algunas de las virtualidades más evidentes de la *artificial intelligence* sobre las que, por cierto, existe cierto grado de consenso internacional. Simplemente se busca poner negro sobre blanco acerca

[1] Proyecto I+D+i: "Inteligencia artificial, Justicia y Derecho: ¿irrupción o disrupción tecnológica en el proceso penal?" (PID2020-119324GB-I00/AEI/10.13039/501100011033); Ayudas para la consolidación y estructuración de unidades de investigación competitivas y otras acciones de fomento en las universidades del Sistema universitario de Galicia, en los organismos públicos de investigación de Galicia y en otras entidades del Sistema gallego de I+D+i (Grupos con potencial crecimiento; ED431B 2022/18). CONSERTECS, "El contrato de prestación de servicios en el actual entorno tecnológico y social", Proyecto de generación de conocimiento en el marco del Programa Estatal para Impulsar la Investigación Científico – Técnica y su Transferencia, del Plan Estatal de Investigación Científica, Técnica y de Innovación 2021 -2023, Referencia: PID2021-122619OB-I00.

de los riesgos de sucumbir ante determinados discursos vencidos ante la tentación eficientista.

Ni que decir tiene lo innecesario de comentar lo obvio: no hay crítica posible a la idea de satisfacer objetivos optimizando recursos, máxime cuando es el ámbito penal el que está en el punto de mira. No podría ser de otra manera: el actual proceso penal español necesita cuantos auxilios sean necesarios para ganar eficiencia y agilidad. Y ello no como algo coyuntural, sino como una necesidad estructural que deriva de la evidencia: con una Ley de Enjuiciamiento Criminal de las postrimerías del siglo XIX y con unas escasísimas posibilidades de aprobar una nueva a corto plazo, el legislador español parece obligado a buscar soluciones de urgencia para salvar las no pocas dificultades que tales circunstancias ocasionan en el día a día de nuestros órganos jurisdiccionales.

A la luz de todo ello, parece éste un buen ecosistema para la incorporación de programas de inteligencia artificial. Sin embargo, a nuestro juicio, esta visión deriva de la necesidad y no de la reflexión. En efecto, superado el espejismo inicial, no cabe duda de que el desarrollo de instrumentos de esa naturaleza no debiera verse impulsado por el apremio de la penuria y la precariedad, sino por un ejercicio de reflexión en el que se ponderen debidamente las virtualidades y los peligros de la inteligencia artificial[2], los umbrales de garantías, los protocolos de minimización de riesgos, etc.

Esta opción que aquí se defiende, obviamente, es menos espectacular, menos triunfalista y menos aparente. Con todo, la dimensión del fenómeno aconseja una prudencia que, aun cuando no es la más estética de las potencialidades, sí acostumbra a ser la más fiel escudera del garantismo más acabado.

2. Algunas experiencias dignas de mención

Así las cosas, una vez efectuadas las prevenciones que la ocasión merece, debemos ahora proceder al análisis de algunos proyectos de la inteligencia

[2] Manifiesta Barona Vilar un temor en voz alta: "(...) *emerge una enorme inquietud acerca de causa-efecto que puede llegar a producir esta transformación social digital en la igualdad social. Puede producir una suerte de desigualdad social creciente. No en vano, 'los grandes beneficiarios de la cuarta revolución industrial son los proveedores de capital intelectual o físico* (...)*, lo cual explica la creciente brecha de riqueza entre las personas que dependen de su trabajo y las que poseen el capital', o dicho de otro modo, la algoritmización de la vida que arrastra esta revolución industrial del 4.0. está mostrando una cada vez mayor concentración de los beneficios y del valor en tan solo un pequeño porcentaje de personas, que generan lo que SCHWAB denomina 'efecto de plataforma', que dominan los mercados, concentrados fundamentalmente en unas pocas manos y que muestran un cada vez más ineficiente ascensor social*", en Barona Vilar, Silvia, "Inteligencia artificial o la algoritmización de la vida y de la justicia: ¿solución o problema?", *in Revista Boliviana de Derecho*, 28, pp. 18-49.

artificial en su interacción con el sistema de Justicia penal (*lato sensu*), principalmente centrados en la actividad policial.

Lo anunciado exige un excurso con naturaleza preliminar. El análisis de la inteligencia artificial, incluso en lo relativo a algunos de sus elementos más fundamentales, no es en absoluto pacífico. Con independencia de la interacción *supra* indicada, lo cierto es que aspectos tales como la conceptualización de la propia inteligencia artificial o algunos de los riesgos a ella inherentes resultan por sí mismos controvertidos. Quiere ello decir que no nos podemos detener en el examen de cuestiones susceptibles de condicionar debates de la literatura especializada, porque ello excedería con creces el objeto de nuestro trabajo y desviaría, sin duda, nuestro punto de mira.

Focalizados, por tanto, en nuestro objeto de estudio, es necesario examinar un maridaje que en los últimos tiempos está ganando peso en el debate doctrinal: el protagonizado por el ámbito penal y la inteligencia artificial. Su estudio resulta capital para optimizar sus incuestionables virtualidades, pero también, por qué no decirlo, para abordar algunas de las dudas que su utilización y desarrollo siguen generando en la actualidad.

Pues bien, como es fácil imaginar, el área de confluencia entre el ámbito de actuación de la policía y la inteligencia artificial pone por encima de la mesa manifestaciones de interés de una importancia más notable.

Cabe destacar, en primer lugar, los denominados mapas criminales que, como es sabido, son empleados para el análisis del espacio en el que los delitos son cometidos a lo largo de un determinado tiempo[3].

Apunta BALCELLS que "*Varias han sido las causas del cambio que ha conducido a una policía eminentemente reactiva a una de naturaleza predictiva, pero una fundamental ha sido el análisis geográfico de los delitos propio de la criminología medioambiental*". Subraya el autor "(...) *la teoría de la geometría del crimen, conforme a la que el delito sucede en espacios que pueden ser predichos porque las oportunidades delictivas coinciden en zonas conocidas por el delincuente. A su vez, hay determinados espacios que se convierten en puntos calientes (hot spots) porque el nivel de convergencia de delincuentes y víctimas es muy elevado, y, por tanto, las oportunidades se disparan*"[4].

El actual desarrollo de este tipo de herramientas muestra una evolución más que relevante. La transición de los mapas en soporte físico a los mapas tecnológicos permitió mejorar la eficiencia de los procesos: se dejó de ope-

[3] Hernández, María, "Inteligencia artificial y Derecho penal", *in Actualidad Jurídica Iberoamericana*, 10 bis, 2019, pp. 792-843.

[4] Balcells, Marc, "Luces y sombras del uso de la inteligencia artificial en el sistema de Justicia penal", *in* Cerrillo, Agustí & Peguera, Miquel (Eds.), *Retos jurídicos de la inteligencia artificial*, Cizur Menor (Navarra), Aranzadi, p. 149.

rar manualmente en favor de una combinación de datos casi automática. Esa mejora exponencial no hizo sino evidenciar de manera aún más clara, la principal potencialidad de estos mapas criminales: su utilización posibilita que se identifiquen las zonas en las que, potencialmente, pueden cometerse ciertos delitos. No es ello una cuestión de orden menor. Si bien se piensa, esa previsión permite que, en atención a lo señalado por el mapa, se ajusten las medidas para la prevención y la lucha contra el delito[5].

Ello no obstante, sin negar las evidentes virtualidades, los reparos aparecen, de igual modo, de manera casi intuitiva: desde la reducción del elemento pacificador para el sistema de Justicia penal, hasta los riesgos de estigmatización de determinados contextos geográficos con todo lo que ello implica, que, infelizmente, no es poco (criminalización social, depauperación urbana, etc.).

En otro orden de cosas, conviene notar, en segundo lugar, el programa "Veripol" para la identificación de denuncias falsas. Fundamentado en un sistema de procesamiento del lenguaje natural, a través de Veripol se coligen determinadas especificidades que son trasladadas a un modelo matemático para valorar la probabilidad de falsedad de la denuncia. Es decir, a partir de esos datos derivados del procesamiento se infieren patrones de comportamiento que –en teoría– identifican las particularidades que singularizan las denuncias falsas.

Al respecto, indica el MINISTERIO DEL INTERIOR que "*El objetivo de este método es el desarrollo de estrategias efectivas de prevención del delito y el aumento de la efectividad de las investigaciones. Se podría definir como un método creado para predecir la veracidad de las declaraciones de las víctimas de delitos graves* (...) *A partir del análisis de las características y coeficientes de VeriPol, es posible sacar conclusiones sobre la veracidad de lo manifestado en una denuncia. De hecho, el modelo es capaz de discernir diferencias significativas en la narración de denuncias verdaderas y falsas que conducen a la mejor separación entre estas dos clases. De este análisis se puede concluir que las denuncias verdaderas y falsas difieren principalmente en tres aspectos principales: modus operandi de la agresión, morfosintaxis de la denuncia y cantidad de detalles*"[6].

De esta lectura un tanto triunfalista del Ministerio, se pueden extraer dos consideraciones de relevancia. Al menos, en el terreno de lo potencial, la herramienta, en primer lugar, requiere de poca información –la ofrecida por el denunciante– para poder operar; asimismo, en segundo lugar, propor-

[5] En relación al debate, *vid*., Miró, Fernando, "Inteligencia artificial y justicia penal: más allá de los resultados lesivos causados por robots", *in Revista de Derecho penal y Criminología*, 20, 102, 2018.

[6] Ministerio del Interior, *La Policía Nacional pone en funcionamiento la aplicación informática VeriPol para detectar denuncias falsas*, Madrid, 2018, *in* https://www.lamoncloa.gob.es/serviciosdeprensa/notasprensa/interior/Paginas/2018/271018veripol.aspx (10.10.2022)

ciona su evaluación inmediatamente. Todo ello permite condicionar la toma de decisiones tales como la distribución de medios materiales, de recursos humanos, etc.[7].

Colegir inferencias a estas alturas del estado del arte –tan incipiente el debate–, resultaría, sin duda, precipitado. Algunas voces han destacado, con base en proyectos que podríamos categorizar como "pilotos", que el programa favorece la identificación de supuestos de denuncias falsas con un notable índice de acierto[8].

Con todo, en nuestra opinión, en su configuración actual, la herramienta admite una nómina de objeciones en absoluto desdeñable[9]. En primer lugar, aunque el instrumento no toma la decisión última, lo cierto es que sí establece un condicionante de relevancia. Justamente por esta relevancia que se apunta, cabe que, con el tiempo y la propia inercia de las cosas, el condicionante pase a convertirse en un determinante con todos los riesgos que ello comporta.

En segundo lugar, es posible destacar de igual modo, dificultades desde una perspectiva criminológica, toda vez que se echa en falta una ponderación adecuada de la incidencia de los procesos de victimización en la información volcada que sirve para nutrir a Veripol.

Por último, a la vista de lo que se viene de apuntar, parece probable que, en buena parte de los supuestos, el criterio que acabará prevaleciendo será el de la herramienta de manera casi automática. Este "automatismo tácito" ¿implicará una particular motivación del agente por separarse del criterio de Veripol? La pregunta tiene enjundia, máxime si se pone en relación con aspectos vinculados con la responsabilidad derivada de las actuaciones policiales.

[7] González-Álvarez, José Luís, Santos-Hermoso, Jorge & Camacho-Collados, Miguel, "Policía predictiva en España. Aplicación y retos de futuro", *in Behavior & Law Journal*, 6 (1), 2020, p. 30.

[8] "*To test the efficacy and effectiveness of VeriPol, a pilot study has been undertaken in the urban areas of Murcia and Málaga, Spain. More in detail, the pilot study was run in Murcia (four police departments involved) from the 5th to the 9th of June 2017, while it took place in Málaga (six police department involved) from the 12th to the 16th of June 2017. In each destination, two agents, experts in false report detection and in VeriPol, were sent to install the software, give a short course on its use to the local agents and inves-tigators, and supervise all the activity. After that, all the new violent robbery reports as well as all the open violent robbery cases of 2017, were analysed by VeriPol*", en Liberatore, Federico, Quijano-Sánchez, Lara & Camacho-Collados, Miguel, "Applications of Data Science in Policing: VeriPol as an Investigation Support Tool", *in European Law Enforcement Research Bulletin-Innovations in Law Enforcement*, 4, pp. 92 y ss.

[9] Alonso Salgado, Cristina, "Acerca de la inteligencia artificial en el ámbito penal: especial referencia a la actividad de las fuerzas y cuerpos de seguridad", *in Ius et Scientia*, volumen 7, número 1, 2021, pp. 32 y s.

3. Para acabar sin concluir

Extraer inferencias a estas alturas del debate y con el actual nivel de incorporación de la inteligencia artificial en el sistema de Justicia penal español resultaría, con certeza, absolutamente prematuro.

Debemos acabar, por tanto, sin concluir, o, al menos, sin asumir conclusiones susceptibles de ser erigidas en los pilares fundamentales de un dogma de fe. El estado del arte recomienda, en efecto, prudencia en las manifestaciones. Y ello porque, más allá de lo acabado de referir, la propia naturaleza de nuestro objeto de estudio –siempre cambiante– aconseja proceder al epílogo a la espera de conclusiones. Únicamente desde esta perspectiva podemos abordar el fenómeno de la inteligencia artificial en su interacción con la Justicia penal.

Hasta el momento, hemos situado el foco en determinados elementos de interés sobre la materia, salpimentadas con un número de interrogantes que, nuevamente, invitan a la prudencia. La nómina de inconvenientes no favorece el optimismo. Así pues, ni que decir tiene que, al menos por el momento, el ámbito procesal penal no es el apropiado para buscar respuestas categóricas.

Con todo, aun cuando se adopten todas las prevenciones, nuestras reflexiones están sometidas a un riesgo mayor que, en ocasiones, pasa desapercibido: el sesgo de la automatización. La complacencia automatizada derivada de una –por qué no decirlo– injustificada e indisimulada fanatización tecnológica distorsiona algunos análisis que de manera acrítica abrazan toda iniciativa tecnológica al precio que fuere, esto es, sin el debido examen de objeciones y potencialidades. Y no genera ello las resistencias que *prima facie* debiera, toda vez que el poder blanqueante del denominado "Mathwashing" juega, por supuesto, muy a favor de obra.

La apuntada conjunción de factores invita a mantenernos vigilantes acerca de algunas cuestiones de relevancia de la intersección inteligencia artificial-sistema de Justicia penal: desde aspectos vinculados con la trazabilidad del razonamiento algorítmico, o la retroalimentación de sesgos discriminatorios, hasta, por supuesto, el impacto sobre los derechos humanos. Pero no sólo por los peligros que directamente implican, sino también, y muy especialmente, por aquéllos que de manera subrepticia acechan por la inercia blanqueadora de esa fanatización tecnológica de la que antes dábamos cuenta. Esta disposición fiscalizadora no surge de la desconfianza o de una suerte de ludismo *new age*, nace por el contrario, de la firme voluntad de que el avance en materia de inteligencia artificial no se produzca en el sistema de Justicia penal a costa de derechos, libertades y garantías, por eficiente (*sic*) que ello pueda parecer.

Referencias bibliográficas

Alonso Salgado, Cristina, "Acerca de la inteligencia artificial en el ámbito penal: especial referencia a la actividad de las fuerzas y cuerpos de seguridad", *in Ius et Scientia*, volumen 7, número 1, 2021, pp. 32 y s.

Balcells, Marc, "Luces y sombras del uso de la inteligencia artificial en el sistema de Justicia penal", *in* Cerrillo, Agustí & Peguera, Miquel (Eds.), *Retos jurídicos de la inteligencia artificial*, Cizur Menor (Navarra), Aranzadi.

Barona Vilar, Silvia, "Inteligencia artificial o la algoritmización de la vida y de la justicia: ¿solución o problema?", *in Revista Boliviana de Derecho*, 28.

González-Álvarez, José Luís, Santos-Hermoso, Jorge & Camacho-Collados, Miguel, "Policía predictiva en España. Aplicación y retos de futuro", *in Behavior & Law Journal*, 6 (1), 2020.

Hernández, María, "Inteligencia artificial y Derecho penal", *in Actualidad Jurídica Iberoamericana*, 10 bis, 2019.

Liberatore, Federico, Quijano-Sánchez, Lara & Camacho-Collados, "Applications of Data Science in Policing: VeriPol as an Investigation Support Tool", *in European Law Enforcement Research Bulletin-Innovations in Law Enforcement*, 4.

Miró, Fernando, "Inteligencia artificial y justicia penal: más allá de los resultados lesivos causados por robots", *in Revista de Derecho penal y Criminología*, 20, 102, 2018.

A Inteligência Artificial como auxiliar das decisões judiciais*

Artificial Intelligence as an auxiliary in the judicial ruling

Elisa Alfaia Sampaio**
Paulo Jorge Gomes**

RESUMO: Assinalamos que a utilização de agentes de Inteligência Artificial (IA) como auxiliares nas decisões judiciais deve ter presente o contexto social e tecnológico do séc. XXI e as experiências já implementadas no ambiente judiciário, sobretudo na Europa. Neste pressuposto, enumeramos alguns desafios éticos e jurídicos, tendo em conta a jurisprudência do Tribunal Europeu dos Direitos Humanos. Também identificamos algumas propostas para uma IA de confiança, nomeadamente quanto ao cumprimento do princípio da explicabilidade na utilização de redes neuronais artificiais.

PALAVRAS-CHAVE: Inteligência Artificial, decisão judicial, Convenção Europeia dos Direitos Humanos, princípio da explicabilidade, redes neuronais artificiais, caixa negra.

ABSTRACT: We point out out that the use of Artificial Intelligence (AI) agents as auxiliaries in judicial decisions must take into account the social and technological context of the 21st century and the experiences already implemented in the judicial environment, especially in Europe. With this assumption, we mention some ethical and legal challenges, bearing in mind the jurisprudence of the

* Os autores não seguem a ortografia do Acordo Ortográfico de 1990. O tema deste texto foi parcialmente desenvolvido por Elisa Alfaia Sampaio, João J. Seixas, Paulo Jorge Gomes, "Artificial Intelligence and the Judicial Rulling", Themis Annual Journal, vol. 1, issue 1, October 2019, pp. 234 e ss.

** Juízes de Direito – jurisdição administrativa e fiscal.

European Court of Human Rights. We also identified some proposals for a trustworthy AI, namely those regarding the observance of the principle of explicability in artificial neural networks.

KEYWORDS: Artificial Intelligence, judicial ruling, European Convention of Human Rights, principle of explicability, artificial neural networks, black box.

1. Contexto

Vivemos uma híper-história descrita por LUCIANO FLORIDI como *"o estádio do desenvolvimento humano em que as relações tecnológicas de terceira ordem se tornam condição necessária para o desenvolvimento, a inovação e o bem-estar"* [1].

Numa primeira ordem, as necessidades humanas são intermediadas pela tecnologia na relação com a natureza[2]. Numa segunda ordem, a componente humana é intermediada por tecnologias na sua relação com outras tecnologias[3]. Nas relações de terceira ordem as tecnologias actuam como utilizadores, mediados por tecnologias que interagem com outras tecnologias[4].

Estima-se que até ao uso generalizado de computadores a Humanidade tenha acumulado aproximadamente 12 *exabytes* de dados ao longo de toda a sua história[5]. No séc. XXI, entre 2006 e 2011 apenas, os dados disponíveis aumentaram para mais de 1600 *exabytes*. Actualmente, vivemos na Era do *zettabyte*[6]. A International Data Corporation previu para 2025 um aumento

[1] LUCIANO FLORIDI, *The Fourth Revolution: how the infosphere is reshaping human reality*, Oxford, Orxford University Press, 2014, pp. 25 e ss.

[2] Por exemplo, alimentação → arma/enxada → caça/cultivo; protecção/conforto → abrigo/casa → meteorologia; saúde → medicamento → aumento da esperança média de vida.

[3] Por exemplo, mão humana → chave de fendas/comando remoto → parafuso/TV; trabalho → motor → máquina; deslocação → estrada → veículo.

[4] Por exemplo, computador → digitalização → Internet; sensor → Internet → robôs automáticos/*Internet of Things*; *Big data* → algoritmo → Inteligência Artificial.

[5] Colocando em perspectiva, 1 *exabyte* corresponde a um vídeo com 50 000 anos de duração e qualidade de um DVD.

[6] 1 *zettabyte* corresponde a 1000 *exabytes*.

do volume total de dados até 175 *zettabytes*[7], com tendência para aumentar, pois a utilização de dados gera mais dados, crescendo de modo exponencial[8].

Neste contexto, a nossa auto-compreensão parece encontrar-se num novo estádio. Com Copérnico, deixámos de ser o centro imóvel do Universo; com Darwin, percebemos que não estamos assim tão afastados do reino animal; com Freud, a nossa consciência deixou de ser cartesianamente transparente; e hoje compreendemos que somos organismos informacionais, que recebemos e transmitimos dados, que fazemos parte de uma infoesfera, de um *Big Data*.

O aumento significativo da capacidade computacional e a drástica diminuição dos custos da memória e armazenamento dos computadores foram os ingredientes perfeitos para no séc. XXI surgir o fenómeno do *Big Data*. Este oceano de dados tem de ser recolhido, conservado, gerido e analisado. Por outras palavras, os dados, para serem *big*, carecem de modelos através dos quais os algoritmos extraem inferências para formular padrões, tendências e correlações. Daqui nasceu a moderna Inteligência Artificial (IA).

2. A IA moderna

Há muitas definições de IA. Para o nosso propósito, adoptamos a definição avançada pelo Grupo de Peritos de Alto Nível sobre a Inteligência Artificial (GPAN-IA):

> "[a IA] *refere-se aos sistemas concebidos por seres humanos que ao receberem um objectivo complexo actuam no mundo físico ou digital percepcionando o seu ambiente, interpretando dados recolhidos, estruturados ou não estruturados, raciocinando sobre o conhecimento resultante desses dados e decidindo qual a melhor acção (ou acções) a adoptar (de acordo com parâmetros pré-definidos). Os sistemas de IA também podem ser concebidos para aprenderem a adaptar o seu comportamento mediante uma análise do modo como o ambiente foi afectado pelas suas acções anteriores*"[9].

[7] IDC White Paper, *The Digitization of the World From Edge to Core*, 2018, disponível em https://www.seagate.com/files/www-content/our-story/trends/files/idc-seagate-dataage-whitepaper.pdf (última consulta, 10-10-2022).

[8] IDC White Paper, *The Digitization of the World From Edge to Core*, 2018, disponível em https://www.seagate.com/files/www-content/our-story/trends/files/idc-seagate-dataage-whitepaper.pdf (última consulta, 10-10-2022). A analogia de Callum Chace, *The Artificial Intelligence and the Two Singularities*, Boca Raton, London, New York, CRC Press, 2018, p. 45, ajuda a compreender o crescimento exponencial: se ao darmos um passo avançarmos 1 metro em 30 passos teremos avançado 30 metros. Mas se avançarmos 30 passos exponenciais, dobrando o comprimento em cada um dos passos, ao 29º passo teremos chegado à Lua e ao 30º regressaremos à Terra.

[9] High-Level Expert Group on AI, *A Definition of AI: Main Capabilities and Disciplines*, 2019.

Partindo dessa definição, devemos distinguir a IA tradicional da IA moderna. A IA tradicional, desenvolvida a partir da década de 70 do séc. XX, assenta no conhecimento codificado, pré-programado, especialista, inspirada em sistemas lógicos, sobretudo dedutiva, mas sem capacidade para resolver situações inesperadas. Foi o caso do computador Deep Blue, que em 1997 desafiou Kasparov no xadrez, e é o que existe actualmente, por exemplo, nos processadores de texto ou nos programas de cálculos que utilizamos nos nossos computadores.

A IA moderna, surgida no dealbar do séc. XXI, é diferente. Suportada no *Big Data*, aprende soluções a partir de princípios, pode generalizar novas tarefas e resolver problemas inesperados, baseia-se em raciocínios indutivos e abdutivos, em suma, inspira-se no cérebro humano e é validada pelas neurociências. É o caso do supercomputador AlphaGo, da Google Deep Mind, que em 2016 venceu (4-1) o campeão mundial Lee Sedol no jogo Go, ou ainda dos algoritmos utilizados em motores de busca ou em redes sociais.

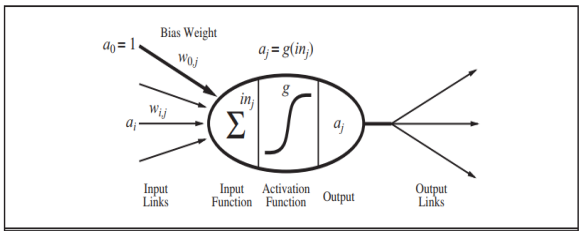
O que torna a IA moderna como um sublime tecnológico[10] tem a ver com o modo da sua aprendizagem. Há diferentes formas de aprendizagem (supervisionada, não supervisionada, por reforço) isoladas ou combinadas, mas a que tem levantado mais questões éticas e jurídicas são as denominadas redes neuronais artificiais (*artificial neural networks*).

Estas redes mimetizam as redes neuronais biológicas, por serem constituídas por nódulos ou unidades, designados neurónios artificiais, que disparam um sinal quando é ultrapassada uma função limite. A saída de cada neurónio é calculada por uma função não linear da soma das suas entradas: cada sinal tem um peso associado, que determina a força e o sinal de conexão à medida que a aprendizagem prossegue. O sinal é depois transmitido para outras unidades ou nódulos, que processam e se conectam a outros neurónios, formando a rede neuronal artificial. Usualmente existem várias camadas (*layers*) escondidas entre as camadas de entrada (*input*) e as camadas de saída (*output*)[11].

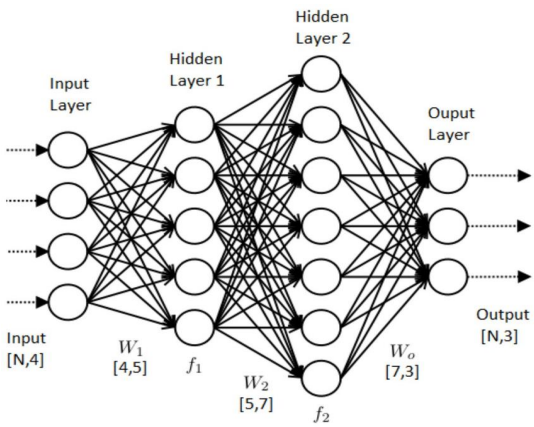
[10] O sublime natural, teorizado por Kant e Edmund Burke, exprime admiração, medo, fascínio, a sensação de insignificância da humanidade perante a Natureza. O sublime tecnológico inspira os mesmos sentimentos, mas por realizações humanas que se equiparam ou mesmo superam os poderes da Natureza: as dimensões dos edifícios, os meios de transporte, os programas espaciais, e agora a inteligência artificial, cf. Hermínio Martins, *Experimentum Humanum*, Lisboa, Relógio D'Água, 2011. pp. 152 e 166.

[11] Stuart J. Russell & Peter Norvig, *Artificial Intelligence: A Modern Approach*, 3rd ed., Pearson Education Limited, England, 2016, pp. 727 e ss.

Modelo matemático de um neurónio artificial[12]



Modelo de uma rede neuronal artificial[13]



Chegamos à pergunta que nos move: deverá o juiz utilizar algoritmos, designadamente redes neuronais artificiais, como auxiliar à sua decisão? A nossa resposta é sim, mas com o necessário *distinguo*.

O agente de IA deve ser de confiança. A União Europeia, através do GPAN-IA, indicou os pressupostos para o conceito "IA de confiança": deve ser legal, garantindo o respeito de toda a legislação e regulamentação aplicáveis; deve ser ética, garantindo a observância de princípios e valores éticos; e deve ser sólida, tanto do ponto de vista técnico como do ponto de vista social, uma vez que, mesmo com boas intenções, os sistemas de IA podem causar danos não intencionais[14].

[12] Imagem retirada de Stuart J. Russell & Peter Norvig, *ob. cit.*, p. 728.

[13] Imagem obtida em https://www.researchgate.net/figure/Artificial-neural-network-of-multiple-layers-and-outputs-31_fig2_331097835 (última consulta, 10-10-2022).

[14] High-Level Expert Group on AI, *Ethics Guidelines for Trustworthy AI*, 2019, disponível em https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai (última

Em bom rigor, estes requisitos aplicam-se a qualquer relação fiduciária, sobretudo quando se trata de alguém com o poder de influenciar as nossas vidas. Mesmo com boas intenções e tecnicamente bem preparado, o juiz pode causar danos não intencionais. Pense-se nas situações de exaustão[15] ou de *burnout* ou nos vieses cognitivos na apreciação da prova originados pela incorrecta utilização de heurísticas, já amplamente estudados pela psicologia comportamental e da decisão[16].

consulta, 11-10-2022). Na proposta (*draft*) divulgada a 18 de Dezembro de 2018, e que na generalidade se manteve na versão final, o conceito de "IA de confiança" foi assumidamente influenciado pelo texto de Luciano Floridi, *et. al.*, "AI4People–An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations", *Minds and Machines*, Springer, 2018.

15 Veja-se o caso relatado por Daniel Khaneman, *Thinking, Fast and Slow*, trad. port. Pedro Vidal, *Pensar, Depressa e Devagar,* Lisboa, Temas e Debates – Círculo de Leitores, 4ª ed., reimpressão, 2021, p.61: *"uma demonstração perturbante dos efeitos do depauperamento no juízo foi recentemente relatada em Proceedings of the National Academy of Sciences. Os desavisados participantes no estudo foram oito juízes de liberdade condicional em Israel. Passaram dias inteiros a examinar pedidos de liberdade condicional.* (...) *(O tempo exacto de cada decisão é registado e as durações dos três intervalos para os juízes comerem durante o dia – intervalo da manhã, do almoço e do lanche – são também registados). Os autores do estudo compararam a proporção de pedidos aprovados com o tempo decorrido desde o último intervalo. A proporção sobe em flecha após cada refeição.* (...) *A melhor justificação possível para os dados fornece más notícias: cansados e com fome, os juízes tendem a acomodar-se à posição mais fácil da negação dos pedidos de liberdade condicional. Tanto a fadiga como a fome desempenharão provavelmente um papel."*

16 Sobre as heurísticas e enviesamentos cognitivos, veja-se Amos Tversky/Daniel Kahneman, "Juízo sob incerteza: heurísticas e enviesamentos", apêndice A in Daniel Khaneman, *Thinking, Fast and Slow*, *cit.*, pp. 551 e ss. Sobre estes vieses na perspectiva da valoração da prova pelos tribunais, cf. Jordi Nieva Fedoll, *La valoración de la prueba*, Madrid, Marcial Pons, 2010; *Idem, Inteligencia artificial y processo judicial*, Madrid, Marcial Pons, 2018; Luís Filipe Pires de Sousa, *Prova Testemunhal*, Coimbra, Almedina, 2013.

As heurísticas são procedimentos mentais simples através dos quais se procuram respostas fáceis para perguntas difíceis, como que atalhos de raciocínio. São úteis, mas podem originar erros graves e sistemáticos, pois actuam de modo inconsciente quanto ao processamento de informação, a que não são alheios processos emotivos (cf. Enrico Altavilla, *Psicologia Judiciária – o processo psicológico e a verdade judicial*, vol. I, 2ª ed., trad. Fernando Miranda, Coimbra, Almedina, 2007, pp. 138 e ss.; Luís Filipe Pires de Sousa, *ob. cit.*, pp. 154 e ss.; Maria Anabela Bento Marinho dos Reis, *A Memória do Testemunho e a Influência das Emoções na Recolha e Preservação da Prova*, dissertação de Doutoramento em Ciências e Tecnologias da Saúde, Especialidade em Desenvolvimento Humano e Social, Universidade de Lisboa, Faculdade de Medicina, 2014, pp. 108 e ss.; Michele Taruffo, *La Prueba*, trad. Laura Manríquez/Ferrer Beltrán, Madrid, Marcial Pons, 2008, pp. 137 e 138). As heurísticas mais comuns são as da (i) representatividade, em que algo é considerado como verificado apenas com base em casos representativos da classe em situações normais, ou se se quiser de outro modo, quando se retira a probabilidade de algo a partir da similaridade a outras classes – o que pode originar a falácia ou viés da conjunção, isto é, um acontecimento específico ou mais pormenorizado não pode ser mais provável do que um evento menos específico ou mais generalizado, subvalorizando-se deste modo a informação estatística;

3. As experiências europeias

3.1. Sistemas implementados e projectados

Existem diversas classificações possíveis de métodos e técnicas de raciocínio da IA. Para precisar de que modo tais categorias técnicas podem ser consideradas ferramentas jurídicas de IA poderiam abordar-se exemplos como os motores de pesquisa avançada de jurisprudência, a resolução de litígios em linha (ou *on-line*, com a sigla RLL), as ferramentas de apoio à elaboração de documentos legais, as ferramentas de análise (preditiva ou escalas), a categorização de documentos (*e.g.* contratos) ou os *chatbots* que disponibilizam informação ou apoio jurídico. Merecem, também, referência alguns projectos académicos que utilizam métodos de raciocínio para prever decisões judiciais, como adiante veremos.

No Reino Unido, é possível encontrar a *Luminance*, uma ferramenta de análise de texto baseada na tecnologia de aprendizagem automática (reconhecimento de padrões, como assinalado pela empresa[17]), que analisa documentos e aprende com a interacção entre advogados e com os documentos; ou o HART (*Harm Assessment Risk Tool*), algoritmo que prevê o nível de risco de os suspeitos cometerem novos crimes num determinado período[18] através de um algoritmo de "floresta aleatória" combinando certos valores, a maior parte dos quais focado nos antecedentes criminais do suspeito, bem como a sua idade, género e área geográfica. Em França, existem algumas ferramentas,

(ii) a heurística da disponibilidade ou acessibilidade, que consiste em julgar a frequência de uma classe ou probabilidade de um evento com base na facilidade com que os casos ocorrem à mente, ou seja, eventos facilmente recuperáveis e imagináveis, ou mais impactantes, tendem a ser sobrestimados, enquanto os eventos difíceis de recuperar ou de imaginar, ou menos impactantes, tendem a ser subestimados (*e.g.* geralmente é maior o receio de ter um acidente de avião do que um acidente de carro, embora este último seja mais provável do que o primeiro); (iii) a heurística da ancoragem e ajuste, que consiste nas pessoas formularem uma ideia inicial a partir de alguns indícios, sendo muito difícil mudarem posteriormente de opinião, apesar de receberem novos dados que na verdade modificam a percepção inicial – as pessoas reinterpretam os dados novos para continuar a defender a opinião inicial e não ter que a modificar, originando o viés ou falácia da confirmação, isto é, procura-se testar a hipótese ajustando os dados compatíveis que confirmem a crença, ignorando os dados incompatíveis, de sinal contrário; (iv) a heurística do afeiçoamento, ou efeito aura, amplamente explorado pela indústria da publicidade, na medida em que os seres humanos deixam-se condicionar por variáveis emocionais introduzidas pela linguagem (incluindo a sua ordem de apresentação) ou pela aparência, independentemente do conteúdo.

[17] Sítio na Internet da empresa em https://www.luminance.com/ [última consulta em 19-10-2022].

[18] Como descreve a Universidade de Cambridge (https://www.cam.ac.uk/research/features/helping-police-make-custody-decisions-using-artificial-intelligence – última consulta em 19-10-2022), este algoritmo ajuda a polícia a decidir, após deter alguém, se deve libertá-la sob fiança ou mantê-la em detenção até ser presente ao tribunal.

como a *Doctrine*, a *LexisNexis* e a *Dalloz*, que são simples motores de pesquisa de decisões judiciais e textos jurídicos. Mais interessantes são as ferramentas de *software Prédictice* e *Case Law Analytics*, ambas ferramentas de análise cujo propósito é prever o resultado de um específico processo judicial[19], também denominadas ferramentas de análise de tendência.

Em 2017, realizou-se em França uma experiência para testar um *software* de justiça preditivo (o *software Prédictice*), aplicando-o a vários recursos nos dois tribunais de recurso em Rennes e Douai. Os resultados não foram os expectáveis. O objectivo da experiência era o de procurar reduzir a variabilidade excessiva das decisões judiciais em nome da igualdade dos cidadãos perante a lei. Embora a experiência não tenha trazido qualquer observação de valor quanto ao papel da IA na tomada de decisão, o estudo concluiu, porém, que o *software* se confundiu entre as ocorrências lexicais e as causalidades que tinham sido decisivas para os juízes nas decisões utilizadas como dados de treino, tendo conduzido a resultados absurdos[20].

Na Áustria, a IA tem sido utilizada como ferramenta para estruturar informação com vista à análise e ao tratamento rápido e eficiente de documentos em contexto judicial[21]: a ferramenta de IA analisa o correio entrado sem qualquer contacto manual por parte dos funcionários do tribunal, extraindo metadados, identificando e reconhecendo procedimentos de arquivo e categorização de documentos. Funciona, também, como ferramenta de gestão de documentos digitais (o que é especialmente importante para a gestão de documentos não estruturados), como ferramenta de análise para investigação de dados (analisando e classificando metadados de qualquer tipo de dados e reconhecendo fluxos de comunicação e relações) e como ferramenta de anonimização automática de decisões judiciais (dados pessoais das partes).

[19] Ver os sítios na Internet https://predictice.com/ e https://www.caselawanalytics.com/ [última consulta em 19-10-2022].

[20] Xávier Rosin & Vasileios Lampos, *In-depth study on the use of AI in judicial systems, notably AI applications processing judicial decisions and data*, *in* European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their environment, Estrasburgo, CEPEJ – Commission Européene pour l'Efficacité de la Justice, 2018, p. 42, disponível em https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c [última consulta em 19-10-2022].

[21] Apresentação de Georg Stawa, *How is Austria approaching AI integration into judicial policies?*, Presidente da the CEPEJ e Chefe do Departamento de Estratégia, Consultoria Organizacional e Gestão da Informação, Ministério Federal da Constituição, Reformas, Desregulamentação e Justiça (2018), disponível em https://rm.coe.int/how-is-austria-approaching-ai-integration-into-judicial-policies-/16808e4d81 [última consulta em 19-10-2022].

O University College of London também realizou uma experiência[22] para prever decisões judiciais do Tribunal Europeu dos Direitos Humanos (TEDH), utilizando apenas informação textual extraída de trechos relevantes de acórdãos do TEDH. Os dados que serviram para treino da IA consistiam em elementos textuais extraídos de determinados casos e o resultado era a própria decisão dos juízes. A ferramenta de IA previu o resultado com 79% de exactidão. Os autores concluíram que "*a informação respeitante ao enquadramento factual do caso tal como formulado pelo Tribunal nas subsecções relevantes dos seus acórdãos constitui a parte mais importante obtendo em média um desempenho preditivo mais forte do resultado da decisão do Tribunal*", e que "*a correlação bastante robusta entre o resultado dos casos e o texto correspondente a padrões factuais contido nas subsecções relevantes é coerente com outro trabalho empírico levado a cabo no que respeita à tomada de decisão em casos difíceis e dá arrimo a intuições básicas pautadas pelo realismo jurídico*".

Outro campo profícuo no que tange à aplicação de soluções de IA é o das acções judiciais de baixo valor. Muitos países europeus instituíram já, ou tencionam instituir, algum tipo de serviço de RLL (em inglês *Online Dispute Resolution*, ODR), entre eles, os Países Baixos, o Reino Unido, a Letónia e a Estónia. Particularmente relevante é o exemplo da Estónia, que tenciona criar um sistema inteiramente independente de intervenção humana que profira decisões em acções de pequeno montante (até 7000€)[23]. Em teoria, as duas partes fariam *upload* dos documentos e outra informação relevante, e a plataforma de RLL proferiria a decisão, que poderia ser objecto de recurso para um juiz.

No Reino Unido também existem plataformas de RLL para acções de baixo valor, mas não são realmente uma solução de IA dado que é um juiz humano a decidir o litígio. A principal diferença entre este método e o método tradicional de decisão é o facto de todos os contactos entre o utilizador e o tribunal se efectuarem através de uma plataforma em linha. Outra diferença face à abordagem tradicional consiste na existência de uma espécie de mediadores em linha. Nas palavras do Professor Richard Susskind, *"indivíduos que olharão para os pedidos e promoverão a negociação entre as partes, possivelmente actuando como mediadores após algum tipo de orientação"*[24].

[22] Nikolaos Aletras, Dimitrios Tsarapatsanis, Daniel Preoţiuc-Pietro, Vasileios Lampos, *Predicting judicial decisions of the European Court of Human Rights: a Natural Language Processing perspective*, 2016, disponível em https://peerj.com/articles/cs-93/ [última consulta em 19-10-2022].

[23] Para mais pormenores, ver o artigo disponível em https://www.wired.com/story/can-ai-be-fair-judge-court-estonia-thinks-so/ [última consulta em 19-10-2022].

[24] Para uma explicação sintética e clara da RLL do Reino Unido, ver a explicação do Professor Richard Susskind em https://www.judiciary.uk/guidance-and-resources/what-is-odr-for-low-value-civil-claims-audio-version-2/ [última consulta em 19-10-2022].

A Letónia tem também uma solução de RLL para pedidos até 2100€. É um procedimento quase inteiramente escrito, apresentado on-line pelo requerente, e só se aplica a acções de pequeno montante para cobrança de montantes pecuniários ou alimentos, sendo que o pedido inicial tem de cumprir regras específicas referentes a estes processos (um determinado formulário-modelo ou, por exemplo, o demandante tem de indicar se requer uma audiência para análise da questão). Tal como na RLL britânica, a decisão é proferida por um juiz e não por qualquer tipo de ferramenta de IA[25].

Também a Comissão Europeia disponibiliza uma plataforma de RLL para ajudar a resolver litígios de consumo com origem em compras em linha sem necessidade de recorrer a um tribunal. Pode ser utilizada para qualquer litígio contratual emergente da aquisição de bens ou serviços em linha em que comerciante e consumidor estejam ambos sediados na União Europeia, na Noruega, na Islândia ou no Liechtenstein. Esta RLL rege-se pelo disposto no Regulamento (UE) nº 524/2013, do Parlamento Europeu e do Conselho, de 21 de Maio de 2013. Trata-se de uma resolução alternativa de litígios e a plataforma limita-se a facilitar a comunicação entre as partes e um órgão de resolução de litígios, sem necessidade de recurso ao tribunal. Uma das maiores vantagens desta RLL reside em disponibilizar traduções automáticas entre todas as línguas da UE, bem como informação e apoio ao longo de todo o procedimento[26].

A extinta RLL dos Países Baixos era a mais antiga na Europa de que temos conhecimento. O *e-Court* era uma iniciativa privada de resolução alternativa de litígios lançada em 2010 e, tal como o modelo pretendido pela Estónia, tratava-se de uma IA de emissão de decisões inteiramente automática para litígios de baixo valor. O credor apresentava a informação necessária (documentos) e a decisão era proferida sem qualquer intervenção humana. Não obstante, para dar início a um processo de execução, os utilizadores do *e-Court* tinham ainda de obter um título executivo emitido por mão humana: as decisões proferidas em linha de modo automatizado eram enviadas a um tribunal público, onde os funcionários de justiça recalculariam manualmente os montantes atribuídos[27].

[25] Para maior análise, ver https://e-justice.europa.eu/content_small_claims-42-lv-en.do?member=1 [última consulta em 19-10-2022].

[26] O endereço electrónico da plataforma de RLL: https://ec.europa.eu/consumers/odr/main/?event=main.home2.show [última consulta em 19-10-2022].

[27] Para mais pormenores, ver H.W.R. (Henriëtte) Nakad-Weststrate, H.J. (Jaap) van den Herik, A.W. (Ton) Jongbloed e Abdel-Badeeh M. Salem, *The Rise of the Robotic Judge in Modern Court Proceedings*, conference paper on the 7th International Conference on Information Technology (2015), pp. 59-67, acessível em https://www.researchgate.net/publication/300720949_The_Rise_of_the_Robotic_Judge_in_Modern_Court_Proceedings [última consulta em 19-10-2022].

Também merece menção o *Rechtwijzer*, outra solução de RLL produzida pelos Países Baixos: tinha por objectivo reduzir a conflitualidade no processo de divórcio ao reduzir a sua natureza contenciosa. O processo começava com uma fase de diagnóstico, seguida de uma fase de recolha de informação junto da parte que dava início ao processo e, por fim, a outra parte era convidada a juntar-se e a participar no processo de recolha de informação. A plataforma funcionava como um canal de comunicação entre as partes para que procurassem alcançar acordos sobre as questões em disputa. Apesar de se tratar de uma solução assente na negociação das partes, estas eram também informadas das normas jurídicas relativas aos acordos negociados (partilha de bens, alimentos, etc.). No final do processo em linha, estes acordos eram revistos por um terceiro imparcial (um advogado). O projecto *Rechtwijzer* terminou em 2017 e não parece existir uma explicação oficial para o seu término[28], mas, à época, os seus criadores encontravam-se a trabalhar com o *Raad voor Rechtsbijstand*[29] para criar uma nova plataforma com um novo tipo de ofertas. Essa nova plataforma existe aos dias de hoje[30], funcionando como uma plataforma de aconselhamento jurídico para variados temas (especialmente de índole social, laboral e familiar), assumindo-se como uma parceria com o *Raad voor Rechtsbijstand* e de coordenação e cooperação com diversas entidades[31].

Em Portugal, foi elaborado muito recentemente[32] um relatório pelo grupo de trabalho criado para delinear uma estratégia para resolver a pendência dos tribunais da jurisdição administrativa e fiscal, onde um dos eixos estratégicos definido foi, precisamente, o da transformação digital. No âmbito daquele relatório, um dos problemas identificados é o tempo que o magistrado perde

[28] Para mais pormenores, ver o texto disponível em https://law-tech-a2j.org/odr/rechtwijzer-why-online-supported-dispute-resolution-is-hard-to-implement/ [última consulta em 19-10-2022].

[29] Uma entidade independente de fins públicos, que tem uma missão, transponível para o nosso ordenamento jurídico, aproximada à da Ordem dos Advogados e à da Ordem dos Solicitadores e Agentes de Execução.

[30] No link https://rechtwijzer.nl/ [última consulta em 17-10-2022].

[31] Na definição da própria plataforma, as respostas são anónimas e não serão arquivadas. O utilizador preencherá um questionário que demora entre 10 e 20 minutos a preencher, de que poderá fazer download no seu próprio computador e enviar por e-mail. A plataforma não assegura um tempo de resposta, uma vez que dependerá da maior ou menor complexidade do caso concreto [https://rechtwijzer.nl/eerste-hulp-oplossingen/ – última consulta em 17-10-2022, com tradução automática para português do Google Chrome].

[32] Datado de 17 de Março de 2022 e disponível em https://justica.gov.pt/Portals/0/Ficheiros/Organismos/JUSTICA/RelatoriosI_IIGrupoTrabalhodosTAF17mar22.pdf?ver=ZNfTt3P419LYdowdOrvKw%3D%3D [última consulta em 19-10-2022].

na prática de actos de expediente no processo, propondo aquele grupo de trabalho alargar a competência da secretaria judicial na prática de actos de expediente, libertando os magistrados para a apreciação do mérito dos litígios. Do ponto de vista da modernização digital, afigura-se-nos que, havendo uma delegação de mais competências na secretaria judicial para a prática de actos processuais relevantes, o momento é oportuno para se automatizar alguns dos actos praticados até então pela secretaria. Criar-se-ia espaço e oportunidade para libertar os funcionários judiciais para o exercício de tarefas mais importantes e dignificantes da sua função e conhecimentos.

Outro dos problemas identificados pelo referido grupo de trabalho são as bagatelas jurídicas, propondo-se naquele relatório a simplificação do regime processual em acções de valor até 5000€. Tendo presente as experiências da Estónia e dos Países Baixos, com sistemas automatizados de resolução de litígios de baixo valor numa plataforma totalmente em linha, o momento é oportuno para ponderar alternativas à mera simplificação processual. É nossa convicção, aliás, que, face à pendência que ainda se verifica na jurisdição administrativa e fiscal[33/34], a mera simplificação processual não é necessariamente sinónimo de uma mais rápida resolução do litígio.

3.2. O que está a ser feito na União Europeia

Para além do que já foi posto em prática, em Abril de 2018 os Estados-Membros da UE assinaram uma declaração sobre a Cooperação em Matéria de

[33] Vejam-se os dados do último relatório anual publicado pelo Conselho Superior dos Tribunais Administrativos e Fiscais atinente ao ano de 2020, [disponível em http://www.cstaf.pt/documentos/Relatorio%20CSTAF%202020.pdf, última consulta em 19-10-2022] onde se refere (p. 7) "*em 31/12/2020 a taxa de congestão processual continuava a situar-se acima dos 200% e as pendências continuavam muito elevadas, com 20.532 processos no contencioso administrativo e 39.912 processos no contencioso tributário, num total de 60.444 processos pendentes*".

[34] Em sequência, refere a Presidente do CSTAF, na nota introdutória àquele relatório: "*Afigura-se-nos, pois, que a resolução, tão rápida e eficaz quanto possível, das pesadas cargas processuais e do volume de pendências acumuladas até 31/12/2020, só pode ser ultrapassada, em tempo útil, mediante medidas excepcionais, em particular através de adequada assessoria jurídica e técnica aos juízes, há muito prevista na lei, o que permitiria superar o desafio de alcançar um sistema de justiça administrativa e fiscal com a qualidade e a celeridade que o século XXI exige.*" Não podíamos estar mais de acordo com a proposta. Poderá ser este o momento para se introduzir mecanismos de assessoria com base em IA como apoio ao exercício da função judicial? É uma solução que tem vindo a ser publicamente apontada pela actual Ministra da Justiça para melhoria do funcionamento dos nossos tribunais (https://cnnportugal.iol.pt/catarina-sarmento-e-castro/ano-judicial/ministra-da-justica-quer-ter-a-inteligencia-artificial-a-ajudar-os-tribunais/20220420/626033d10cf26256cd1f96fc e https://www.cmjornal.pt/politica/detalhe/ministra-da-justica-defende-inteligencia-artificial-e-mais-recursos-para-justica-administrativa-e-fiscal, última consulta em 11-10-2022).

Inteligência Artificial[35], na qual os países acordaram em construir uma UE voltada para o progresso e o investimento em matéria de IA, assim como para o progresso no sentido da criação de um Mercado Único Digital.

Nesse mesmo mês, a Comissão Europeia emitiu uma comunicação sobre Inteligência Artificial para a Europa[36]. Nessa comunicação, a Comissão afirma que a UE *"deve adoptar uma abordagem coordenada, a fim de tirar o máximo partido das oportunidades oferecidas pela IA e fazer face aos novos desafios que esta acarreta"*[37], apoiando explicitamente a investigação em IA, nomeadamente quanto à "*administração pública (nomeadamente a justiça)*"[38]. Mais tarde, nesse ano, a Comissão para a Eficácia da Justiça na Europa, do Conselho da Europa (Council of Europe European Commission for the efficiency of justice – CEPEJ) lançou a *Carta Europeia de Ética sobre o Uso da Inteligência Artificial em Sistemas Judiciais e seu ambiente*[39].

Apesar do caminho trilhado pela UE ao longo dos últimos anos, existe ainda um longo percurso no que respeita à utilização da tecnologia de IA nas decisões judiciais no âmbito da UE.

4. Desafios éticos e legais

Tendo presente que a implementação da IA não é algo para um futuro distante, mas para o nosso tempo, a CEPEJ adoptou formalmente cinco princípios fundamentais relativos à utilização da IA nos sistemas judiciais e seu ambiente. Estes princípios visam assegurar o respeito pela Convenção Europeia dos Direitos Humanos (CEDH) e pela Convenção relativa à Protecção dos Dados Pessoais (CPDP), ao enquadrarem as políticas públicas nessa área e assegurarem que o processamento da IA respeita princípios como a *transparência*, a *imparcialidade* e a *igualdade*, com certificação através de uma avaliação levada a cabo por peritos externos e independentes.

Estes princípios, porém, não devem ser definidos de forma rígida. A CEPEJ pretende sujeitá-los a monitorização e supervisão, com um objectivo de contínua melhoria das práticas. Por ora, os cinco princípios são os seguintes: (1) respeito pelos direitos fundamentais: assegurar que a concepção e a imple-

[35] Disponível em https://ec.europa.eu/jrc/communities/en/node/1286/document/eu-declaration-cooperation-artificial-intelligence [última consulta em 19-10-2022].

[36] Disponível em https://ec.europa.eu/digital-single-market/en/news/communication-artificial-intelligence-europe [última consulta em 19-10-2022].

[37] *Supra* nota 36, p. 3.

[38] *Supra* nota 36, p. 8.

[39] CEPEJ, *European Ethical Charter on the use of Artificial Intelligence in judicial systems and their environment* (2018), disponível em https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c [última consulta em 19-10-2022].

mentação de ferramentas e serviços de IA são compatíveis com os direitos fundamentais; (2) não discriminação: especificamente, impedir o surgimento ou a intensificação de qualquer discriminação entre pessoas ou grupos de pessoas; (3) princípio da qualidade e segurança: no que respeita ao tratamento de decisões e dados judiciais, utilizar fontes certificadas e dados intangíveis com modelos elaborados de modo multidisciplinar e num ambiente tecnológico seguro; (4) princípio da transparência, imparcialidade e equidade: tornar os métodos de tratamento de dados acessíveis e compreensíveis e autorizar auditorias externas; e (5) princípio "sob controlo do utilizador": prevenir uma abordagem prescritiva e assegurar que os utilizadores sejam actores informados e com controlo sobre as escolhas feitas. Estes cinco princípios abordam algumas das principais questões éticas suscitadas pela utilização de ferramentas de IA num sistema judicial e seu ambiente, bem como os princípios e obstáculos jurídicos que rodeiam esta área no âmbito da UE.

Em Portugal, foi publicada a Carta Portuguesa de Direitos Humanos na Era Digital pela mão da Lei nº 27/2021, de 17 de Maio. Destacamos o artigo 9º daquela Carta, que consagra os principais princípios e direitos no uso da inteligência artificial e de robôs. A norma contida no nº 1 daquele artigo destaca, à semelhança da Carta Europeia, o respeito pelos direitos fundamentais, garantindo-se *"um justo equilíbrio entre os princípios da explicabilidade, da segurança, da transparência e da responsabilidade, que atenda às circunstâncias de cada caso concreto e estabeleça processos destinados a evitar quaisquer preconceitos e formas de discriminação"*. Consagra-se, ainda, uma forma do princípio *user under control* quando se prevê no nº 2 daquele artigo que as *"decisões com impacto significativo na esfera dos destinatários que sejam tomadas mediante o uso de algoritmos devem ser comunicadas aos interessados"*. O nº 3 do artigo 9º da Carta Portuguesa de Direitos Humanos na Era Digital, por sua vez, estabelece que são aplicáveis à criação e uso de robôs, especificamente, os princípios da beneficência, da não-maleficência, do respeito pela autonomia humana e pela justiça, bem como os princípios e valores consagrados no artigo 2º do Tratado da União Europeia, designadamente a não discriminação e a tolerância[40].

[40] Princípios, de resto, que não são uma total novidade. Os grandes princípios da bioética tiveram como precursor o Relatório Belmont, de 1979, redigido pela Comissão Nacional para a Protecção de Sujeitos Humanos de Pesquisa Biomédica e Comportamental dos Estados Unidos, onde se identificam os princípios éticos básicos e directrizes que abordam questões éticas decorrentes da condução de pesquisas com seres humanos. O relatório completo pode ser encontrado, no seu original, em https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/index.html [última consulta em 19-10-2022].

4.1. A jurisprudência do TEDH e suas directrizes (*guidelines*)

Um dos maiores desafios criados pela utilização da IA nos sistemas judiciais e seu ambiente é o do respeito pelos direitos fundamentais e princípios consagrados, sejam os da Constituição da República Portuguesa, sejam os da CEDH. Como se afirma na Carta de Ética e na Carta Portuguesa, tais soluções têm de respeitar direitos individuais como o *"direito a um processo equitativo (em especial o direito ao juiz natural estabelecido pela lei, o direito a um tribunal independente e imparcial e à igualdade de armas em processo judicial) e, onde não tenha sido adoptada diligência bastante para proteger dados comunicados em dados abertos, o direito ao respeito pela vida privada e familiar"*[41].

O respeito pelos direitos fundamentais implica, desde logo, o respeito pelo direito a um processo equitativo (artigo 6º da CEDH). A densificação daquele direito tem sido feita ao longo do tempo pelo TEDH, estabelecendo importantes directrizes para todos os Estados subscritores da CEDH[42].

Em primeiro lugar, o TEDH tem vindo a interpretar que apenas existe respeito pelo direito a um processo equitativo quando seja cumprido o direito a uma audiência pública. Em *Vernes c. França*[43], o TEDH considerou existirem diversas violações do nº 1 do artigo 6º, uma das quais consistindo na violação resultante da impossibilidade de os requerentes solicitarem uma audiência pública[44] e uma outra resultante da falta de imparcialidade do órgão administrativo como consequência da ausência de indicação da sua composição[45]. O Tribunal recordou que a audiência pública é um princípio fundamental consagrado no nº 1 do artigo 6º da Convenção. Este princípio pode sofrer compressões justificadas em especial pelos interesses da vida privada das partes ou da salvaguarda da justiça (*Diennet c. França*, nº 18160/91, 26 de Setembro de 1995) ou pela natureza das questões submetidas ao tribunal no âmbito do caso concreto (*Miller c. Suécia*, nº 55853/00, 8 de Fevereiro de 2005; *Göç c. Turquia*, nº 36590/97, 11 de Julho de 2002). O TEDH concluiu que, na ausência de uma audiência pública, o direito do requerente a um processo equitativo não fora assegurado. Naquele aresto, o TEDH recordou também que, para os efeitos do artigo nº 1 do artigo 6º, a imparcialidade deve ser avaliada com base

[41] *Supra* nota 39, p. 8.

[42] Veja-se também a súmula sobre o direito a um processo equitativo e público do artigo 6º, nº 1, da CEDH de Marco Carvalho Gonçalves, "Direito a um processo equitativo e público", *in* AA. VV., *Comentário da Convenção Europeia dos Direitos Humanos e dos Protocolos Adicionais*, org. P. Pinto de Albuquerque, vol. II, Lisboa, Universidade Católica Editora, 2019, pp. 931-964.

[43] Queixa nº 30183/06, Acórdão de 20 de Janeiro de 2011.

[44] *Supra* nota 43, §§ 30-31.

[45] *Supra* nota 43, §§ 41-44.

numa abordagem subjectiva, permitindo determinar a convicção do julgador em tal ocasião, e também de acordo com uma abordagem objectiva, de forma a assegurar que o mesmo oferece garantias suficientes para excluir qualquer dúvida legítima a esse respeito[46]. E, assim sendo, o Tribunal concordou com o requerente quanto ao facto de a falha do Estado em identificar todos os membros do órgão administrativo que haviam deliberado era de molde a pôr em causa a sua imparcialidade, o que implica que a imparcialidade é também assegurada pela identificação dos juízes que venham a proferir a decisão. Este é um factor de relevância inegável se um caso for julgado por uma IA.

Em *Golder c. Reino Unido*, o TEDH reconheceu o direito de acesso a um tribunal, mas afirmou também que este não é absoluto, admitindo algumas limitações implícitas [47/48]. A jurisprudência firmada em *Deweer c. Bélgica*[49] tornou claro que o direito a um tribunal é visto como um elemento do direito a um processo equitativo, consagrado no nº 1 do artigo 6º, e não é mais absoluto em matéria penal do que em matéria civil.

Em *Kontalexis c. Grécia*[50], o TEDH recorda que, ao abrigo do nº 1 do artigo 6º, um tribunal deve ser sempre estabelecido por lei, o que reflecte o princípio do Estado de Direito, inerente a todo o sistema da Convenção e seus protocolos. O Tribunal afirma que a um órgão que não tenha sido estabelecido de acordo com a vontade do legislador faltará necessariamente a legitimação necessária numa sociedade democrática. Em *DMD GROUP, a.s, c. Eslováquia*, o TEDH reiterou a noção de um tribunal estabelecido por lei e da independência judicial e certeza jurídicas para o Estado de Direito exigirem uma particular clareza das regras aplicadas em qualquer caso e garantias claras para assegurar objectividade e transparência, de modo a evitar qualquer aparência

[46] Neste sentido, também a decisão proferida em Ivanovski c. *"Antiga República Jugoslava da Macedónia"*, queixa nº 29908/11, Acórdão de 21 de Janeiro de 2016. É jurisprudência assente do TEDH que a imparcialidade tem de ser determinada *"de acordo com um teste subjectivo, em que se deve atentar na convicção e comportamento pessoais de um determinado juiz, isto é, se o juiz tinha qualquer preconceito ou tendenciosidade pessoal num determinado caso* [a imparcialidade tem de ser presumida até prova em contrário]*; e também de acordo com um teste objectivo, isto é, determinando se o tribunal propriamente dito e, entre outros aspectos, a sua composição, ofereciam garantias suficientes para afastar qualquer dúvida legítima a respeito da sua imparcialidade"*.

[47] Queixa nº 4451/70, Acórdão de 21 de Fevereiro de 1975, § 38.

[48] Mas mesmo quando existem limitações implícitas, outros aspectos do direito consagrado no nº 1 do artigo 6º devem ser respeitados, como o direito a ser ouvido por um tribunal dentro de um prazo razoável (cf. *Kart c. Turquia*, nº 8917/05, 3 de Dezembro de 2009).

[49] Queixa nº 6903/75, Acórdão de 27 de Fevereiro de 1980, § 49.

[50] Queixa nº 59000/08, Acórdão de 31 de Maio de 2011.

de arbitrariedade na distribuição de processos a juízes específicos[51]. Noutra decisão[52], o Tribunal considerou que os casos de distribuição discricionária de um processo (no sentido de uma distribuição cujas modalidades não estão previstas em lei), põem em risco a aparência de imparcialidade, ao dar azo a especulação sobre a influência política ou de outro tipo sobre o tribunal destinatário da distribuição e o juiz competente, ainda que a distribuição do caso a um juiz específico em si mesma tenha seguido critérios transparentes.

Quanto ao princípio da igualdade de armas, o Tribunal afirmou mais recentemente[53] que *"o princípio do contraditório e o princípio da igualdade de armas, os quais estão intimamente interligados, são componentes fundamentais do conceito de um 'processo equitativo' no sentido que lhe é dado pelo nº 1 do artigo 6º da Convenção. Eles exigem um 'justo equilíbrio' entre as partes: a cada parte deve ser dada oportunidade razoável de expor o seu caso em circunstâncias que não a coloquem em especial desvantagem face ao(s) seu(s) oponente(s)"*.

Um dos mais importantes direitos consagrados no nº 2 do artigo 6º é o da presunção de inocência. O TEDH interpreta-o como o direito a ser-se presumido inocente até prova da culpabilidade em conformidade com a lei. É *"visto como uma garantia processual no contexto do próprio julgamento penal"*, mas a presunção da inocência também *"impõe requisitos relativos, nomeadamente, ao ónus da prova; à presunção legal de facto e de direito; à protecção contra a auto-incriminação; à publicidade na fase pré-julgamento; e a expressões prematuras, por parte do tribunal de julgamento ou de outros funcionários públicos, quanto à culpabilidade de um arguido"*[54].

4.2. O acórdão Sigurdur Einarsson "e.o.c" (e outros contra) Islândia

Conjugando quase todas as disposições sobre direitos consagrados no artigo 6º, o TEDH foi confrontado em *Sigurdur Einarsson a. o. c. Islândia*[55] quanto a potenciais violações do referido artigo num processo penal em que os arguidos alegavam ter-lhes sido negado acesso pleno ao processo detido pela acusação.

O processo penal em causa dizia respeito a uma conduta alegadamente criminosa ligada ao colapso de um dos maiores bancos do país durante a

[51] Queixa nº 19334/03, Acórdão de 5 de Outubro de 2010, § 66.

[52] *Miracle Europe KFT c. Hungria*, Queixa nº 57774/13, Acórdão de 12 de Janeiro de 2016, § 58

[53] *Prebil c. Eslovénia*, Queixa nº 29278/16, Acórdão de 19 de Março de 2019.

[54] *Kangers c. Letónia*, Queixa nº 35726/10, Acórdão de Março de 2019. Igualmente, *Lolov c. Bulgária*, Queixa nº 6123/11, Acórdão de 21 de Fevereiro de 2019; *Allenet de Ribemont c. França*, Queixa no 15175/89, Acórdão de 10 de Fevereiro de 1995; *Viorel Burzo c. Roménia*, Queixas nº 75109/01 e 12639/02, Acórdão de 30 de Junho de 2009; *Lizaso Azconobieta c. Espanha*, Queixa nº 28834/08, Acórdão de 28 de Junho de 2011.

[55] Queixa nº 39757/15, Acórdão de 4 de Junho de 2019.

crise financeira que atingiu a Islândia em 2008. A investigação durou quase três anos e levou a uma vasta recolha de dados, incluindo dados apreendidos ao abrigo de um mandado judicial de busca. Para levar a cabo a pesquisa nos dados electrónicos (*e-discovery*), a acusação utilizou uma ferramenta de IA chamada *Clearwell*, cujos resultados foram exportados e marcados como "documentos de investigação". Os requerentes queixaram-se de que nunca tinham tido a oportunidade de analisar os documentos submetidos ao tribunal e que lhes fora negada a possibilidade de pesquisar os mesmos dados utilizando o sistema electrónico empregue pela acusação. Isto consubstanciava, na sua opinião, uma violação do princípio da igualdade de armas (estabelecido no nº 1 e na alínea b) do nº 3 do artigo 6º), dado que eles deveriam ter tido as mesmas oportunidades que a acusação de aceder à prova e seleccioná-la a partir do conjunto de documentos recolhidos pela polícia durante a investigação.

O Tribunal considerou não existir qualquer violação do artigo 6º quanto a dados em massa não marcados, afirmando que nesse aspecto a acusação não detinha qualquer vantagem face à defesa, por não se tratar de um caso de *não divulgação*. Quanto aos dados marcados, tinham sido analisados pelos investigadores, manualmente e através do *Clearwell*, com vista a escolher o material que deveria integrar o processo de investigação e a acusação. O Tribunal reconheceu que esta selecção fora feita somente pela acusação, sem o envolvimento da defesa e sem qualquer supervisão judicial, e também que novas pesquisas nos dados pela defesa teriam sido tecnicamente possíveis e apropriadas para uma busca por prova potencialmente ilibatória. O Tribunal concluiu assim que *"qualquer recusa em autorizar a defesa a realizar novas buscas nos documentos 'marcados' suscitaria em princípio um problema à luz da alínea b) do nº 3 do artigo 6º no que respeita à concessão de meios adequados para a preparação da defesa"*[56].

A jurisprudência *Sigurdur Einarsson a. o. c. Islândia* é fundamental para a conjugação das ferramentas de IA com os direitos consagrados no artigo 6º:

[56] *Supra* nota 55, §§ 85-91. Apesar da conclusão mencionada, o Tribunal afastou a violação da alínea b) do nº 3 do artigo 6º da CEDH por ter considerado que os requerentes não tinham formalmente procurado um tribunal com vista a ter acesso àqueles documentos e, nessa medida, não lhes tinha sido negado globalmente um processo equitativo. O acórdão mereceu uma declaração de voto por parte do Juiz Pavli. Com especial interesse, veja-se o § 10 (tradução livre): *"Vale a pena recordar neste ponto que o que está em causa é um pilar fundamental do processo penal equitativo, nomeadamente a igualdade de armas. À luz deste princípio cardeal, a abordagem global da maioria parece insuficientemente harmonizada com a complexidade da revelação electrónica em processo penal (ou civil) envolvendo elevados volumes de dados; à utilização de ferramentas tecnológicas modernas neste contexto; e às suas implicações combinadas para a igualdade de armas. O pressuposto de que as regras de revelação padrão devem aplicar-se sem alterações neste contexto carece, no mínimo, de ser testada."*

estabeleceu um claro princípio de que, onde sejam usadas ferramentas de IA para tratar dados em massa e exista informação extraída através desse mecanismo, o princípio da igualdade de armas (nº 1 do artigo 6º) e o direito a dispor de tempo e meios adequados para a preparação da defesa exigem que o arguido/demandado (num processo penal ou civil, como afirmado em *Deweer c. Bélgica*) tenha o direito de participar na escolha da informação e tenha o direito de realizar a sua própria pesquisa nos dados, utilizando a mesma ferramenta que a acusação.

Para garantir o direito a um processo judicial equitativo, a utilização de uma ferramenta de IA nesse contexto terá sempre de garantir o direito a uma audiência pública; a imparcialidade do julgador, avaliada com base numa abordagem *subjectiva* e também de acordo com uma abordagem *objectiva*; a identificação do decisor; o direito de acesso a um tribunal (embora não seja um direito absoluto); a garantia de legitimação mediante criação por lei (princípio do Estado de Direito); a clareza das regras aplicadas em qualquer caso e garantias para assegurar objectividade e transparência, de modo a evitar qualquer aparência de arbitrariedade na distribuição de processos a juízes específicos, em nome do princípio da transparência; o princípio do contraditório; a igualdade de armas, seja na acepção de que a cada parte deve ser dada oportunidade razoável de expor o seu caso em circunstâncias que não a coloquem em especial desvantagem face ao(s) seu(s) oponente(s), seja no direito a dispor de tempo e meios adequados para a preparação da defesa, utilizando os mesmos meios de pesquisa, caso uma IA tenha sido utilizada no âmbito na investigação criminal.

4.3. A opacidade das caixas negras

Conexo com as questões atrás mencionadas está o problema das caixas negras (*black boxes*), isto é, quando um agente de IA produz um resultado que os humanos ou mesmo os seus criadores não conseguem perceber nem explicar como foi obtido, embora a sua exactidão ultrapasse as melhores decisões ou previsões humanas[57]. O resultado de uma função algorítmica inexplicável poderá ser utilizado numa decisão judicial?

[57] Veja-se o estudo de Alvin Rajkomar, Eyal Oren *et al.*, denominado "Scalable and accurate deep learning with electronic health records", *Npj* (Nature Partner Journals), *Digital Medicine*, 1, nº 18, 2018, disponível em https://www.nature.com/articles/s41746-018-0029-1 (última consulta, 19-10-2022). Neste estudo foi apresentado o surpreendente desempenho de um algoritmo preditivo de aprendizagem profunda, que analisou os relatórios clínicos de 216.221 pacientes, incluindo notas soltas dos médicos assistentes, sem intervenção humana na selecção dos dados. O algoritmo criou quatro categorias e previu, com um nível de exactidão de 75% a 94% – superando o melhor

Uma decisão judicial sem especificação dos seus fundamentos é nula[58]. Pelo que não é admissível a resposta *"computer says no"*. Mas o princípio da explicabilidade não impõe necessariamente que saibamos como um agente de IA chegou aos seus resultados, nem implica que se abra a caixa negra. Impõe apenas – o que já não é pouco – que o resultado obtido tenha suficiente poder explicativo. Uma proposta para o alcançar poderá ser através da utilização de justificações contrafactuais[59].

Na lógica formal, uma condicional contrafactual é uma proposição no modo subjuntivo, tal como *"se o osso estivesse partido, o raio X teria resultado diferente"*. São condicionais contrárias aos factos, implicando a sugestão de que o antecedente de tal condição é falso. Uma vez que as contrafactuais podem estar relacionadas com todo o tipo de mundos possíveis, é importante que o mundo que utilizamos seja próximo do mundo real, isto é, deve ser o mundo mais próximo possível[60].

A explicabilidade dos resultados provenientes de uma caixa negra pode ser alcançada através de algoritmos criados para confundir classificadores, gerando um ponto de dados sintéticos próximo a um já existente, de modo a que esse novo ponto de dados sintéticos possa ser classificado de forma diferente do original, descrevendo as dependências entre uma decisão particular e factos externos específicos. O agente de IA vai então pesar esse novo dado sintético ou "contra-o-facto" que resultou do dado original. É importante, por exemplo, para aferir a robustez do nexo de causalidade[61].

Imaginemos uma situação de rebentamento de uma conduta de água que provoca danos. O teste do algoritmo para aferir o nexo de causalidade com o auxílio de justificações contrafactuais será o seguinte: num cenário próximo ao que está a ser julgado (mundo possível), em que a conduta de água não

prognóstico médico humano – os pacientes que se encontravam em risco de mortalidade hospitalar (93%-94%), os que seriam submetidos a reinternamento não planeado a 30 dias (75-76%), os casos de hospitalização prolongada (85-86%) e as situações de alta médica (90%). Já com o conhecimento da previsão do algoritmo, a informação clínica foi reexaminada pelos médicos que continuaram sem perceber ou conseguir explicar a elevada percentagem de acerto das previsões do algoritmo. Esta função algorítmica foi designada como não-explicável (*non-explainable algorithmics*).

[58] Artigo 615º, nº 1, alínea *c*) do Código de Processo Civil.

[59] Sandra Wachter, Brent Mittelstadt, Chris Russel, "Counterfactual Explanations Without Opening the Black Box: Automated Decisions and the GDPR"*, in Harvard Journal of Law & Technology,* 31 (2), 2019.

[60] Simon Blackburn, *The Oxford Dictionary of Philosophy*, Oxford, Oxford University Press, 1996, pp. 85-86.

[61] Veja-se o caso do programa Resnet, mencionado por Sandra Wachter, Brent Mittelstadt, Chris Russel, *ob. cit.*

rebentou (condicional ou dado sintético), o lesado não iria sofrer os danos que alega (perturbação adversária).

Trata-se de uma proposta de solução mais realista para ultrapassar a opacidade das caixas negras do que um direito à demonstração plena do funcionamento interno do algoritmo. As decisões de redes neuronais artificiais partem de milhares de funções e milhões de parâmetros que controlam o seu comportamento. Já a memória humana de curto prazo consegue reter, em média, cerca de sete objectos (números, palavras, letras) em simultâneo. A transparência total do agente de IA dificilmente será compreensível para a maioria das pessoas, a que teríamos de adicionar os direitos de propriedade intelectual sobre os algoritmos.

Do mesmo modo que não é possível pedir a um juiz humano que abra o seu cérebro e descreva como chegou à decisão, também não podemos esperar total transparência dos algoritmos de IA, a qual só poderia ser obtida à custa do seu próprio desempenho. O que não significa uma confiança cega e acrítica no agente de IA.

5. Conclusão

A aversão a algoritmos susceptíveis de afectar os seres humanos inclui-se na frequente hostilização ao que é sintético ou artificial. Esta tendência verifica-se, por exemplo, na alimentação, na energia e, no caso que nos ocupa, também no processo de decisão. Todavia, não se prescindindo dos juízos humanos, mais ou menos intuitivos, é eticamente muito questionável negar o auxílio de algoritmos que cometerão menos erros no processo decisório, sobretudo em ambientes regulares e estáveis.

A IA é uma realidade actual e muito seguramente, num futuro próximo, será um preciso auxiliar do juiz, quer na tramitação processual quer na eficiência e qualidade das decisões. A sua utilização no processo judicial deve ser encorajada, começando pelas situações que suscitem menor desconfiança. Mas não devemos apostar numa singularidade jurídica, em que o auxílio da IA à decisão judicial acertará em todos os casos e eliminará toda a incerteza jurídica ou a própria intervenção humana.

A posição prudentemente optimista por nós defendida pode ser sintetizada na vetusta sentença do poeta grego Arquíloco, popularizada por Isaiah Berlin, de que "*a raposa sabe muitas coisas, mas o ouriço sabe uma coisa muito importante*"[62].

[62] Isaiah Berlin, "The Hedgehog and the Fox", 1953, *in The Proper Study of Mankind: An Anthology of Essays*, New York, Farrar, Straus and Giroux, 2000, pp.436 e ss.

Para a boa decisão da causa, a IA será a raposa, pois sabe muitas coisas: armazenar informação, retirar taxas base, alcançar juízos probabilísticos, compilar o acervo jurisprudencial, executar automatismos sem dificuldade, reduzir vieses cognitivos. Mas para as partes em litígio, o juiz será o ouriço, pois tem a obrigação de saber uma coisa muito importante: o caso particular submetido a julgamento.

Bibliografia

Aletras, Nikolaos/Tsarapatsanis, Dimitrios/Preoțiuc-Pietro, Daniel/Lampos, Vasileios, *Predicting judicial decisions of the European Court of Human Rights: a Natural Language Processing perspective*, 2016.

Altavilla, Enrico, *Psicologia Judiciária – o processo psicológico e a verdade judicial*, vol. I, 2ª ed., trad. Fernando Miranda, Coimbra, Almedina, 2007.

Berlin, Isaiah, "The Hedgehog and the Fox" (1953), in *The Proper Study of Mankind: An Anthology of Essays*, New York, Farrar, Straus and Giroux, 2000.

Blackburn, Simon, *The Oxford Dictionary of Philosophy*, Oxford, Oxford University Press, 1996.

Chace, Callum, *The Artificial Intelligence and the Two Singularities*, Boca Raton, London, New York, CRC Press, 2018.

Fedoll, Jordi Nieva, *La valoración de la prueba*, Madrid, Marcial Pons, 2010.

Fedoll, Jordi Nieva, *Inteligencia artificial y processo judicial*, Madrid, Marcial Pons, 2018.

Floridi, Luciano, *The Fourth Revolution: how the infosphere is reshaping human reality*, Oxford, Oxford University Press, 2014.

Floridi, Luciano *et. al.*, "AI4People–An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations", *Minds and Machines*, Springer, 2018.

Gonçalves, Marco Carvalho, "Direito a um processo equitativo e público", in AA. VV., *Comentário da Convenção Europeia dos Direitos Humanos e dos Protocolos Adicionais*, org. P. Pinto de Albuquerque, vol. II, Lisboa, Universidade Católica Editora, 2019.

Khaneman, Daniel, *Thinking, Fast and Slow*, trad. port. Pedro Vidal, *Pensar, Depressa e Devagar*, 12ª ed., Lisboa, Temas e Debates – Círculo de Leitores, reimpressão, 2021.

Martins, Hermínio, *Experimentum Humanum*, Lisboa, Relógio D'Água, 2011.

Nakad-Weststrate, H.W.R. (Henriëtte)/Van den Herik, H.J. (Jaap)/Jongbloed, A.W. (Ton)/ Salem, Abdel-Badeeh M., The Rise of the Robotic Judge in Modern Court Proceedings", conference paper on the *7th International Confer-ence on Information Technology*, 2015, acessível em https://www.researchgate.net/publication/300720949_The_Rise_of_the_Robotic_Judge_in_Modern_Court_Proceedings (última consulta em 19-10-2022).

Rajkomar, Alvin/Oren, Eyal/*et. al.*, "Scalable and accurate deep learning with electronic health records", *Npj (Nature Partner Journals)*, Digital Medicine, 1, nº 18, 2018, disponível em https://www.nature.com/articles/s41746-018-0029-1 (última consulta, 19-10-2022).

Reis, Maria Anabela Bento Marinho dos, *A Memória do Testemunho e a Influência das Emoções na Recolha e Preservação da Prova*, dissertação de Doutoramento em Ciências e Tecnologias da Saúde Especialidade em Desenvolvimento Humano e Social, Universidade de Lisboa, Faculdade de Medicina, 2014.

Rosin, Xávier/Lampos, Vasileios, "In-depth study on the use of AI in judicial systems, notably AI applications processing judicial decisions and data", in *European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their environment*, Estrasburgo, CEPEJ – Commission Européene pour l'Efficacité de la Justice, 2018.

Russell, Stuart J./Norvig, Peter, *Artificial Intelligence: A Modern Approach*, 3rd ed., Pearson Education Limited, England, 2016.

Sampaio, Elisa Alfaia/ Seixas, João J./ Gomes, Paulo Jorge, "Artificial Intelligence and the Judicial Rulling", *Themis Annual Journal*, vol. 1, issue 1, October 2019, disponível em https://www.ejtn.eu/MRDDocuments/TAJ%202019.pdf (última consulta em 10-10-2022).

Sousa, Luís Filipe Pires de, *Prova Testemunhal*, Coimbra, Almedina, 2013.

Stawa, Georg, "How is Austria approaching AI integration into judicial policies?", 2018, disponível em https://rm.coe.int/how-is-austria-approaching-ai-integration-into-judicial-policies-/16808e4d81 (última consulta em 19-10-2022).

Taruffo, Michele, *La Prueba*, trad. Laura Manríquez/Ferrer Beltrán, Madrid, Marcial Pons, 2008.

Tversky, Amos/Khaneman, Daniel, "Juízo sob incerteza: heurísticas e enviesamentos", apêndice A in Daniel Khaneman, *Thinking, Fast and Slow*, trad. port. Pedro Vidal, *Pensar, Depressa e Devagar,* 12ª ed., Lisboa, Temas e Debates – Círculo de Leitores, reimpressão, 2021.

Watcher, Sandra/ Mittelstadt, Brent/Russel, Chris, "Counterfactual Explanations Without Opening the Black Box: Automated Decisions and the GDPR", *in Harvard Journal of Law & Technology*, 31 (2), 2019.

Jurisprudência do TEDH

Acórdão de 21 de Fevereiro de 1975, queixa nº 4451/70, *Golder c. Reino Unido*
Acórdão de 27 de Fevereiro de 1980, queixa nº 6903/75, *Deweer c. Bélgica*
Acórdão de 10 de Fevereiro de 1995, queixa nº 15175/89, *Allenet de Ribemont c. França*
Acórdão de 30 de Junho de 2009, queixas n. os 75109/01 e 12639/02, *Viorel Burzo c. Roménia*
Acórdão de 5 de Outubro de 2010, queixa nº 19334/03, *DMD GROUP, a.s, c. Eslováquia*
Acórdão do 20 de Janeiro de 2011, queixa nº 30183/06, *Vernes c. França*
Acórdão de 31 de Maio de 2011, queixa nº 59000/08, *Kontalexis c. Grécia*
Acórdão de 28 de Junho de 2011, queixa nº 28834/08, *Lizaso Azcono-bieta c. Espanha*
Acórdão de 12 de Janeiro de 2016, queixa nº 57774/13, *Miracle Europe KFT c. Hungria*
Acórdão de 21 de Janeiro de 2016, queixa nº 29908/11, *Ivanovski c. "Antiga República Jugoslava da Macedónia"*
Acórdão de 21 de Fevereiro de 2019, queixa nº 6123/11, *Lolov c. Bulgária*
Acórdão de 14 de Março de 2019, queixa nº 35726/10, *Kangers c. Letónia*
Acórdão de 19 de Março de 2019, queixa nº 29278/16, *Prebil c. Eslovénia*
Acórdão de 4 de Junho de 2019, queixa nº 39757/15, *Sigurdur Einarsson a. o. c. Islândia*
Todos disponíveis em https://hudoc.echr.coe.int.

Relatórios/Grupos de trabalho

CEPEJ, *European Ethical Charter on the use of Artificial Intelligence in judicial systems and their environment* (2018), disponível em https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c (última consulta em 19-10-2022).

Conselho Superior dos Tribunais Administrativos e Fiscais, *Relatório Anual*, 2020, disponível em http://www.cstaf.pt/documentos/Relatorio%20CSTAF%202020.pdf, (última consulta em 19-10-2022).

Grupo de Trabalho dos Tribunais Administrativos e Fiscais, *I e II Relatórios Intercalares, Objectivos | Medidas*, 17-03-2022, disponível em https://justica.gov.pt/Portals/0/Ficheiros/Organismos/JUSTICA/RelatoriosI_IIGrupoTrabalhodosTAF17mar22.pdf?ver=ZN-fTt3P419LYdowdOrvKw%3D%3D, (última consulta, 19-10-2022).

High-Level Expert Group on AI, *A Definition of AI: Main Capabilities and Disciplines*, 2019.

High-Level Expert Group on AI, *Ethics Guidelines for Trustworthy AI*, 2019, disponível em https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai (última consulta em 11-10-2022).

IDC White Paper, *The Digitization of the World From Edge to Core*, 2018, disponível em https://www.seagate.com/files/www-content/our-story/trends/files/idc-seagate-dataage-whitepaper.pdf (última consulta em 10-10-2022).

Páginas *web* consultadas

https://www.luminance.com/ (última consulta em 19-10-2022).

https://www.researchgate.net/figure/Artificial-neural-network-of-multiple-layers-and--outputs-31_fig2_331097835 (última consulta, 10-10-2022).

https://www.cam.ac.uk/research/features/helping-police-make-custody-decisions-using--artificial-intelligence (última consulta em 19-10-2022).

https://predictice.com/ (última consulta em 19-10-2022).

https://www.caselawanalytics.com (última consulta em 19-10-2022).

https://www.wired.com/story/can-ai-be-fair-judge-court-estonia-thinks-so/ (última consulta em 19-10-2022).

https://e-justice.europa.eu/content_small_claims-42-lv-en.do?member=1 (última consulta em 19-10-2022).

https://ec.europa.eu/consumers/odr/main/?event=main.home2.show (última consulta em 19-10-2022).

https://law-tech-a2j.org/odr/rechtwijzer-why-online-supported-dispute-resolution-is--hard-to-implement (última consulta em 19-10-2022).

https://rechtwijzer.nl (última consulta em 17-10-2022).

https://cnnportugal.iol.pt/catarina-sarmento-e-castro/ano-judicial/ministra-da-justica--quer-ter-a-inteligencia-artificial-a-ajudar-os-tribunais/20220420/626033d10cf26256cd1f96fc (última consulta em 11-10-2022).

https://www.cmjornal.pt/politica/detalhe/ministra-da-justica-defende-inteligencia-artificial-e-mais-recursos-para-justica-administrativa-e-fiscal (última consulta em 11-10-2022).

https://ec.europa.eu/jrc/communities/en/node/1286/document/eu-declaration-cooperation-artificial-intelligence (última consulta em 19-10-2022).
https://ec.europa.eu/digital-single-market/en/news/communication-artificial-intelligence-europe (última consulta em 19-10-2022).

A prova resultante de "software de aprendizagem automática"

Machine Learning Evidence[1]

Fernando Silva Pereira[*]

RESUMO: A "aprendizagem automática" é um campo da inteligência artificial que dá aos computadores a capacidade de aprenderem sem serem explicitamente programados, colocando-se o problema da utilização de *outputs* de softwares de aprendizagem automática como prova num processo judicial. Tendo por base o processo civil, este artigo reflete sobre este problema, do ponto de vista da admissibilidade e da valoração da prova, olhando muito de perto para o modo como o mesmo é tratado no sistema jurídico norte-americano, dada a existência, neste sistema jurídico, de uma larga discussão, doutrinal e jurisprudencial, sobre o problema do uso probatório da prova técnico-científica.

PALAVRAS-CHAVE: Inteligência Artificial; Máquinas de Aprendizagem Automática; Prova; Processo Civil.

ABSTRACT: Machine learning is a field of artificial intelligence that gives computers the ability to learn without being explicitly programmed, posing the problem of using the outputs of deep learning software as evidence in a judicial

[*] Professor Auxiliar Convidado da Faculdade de Direito da Universidade do Porto. Membro do Centro de Investigação Jurídica (CIJ/FDUP).

[1] O presente texto, que serviu de inspiração à intervenção no Colóquio sobre Inteligência Artificial, com intervenção com o mesmo título, corresponde, na integralidade, a um artigo publicado em RED Ano 2020, Nº 2, Volume 22, ISSN 2182-9845, pp. 79-98.

process. Focusing on Civil Procedure Law, this article aims to reflect on this problem, from the point of view of the admissibility and weight of such an evidence, giving close attention to the north-American experience, where the problem of the use of scientific and technic evidence has been largely discussed.

KEYWORDS: Artificial Intelligence; Machine Learning Evidence; Evidence; Civil Procedure Law.

1. Introdução

O presente trabalho tem por objeto a reflexão sobre o uso de *outputs* de "software de aprendizagem automática", um campo da inteligência artificial, como prova num processo civil (temos em mente os casos em que este software seja utilizado num processo judicial como instrumento de apreensão do registo ou indícios de uma fonte de prova[2], o que subsume o problema da sua aplicação ao âmbito da prova técnica[3]). O nosso sistema processual civil parte, com

[2] José Lebre de Freitas, *A Ação declarativa comum, À luz do Código de Processo Civil de 2013*, 4º ed., Gest-legal, 2017, p. 238. Nota o autor, a propósito da distinção entre fonte de prova e fator probatório, que esta distinção se torna nítida quando, entre a fonte de prova e o juiz, se verifica a intermediação do perito (artigo 388º CC), necessária em virtude dos seus conhecimentos técnicos: "apreendendo e (ou) apreciando o registo ou os indícios da fonte de prova, o perito intervém no processo da sua manifestação como fator probatório". O perito não é fonte de prova pessoal; intervém, sim, no processo de manifestação da fonte de prova, pessoal ou real (ob. cit., p. 238, nota 9). Imagine-se, por ex., o caso de se utilizar um software de aprendizagem automática para decifrar o conteúdo de uma gravação que se encontra danificada, e que não é percetível ao ouvido humano. Sendo a gravação (documento) a fonte de prova, o *software* seria utilizado para a apreensão do seu conteúdo.

[3] Claro está que o *software* deverá seguir determinados critérios de natureza metodológica para que a respetiva a respetiva prova possa ser considerada como um produto de um instrumento técnico fiável. É este um dos problemas de fundo do presente artigo, que será tratado tendo como pano de fundo a Regra 702 das *Federal Rules of Evidence*. Note-se que a criação de softwares deste tipo

exceções (por ex., artigos 875º, e 393º a 395º do Código Civil – CC), de um princípio de prova livre[4]. No que diz respeito à aquisição de conhecimentos técnico-científicos, a mesma tem lugar com recurso a prova pericial, regulada, formal e materialmente, pelas normas de direito probatório respetivamente consagradas nos artigos 467º a 489º do Código de Processo Civil (CPC), e pelos artigos 388º e 389º CC. O resultado da prova pericial deve ser livremente valorado pelo juiz (artigos 389º CC, e 607º, nº 5 CPC)[5].

(como, aliás, se verifica, em geral, em relação a toda a profissão de tecnologia da informação) não é uma profissão inteiramente regulamentada (não é necessário que o criador do software tenha uma formação teórica específica ou esteja inscrito num órgão profissional de controlo de qualidade), o que reforça a importância da função de *gatekeeper* que deve ser realizada pelo tribunal. No futuro, pode imaginar-se a criação de uma lista de *softwares* aceitos como prova.

[4] Sobre este aspeto pode ver-se Manuel Antunes Varela, José Miguel Bezerra, Sampaio e Nora, *Manual de Processo Civil de acordo com o Dec.-Lei 242/85* (2ª Edição Reimpressão), Coimbra Editora, 2004, pp. 467-470.

[5] Não nos interesse abordar alguns aspetos dogmáticos relativos à figura da prova pericial, em particular, o problema de saber se a perícia constitui um meio de prova ou se, ao invés, o perito deve ser visto como um auxiliar do tribunal (pode ver-se, sobre este aspeto, entre outros, José Lebre de Freitas, *A Ação declarativa comum, À luz do Código de Processo Civil de 2013*, cit., p. 238, nota 9). Sobre o diferente tratamento da figura, em sistemas jurídicos de países da União Europeia, veja-se José Lebre de Freitas, "La Preuve dans L'Union Européenne: Différences et Similitudes", *in Estudos sobre Direito Civil e Processo Civil*, vol. I, 2ª edição, Coimbra Editora, pp. 573-609. Outro aspeto que não abordaremos diz respeito à figura do "técnico" (veja-se, por todos, Maria José Capelo, "A enigmática figura do técnico no Código de Processo Civil", *in Estudos em Homenagem ao Prof. Doutor José Lebre de Freitas*, vol. I, Coimbra Editora, pp. 1045-1067). Nos termos dos artigos 492º e 601º, nº 1 CPC, o juiz pode designar um técnico que o auxilie nos atos de produção de prova e de discussão da matéria de facto em audiência, a eles assistindo e prestando os esclarecimentos necessários, tal como pode, em qualquer estado da causa, requisitar os pareceres técnicos indispensáveis ao apuramento da matéria de facto. Considerando que o técnico é mero auxiliar do tribunal, pode ver-se José Lebre de Freitas/Montalvão Machado/Rui Pinto, *Código Processo Civil Anotado*, 2º Volume, 2ª ed., Coimbra Editora, artigo 649º, anotação 2. Já aquela autora (Maria José Capelo, "A enigmática figura do técnico no Código de Processo Civil", cit., pp. 1051-1053) interroga-se sobre a conjugação desta figura com a da perícia. Refere a autora o caso do sistema italiano, que alimenta estas dúvidas, no qual a figura do "técnico" merece outro enquadramento teórico, abarcando e diluindo a categoria de "perito" como fonte de prova. Interroga-se, assim, a autora: "Não padecerá o nosso Código de alguma imprecisão ou ambiguidade dogmática no que diz respeito à figura do técnico? (...) Dever-se-á optar pelo modelo do sistema italiano retirando o "perito" do sistema preclusivo de prova e reconduzindo a oportunidade da intervenção do técnico a todos aqueles momentos em que a experiência técnico-científica seja relevante para a resolução do litígio? Ou, diversamente, não seria mais razoável um regime, como o francês, que prevê o recurso a um técnico (como medida instrutória) através de três modalidades de participação com amplitude e complexidade progressivas: a "constatation", a "consulation" e, reservada a casos mais complicados, a "expertise"" (ob. cit. pp. 1054-1055).

O problema do uso probatório da ciência coloca, entretanto, algumas questões de ordem geral, que aqui não serão desenvolvidas, e que dizem respeito ao problema de saber se o processo civil deve estar orientado para o apuramento da verdade, e qual a função da prova (civil)[6]. Entendemos que a verdade desempenha uma função axiológica, ou seja, constitui um valor do processo, e que a prova constitui o meio destinado a fornecer ao juiz os instrumentos cognitivos necessários para a correta e racional reconstrução processual dos factos[7]. A verdade judicial dos factos corresponde, assim, ao êxito de operações racionais, tanto mais confiáveis quanto fundadas em conhecimentos exatos. Donde, a utilidade da ciência, e, em geral, da técnica, para o processo[8].

Os "softwares de aprendizagem automática" têm aberto novas possibilidades no campo da automatização de tarefas humanas, permitindo ultrapassar os limites da codificação tradicional[9]. O problema do uso probatório de *outputs* de máquinas, na realidade, não é novo (pense-se, por ex., na utilização de radares, bafômetros, software de análise de DNA, GPS, software de análise de risco, etc.). Mas, o uso probatório de *outputs* daquele tipo de softwares coloca problemas específicos, dada a metodologia de programação utilizada, o volume de dados e capacidade de processamento envolvidos, etc., assim como o facto de não ser, em muitos casos, explicável o modo como a máquina chegou a um determinado resultado[10].

[6] Michele Taruffo, *A prova*, trad. João Gabril Couto, 1ª ed., São Paulo, Marcial Pons, 2014, pp. 299 ss.

[7] Luigi Paolo Comoglio, *Le Prove Civili*, 3ª ed., Torino: UTET, 2010, pp. 11 ss. Nota o autor que esta reconstrução é controlada, *a posteriori*, pela demonstração lógico-dedutiva.

[8] Michele Taruffo, *A prova*, cit., pp. 299-230.

[9] De entre a multiplicidade de tarefas que as máquinas podem aprender conta-se, entre as mais comuns, a classificação (por ex., reconhecimento de uma imagem ou reconhecimento facial), classificação com parâmetros ausentes (por ex., o reconhecimento de um objeto ou de um rosto a partir de uma imagem danificada ou incompleta), regressão (por ex., a predição de um valor numérico dadas determinadas condições), transcrição (por ex., software de reconhecimento de voz), deteção de anomalia (por ex., deteção de fraude com cartão de crédito), imputação de valores em falta (i.e., predizer determinados pontos de informação dados outros pontos de informação), etc. Muitas das aplicações emergentes de aprendizado de máquina como prova subsumem-se a estas categorias gerais, como é o caso da identificação de um arguido através de um vídeo ou de uma fotografia danificados, ou a deteção de anomalias em arquivos de empresas para prova de cometimento de atos ilícitos (Patrick W. Nutter, "Machine Learning Evidence: Admissibility and Weight", University of Pennsylvania Journal of Constitutional Law 21, no. 3, February 2019, pp. 919-958 (pp. 920-930).

[10] Patrick W. Nutter, "Machine Learning Evidence: Admissibility and Weight", cit., pp. 921 ss.

Nos EUA, o problema da admissibilidade da prova técnico-científica – adquirida através de testemunhas-peritos[11] – tem sido objeto de larga discussão doutrinal e jurisprudencial, à luz da Regra 702 das *Federal Rules of Evidence*, e, entre outras, da decisão do Supremo Tribunal Norte-Americano no caso *Daubert*. Nesta, foi estabelecido um conjunto de critérios que devem ser utilizados na apreciação da admissibilidade deste tipo de prova, consagrando-se a função de *gatekeeper* do juiz[12]. Ao juiz cabe distinguir boa e má ciência[13], controlando os métodos e procedimentos que presidiram a formação da prova. Dada a presença do sistema de júri (pese embora residual nas ações civis[14]), o controlo da fiabilidade da prova afigura-se muito importante[15], controlo

[11] Deve notar-se que no sistema jurídico norte-americano o perito técnico depõe como testemunha (testemunha-perito). Conforme nota Michele Taruffo (Michele Taruffo, "Le prove scientifiche nella recente esperienza statunitense", Rivista Trimestrale di Diritto e Procedura Civile, Anno L (1996), pp. 241-244): "*L`acquisizione della prova scientifica avviene di regola per mezzo di consulenze tecniche, il che significa – negli Stati Uniti – la testimonianza di expert witnesses. (...) L`aspetto forse più importante, e che maggiormente diversifica il sistema nordamericano dai sistemi di civil law, è che l`esperto viene considerato a tutti gli efetti un testimone. Ciò significa in particolare che gli esperti vengono dedotti dalle parti, che ovviamente li scelgono e li portano in giudizio come fonti di conoscenza dei dati scientifici che le parti stesse ritengono utili per la decisione. Sono poi le parti che pagano gli esperti di cui si servono, e li "preparano" per la testimonianza. (...) Nessuno si aspetta che l`expert witness, essendo istituzionalmente partisan, fornisca conoscenze oggettive formulate in modo neutrale ed imparziale*".

[12] Jack V. Matson/Suha F. Daou/Jeffrey G. Soper, *Effective Expert Witnessing*, Fourth Edition, CRC Press, Boca Raton, London, Ney York, Washington D.C., 2004, pp. 13- 34 (p. 27): "*Daubert established the role of the trial judge as "gatekeeper" for the admission of expert evidence and testimony. This gatekeeper role guards the jury from considering evidence that was purely speculative but offered under the guise of legitimate, scientifically based expert opinion. Although judges are not expected to be scientists, they must demonstrate the ability to think like scientists. They must understand the philosophical and practical standards of scientific method. In cases of doubt or inordinate complexity, judges can engage experts in the field to held them understand and ultimately decide an issue, but judges make the final decision*".

[13] Sobre o significado deste conceito (no inglês, *junk science*), pode ver-se Jack V. Matson/Suha F. Daou/Jeffrey G. Soper, *Effective Expert Witnessing*, cit., pp. 31-32; Gary Edmond/David Mercer, "Trashing "Junk Science", Stanford Technology Law Review 3 (1998), disponível em http://stlr.stanford.edu/STLR/Articles/98_STRL_3. Nos EUA, o problema assume grande importância em *Cass Actions*, relativas a *toxic tort cases*, onde se trata de estabelecer a causalidade.

[14] Nos últimos anos, a realização de processos civis com júri os 1-2% de todos os casos, e 4-5% no Tribunal Federal (Michele Taruffo, "Le prove scientifiche nella recente esperienza statunitense", cit., p. 244, nota 117).

[15] Sobre este aspeto pode ver-se Giovanni Canzio, "Prova scientifica, ricerca della "verità" e decisione giudiziaria nel processo penale", Quaderni della Rivista Trimestrale di Diritto e Procedura Civile, 8, Decisione giudiziaria e verità scientifica, Milano – Dott. A. Giuffrè Editore, Milão, 2005, pp. 55-79 (pp. 58 ss.): "*Il paelse intento della disciplina è quello di sterilizzare tempestivamente il rischio che le caratteristiche dello stile adversary (...) possano essere inquinate da operazioni técnico-scientifiche incomprensibili, confuse, non verificabili, suggestive e pregiudizievoli per il corretto esame da parte della guiria, cui è attribuito il ruolo effettivo di trier of fact*". Pode ver-se também Michele Taruffo, "Le prove

este que se exerce sobre os métodos e procedimentos utilizados, não sobre o resultado da prova. Por outras palavras, o tribunal não deve substituir-se ao júri, nos casos em que a este compete o julgamento da matéria de facto.

Por sua vez, quando este julgamento compete ao juiz, *peritus peritorum*[16], o mesmo deve valorar o resultado da prova científica, determinando o seu específico valor probatório, e realizando as inferências necessárias em relação aos factos da causa. Assim acontece sempre no caso do nosso sistema jurídico.

Pode questionar-se o facto de a nossa análise se centrar num sistema, como o norte-americano, que pertence a uma diferente família de Direito, e que possui uma natureza diversa, marcadamente adversarial. Tal, com efeito, confere à prova técnico-científica um diferente enquadramento jurídico-dogmático. No entanto, e conforme afirma João Henrique Gomes de Sousa, apesar de a Regra 702 das *Federal Rules of Evidence* ter sido pensada para um sistema adversarial puro de apresentação de juízos científicos contraditórios pelas partes, "isso não invalida o seu acerto metodológico mesmo num sistema diverso, pois que expõe boa metodologia de apreciação e de racionalização das perícias técnicas ou científicas apresentadas ao tribunal e que se torna independente do sistema de perícias utilizado"[17]. Donde, e uma vez que é naquele país que tem sido levada mais longe a discussão sobre o problema do uso probatório da prova técnico-científica, a experiência dele recolhida, não obstante as importantes diferenças procedimentais quanto ao modo de aquisição desta prova, tem muito interesse para os sistemas do *civil law*.

scientifiche nella recente esperienza statunitense", cit., p. 227; Petra Claudia Meyer, *Der Sachverständigenbeweis zwischen Partei und Richter – Rechtsvergleich zum US- amerikanischen Zivilprozess und Reformansätze um deutschen Recht*, Münster, Nomos, 2013, p. 167 ("*Während im deutschen Zivilprozess der Sachverständigenbeweis meist schriftlich eingebracht wird, ist die US-am. Hauptverhandlung einer strikteren Handhabung des Mündlichkeitsprinzips verhaftet*"); Jack V. Matson/Suha F. Daou/ Jeffrey G. Soper, *Effective Expert Witnessing*, cit., p. 27 ("*Daubert established the role of the trial judge as "gatekeeper" for the admission of expert evidence and testimony. This gatekeeper role guards the jury from considering evidence that was purely speculative but offered under the guise of legitimate, scientifically based expert opinion. Although judges are not expected to be scientists, they must demonstrate the ability to think like scientists. They must understand the philosophical and practical standards of scientific method. In cases of doubt or inordinate complexity, judges can engage experts in the field to held them understand and ultimately decide an issue, but judges make the final decision*").

[16] Utilizando a expressão "o perito dos peritos" pode ver-se, entre nós, Manuel A. Domingues de Andrade, *Noções Elementares de Processo Civil*, Coimbra Editora, Coimbra, 1979, p. 263.

[17] João Henrique Gomes de Sousa, A "Perícia" Técnica ou Científica Revisitada Numa Visão Prático-Judicial, Julgar – Nº 15 – 2011, pp. 27-52, p. 43.

2. Inteligência artificial e "aprendizagem automática"

A "aprendizagem automática" consiste numa capacidade de aprender (extraindo padrões a partir de dados brutos) sem que o computador tenha sido explicitamente programado para isso[18]. Trata-se de programas dinâmicos, com a capacidade de se ajustar ou modificar em resposta aos dados a que são expostos, sem necessidade de intervenção humana[19]. Por sua vez, a "aprendizagem profunda"[20] consiste num tipo de software capaz de atingir resultados muito precisos na realização de uma tarefa[21]. Trata-se de softwares que funcionam através do método de programação computacional, mas que são diferentes dos métodos de programação tradicionais. Nestes, os programadores escrevem código daquilo que pretendem exatamente que a máquina realize. Trata-se de um método de programação que alimentou uma enorme variedade de aplicações computacionais durante o século XX, mas que não permite automatizar muitas tarefas que os seres humanos realizam, e que não podem ser reconduzidas a um "conjunto de regras", como acontece, por exemplo, com o reconhecimento facial[22].

O próprio dos "softwares de aprendizagem automática", conforme referido, consiste no facto de o computador aprender a realizar uma tarefa sem que o programador lhe possa exatamente explicar como a realizar[23]. O método de programação consiste no seguinte: depois de serem exibidos ao computador milhares, ou mesmo milhões de dados, o mesmo aprende determinados padrões, correlações ou regras. Por vezes, estas regras são as mesmas que os humanos utilizam na realização da respetiva tarefa, mas pode tratar-se de

[18] Chris Nicholson, "Artificial Intelligence (AI) vs. Machine Learning vs. Deep Learning, disponível online: https://pathmind.com/wiki/ai-vs-machine-learning-vs-deep-learning. Sobre o futuro da inteligência artificial, e concretamente, sobre a matéria dos chamados algoritmos evolucionários, pode ver-se também Matthew Hutson, "Computers Evolve a New Path Toward Human Intelligence", Quantamagazine, November, 6, 2019.

[19] Chris Nicholson, "Artificial Intelligence (AI) vs. Machine Learning vs. Deep Learning, cit.

[20] Em língua inglesa usa-se a expressão "*deep learning machines*".

[21] Patrick W. Nutter, "Machine Learning Evidence: Admissibility and Weight", cit., p. 927.

[22] *Ibidem*. O autor dá o exemplo de um investigador de inteligência artificial que comenta: eu consigo facilmente reconhecer o rosto da minha mãe, mas não seria capaz de escrever um código para uma máquina fazer isso. Este exemplo ilustra o chamado paradoxo de Polanvi, de acordo com o qual existem limites fundamentais em relação a quanto conhecimento os seres humanos podem transmitir às máquinas. As máquinas de aprendizagem automática surgiram, no entanto, como um subcampo da inteligência artificial capaz de contornar esta limitação.

[23] *Ibidem*, pp. 923 ss. O termo "aprender" é aqui utilizado em sentido impróprio. "Aprender", no contexto em que o termo vem utilizado, refere-se ao aperfeiçoamento na realização de uma tarefa ao longo do tempo (ob. cit., p. 929).

regras que os humanos não são capazes de observar, ou que não foram por si anteriormente utilizadas. Muitas vezes, o próprio programador não é capaz de explicar o modo como o computador chegou ao resultado, ou seja, o tipo de regras por ele inferidas, ainda que o resultado seja correto.

Para realizar a tarefa o computador começa por aprender a partir de dados fornecidos pelo programador, que constituem o conjunto de "dados quantificados". Quando se trata de informação numérica, a quantificação é direta. Noutros casos, o modo de quantificação é menos evidente, ou pode depender de decisão do programador (por ex., a imagem de um rosto é quantificada com base em valores pixel que um ecrã utiliza para exibir a imagem). Tendo a informação sido traduzida para números, o programador utiliza alguma da informação, cujas propriedades são conhecidas (referidas como "informação de treino"), e ensina o computador regras ou associações que lhe serão úteis, quando o mesmo analisar dados cujas propriedades são desconhecidas. A este processo dá-se o nome de "aprendizagem supervisionada"[24]. O computador faz a partir daí as suas próprias inferências.

Depois de a "máquina ter aprendido" a partir da informação de treino, e deduzido um conjunto de regras, a sua *performance* é testada e aperfeiçoada através de um novo conjunto de dados, chamado "conjunto teste", as propriedades dos quais também são conhecidas. Com base nisso, o programador avalia a precisão da máquina e as taxas de erro, podendo fazer novos ajustes. Se a máquina atingiu um grau de precisão que o investigador considera satisfatório, ela pode ser utilizada para analisar informação extraída do mundo real.

3. O problema da admissibilidade da prova

3.1. Introdução

Dada a especificidade do modo de programação e de funcionamento dos "softwares de aprendizagem automática", colocam-se problemas específicos de fidedignidade dos seus resultados, diferentes daqueles que se colocam a propósito dos softwares tradicionais.

[24] *Ibidem*, pp. 929 ss. Utilizando o mesmo exemplo acima dado (cf., *supra*, nota 17), um programador, nesta fase, irá fornecer à máquina um conjunto de imagens da sua mãe (imagens que o programador sabe que são da sua mãe), dizendo à máquina para associar a imagem desta face à sua mãe (por ex., marcando cada imagem com o seu nome). Neste ponto, a máquina sabe que estas imagens são da mãe do investigador, não por qualquer inferência computacional, mas porque o programador lho disse explicitamente. Depois disso, a máquina analisa as imagens e, ela própria, estabelece associações, correlações, ou regras que lhe permitirão reconhecer a mãe do programador em novas imagens que ela não viu antes. Por exemplo, a máquina pode estabelecer regras sobre o tom de pele, a distância entre os olhos, a altura ou largura do rosto.

Uma vez que a problemática se insere no campo da admissibilidade da prova técnico-científica, começaremos por abordar o problema de um modo geral. Para isso, teremos em conta o sistema jurídico norte-americano, em especial a Regra 702 das *Federal Rules of Evidence* e a decisão do Supremo Tribunal de Justiça no caso *Daubert*, por se tratar do sistema onde o problema da admissibilidade da prova técnico-científica mais profundamente tem sido debatido[25]. Só depois disso analisaremos a aplicação destes critérios ao caso da prova resultante de "software de aprendizagem automática". Assim.

3.2. A jurisprudência do Supremo Tribunal Norte Americano sobre o problema da admissibilidade da prova técnico-científica e a Regra 702 das *Federal Rules of Evidence*

a) O general acceptance test, ou Frye test

Tradicionalmente, os tribunais norte-americanos utilizavam como critério para a aquisição de conhecimentos técnico-científicos o recurso a peritos qualificados. Ou seja, era a presença do perito no respetivo campo profissional que garantia a fiabilidade do seu depoimento, e a possibilidade de o juiz se servir da sua colaboração[26]. Entretanto, um passo importante na evolução do problema foi dado em 1923, quando o Tribunal Regional do Distrito de Columbia decidiu o caso *Frye v. United States*[27]. Estava em causa a admissibilidade de um dos primeiros modelos de "máquina da verdade", num caso de homicídio, tendo o Tribunal formulado o critério segundo o qual um engenho ou princípio científico pode ser admitido como prova quando o mesmo esteja "*sufficiently established to have gained general acceptance in the particular field in which it belongs*".

Estabeleceu-se, assim, o *general acceptance test*, ou *Frye test*, segundo o qual a admissibilidade da prova científica depende do "mercado intelectual", ou seja, da existência de um consenso difuso e consolidado, na respetiva área científica, sobre a validade da mesma. Apesar de aquela decisão não ter força de precedente, este critério foi utilizado pelos tribunais durante muitos anos, e ainda hoje é utilizado por alguns juízes, por ter a vantagem de permitir aos

[25] Existe outro tipo de considerações, no que diz respeito à admissibilidade deste tipo de prova, que são relevantes no que diz respeito ao sistema jurídico norte-americano, como aquelas que dizem respeito à aplicação da Quinta e Sexta Emendas Constitucionais, mas esses aspetos não serão abordados no presente artigo.

[26] Michele Taruffo, "Le prove scientifiche nella recente esperienza statunitense", cit., pp. 232 ss.

[27] Frye v. United States, 293 Fed. 1013, 1014 (D.C.Cir. 1923).

juízes não enfrentarem diretamente o problema da validade científica da prova, escudando-se naquilo que é aceite pela própria comunidade científica[28].

b) Evolução posterior e entrada em vigor das *Federal Rules of Evidence*

Em anos mais recentes, o *general acceptance test* começou, porém, a reunir menor consenso doutrinal. Com efeito, a par da posição dos autores que sustentam a racionalidade do critério segundo o qual apenas deve ser admitida como prova científica aquela que é reconhecida como tal pela comunidade científica, críticas começaram a surgir a este critério. Entre as mais relevantes e frequentes, estão aquelas que consideram tratar-se de um critério demasiado restritivo e excessivamente conservador, podendo resultar da sua aplicação a exclusão de prova científica fundada em métodos e princípios válidos, embora não aceites pela generalidade da comunidade científica por serem novos e originais. Observa-se que, em muitos casos, são necessárias décadas para que a opinião do ambiente científico se consolide e, por outro lado, que em muitas áreas ou sobres muitos dados científicos, embora válidos, não existe uma *communis opinio*. Do mesmo passo, mas em sentido inverso, pode suceder que o teste seja aplicado de modo excessivamente elástico, consoante o modo como o juiz defina o âmbito profissional em relação ao qual a aceitação generalizada é considerada relevante[29].

Outro aspeto muito importante consistiu na entrada em vigor, em 1975, das *Federal Rules of Evidence*, depois tomadas como modelo pela legislação da maior parte dos Estados. Esta lei não faz qualquer referência ao *general acceptance test*, ou ao caso Frye, e as Regras 702-706, que se ocupam da prova por testemunha-perito, não usam aquele teste[30]. Esta situação originou uma divisão doutrinal e jurisprudencial, entre aqueles que continuaram a utilizar o *Frye test*, e os que entendiam que este teste se encontrava ultrapassado pela entrada em vigor daquela lei.

c) Decisão do caso Daubert v. Merrel Dow Pharmaceuticals, Inc. e função de *gatekeeper* do tribunal

Foi este o contexto em que o Supremo Tribunal dos EUA se pronunciou, em 1993, no caso Daubert v. Merrel Dow Pharmaceuticals, Inc.[31]. No Acórdão, o

[28] Michele Taruffo, "Le prove scientifiche nella recente esperienza statunitense", cit., pp. 232 ss.; Michel H. Graham, *Federal Rules of Evidence in a nutshell*, 10th edition, West Academic Publishing, 2018, United States of American, comentário à Rule 702, pp. 367-401 (pp. 374 ss).

[29] Michele Taruffo, "Le prove scientifiche nella recente esperienza statunitense", cit., pp. 232 ss.

[30] *Ibidem*.

[31] Daubert v. Merrell Dow Pharmaceuticals, Inc., 509 U.S. 579 (1993).

Tribunal assumiu uma posição muito clara, embora algumas das afirmações contidas na decisão tenham dado azo a dúvidas e discussões[32]. Em primeiro lugar, o Tribunal refere que as *Federal Rules of Evidence* contêm uma disciplina completa e exaustiva do direito da prova, devendo ser com recurso às mesmas, e não à jurisprudência anterior, que devem resolver-se os problemas da admissibilidade do testemunho pericial e da prova científica. Ou seja, o Tribunal adere à tese segundo a qual o *Frye test* não foi recebido pelas *Federal Rules of Evidence*, sendo incompatível com a orientação liberal das mesmas[33]. Entende o Tribunal que esta lei adotou critérios mais elásticos no que diz respeito à admissibilidade da prova em geral, e da prova técnico-científica em particular, não podendo o "teste da aceitação generalizada" constituir o único critério para admitir ou excluir esta prova. Assim, e embora reafirme que só deve ser admitida como prova aquela que for cientificamente válida, o Tribunal sublinha que a fiabilidade da prova deve ser avaliada de acordo com vários critérios, enumerando quatro critérios principais: a) a controlabilidade ou falsificabilidade da teoria ou da técnica que se encontram na base da prova (*testability*[34]), b) o percentual de erro conhecido ou potencial e o respeito pelos *standards* relativos à técnica empregue, c) o facto de a teoria ou técnica em causa terem sido objeto de publicação científica e, portanto, controlada por outros peritos (*peer review*), e, d) a aceitação geral por parte da comunidade científica interessada. Ou seja, o *Frye test* sobrevive, mas apenas como um entre vários critérios possíveis de valoração. Para além disso, o Tribunal admite que outros critérios possam ser empregues, sem fazer um elenco taxativo dos mesmos, e deixando essa apreciação à discricional determinação do juiz[35]. Por fim, sublinha o Tribunal que a prova científica apenas

[32] Seguimos Michele Taruffo, "Le prove scientifiche nella recente esperienza statunitense", cit., pp. 236 ss.

[33] Sobre o fundo cultural desta decisão, pode ver-se Angelo Dondi, "Paradigmi processuali ed "expert witness testimony" nel diritto statunitense", Rivista Trimestrale di Diritto e Procedura Civile, Anno L (1996), pp. 261-285 (p. 268), e David Goodstein , "How Science Works", *Reference Manual on Scientific Evidence*, Third Edition, Federal Judicial Center, The National Academies Press, Washington, United States of American, 2011, pp. 37-54 (pp. 52-54).

[34] Sobre este aspeto, pode ver-se: Michel H. Graham, *Federal Rules of Evidence in a nutshell*, 10th edition, West Academic Publishing, 2018, United States of American, comentário à Rule 702, pp. 367-401 (pp. 374 ss); Margaret A. Berger, "The Admissibility of Expert Testimony", *Reference Manual on Scientific Evidence*, Third Edition, Federal Judicial Center, The National Academies Press, Washington, United States of American, 2011, pp. 11-36 (p. 13).

[35] Jack V. Matson/Suha F. Daou/Jeffrey G. Soper, *Effective Expert Witnessing*, cit., p. 28: "*While the Daubert test provides the basis for challenging proposed expert testimony, trial judges can use, reject, or modify any of the Daubert factors at their own discretion*".

deve ser admitida quando seja diretamente útil e relevante para estabelecer os concretos factos da causa[36].

Segundo o *Daubert test* o juiz deve, assim, desempenhar um papel de efetivo *gatekeeper*, sendo responsável por controlar a fiabilidade da prova científica, através do controlo dos *métodos* e *procedimento*s que presidem à sua formação, segundo os múltiplos, mas não necessariamente cumulativos, critérios enunciados pelo Supremo Tribunal[37]. Ou seja, não se requer um controlo sobre a atendibilidade dos resultados específicos da prova científica, mas a verificação preliminar dos métodos que esta emprega[38].

d) A Regra 702 das Federal Rules of Evidence e o Daubert/ Kumho's criterion

Segundo Michel H. Graham[39], apesar da orientação liberal seguida pelo Supremo Tribunal, no sentido de uma maior flexibilização dos critérios de admissibilidade da prova, a decisão do caso *Daubert* criou, na prática, um teste mais rigoroso, especialmente nos casos civis[40], tendo dado lugar a opiniões muito divergentes. As dificuldades dizem respeito ao âmbito e ao modo de aplicação do teste. Com efeito, o Tribunal não especificou se o mesmo se aplica a todos os casos de prova pericial, ou apenas à prova científica, e dentro desta, se o mesmo se aplica a toda a prova científica ou apenas a *técnicas científicas inovadoras*. Por outro lado, a decisão podia deixar a dúvida sobre se os critérios nela enumerados devem ser aplicados cumulativamente, ou não.

[36] Seguimos Michele Taruffo, "Le prove scientifiche nella recente esperienza statunitense", cit., pp. 236 ss.

[37] Veja-se, Giovanni Canzio, "Prova scientifica, ricerca della "verità" e decisione giudiziaria nel processo penale", cit., pp. 58 ss.: "*In altri termini, il passaggio è dalla tendenza a definire in generale, genericamente e una volta per tutte il significato di scientific evidence – come avveniva nella decisione Frye – alla tendeza a qualificare le metodologie di acquisizione di questi tipi di conoscenze, giungendo a determinare secondo l`espressione della stessa Corte suprema, alcuni "standards di affidabilità probatória*"); Michel H. Graham, *Federal Rules of Evidence in a nutshell*, cit., pp. 374 ss.; Margaret A. Berger, "The Admissibility of Expert Testimony", cit., ps. 11-36 (*"Although there was nothing particularly novel about the Supreme Court finding that a trial judge has the power to make an admissibility determination–Federal Rules of Evidence 104(a) and 702 pointed to such a conclusion–and federal trial judges had excluded expert testimony long before Daubert, the majority opinion in Daubert stated that the trial court has not only the power but the obligation to act as gatekeeper*").

[38] Michele Taruffo, "Le prove scientifiche nella recente esperienza statunitense", cit., p. 239; Stephen Breyer, "Introduction", *Reference Manual on Scientific Evidence*, Third Edition, Federal Judicial Center, The National Academies Press, Washington, United States of American, 2011, pp. 1- 9 (p. 6).

[39] Michel H. Graham, *Federal Rules of Evidence in a nutshell*, cit., pp. 380 ss.

[40] *Ibidem* (veja-se, em particular, as páginas 382-384, 386-387).

Algumas destas dúvidas foram esclarecidas pelo Supremo Tribunal no caso *Kumho Tire Company, Ltd. v. Carmichael*[41], onde foi declarado que a obrigação de *gatekeeping* se aplica a toda a prova especializada (*i.e.,* a toda a prova baseada em conhecimentos técnicos, e não apenas àquela que se baseia em conhecimentos científicos), e que o juiz, ao exercer esta função, pode considerar um ou mais dos critérios enunciados no teste *Daubert*, quando isso ajude a determinar se a prova é fiável[42].

Entretanto, a Regra 702 da Lei Federal da Prova foi alterada, com efeitos a partir de 1 de dezembro de 2000, em conformidade com a decisão do Supremo Tribunal no caso *Daubert*. De acordo com a norma reformulada, para além do critério de relevância da prova para a demonstração dos "factos controvertidos", o juiz deve determinar, antes de aceitar o depoimento da testemunha-perito, que "(1) o depoimento se baseia em factos ou dados suficientes, (2) o depoimento é resultado de princípios e métodos fidedignos, e (3) o perito aplicou corretamente os princípios e métodos aos factos da causa"[43].

3.3. A Regra 702 e a prova resultante de "software de aprendizagem automática"

A utilização como prova do resultado de "softwares de aprendizagem automática" conduz-nos, em princípio, ao depoimento de testemunha-perito, regulado pela Regra 702º das *Federal Rules of Evidence* e pelo teste *Daubert*[44]. Como vimos, para que uma testemunha-perito possa depor, a parte deve ser capaz de demonstrar que o depoimento é relevante para a decisão, que o mesmo

[41] Kumho Tire Company, Ltd. v. Carmichael, 526 U.S. 137 (1999).

[42] Michel H. Graham, *Federal Rules of Evidence in a nutshell*, cit., pp. 380 ss.: Overall, *Kuhmo* instructs that the test of "reliable" is "flexible" and that "*Daubert`s* list of specific factors neither necessarily nor exclusively applies to all experts in every case." "Rather the law grants a district court the same broad latitude when it decides how to determine reliability as it enjoys in respect of its ultimate reliability determination."

[43] Existe uma nota do Comité Consultivo na qual se afirma que a Regra 702 modificada é consistente com a interpretação do *Daubert test* feito na decisão *Kumho*. A estrutura da norma conduziria normalmente à interpretação de que cada um dos requisitos deve ser considerado separada e distintamente. Ou seja, parece resultar da mesma que, ao determinar a admissibilidade do *depoimento por testemunha especializada*, o tribunal deve declarar que cada um dos três requisitos se encontra satisfeito, ou que um ou mais não se encontra suficientemente estabelecido. Na prática, contudo, a linha divisória entre estes requisitos é muitas vezes pouco clara, sendo os mesmos, na verdade, parte e parcela de uma única determinação (*ibidem*, p. 390).

[44] Patrick W. Nutter, "Machine Learning Evidence: Admissibility and Weight", cit., pp. 931 ss. O autor nota, no entanto, que o modo como a prova pode ser usada em tribunais estaduais depende da lei vigente em cada Estado.

se baseia em factos ou dados suficientes, que ele é o produto de princípios e métodos credíveis, e que estes princípios ou métodos são corretamente aplicados aos factos da causa[45]. Em princípio, não resulta da aplicação destes critérios a exclusão daquela prova, embora o modo como o algoritmo foi criado, ou o modo como o mesmo é utilizado em tribunal possam torná-la inadmissível[46]. Vejamos.

À partida, esta prova satisfaz três dos critérios enunciados pelo Supremo Tribunal no caso *Daubert*, a saber, o requisito da "testabilidade" (trata-se de processos cujos resultados podem, teoricamente, ser demonstrados como sendo falsos), o requisito da *peer review*[47], e a aceitação generalizada no seio da comunidade científica. Mais delicada é a aplicação do critério da existência de taxas de erro conhecidas ou potenciais.

Na verdade, os "softwares de aprendizagem automática" têm margens de erro calculáveis, mas a relevância das mesmas, no caso concreto, é muitas vezes questionável. Os algoritmos deste tipo podem ter duas margens de erro, aquela que diz respeito à atuação do algoritmo em relação aos dados do "conjunto teste"[48], cujas propriedades são conhecidas, e, eventualmente, uma segunda taxa, respeitante à *performance* da máquina com dados provenientes do mundo real, cujas propriedades são desconhecidas. Ambas as taxas aparecem normalmente como um número singular, mascarando, no entanto, outros valores importantes, como o que se refere à maior ou menor probabilidade de o algoritmo dar falsos positivos, ou falsos negativos[49]. Por outro lado, o crité-

[45] Como vimos, quando o juiz determina a admissibilidade do depoimento pericial, ele está apenas a decidir preliminarmente se o raciocínio ou metodologia subjacente ao depoimento é cientificamente válido e se o raciocínio ou metodologia pode ser apropriadamente aplicado aos casos da causa. O foco está não nas conclusões, mas nos métodos utilizados.

[46] Seguimos de perto Patrick W. Nutter, "Machine Learning Evidence: Admissibility and Weight", cit., pp. 931 ss.

[47] Com efeito, a literatura com revisão de pares sobre a matéria proliferou nos últimos anos, com alguns dos seus princípios científicos datando de meados do século XX.

[48] Trata-se da informação de treino (dados cujas propriedades são conhecidas, e que são a base para o algoritmo melhorar a sua *performance* ao longo do tempo).

[49] Sobre este ponto veja-se Patrick W. Nutter, "Machine Learning Evidence: Admissibility and Weight", cit., pp. 931 ss., e p. 951. Partindo do exemplo de um programador que ensina a máquina a reconhecer o rosto da sua mãe, o autor nota: "*..if the machine has only ever learned from images in which the mother was photographed with flash on, the machine may use the brightness of the image as a basis to identify the mother, and with more weight than any attribute about her face. If this were the case, when the machine later must confront an image of the mother in which she was not photographed with flash, the machine might not be able to identify her (a false negative), even though humans would not be confused by such a situation. Conversely, the machine might mistakenly identify as the mother an entirely different woman who was photographed with flash (a false positive)*" (ob. cit., p. 951).

rio do programador pode influenciar as taxas de erro, por ex., nos casos em que a máquina tenha um resultado apenas parcialmente bem-sucedido[50]. Por último, pode suceder que a taxa de erro oculte uma taxa superior, quando o software seja aplicado a dados reais com caraterísticas diferentes daquelas que foram utilizados no treino inicial da máquina[51].

Existem, ainda, outros aspetos muito relevantes para a apreciação da admissibilidade desta prova, como aquele que diz respeito à extensão, modo de seleção, e manuseamento dos dados utilizados para treinar a máquina[52]. Em primeiro lugar, quanto mais complicada for a tarefa, mais extenso deve ser o "conjunto teste" utilizado. Ou seja, o tamanho dos dados influencia a *performance* da máquina, devendo ser demonstrado que os dados utilizados no treino são suficientes, segundo o que é aceite pela comunidade científica.

Em segundo lugar, o modo de recolha ou de geração dos dados pode ter ocorrido de um modo que conduziu à produção de amostras tendenciosas[53]. Trata-se, pois, de apurar a qualidade dos dados utilizados, e de determinar a medida em que os mesmos podem ser tendenciosos no caso particular: qual a proveniência dos dados, qual o método de reunião ou geração dos mesmos[54],

[50] O autor dá o exemplo de uma máquina destinada à leitura de lábios. Neste caso, uma tradução não elegante, mas apesar de tudo compreensível, deve ser considerada um bom ou mau resultado? Este exemplo demonstra que, por vezes, não é, sequer, fácil saber o que é um bom ou mau resultado. Ora, a resposta, que entrará na estatística da taxa de erro, constituirá em última análise uma decisão de um ser humano, e pode não existir consistência de um programador para outro.

[51] O autor dá o exemplo de uma máquina para identificação de um arguido. Pode suceder que a máquina tire conclusões sobre um arguido que não partilha as caraterísticas da informação inicial de treino. Por ex., uma taxa de erro para uma máquina que foi treinada com dados provenientes de sujeitos de raça diferente pode ser menos fidedigna para a identificação de pessoas de uma concreta raça do que para outra. Problema semelhante se prende com a geografia dos dados utilizados para o treino da máquina.
Esta prova é, assim, particularmente suscetível de violar o critério da "correta aplicação dos princípios e métodos utilizados aos factos da causa".

[52] *Ibidem*, p. 935. Este aspeto relaciona-se com a imposição da Regra 702 de que a prova oferecida se baseie em factos ou dados suficientes e que a mesma seja o produto de princípios e métodos credíveis.

[53] *Ibidem*, pp. 936 ss.

[54] Quando uma base de dados não é suficientemente extensa, existem várias técnicas que os programadores podem utilizar para a manipular de modo a criar artificialmente uma base de dados mais extensa. Por ex., o algoritmo pode recolher exemplos de dados de um modo aleatório e criar outras bases de dados mais pequenas. O programador pode também modificar intencionalmente os dados, por ex., distorcendo imagens ou adicionando ruídos de um modo aleatório. Os modos de manipulação são largamente influenciados pelas decisões do programador e pelas normas no respetivo campo de atuação (*ibidem*, ps. 938ss).

como atestou o programador a sua qualidade no caso de os ter recebido de um terceiro[55]?

Em terceiro lugar, também o modo como os dados foram rotulados se afigura muito relevante[56]. Ou seja, mesmo sendo a amostra suficientemente extensa, e os dados recolhidos ou gerados de acordo com técnicas *standard*, os mesmos devem ser devidamente rotulados e organizados, tarefa muito difícil, mas crucial, no caso de "conjuntos de teste" com milhões de exemplos. As "máquinas inteligentes" aprendem com os dados que lhes são fornecidos, e de acordo com o modo como os mesmos lhes são apresentados (é o programador que, num momento inicial, diz à máquina aquilo que os dados que lhe são fornecidos significam). Quem identifica os dados, e o modo como os mesmos são identificados afiguram-se, assim, aspetos críticos.

Por último, podem surgir problemas com o código base. Um exame do código do software pode revelar detalhes que afetam a admissibilidade da prova, de acordo com a segunda parte da Regra 702. Com efeito, se a própria programação contém erros, é possível que as conclusões da máquina não constituam o resultado de "princípios e métodos fidedignos". A análise do código pode revelar simples erros ou falsas assunções na criação do programa, suficientes para afetar o resultado da máquina[57].

4. O problema da valoração da prova

O juízo sobre a admissibilidade da prova não resolve o problema da sua valoração. Em termos gerais, pode notar-se que a prova científica não fornece, por regra, certeza dedutiva às conclusões relativas aos "factos controvertidos"[58], mas apenas um determinado grau de probabilidade[59] (como acontece no caso presente). Ao juiz cabe, assim, apreciar o resultado da prova científica, de acordo com o princípio da livre apreciação da prova (artigos 607º, nº 5 CPC e 389º CC) e segundo um critério de probabilidade prevalente[60], em conjunto

[55] Note-se que, no caso de métodos de reunião de dados através de fontes abertas ou de dados resultantes de *crowdsorces*, que são comuns no campo de máquinas inteligentes, essa verificação pode nem ser possível.

[56] *Ibidem*, pp. 938 ss.

[57] *Ibidem*, pp. 939 ss.

[58] Michele Taruffo, *A prova*, cit., pp. 223-224.

[59] Michele Taruffo, "Conoscenza scientifica e decisione giudiziaria: profili generali", Quaderni della Rivista Trimestrale di Diritto e Procedura Civile, 8, Decisione giudiziaria e verità scientifica, Milano – Dott. A. Giuffrè Editore, Milão, 2005, pp. 3-23.

[60] Sobre o conceito de probabilidade e de standards de prova pode ver-se Michele Taruffo, *A prova*, cit., p. 303; Michele Taruffo, "Conoscenza scientifica e decisione giudiziaria: profili generali", cit., p. 20. Segundo o autor, o critério de probabilidade prevalente representa um critério

com as demais provas existentes. Nisto consiste o, por vezes, chamado paradoxo da prova científica[61]. Com efeito, tendo decidido nomear um perito, por entender que não dispõe de conhecimentos científicos ou técnicos necessários para julgar a causa, o juiz deve ser *peritus peritorum*[62], valorando as conclusões do perito, e determinando inclusive o seu grau de fiabilidade e de validade científica[63].

Contudo, o facto de a força probatória do resultado do laudo dos peritos ser apreciada livremente pelo tribunal não significa, como é óbvio, que o tribunal a possa considerar arbitrária ou discricionariamente, mas apenas que não está vinculado a regras ou critérios legais[64]. Com efeito, são critérios epistemológicos, e não jurídicos, aqueles que, uma vez estabelecido o grau de confirmação lógica que os elementos de prova atribuem a um dado "facto", determinam a escolha da hipótese racionalmente preferível[65]. Ou seja, é com

de racionalidade de adoção necessária para preencher o vazio normativo que resulta da adoção do princípio do livre convencimento do juiz. Este princípio desvincula o juiz de regras de prova legal, porém, não o desvincula dos critérios da lógica, da razão e da confiabilidade intersubjetiva da valoração das provas. Portanto, aquilo que não é mais coberto por normas relativas ao valor das provas deve ser regulado pro critérios racionais, sob pena de legitimar a existência de decisões arbitrárias. Nota, ainda, o autor que não se trata da aceitação de uma noção estatística de probabilidade, teorias estas que não têm aplicação ao fenómeno da prova. O problema é, assim, o de saber e em que medida o conhecimento científico pode ajudar o juiz a aplicar o standard de decisão requerido, ou seja, a determinar o nível de probabilidade de uma hipótese relativa a um facto e estabelecer se, pelo menos, uma hipótese atinge o standard da probabilidade prevalente.

[61] Michele Taruffo, "Le prove scientifiche nella recente esperienza statunitense", cit., pp. 244 ss.. Veja-se sobre este problema Denti, V., "Scientificità della prova e libera valutazione del giudice", in Riv. dir. proc., 1972, pp. 414 ss.

[62] Michele Taruffo, "Le prove scientifiche nella recente esperienza statunitense", cit., pp. 244 ss. (p. 244); Michele Taruffo, *A prova*, cit., pp. 223-224; Giovanni Canzio, "Prova scientifica, ricerca della "verità" e decisione giudiziaria nel processo penale", cit., pp. 64-65. Nota este último autor: "*(...) sembra consolidarsi l`orientamento giurisprudenziale per il quale spetta comunque al giudice l`impegnativo compito di verificare con particolare rigore la validità scientifica dei criteri e dei metodi di indagni utilizzati dal perito, allorché essi si presentinho come "nuovi" e perciò non ancora sottoposti al vaglio di una pluralità di casi ed al reiterato confronto critico tra gli esperti del settore, sì da non potersi considerare ancora acquisiti al património della comunità scientifica".*

[63] Michele Taruffo, *A prova*, cit., pp. 223-224. "*A força probatória das respostas dos peritos não é vinculativa para o tribunal, que pode afastar-se livremente do parecer dos peritos, quer porque tenha partido de factos diferentes dos que aceitou o perito, quer porque discorde das conclusões deles ou dos raciocínios em que eles se apoiam, quer porque os demais elementos úteis de prova existentes nos autos sejam mais convincentes, em seu entender, que o laudo dos peritos*" (Fernando Pereira Rodrigues, *Os meios de prova em processo civil*, 3ª edição, 2017, Almedina, p. 141). Veja-se também José Lebre de Freitas, *A Ação declarativa comum, À luz do Código de Processo Civil de 2013*, cit., p. 344.

[64] Fernando Pereira Rodrigues, *Os meios de prova em processo civil*, cit., p. 141

[65] Michele Taruffo, *A prova*, cit., p. 319.

argumentos científicos que o juiz deve apreciar o resultado da prova científica, recaindo sobre o mesmo um acrescido dever de fundamentação no caso de dela divergir[66].

Dito de outro modo, o juízo técnico-científico, propriamente dito, não está sujeito a livre apreciação, mas o juiz deve controlar a validade científica do mesmo. Para tal, o juiz não precisa de ser um cientista (é aparente aquele paradoxo[67]), mas deve ser capaz de apreciar a validade dos métodos de que o perito se serviu, determinando o grau de confiabilidade das suas conclusões[68]. Recolocam-se, assim, os problemas que surgem quando o juiz se pronuncia sobre a admissibilidade da prova, agravados pela circunstância de o mesmo, sendo julgador da matéria de facto, dever fundar nesta prova o respetivo juízo probatório. Ou seja, ao juiz cabe não apenas controlar a correção dos métodos e procedimentos utilizados na formação da prova, como determinar o seu específico valor probatório, realizando as inferências necessárias em relação aos concretos "factos controvertidos"[69].

Concretamente no que diz respeito à prova resultante de "software de aprendizagem automática", pese embora a sua aparência mecânica, e a simplicidade do seu resultado – que lhe conferem um véu de objetividade, certeza e neutralidade –, os mesmos são produto da criação humana, estando sujeitos às suas decisões, e à sua tendência para cometer erros[70]. Os aspetos

[66] Neste sentido, pode ver-se, entre outros, os seguintes Acórdãos: Acórdãos do Supremo Tribunal de Justiça de 05.05.1993, proc. nº 044111, relator Ferreira Dias ("*A própria lei estabelece excepções ao princípio da livre apreciação da prova, respeitantes ao valor probatório dos documentos autênticos ou autenticados, ao caso julgado, à confissão integral e sem reservas no julgamento e à prova pericial. O juízo técnico-científico dos peritos deve ser acatado pelos julgadores, a não ser que estes dele divergirem, caso em que têm de fundamentar cientificamente a divergência*"), e de 06.07.2011, proc. nº 3612/07.6TBLRA.C2.S1, relator Hélder Roque; Acórdão do Tribunal da Relação de Lisboa de 11.03.2010, proc. nº 949/05.4TBOVR-A.L.1-8, relator Bruto da Costa; Acórdãos do Tribunal da Relação de Guimarães de 07.06.2018, proc. nº 3/14.8TJVNF.G2, relatora Maria Cristina Cerdeira, e de 25.10.2018, proc. nº 6166/15.8T8GMR-A.G1, relatora Maria Cristina Cerdeira, e Acórdão do Tribunal da Relação de Évora de 18.10.2018, proc. nº 803/06-2TBVNO-A.E1, relatora Ana Margarida Leite, todos disponíveis em www.dgsi.pt.

[67] Michele Taruffo, *A prova*, cit., p. 319.

[68] *Ibidem*. Segundo David Goodstein ("How Science Works", cit., pp. 52-54)*:* "*The presentation of scientific evidence in a court of law is a kind of shotgun marriage between the two disciplines. Both are obliged to some extent to yield to the central imperatives of the other's way of doing business, and it is likely that neither will be shown in its best light. The Daubert decision is an attempt (not the first, of course) to regulate that encounter. Judges are asked to decide the "evidential reliability" of the intended testimony, based not on the conclusions to be offered, but on the methods used to reach those conclusions*".
Em caso de dificuldade, o juiz pode servir-se do auxílio de um consultor técnico (artigo 601º CPC).

[69] Michele Taruffo, "Le prove scientifiche nella recente esperienza statunitense", p. 248.

[70] Patrick W. Nutter, "Machine Learning Evidence: Admissibility and Weight", cit., p. 924.

anteriormente assinados (quantidade, qualidade, organização e adequação dos dados de treino, perfeição do código base, etc.) são preponderantes para determinar o grau de confiabilidade dos resultados obtidos, devendo ser tidos em conta na apreciação crítica da prova.

Mas, mesmo sendo fiáveis estes resultados, à luz referidos aspetos, deve ter-se presente uma caraterística dos "softwares de aprendizagem automática", que diz respeito à inexplicabilidade das suas conclusões. Como nota Patrick W. Nutter, esta caraterística deve-se ao enorme número de dados utilizados, assim como à avalanche de probabilidade estatística envolvida[71]. Acontece, assim, muitas vezes, que ninguém, nem mesmo o próprio programador, pode explicar como ou porquê a máquina atingiu um determinado resultado, o que pode, ou não, reduzir o peso que lhe é atribuído pelo juiz. Descobrir as regras e correlações que o computador utilizou pode implicar uma investigação adicional, e pode permanecer para sempre um mistério[72].

Não existe um critério lógico evidente para responder a este problema. Sendo inexplicável o tipo de inferências realizadas pela máquina, o juiz pode apenas confiar na taxa de erro relativa ao seu funcionamento (não obstante as advertências acima feitas acerca deste aspeto[73]), retirando daí o grau de probabilidade (epistémica) que, em seu entender, pode ser atribuído ao seu

[71] Ibidem, p. 953. O autor refere a existência de esforços no sentido de criar tecnologias de inteligência artificial capazes de explicar os resultados por si obtidos, embora esta investigação ainda esteja numa fase embrionária. Entretanto, nas pp. 954-955 o autor compara esta prova com outros exemplos de prova inexplicável, como é o caso daquela que resulta da utilização de cães farejadores de droga. O homem treina o cão, e confia no seu resultado, mas não sabe explicar como o cão chega a esse resultado.

[72] Patrick W. Nutter, "Machine Learning Evidence: Admissibility and Weight", cit., pp. 949 ss. Veja-se o seguinte exemplo dado autor (ob. cit., p. 952): "*This principle was at work in a recent Stanford University study that aimed to build a machine learning algorithm that could analyze a person`s face and determine that person`s sexual orientation. Specifically, researchers compiled images of 75,000 user`s faces from various dating sites and used the profils` self-reported gay or straight identification to train the algorithm. From this pool of data, the algorithm focused on 35,000 images of 15,000 users to learn a set of correlations between the content of the images and the labels "gay" or "straight". Later, in a test set of different images that the machine had never seen before, the algorithm would make its best guess. The program was remarkably accurate at determining straight versus gay men, at eight-one percent accuracy, and slightly less accurate at sorting gay versus straight women, at seventy-one percent. Meanwhile, the machine was far more accurate than humans, who only correctly determined male sexual orientation sixty-one percent of the time and that of women fifty-four percent of the time. As to how the algorithm was making its relatively accurate determinations, the researchers could only speculate. One of their hypotheses was that the levels of different hormones in gay versus straight users (the prenatal hormone theory of sexual orientation) might have manifested some minute differences in their respective facial structures, differences unseen by the human eye but detectable by the algorithm. (...)*".

[73] Cf., *supra*, ponto 3.3, pp. 12-13.

resultado. Não há outro critério objetivo seguro de valoração. Com certeza a íntima convicção do juiz sobre o funcionamento das máquinas não constitui um elemento de apreciação da prova, dado o mesmo não ser racionalmente controlável ou sindicável pelo tribunal de recurso. Resta-nos enquadrar o problema em termos gerais, ou seja, no plano da livre, e racional, valoração da prova, sendo revelante a existência de outras provas que possam corroborar o resultado da máquina, tornando mais provável, de acordo com um critério de probabilidade prevalente, a respetiva hipótese de facto.

Reformulando o raciocínio: a valoração (racional) da prova envolve a utilização, pelo juiz, de raciocínios indutivo-abdutivos, e a utilização de regras de experiência comum. Quando o juiz valora, por ex., o depoimento de uma testemunha, considerando aspetos como a postura da mesma, o seu tom de voz, etc., é com base nestas regras de experiência que ele formula o seu juízo probatório. No futuro, imaginando a existência de máquinas-pessoas, conscientes (?), podem antever-se problemas interessantíssimos. Imagine-se, por ex., que uma máquina é chamada a depor por ter testemunhado um facto com relevo para a decisão da causa[74]. Como funciona a utilização de regras de experiência em relação às máquinas? Ou melhor: podem utilizar-se regras de experiência em relação ao comportamento de máquinas? Faz sentido impor à testemunha-máquina a prestação de juramento? São aspetos que, por enquanto, não podemos senão imaginar[75]. Entretanto, cremos que as máquinas inteligentes do presente são ainda e só produto humano, sem auto-realização. O controlo sobre o resultado da prova deve, assim, centrar-se nos *métodos e procedimentos* utilizados pelos humanos na programação da máquina, e nas taxas de erro que o programador for capaz de determinar. Com base nestes elementos, e nas demais provas existentes, o juiz deve determinar o grau de probabilidade que a prova atribui à respetiva hipótese de facto. A inexplicabilidade do resul-

[74] Claro está que o exemplo só é bom, se imaginarmos que o facto em causa não se encontra registado na máquina, através de um suporte que possa ser exibido em tribunal (nesse caso, sendo exibido o suporte, a valoração da prova centrar-se-ia, evidentemente, no controlo da autenticidade e fidedignidade do mesmo). Outros problemas podem imaginar-se. Na verdade, estando o facto documentado na máquina, pode supor-se a existência de um regime jurídico de proteção da máquina, que impede o acesso forçado ao mesmo. Então, sim, a obtenção da fonte de prova – o registo do facto histórico na memória da máquina (como o registo na memória de uma testemunha) –, implicaria o depoimento da mesma. O depoimento é o meio através do qual a fonte de prova se revela.

[75] Não é este o objeto do presente artigo, que trata apenas das hipóteses em que o software de aprendizagem automático é utilizado num processo, como instrumento para a apreensão do registo ou dos indícios de uma fonte de prova, o que subsume o problema da sua aplicação ao âmbito da prova técnica.

tado, não permitindo o controlo racional do mesmo, constitui, segundo cremos, um elemento que contribui para reduzir o grau de fiabilidade da prova.

5. Conclusão

Embora os "softwares de aprendizagem automática" estejam revestidos de um véu de neutralidade e de certeza, e pese embora os algoritmos utilizados na sua programação funcionem de um modo diferente dos programas tradicionais[76], eles ainda são criados como qualquer outro software: como produto de uma decisão humana, com código a correr em conjunto com outros softwares, e em hardware que se degrada com o tempo[77]. Podem, assim, ocorrer falhas humanas na programação da máquina. Alguns aspetos cruciais para apreciação da fiabilidade dos seus resultados dizem respeito à quantidade, qualidade, organização e adequação dos dados de treino, e à perfeição do código base.

Para além destes aspetos, que podem conduzir à exclusão da prova, ou condicionar o grau de confiabilidade dos seus resultados, um outro subsiste, que não deve perder-se de vista, e que diz respeito à característica de inexplicabilidade. Com efeito, acontece muitas vezes que ninguém, nem mesmo o próprio programador, é capaz de explicar o tipo de inferências realizadas pela máquina, ou seja, as regras que a mesma utilizou para alcançar um determinado resultado. Existem campos de investigação destinados à criação de máquinas inteligentes capazes de explicar os próprios resultados, em fase embrionária, mas no estado atual de desenvolvimento aspetos há que podem permanecer inexplicáveis.

Ao proceder à valoração da prova, o juiz deve determinar o grau de confirmação lógica que a mesma atribui a uma determinada hipótese de facto, reconstruindo racionalmente o respetivo facto histórico. A íntima convicção do juiz sobre o funcionamento das máquinas inteligentes não constitui um elemento que possa ser utilizado na valoração da prova. Ao juiz cabe atender a elementos objetivos de valoração, como as taxas de erro conhecidas ou potenciais da máquina, inferindo, a partir daí, em conjunto com as demais provas produzidas, o grau de probabilidade (epistémica) que a prova atribui ao *factum probandum*. A inexplicabilidade da prova, impedindo o controlo racional da mesma, constitui um elemento que reduz o grau de fiabilidade dos seus resultados.

[76] Implicando maiores conjuntos de dados, um maior poder de processamento, e diferentes metodologias.

[77] Patrick W. Nutter, "Machine Learning Evidence: Admissibility and Weight", cit., p. 941.

Bibliografia

Andrade, Manuel A. Domingues de, *Noções Elementares de Processo Civil*, Coimbra, Coimbra Editora, 1979.

Berger, Margaret A., "The Admissibility of Expert Testimony", *Reference Manual on Scientific Evidence*, Third Edition, Federal Judicial Center, The National Academies Press, Washington, United States of American, 2011, pp. 11-36.

Billauer, Barbara Pfeffer, "Admissibility of scientific evidence under Daubert: The fatal flaws of "Falsifiability" and "Falsification", Boston University Journal of Science and Technology Law, 2016, vol. 22, pp. 23-85.

Breyer, Stephen, "Introduction", *Reference Manual on Scientific Evidence*, Third Edition, Federal Judicial Center, The National Academies Press, Washington, United States of American, 2011, pp. 1- 9.

Canzio, Giovanni, "Prova scientifica, ricerca della "verità" e decisione giudiziaria nel processo penale", *in* Quaderni della Rivista Trimestrale di Diritto e Procedura Civile, 8, Decisione giudiziaria e verità scientifica, Milano – Dott. A. Giuffrè Editore, Milão, 2005, pp. 55-79.

Coimbra, Francisco Jorge Gemaque, *Juiz, Prova e Instrução probatória nos processos, À luz do Civil Law e do Common Law*, Porto, Juruá Editorial, 2018.

Capelo, Maria José, "A enigmática figura do técnico no Código de Processo Civil", *in* Estudos em Homenagem ao Prof. Doutor José Lebre de Freitas, vol. I, Coimbra Editora, pp. 1045-1067.

Comoglio, Luigi Paolo, *Le Prove Civili*, 3ª ed., Torino, UTET, 2010.

Denti, V., "Scientificità della prova e libera valutazione del giudice", *in* Riv. dir. proc., 1972, pp. 414 ss.

Dondi, Angelo, "Paradigmi processuali ed "expert witness testimony" nel diritto statunitense", Rivista Trimestrale di Diritto e Procedura Civile, Anno L (1996), pp. 261-285.

Edmond, Gary / Mercer, David, "Trashing "Junk Science", Stanford Technology Law Review 3 (1998), disponível em http://stlr.stanford.edu/STLR/Articles/98_STRL_3.

Freitas, José Lebre de, "La Preuve dans L`Union Européenne: Différences et Similitudes", *in* Estudos sobre Direito Civil e Processo Civil, vol. I, 2ª edição, Coimbra Editora, pp. 573-609.

__ *A ação declarativa comum, À luz do código de processo civil de 2013*, 4ª edição, GestLegal, 2017.

Freitas, José Lebre de/Machado, Montalvão/Pinto, Rui, *Código Processo Civil Anotado*, 2º Volume, 2ª ed., Coimbra Editora.

Goodstein, David, "How Science Works", *Reference Manual on Scientific Evidence*, Third Edition, Federal Judicial Center, The National Academies Press, Washington, United States of American, 2011, pp. 37-54.

Graham, Michael H. "The Expert Witness Predicament: Determining Reliable under the Gatekeeping Test of Daubert, Kumho, and Proposed Amended Rule 702 of the Federal Rules of Evidence", University of Miami Law Review 54, no. 2, January 2000, pp. 317-358.

__ *Federal Rules of Evidence in a nutshell*, 10th edition, West Academic Publishing, 2018, United States of American, comentário à Rule 702, pp. 367-401.

HUTSON, MATTHEW, "Computers Evolve a New Path Toward Human Intelligence", Quantamagazine, November, 6, 2019.

MATSON, JACK V. / DAOU, SUHA F. / SOPER, JEFFREY G., *Effective Expert Witnessing*, Fourth Edition, CRC Press, Boca Raton, London, Ney York, Washington D.C., 2004.

MEYER, PETRA CLAUDIA, *Der Sachverständigenbeweis zwischen Partei und Richter – Rechtsvergleich zum US- amerikanischen Zivilprozess und Reformansätze um deutschen Recht*, Münster, Nomos, 2013.

NUTTER, PATRICK W, "Machine Learning Evidence: Admissibility and Weight", University of Pennsylvania Journal of Constitutional Law 21, no. 3, February 2019, pp. 919-958.

RODRIGUES, FERNANDO PEREIRA, *Os meios de prova em processo civil*, 3ª edição, Almedina, 2017.

TARUFFO, MICHELE, "Le prove scientifiche nella recente esperienza statunitense", Rivista Trimestrale di Diritto e Procedura Civile, Anno L (1996), pp. 219-249.

__ "Conoscenza scientifica e decisione giudiziaria: profili generali", in Quaderni della Rivista Trimestrale di Diritto e Procedura Civile, 8, Decisione giudiziaria e verità scientifica, Milano – Dott. A. Giuffrè Editore, Milão, 2005, pp. 3-23.

__ *A prova*, trad. João Gabril Couto, 1ª ed., São Paulo, Marcial Pons, 2014.

SOUSA, JOÃO HENRIQUE GOMES DE, A "Perícia" Técnica ou Científica Revisitada Numa Visão Prático-Judicial, Julgar – Nº 15 – 2011, pp. 27-52.

VARELA, MANUEL ANTUNES/BEZERRA, JOSÉ MIGUEL/NORA, SAMPAIO E, *Manual de Processo Civil de acordo com o Dec.-Lei 242/85* (2ª Edição Reimpressão), Coimbra Editora, 2004.

L'intelligenza artificiale, le professioni legali e il dovere di competenza tecnologica dell'avvocato

Artificial intelligence, legal professions and lawyer's duty of technological competence

Giorgia Anna Parini*

RESUMO: A inovação tecnológica coloca vários desafios ao direito, especialmente em relação à crescente utilização de formas de inteligência artificial. Deve-se notar que o uso generalizado de sistemas de IA não apenas apoia os profissionais em atividades repetitivas, mas até os substitui totalmente. Este aspecto requer alguma reflexão, especialmente no que diz respeito ao desempenho de atividades que tradicionalmente foram reservadas a "profissões protegidas", com problemas ainda relacionados ao processo contratual e responsabilidade. Além disso, o profissional passa a ser obrigado a conhecer as tecnologias que podem apoiá-lo na realização da tarefa em cumprimento de um dever de competência tecnológica. Nesses casos, o profissional precisa saber acompanhar o trabalho da inteligência artificial, pois de qualquer forma será o profissional quem será responsabilizado por eventual inumprimento de contrato com o cliente. O profissional tem, portanto, uma obrigação de competência tecnológica: no entanto, é preciso compreender os limites desse compromisso.

PALAVRAS-CHAVE: inteligência artificial, profissões protegidas, atividades reservadas, dever de competência tecnológica, advogados, responsabilidade.

ABSTRACT: The law is currently facing various challenges related to technological innovation, in particular the increasing utilization of forms of Artificial

* Ricercatore a tempo determinato di tipo b senior. Università degli Studi di Verona . Dipartimento di Scienze Giuridiche. Via Montanari n. 9 Verona (VR). Giorgiaanna.parini@univr.it

Intelligence. It should be noted that the widespread usage of AI systems does not only support professionals in repetitive tasks but even replaces them entirely. This aspect requires some reflection, especially in respect of the performance of activities which have traditionally been reserved to "protected professions", with further problems related to contractual process and liability. Furthermore, the professional is now required to know the technological tools that can support him in carrying out the assignment in compliance with a duty of technological competence. In these cases, the professional need to know how to monitor the work of the technology tool, because in any case it will be the professional who will be liable for any breach of contract with the client. The professional therefore has an obligation of technological competence: however, it is necessary to understand the limits of this commitment.

KEYWORDS: artificial intelligence, protected professions, reserved activities, duty of technological competence, lawyers, liability.

1. Introduzione

La diffusione di macchine autonome e intelligenti impone riflessioni circa gli effetti del fenomeno sulla società nel suo complesso e sugli assetti del mondo del lavoro: da diversi lustri, tali meccanismi vengono impiegati nella produzione di beni in sostituzione degli esseri umani e sempre più di frequente sono adoperati nella prestazione di servizi. In tale ambito, si rivelano, infatti, preziosi per la loro rapida capacità di elaborare dati e – grazie all'accesso al web e all'analisi dei *big data* – di disporre di un bagaglio di conoscenze impensabile per un essere umano.

Proprio in virtù di tali caratteristiche si assiste a un sempre più significativo e pressoché quotidiano impiego dei sistemi – più o meno intelligenti – come "sostegno" ai professionisti per eseguire i compiti più ripetitivi: da tempo la tecnologia ha fatto il proprio ingresso negli studi professionali, agevolando parte del lavoro[1] e producendo significativi impatti sullo svolgimento e l'or-

[1] Rileva tale tendenza P. Moro, *Intelligenza artificiale e professioni legali. La questione del metodo*, in *Journal of Ethics and Legal Technologies*, 2019, I, p. 25.

ganizzazione dello stesso. Tale tendenza, tuttavia, impone ai professionisti, che vogliano rimanere competitivi nel mercato e adempiere diligentemente al proprio incarico, un costante aggiornamento e l'assunzione di competenze basilari riguardanti gli strumenti tecnologici a disposizione e il loro funzionamento, con la complessità di comprendere sino a che punto ciò è esigibile.

Non solo però la macchina viene impiegata a "supporto" del prestatore d'opera: vi sono, infatti, ipotesi ove si realizza una vera e propria "sostituzione" dell'uomo con la macchina, situazione che desta più di una preoccupazione per le ricadute che comporta e che è opportuno tenere in debita considerazione. Tale evidenza, peraltro, apre numerosi scenari di indagine con riferimento allo svolgimento di quelle attività tradizionalmente riservate alle cc.dd. professioni protette: pensiamo, a mero titolo esemplificativo, ai professionisti forensi, ai commercialisti, agli psicologi, ai medici e agli ingegneri. Tale scenario, inevitabilmente, offre l'occasione per valutare i limiti entro i quali l'impiego di tali tecnologie in ambiti di tale fatta è conforme alla legge e di confrontarsi anche sulle ricadute che il fenomeno produce sul versante della patologia del contratto[2].

Necessariamente, la questione si inserisce in un panorama più ampio che vede il giurista chiamato ad affrontare le molteplici sfide connesse all'evoluzione tecnologica, allo sviluppo di macchine dotate della capacità di apprendere automaticamente (c.d. *machine learning*)[3] e – più in particolare – al sempre maggiore ricorso a forme d'intelligenza artificiale[4]: inevitabilmente, tale tendenza porta a meditare sulle stesse categorie in ragione delle ricadute di tale *trend* e valutare l'idoneità dell'ordinamento a rispondere alle nuove istanze di tutela[5].

[2] Sul punto v. tra i tanti R. Clarizia, *I contratti informatici*, in *Trattato dei contratti*, Giappichelli, 2007, *passim*.

[3] Rivelano la complessità sottesa a individuare quando un robot o comunque un meccanismo di intelligenza artificiale abbia tali crismi G. Sartor e A. Omicini, *The autonomy of technological systems and responsibilities for their use*, in *Autonomous Weapon Systems. Law, Ethics, Policy*, Cambridge University Press, 2016, p. 39; A. Santosuosso, *Diritto scienza e nuove tecnologie*, Cedam, 2016 (2° ed.), p. 330.

[4] Sul punto v. A. Turing, *Computing Machinery and Intelligence*, in *Mind: new series*, 1950, 239, p. 443. Definire cosa si intenda per intelligenza artificiale e per robot è questione complessa. Sul punto v. A. Santosuosso, C. Boscarato, F. Caroleo, *Robot e diritto: una prima ricostruzione*, in *Nuova giur. civ. comm.*, 2012, II, p. 497; G. Sartor, *L'informatica giuridica e le tecnologie dell'informatica. Corso d'informatica giuridica*, Giappichelli, 2016, *passim*.

[5] In generale sull'influenza che l'evoluzione tecnologica ha sul diritto v. A. Santosuosso, *A general theory of law and technology or a general reconsideration of law?*, in E. Palmerini e E. Stradella (a cura di), *Law and Technology. The Challenge of regulating technological development*, Pisa, Pisa University Press, 2013, p. 146.

2. Il diritto dei contratti e il ricorso ai sistemi di intelligenza artificiale

Prima di affrontare le problematiche giuridiche sottese al demandare l'esecuzione di una prestazione d'opera intellettuale a una macchina, occorre in prima battuta chiarire quali siano le parti del contratto, in caso di conferimento di incarichi professionali di tale fatta.

La diffusione degli strumenti tecnologici ha avuto nel tempo un impatto significativo sulla dimensione dei rapporti contrattuali in ragione del fatto che la negoziazione si è trasferita *online*. Tale fenomeno porta nuovi interrogativi all'interprete, in quanto influisce sul procedimento di formazione del contratto: nello specifico, ha posto nuovi dilemmi circa l'attuale valenza del principio del consenso e della compatibilità con la disciplina prevista dall'art. 1326 c.c. e ss. [6], conducendo autorevole dottrina ad affermare trattarsi di contratti senza accordo, mera combinazione di decisioni unilaterali[7]; soluzione che però non appare del tutto convincente, giacché la contrattazione appare comunque riconducibile nelle modalità tradizionali di perfezionamento dell'intesa[8].

Lo spostamento della contrattazione *online* rende, inoltre, talvolta difficoltosa la concreta individuazione delle parti del contratto: sono evidenti, infatti, la complessità e le problematiche sottese al profilo dell'identificazione di chi opera *online* e che sovente cela la propria identità, aspetto determinante per comprendere la disciplina applicabile[9].

[6] Tra i numerosi saggi che evidenziano tali problematiche v.: V. Pasquino, *La vendita attraverso reti telematiche. Profili civilistici*, in *Dir. inform. Informatica*, 1990, p. 697; S. Nepor, *Internet e la legge*, Milano, Giuffrè, 1999, p. 217.

[7] Così N. Irti, *Scambi senza accordo*, in *Riv. trim. dir. proc. civ.*, 1998, p. 347; N. Irti, *«È vero ma...» (Replica a Giorgio Oppo)*, in *Riv. dir. civ.*, 1999, I., p. 273. *Contra*, v. G. Oppo, *Disumanizzazione del contratto?*, in *Riv. dir. civ.*, 1998, I, p. 525. Al riguardo, C. Camardi, *Contratto e rapporto nelle reti telematiche. Un nuovo modello di scambio*, in V. Ricciuto e N. Zorzi (a cura di), *Il contratto telematico*, Cedam, 2002, p. 14, osserva che il contratto delle reti telematiche diviene un atto sempre più governato da una procedura standardizzata irreversibile, decontestualizzato rispetto al tempo e allo spazio; un atto non importa se bilaterale o risultante dalla combinazione di due atti unilaterali, comunque lontano dal concetto di negozio che ha fin qui nutrito le riflessioni del civilista. Sul punto v. anche C.M. Bianca, *Acontrattualità dei contratti di massa?*, in *Vita not.*, 2001, p. 1120; P. Perlingieri, *Metodo, categorie, sistema nel diritto del commercio elettronico*, in *Il diritto dei contratti fra persona e mercato*, Napoli, Esi, 2003, p. 652; F. Galgano, *La categoria del contratto alla soglia del terzo millennio*, in *Contr. impr.*, 2000, p. 223.

[8] Sul punto sia consentito il rimando G.A. Parini, *Riflessioni sul ricorso all'intelligenza artificiale nelle professioni intellettuali*, in *Comp. dir. civ.*, 2019, p. 3, ove sono presenti ulteriori richiami dottrinali e giurisprudenziali.

[9] Sul punto v.: E. Roppo, *Behavioural Law and Economics, regolazione del mercato e sistema dei contratti*, in *Riv. dir. priv.*, 2013, p. 168 ss.; G. Dore, *I doveri di informazione nella rete degli scambi commerciali telematici*, in *Giur. merito*, 2013, p. 2569; S. Nardi, *Accordo concluso* online *dal minore di età*, in *Comp. dir. civ.*, 1 ss.

L'individuazione delle parti del contratto, peraltro, è resa maggiormente complessa in ragione dell'evidenza per la quale la macchina sta vedendo nel tempo progressivamente mutare il suo ruolo nella contrattazione da strumento che agevola la comunicazione a partecipante "attivo" nella negoziazione[10]: ad esempio si consideri il *software* che effettua un ordine di merce – poiché è stato programmato per farlo al verificarsi di talune condizioni –, contingenza che ha portato taluni a sostenere che esso concorra in una certa misura al procedimento di formazione della volontà[11].

La questione non è di poco conto se si considera la sempre maggiore diffusione del fenomeno dell'*Internet of Things*[12] e delle sue possibili implicazioni; ancora, sotto tale versante non si può non fare cenno al fenomeno degli *smart contracts*[13] – le cui potenzialità sono senz'altro accresciute grazie allo sviluppo della tecnologia *blockchain* – nei quali un algoritmo esegue automaticamente al verificarsi di determinate condizioni una funzione per la quale è stato programmato nel rispetto di regole preimpostate[14] e secondo il meccanismo *if-then*[15].

[10] T. Allen e R. Widdison, *Can computers make contracts?*, in *Harvard journal of law & technology*, 1996, 9, I , p. 26. Sul punto v. G. Sartor, *Gli agenti software. Nuovi soggetti di ciberdiritto*, in *Contr. impr.*, 2002, p. 465 ss.

[11] Così, tra i tanti, G.A. Cavaliere e M. Iaselli, *Contratto telematico*, in G. Buffone, I. De Giovanni, A.I. Natale, *Il contratto*, Padova, Cedam, 2013, p. 1571 ss., sostengono che nei contratti cibernetici – conclusi tra persone e computer come parti contraenti contrapposte – il processo di formazione della volontà sia riconducibile alla macchina.

[12] S. Greengard, *Internet delle cose*, Bologna, Il Mulino, 2017, passim; N. Bouhai e I. Saleh, *Internet of Things: Evolutions and Innovations*, Wiley, 2017, *passim*.

[13] Mediante legge, 11 febbraio 2019, n. 12 che ha convertito il decreto legge, 14 dicembre 2018, n. 135 (il c.d. D.L. semplificazioni), gli *smart contracts* sono stati espressamente riconosciuti dal legislatore, che all'art. 8 *ter*, secondo comma, ha previsto che «Si definisce *smart contract* un programma per elaboratore che opera su tecnologie basate su registri distribuiti e la cui esecuzione vincola automaticamente due o più parti sulla base di effetti predefiniti dalle stesse. Gli *smart contracts* soddisfano il requisito della forma scritta previa identificazione informatica delle parti interessate, attraverso un processo avente i requisiti fissati dall'Agenzia per l'Italia digitale con linee guida da adottare entro novanta giorni dalla data di entrata in vigore della legge di conversione del presente decreto». Peculiare appare la definizione fornita dal legislatore, secondo il quale l'esecuzione di tale protocollo "vincola" le parti agli effetti dalle stesse prestabiliti, espressione che porta a stimare che questi abbia ritenuto gli *smart contracts* veri e propri contratti: plurime riflessioni sorgono ispirate da tale aspetto, nonché dal riconoscimento dell'idoneità a integrare forma scritta, ma non è questa la sede per intervenire sul punto. Critica la norma nel momento in cui descrive la tecnologia della *blockchain* e le applicazioni di *smart contract*, cristallizzandoli e così non rispettando il principio della neutralità tecnologica G. Finocchiaro, *Intelligenza Artificiale e protezione dei dati personali*, in *Giur. it.*, 2019, p. 670 s.

[14] Sugli *smart contracts* v. N. Szabo, *Formalizing and Securing Relationships on Public Networks*, in *First Monday*, 2015, p. 2; M. Raskin, *The law and the legality of smart contracts*, in *Georgetown law Technology*

In realtà, a bem vedere, anche in dette ipotesi, la macchina non è parte del contratto e non assume la qualifica di contraente[16], in quanto, in prima battuta, non è dotata – allo stato – di soggettività, con la conseguenza che non integra un centro di imputazione di diritti e obblighi. Inoltre, la volontà è pur sempre riconducibile all'utilizzatore[17], che – manifestando a monte il proprio intento negoziale – ha dato impulso al procedimento e fornito allo strumento di intelligenza artificiale tutte le indicazioni e le opzioni che questo procede poi ad applicare. Quanto, ad esempio, agli *smart contracts*, basti considerare che il susseguirsi di processi automatici è conseguenza dell'intento negoziale manifestato in precedenza, all'inizio del processo, con la conseguenza che quello che viene chiamato contratto intelligente non è in realtà un contratto, quanto una modalità di attuazione dell'intesa[18]: all'interno di tale schema, infatti, la macchina si limita a compiere un automatismo, senza spazi di autonomia.

In quest'ottica, va censurata, altresì, quella tesi secondo la quale nel caso concreto si attuerebbe un meccanismo di rappresentanza[19], giacché non assistiamo a un fenomeno di sostituzione nell'attività contrattuale: il sistema di intelligenza artificiale non manifesta, infatti, una volontà propria e – alla base – non è un soggetto di diritto e, dunque, non è un autonomo centro di imputazione e di responsabilità, con la conseguenza che non si assiste a una scissione tra parte in senso formale e parte in senso sostanziale. Piuttosto, siccome la macchina si limita a fungere da "mezzo di trasmissione" della volontà

Revue, 2017, p. 305. Per la dottrina italiana v. Cuccuru P., Blockchain *e automazione contrattuale. Riflessioni sugli* smart contracts, in *Nuova giur. civ. comm.*, 2017, II, p. 107; nonché Castellani G., Smart contracts *e profili di diritto civile*, in *Comp. dir. civ.*, 2019, p. 1.

[15] Sul punto v. D. Di Sabato, *Gli* smart contracts*: robot che gestiscono il rischio contrattuale*, in *Contr. impr.*, 2017, p. 378 ss. V. anche L. Parola, P. Merati, G. Gavotti, Blockchain *e* smart contract*: questioni giuridiche aperte*, in *Contratti*, 2018, p. 681.

[16] V. G. Sartor, *Gli agenti software. Nuovi soggetti di ciberdiritto*, cit., p. 465; G. Finocchiaro, *La conclusione del contratto mediante i* "software agents"*: un falso problema giuridico?*, in *Contr. impr.*, 2002, p. 500.

[17] Così G. Finocchiaro, *I contratti informatici*, in *Tratt. dir. comm. dir. pub. ec.*, a cura di F. Galgano, Cedam, 1997, p. 60.

[18] V. Caggiano I.A., *Il contratto nel mondo digitale*, in *Nuova giur. civ. comm.*, 2018, II, p. 1152 ss. Peraltro, evidenzia come il meccanismo si connoti per l'automaticità e sia poco "*smart*", F. Di Giovanni, *Sui contratti delle macchine intelligenti*, in *Intelligenza artificiale. Il diritto, i diritti e l'etica*, Giuffrè, 2020, p. 257.

[19] Così R. Borruso, *Computer e diritto*, II, Milano, Giuffrè, 1988, p. 259. Diversamente, F. Bravo, *Contratto cibernetico*, in *Dir. inform. Informatica*, 2011, 169 ss., sostiene che tali tecnologie sono usate dall'utilizzatore non per veicolare una volontà contrattuale già formata, bensì per integrare la dichiarazione negoziale e giungere alla definitiva formazione del regolamento contrattuale.

altrui, si potrebbe assimilare a un peculiare *nuncius* (privo però di soggettività e di capacità d'agire) che non partecipa al procedimento di formazione della volontà, con la conseguenza che, sotto il profilo dei vizi della volontà e della rilevanza degli stati di capacità, l'attenzione va sempre posta sull'utilizzatore.

Occorre però considerare – soprattutto con lo sguardo rivolto al futuro – che la creazione di meccanismi di intelligenza artificiale sempre più sofisticati, dotati di capacità di apprendere dall'esperienza e interagire col mondo esterno in termini sempre più autonomi e attuare finanche una propria strategia, rende certo ancora più complesso ricondurre la vicenda nei paradigmi tradizionali e giustificare le ragioni che portano eventualmente il soggetto utilizzatore a rispondere ed essere vincolato dal contratto concluso dalla macchina seguendo una strategia propria con esiti non del tutto prevedibili a monte e non preordinati[20].

L'esercizio dell'autonomia privata dell'utilizzatore in questo caso appare affatto peculiare poiché questi – scegliendo di ricorrere a tali strumenti – diviene parte contrattuale e si vincola anche in relazione a contratti il cui contenuto – in parte – potrebbe sfuggire dal suo controllo in virtù del principio di autoresponsabilità. E ciò in quanto – giova rilevarlo – l'aspetto centrale rimane la tutela dell'affidamento di chi ricevendo la dichiarazione ha attuato comportamenti coerenti con la stessa[21].

3. L'esercizio delle professioni protette e l'utilizzo dei sistemi di intelligenza artificiale

Come accennato, numerosi meccanismi dotati di intelligenza artificiale sono in grado di fornire all'utente risposte a domande specifiche, così consentendo l'automatizzazione della prestazione di taluni servizi. Nello specifico, sempre più di frequente tramite l'intelligenza artificiale vengono forniti servizi

[20] A T. Allen e R. Widdison, *Can computers make contracts?*, in *Harvard journal of law & technology*, 1996, 9, I, p. 28. Sul punto F. Di Giovanni, *Attività contrattuale e Intelligenza Artificiale*, in *Giur. it.*, 2019, p. 1683. Sul punto, A. Musio, *La storia non finita dell'evoluzione del contratto tra novità tecnologiche e conseguenti esigenze di regolazione*, in *Nuova giur. civ. comm.*, 2021, II, p. 233 s., evidenzia come nelle ipotesi in cui l'investitore demandi a un sistema di intelligenza artificiale la scelta del tipo di investimento patrimoniale da effettuare, dando quale unica indicazione semplicemente quella di spuntare il miglior rendimento possibile, il ruolo ascritto al medium elettronico nella formazione dell'accordo è destinato a cambiare, in quanto da mero mezzo di trasmissione svolto nella c.d. contrattazione telematica, diventa strumento di integrazione della volontà negoziale già formatasi.

[21] Evidenzia che al principio di affidamento, quale tutela posta a favore di chi riceve la dichiarazione, corrisponda il principio di autoresponsabilità di chi con la sua condotta ha determinato il sorgere di tale affidamento F. Ruscello, *Istituzioni di diritto civile*, Bari, Cacucci, 2017, p. 128. Nei medesimi termini anche S. Monti, *Il contratto dall'uomo alla macchina e... viceversa?*, in *Contr.*, 2021, p. 459.

riconducibili a quelle attività tradizionalmente svolte dai prestatori d'opera intellettuale.

Tenendo fermo l'assunto – in precedenza affermato – secondo il quale parte del contratto sono comunque esclusivamente le persone ed è solamente l'esecuzione dello stesso che viene demandata alla macchina, occorre chiarire entro quali confini è possibile per coloro che non sono professionisti (e che magari sono per giunta privi di competenze specifiche in quel settore) esercitare una professione intellettuale, magari demandandone l'esecuzione a un sistema di intelligenza artificiale: vi sono, infatti, precisi limiti per l'esercizio di tali attività, circostanza che appare dall'analisi della disciplina relativa al contratto d'opera intellettuale di cui agli artt. 2229 c.c. e ss.[22].

Tale tipologia contrattuale si connota per le particolari modalità tramite le quali la prestazione è svolta, nonché il carattere intellettuale dell'attività, postulante una specifica competenza in capo al professionista, chiamato all'adempimento dell'incarico ricevuto nell'interesse del cliente-creditore[23].

Elemento che connota tale fattispecie è, peraltro, la circostanza che il commissionario è un libero professionista, che esercita la propria opera in condizioni d'indipendenza e in piena autonomia, con la precisazione che per esercitare alcune professioni (le cc.dd. professioni protette)[24] la legge richiede l'iscrizione in appositi albi ed elenchi, assoggettando i soggetti al controllo e al potere disciplinare delle associazioni professionali[25].

[22] Sui contratti d'opera intellettuale in generale v., tra i tanti: F. Santoro Passarelli, voce *Professioni intellettuali*, in *Noviss. dig. it.*, Torino, Utet, 1968, XIV, p. 24 ss.; C. Lega, *Le libere professioni intellettuali*, Milano, Giuffrè, 1974, *passim*; A. Perulli, *Il lavoro autonomo*, in *Tratt. Cicu – Messineo*, I, Milano, Giuffrè, 1996, p. 351; A. Anastasi, *Professioni intellettuali – Dir. lavoro*, in *Enc. giur.*, Roma, Ed. Enc. it., 1991, XXIV, p. 3; G. Musolino, *Contratto d'opera professionale. Artt. 2229 – 2238 c.c.*, Milano, Giuffrè, 2016 (2° ed.), *passim*; R. Salomone, *Le libere professioni intellettuali*, Padova, Cedam, 2010, *passim*.

[23] V., sul punto, L. Riva Sanseverino, *Lavoro autonomo*, in *Del lavoro autonomo*, Zanichelli, 1963, p. 191 ss.

[24] Tale definizione è riportata nella nota sentenza Corte Cost., 22 gennaio 1976, n. 17, *Riv. dir. lav.*, 1976, II, 47 ss.

[25] V. L. Riva Sanseverino, *Lavoro autonomo*, cit., p. 194. Rileva, condivisibilmente, come la mera iscrizione a un albo non costituisca la *condicio sine qua non* per assicurare la competenza di un professionista, R. Mazzariol, *Attività di psicoanalista e professioni intellettuali «protette»: spunti per una riflessione critica*, in *Nuova giur. civ. comm.*, 2013, II, p. 423. Più in generale, evidenzia che è l'idea dell'attività professionale come diretta a tutelare un interesse pubblico – l'idea cioè dell'esercizio di un ufficio di diritto privato – ad attrarre le professioni intellettuali nell'ambito di una disciplina pubblicistica che incide sulla regolamentazione del rapporto a tutela di un interesse collettivo, A. Perulli, *Il lavoro autonomo*, cit., p. 356.

La conseguenza dell'esercizio di una professione protetta senza le prerogative citate è – di là dagli importanti risvolti sul piano penale, quali l'integrazione degli estremi del reato di esercizio abusivo di una professione *ex* art. 348 c.p., e sotto il versante deontologico – l'esclusione del diritto al compenso *ex* art. 2231 c.c.[26], che trova la propria ragion d'essere nella nullità del contratto, che si pone in contrasto con norme imperative[27].

Tali conseguenze, però, – nonostante alcune opinioni contrarie[28] – si verificano solamente laddove la legge stabilisca un'esclusiva a favore di coloro che possiedono tali requisiti, circostanza che spinge a indagare quali siano le attività espressamente riservate dal legislatore a tali soggetti.

I contratti d'opera intellettuali si caratterizzano, inoltre, per il carattere fiduciario del rapporto, basato sull'*intuitu personae*[29], fondato sull'affidamento che il cliente pone nei confronti del professionista: da tale elemento deriva, secondo il disposto di cui all'art. 2232 c.c., che il professionista deve eseguire la prestazione personalmente, potendo avvalersi – sotto la propria direzione e responsabilità – di sostituti e ausiliari esclusivamente laddove tale collaborazione sia consentita dal contratto o dagli usi e non sia incompatibile con l'oggetto della prestazione[30]. Pure sussistendo voci contrarie al riguardo[31], si

[26] La Suprema Corte non ha ritenuto esperibile nel caso di specie neppure l'azione generale per ingiustificato arricchimento di cui all'art. 2041 c.c. (Cass., 2 ottobre 1999, n. 10937, in *Plurisonline.it*). Nei medesimi termini anche Cass., 28 maggio 2021, n. 15004, in *Plurisonline.it*.; Cass., 3 novembre 2000, n. 14381, in *Plurisonline.it*; Cass., 28 marzo 2019, n. 8683, in *Dir. giust*., 60, 2019, 2 ss.; nonché Cass., 11 giugno 2010, n. 14085, in *Giust. civ*., 2011, I, 987 ss.

[27] V. sul punto e tra i tanti, F. Carnelutti, *Nullità del contratto di patrocinio per difetto del titolo professionale*, in *Riv. dir. proc*., 1953, I, p. 313 ss. Nei medesimi termini in giurisprudenza Cass., 3 novembre 2000, n. 14381, cit.

[28] Le Sezioni Unite della Suprema Corte affermano che «concreta esercizio abusivo della professione, punibile ai sensi dell'art. 348 c.p., non solo il compimento senza titolo, anche se posto in essere occasionalmente e gratuitamente, di atti da ritenere attribuiti in via "esclusiva" ad una determinata professione, ma anche il compimento senza titolo di atti che, pur non attribuiti singolarmente in via esclusiva, siano univocamente individuati come di competenza "specifica" di una data professione, allorché lo stesso compimento venga realizzato con modalità tali, per continuatività, onerosità e (almeno minimale) organizzazione, da creare, in assenza di chiare indicazioni diverse, le oggettive apparenze di un'attività professionale svolta da soggetto regolarmente abilitato» (Cass. Pen. Sez. Unite, 23 marzo 2012, n. 11545, in *Plurisonline.it*). Nei medesimi termini Cass., 18 luglio 2018, n. 33464, in *Plurisonline.it*; nonché Trib. Bologna, 2 marzo 2021, n. 189, in *Plurisonline.it*.

[29] Sulla rilevanza di tale profilo nei rapporti contrattuali, v. A. Cataudella, Intuitus personae *e tipo negoziale*, *Studi in onore di F. Santoro Passarelli*, 1972, Esi, p. 631.

[30] V., tra i tanti, G. Giacobbe, voce *Professioni intellettuali*, in *Enc. dir.*, Milano, Giuffrè, 1987, XXXVI, p. 1074.

[31] Siccome i sostituti e gli ausiliari non diventano parte del rapporto di clientela, restando invece la loro attività giuridicamente assorbita da quella del prestatore d'opera che ha concluso il contratto

ritiene che anche tali soggetti debbano possedere i requisiti richiesti per l'espletamento dell'incarico, poiché la mancata iscrizione degli ausiliari all'albo o ruolo professionale è stimata equiparabile all'esercizio di attività del professionista incaricato non iscritto[32]. A ogni modo, la facoltà per il professionista di servirsi, ai sensi dell'art. 2232 c.c., della collaborazione di sostituti ed ausiliari, non rende tali soggetti parte del contratto d'opera.

Applicando le riflessioni ora svolte alla fattispecie in esame, il problema è, dunque, individuare – parlando di professioni protette – quale attività è strettamente riservata ai professionisti, questione che è delineata dalle diverse leggi professionali. Proprio le leggi professionali indicano – seppure talvolta in termini sibillini – i limiti al dilagare del fenomeno, travalicati i quali il contratto sarebbe nullo e verrebbe meno il diritto al corrispettivo secondo quanto sancito dall'art. 2231 c.c.

L'efficacia deterrente e dissuasiva alle conseguenze negative sotto il versante civilistico trova, tuttavia, un freno nell'evidenza per la quale molti di tali servizi sono offerti (apparentemente) gratuitamente[33].

Di contro invece, il professionista, nell'ambito professionale di riferimento, può certo scegliere di avvalersi nello svolgimento dell'incarico anche di strumenti – più o meno intelligenti – in grado di prestare un importante supporto allo stesso. In questo caso, la macchina, non avendo una soggettività, non può essere stimata sostituto o ausiliario del professionista – circostanza che, come accennato, comporterebbe talune complicazioni – ma mero mezzo di supporto: profilo estremamente rilevante diviene in tale ipotesi la capacità del professionista di servirsi dello strumento tecnologico e controllare l'operato dello stesso, aspetto sul quale occorre meditare anche considerando che comunque sarà il professionista a rispondere dell'eventuale inattuazione degli impegni assunti mediante il contratto. Siccome i sistemi di intelligenza

con il cliente, poco importa che la essi siano o meno abilitati, ciò che conta è che sia abilitato il professionista incaricato secondo Cass., 9 luglio 2021, n. 24374, in *Plurisonline.it*, nonché Cass., 18 ottobre 2018, n. 26264, in *Plurisonline.it*.

[32] Pretura di Torino, 23 ottobre 1998, in *Foro it.*, 1999, I, c. 710. In dottrina v. G. Musolino, *Contratto d'opera professionale. Artt. 2229 – 2238 c.c.*, cit., p. 276 ss.

[33] In realtà chi fornisce tali servizi, pur non pretendendo un corrispettivo, ottiene vantaggi indiretti: ad esempio, così facendo, chi fornisce tali servizi entra in possesso di una serie di informazioni – quali i gusti e le preferenze degli utenti – dal significativo valore commerciale, ottenendo così vantaggi economici sia pure indiretti. Addirittura, G. Resta e V. Zeno-Zencovich, *Volontà e consenso nella fruizione dei servizi in rete*, in *Riv. trim. dir. proc. civ.*, 2018, p. 411, sostengono che, anche senza il pagamento di un corrispettivo, si tratterebbe di contratti sinallagmatici, poiché a fronte del servizio di cui fruisce, l'utente fornisce i propri dati, i quali insieme ai "metadati" a essi associati, costituiscono un bene oggetto di relazioni economiche e giuridiche.

artificiale non hanno una soggettività e non integrano un autonomo centro di imputazione, chiamato a rispondere di eventuali pregiudizi sarà, infatti, solamente il debitore (ovvero la persona fisica o giuridica), che si era impegnato ad adempiere mediante il contratto e che ha demandato l'esecuzione della prestazione al robot o al *software*.

Troverà, dunque, applicazione la disciplina generale di cui agli artt. 1218 c.c. e ss., la quale, pur risalente, è dotata di una certa flessibilità. Nulla vieta che poi, successivamente, il professionista si attivi verso altri soggetti per chiamarli a rispondere dell'operato della macchina, con la precisazione che, non sussistendo allo stato alcuna previsione dettata *ad hoc* per regolare la materia, si dovrà fare ricorso alla disciplina generale[34]. Al riguardo, si ricorda che per fare fronte a esigenze di certezza,sono state ipotizzate diverse soluzioni[35], quali la possibile introduzione di un nuovo sistema di responsabilità *ad hoc* di

[34] Sul punto v.: U. Ruffolo, *Per i fondamenti di un diritto della robotica*, in *Intelligenza artificiale e responsabilità,* Milano, Giuffrè, 2018, p. 8; Id., *Intelligenza Artificiale,* machine learning *e responsabilità da algoritmo*", in *Giur. it.*, 2019, p. 1689; C. Boscarato, *Who is responsible for Robot's actions? An initial examination of Italian law within a European perspective*, in B. Van Berg e L. Klaming (a cura di), *Technologies on the stand: legal and ethical questions in neuroscience and robotics*, Wolfpublisher, 2011, p. 383 ss.; M. Costanza, *L'intelligenza artificiale e gli stilemi della responsabilità civile*, in *Giur. it.*, 2019, p. 1687 ss. In considerazione della peculiarità della materia vi è chi ha ipotizzato l'impiego del principio di precauzione per giustificare la responsabilità del produttore (G. Capilli, *Responsabilità e robot*, in *Nuova giur. civ. comm.*, 2019, II, p. 623).

[35] V. sul punto, tra i diversi interventi, la Risoluzione del Parlamento europeo del 16 febbraio 2017 recante raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica. Si segnala poi la Proposta di Regolamento del Parlamento Europeo e del Consiglio che mira a stabilire regole armonizzate sull'intelligenza artificiale e a modificare alcuni atti legislativi dell'Unione del 21 aprile 2021. Si è arrivati poi nel giugno 2023 alla approvazione della bozza di regolamento (c.d. Artificial Intelligence act) da parte del Parlamento europeo: ora verrà coinvolto il Consiglio per decidere sulla questione e arrivare alla versione finale. Sul punto v. F. Costantini, *Intelligenza artificiale, design tecnologico e futuro del lavoro nell'UE: i presupposti e il contesto*, *Lav. nella giur.*, 2021, p. 807. V. anche: European Parliament 2019-2024: Special Committee on Artificial Intelligence in a Digital Age, Draft Report on Atificial Intelligence in a Digital Age (2020/2266(INI), del 2 novembre 2021. Nell'ottobre del 2020 è intervenuta, inoltre, una risoluzione del Parlamento europeo recante raccomandazioni alla Commissione tese a spingerla a intervenire sul regime di responsabilità civile per il funzionamento dell'intelligenza artificiale. Nella bozza sono contenute interessanti previsioni in punto di individuazione del soggetto chiamato a rispondere per il danno cagionato dall'intelligenza artificiale. Abbandonata l'aspirazione all'introduzione di una personalità elettronica, si punta sull'introduzione di un regime comune di responsabilità oggettiva per i sistemi di IA autonomi ad alto rischio. Inoltre, tutti gli operatori di sistemi di IA ad alto rischio elencati nell'allegato al regolamento proposto dovrebbero essere in possesso di un'assicurazione per responsabilità civile. Successivamente è intervenuta la proposta della AI Liability Directive dell'ottobre 2022 tesa ad armonizzare il regime di responsabilità applicabile laddove un danno sia causato da sistemi di intelligenza artificiale. Si cerca, in particolare, di introdurre una "presumption of causality" al fine di agevolare il danneggiato.

tipo oggettivo o comunque ispirato alla gestione dei rischi, l'istituzione di un regime assicurativo obbligatorio per categorie specifiche di robot, nonché la costituzione di un fondo di risarcimento per garantire il risarcimento nelle ipotesi di inoperatività della copertura assicurativa e l'istituzione di uno *status* giuridico specifico per i *robots* e i *software agents* di modo che quelli più autonomi possano essere considerati come "persone elettroniche"; e ciò al fine di chiamarli a rispondere dei danni dagli stessi cagionati[36].

Tale ultima soluzione non è stata ulteriormente avvallata nelle proposte di regolamenti e direttive successive. E' ciò in quanto, del tutto condivisibilmente, pare eccessiva rispetto alle finalità che si intendono perseguire e non considera che il riconoscimento della qualità di soggetto postula – quale altro lato della medaglia – anche il riconoscimento di diritti[37]: se il fine è meramente superare il problema concernente l'imputazione della responsabilità e il garantire al terzo una sicurezza in ordine alla possibilità di soddisfare le proprie pretese risarcitorie, si dovrebbe piuttosto intervenire sul sistema assicurativo, con la consapevolezza che sarà necessario adattare i prodotti assicurativi per renderli idonei ad affrontare tale evoluzione tecnologica.

4. L'intelligenza artificiale utilizzata in sostituzione o a supporto degli avvocati

Allo stato attuale, le nuove tecnologie hanno fatto ingresso anche all'interno del settore legale. Non solamente si prospetta l'utilizzo degli stessi in sostituzione dei Giudici[38], ma anche all'interno degli studi legali, ove da tempo

[36] Come è noto, la qualità di soggetto di diritto non è propria solamente degli esseri umani, ma è riconosciuta anche agli enti collettivi che presentino determinati presupposti, trattandosi di un concetto strumentale al riconoscimento di diritti e obblighi. Sulla soggettività v., tra i tanti: P. Perlingieri, *La personalità umana nell'ordinamento giuridico*, Esi, 1972, *passim*; P. Zatti, *Persona giuridica e soggettività*, Cedam, 1975, *passim*.

[37] V. G. Teubner, *Soggetti giuridici digitali? Sullo* status *privatistico degli agenti software autonomi*, a cura di P. Femia, Esi, 2019, *passim*.

[38] Ai sensi dell'art. 22 Regolamento n. 679/2016, l'interessato ha diritto a non essere sottoposto a una decisione basata unicamente sul trattamento automatizzato, compresa la profilazione, che produca effetti giuridici che lo riguardano o che incida in modo analogo significativamente sulla sua persona. Sul punto è intervenuta anche la Direttiva n. 680/2018 del Parlamento europeo e del Consiglio del 27 aprile 2016 relativa alla protezione delle persone fisiche con riguardo al trattamento dei dati personali da parte delle autorità competenti a fini di prevenzione, indagine, accertamento e perseguimento di reati o esecuzione di sanzioni penali, nonché alla libera circolazione di tali dati, che è stata attuata in Italia con D.Lgs., 18 maggio 2018, n. 51. L'art. 8 "Processo decisionale automatizzato relativo alle persone fisiche" di tale decreto legislativo sancisce che «1. Sono vietate le decisioni basate unicamente su un trattamento automatizzato, compresa la profilazione, che producono effetti negativi nei confronti dell'interessato, salvo che siano autorizzate dal diritto

vengono impiegati strumenti più o meno intelligenti per effettuare attività indispensabili e ricorrenti come il deposito di atti giudiziari, la notifica a mezzo posta elettronica certificata, la sottoscrizione digitale di atti, l'archiviazione di documenti, le ricerche di dottrina e giurisprudenza, la fatturazione, l'organizzazione dello studio, etc.

Ancora, sussistono diversi sistemi automatici per calcolare l'ammontare dell'assegno di mantenimento per il coniuge o per i figli; di particolare pregio risulta, poi, l'impiego di tali meccanismi nell'ambito della *due diligence* e del *risk management*, nonché per generare e aggiornare la *cookies policy*. Si pensi, inoltre, ai *software* di *contract analysis*, che estrapolano automaticamente informazioni dai documenti, e agli strumenti che consentono la redazione automatizzata di contratti.

Sempre più importante è lo sviluppo di *start up* che si occupano di *legal tech*, fornendo servizi alle imprese o agli altri professionisti[39] e che impiegano

dell'Unione europea o da specifiche disposizioni di legge. 2. Le disposizioni di legge devono prevedere garanzie adeguate per i diritti e le libertà dell'interessato. In ogni caso è garantito il diritto di ottenere l'intervento umano da parte del titolare del trattamento. 3. Le decisioni di cui al comma non possono basarsi sulle categorie particolari di dati personali di cui all'articolo 9 del regolamento UE, salvo che siano in vigore misure adeguate a salvaguardia dei diritti, delle libertà e dei legittimi interessi dell'interessato. 4. Fermo il divieto di cui all'articolo 21 della Carta dei diritti fondamentali dell'Unione europea, è vietata la profilazione finalizzata alla discriminazione di persone fisiche sulla base di categorie particolari di dati personali di cui all'articolo 9 del regolamento UE». Per dare uno sguardo alla giurisprudenza amministrativa sorta in ragione del sempre maggiore utilizzo del meccanismo nella pubblica amministrazione, si segnala la sentenza Cons. Stato, Sez. VI, 8 aprile 2019, n. 2270, in banca dati *Onelegale*, secondo la quale «La decisione amministrativa automatizzata, di per sé legittima, impone al giudice di valutare in primo luogo la correttezza del processo informatico in tutte le sue componenti: dalla sua costruzione, all'inserimento dei dati, alla loro validità, alla loro gestione. Da qui, come si è detto, si conferma la necessità di assicurare che quel processo, a livello amministrativo, avvenga in maniera trasparente, attraverso la conoscibilità dei dati immessi e dell'algoritmo medesimo. In secondo luogo, il giudice deve poter sindacare la stessa logicità e ragionevolezza della decisione amministrativa robotizzata, ovvero della "regola" che governa l'algoritmo, di cui si è ampiamente detto». V. anche Cons. Stato, Sez. VI, 4 febbraio 2020, n. 881, in banca dati *Onelegale*, ove si rileva come, ai fini della ammissibilità del ricorso ad algoritmi informatici nel procedimento di formazione della decisione amministrativa, assumono rilievo fondamentale due aspetti preminenti, quali elementi di minima garanzia per ogni ipotesi di utilizzo di algoritmi in sede decisoria pubblica: a) la piena conoscibilità a monte del modulo utilizzato e dei criteri applicati; b) l'imputabilità della decisione all'organo titolare del potere, il quale deve poter svolgere la necessaria verifica di logicità e legittimità della scelta e degli esiti affidati all'algoritmo». Sul punto v. G. Di Rosa, *Quali regole per i sistemi automatizzati "intelligenti"?*, in *Riv. dir. civ.*, 2021, p. 828.

[39] V. anche G. Vaciago, *Opportunità e cambiamenti dei servizi* legal tech *all'interno degli studi legali all'interno del contesto italiano ed europeo*, *Il diritto dell'internet nell'era digitale*, a cura di G. Cassano e S. Previti, Giuffrè, 2020, p. 242.

talvolta e sempre più di frequente un approccio al fenomeno del *legal design*, teso a consentire una maggiore comprensione e fruizione dei contenuti.

Anche in tale ambito – soprattutto oltreoceano – si assiste alla creazione di meccanismi che, lungi dal limitarsi a "supportare" il professionista legale nello svolgimento dell'incarico, tendono a sostituirlo in *toto*, generando spontanee riflessioni sull'individuazione degli argini al dilagare del fenomeno.

Al riguardo, si consideri che, secondo quanto previsto dall'art. 2, comma 3, della legge, 31 dicembre 2012, n. 247 "*Nuova disciplina dell'ordinamento della professione forense*", l'iscrizione a un albo circondariale è condizione per l'esercizio della professione di avvocato[40], con la conseguenza che l'attività che la legge riserva a tali professionisti non può essere svolta da chi non abbia tali requisiti – salvo incorrere nelle conseguenze in precedenza evidenziate – né può essere tantomeno da questi demandata a un sistema di intelligenza artificiale.

Di contro, invece, ben potrebbe un professionista legale scegliere di farsi supportare nell'esecuzione dell'incarico da tali meccanismi: come accennato, non trattandosi di soggetti, non si potrebbe stimare la macchina ausiliario, circostanza che escluderebbe l'ulteriore interrogativo connesso al fatto che l'ausiliario debba o meno possedere le specifiche caratteristiche richieste al professionista.

Quanto all'individuazione dell'attività riservata, la già citata legge sancisce espressamente all'art. 2 comma 5, che «*Sono attività esclusive dell'avvocato, fatti salvi i casi espressamente previsti dalla legge, l'assistenza, la rappresentanza e la difesa nei giudizi davanti a tutti gli organi giurisdizionali e nelle procedure arbitrali rituali*». Ne deriva che tutto ciò che riguarda l'agire o il resistere davanti agli organi giurisdizionali o nelle procedure arbitrali rituali è riservato agli iscritti all'albo.

In particolare, secondo la giurisprudenza l'esercizio abusivo della professione legale non implica necessariamente la spendita al cospetto del giudice della qualità indebitamente assunta, con la conseguenza che il reato si perfeziona per il solo fatto che l'agente curi pratiche legali dei clienti o predisponga atti giudiziari, facendoli poi firmare a un professionista legale[41], anche senza comparire in udienza qualificandosi come avvocato[42]. E ciò, in quanto, laddove fosse ritenuto sufficiente un siffatto banale *escamotage* per consentire ad un soggetto non abilitato di operare in un settore attribuito in via esclusiva a una determinata professione, risulterebbe vanificato il principio della

[40] Tale concetto è affermato anche all'art. 5 del codice deontologico forense.

[41] Così Cass. pen., 16 dicembre 2020, n. 1931, in *Plurisonline.it.*

[42] Così Cass. pen., 6 novembre 2013, n. 646, in *Plurisonline.it.*

generale riserva riferita alla professione in quanto tale, con correlativo tradimento dell'affidamento dei terzi[43].

Alla luce di tale orientamento non solo i bot non potranno di regola rappresentare e assistere in giudizio le parti, ma neppure provvedere alla predisposizione di atti giudiziari. E ciò in quanto verrebbe meno il diritto al corrispettivo[44] e sussisterebbero le conseguenze sotto il versante penalistico, evidentemente non per la macchina, che non ha soggettività, ma per il soggetto che ha fornito tale servizio demandandone l'esecuzione al meccanismo di intelligenza artificiale. Come già evidenziato, ben potrebbe, invece, l'avvocato impiegare tali strumenti – che non avendo soggettività non possono essere stimati ausiliari – per adiuvarlo nello svolgimento dell'incarico e, dunque, nella redazione di atti giudiziari.

Potranno invece essere impiegati tali strumenti per supportare i privati nella redazione di quegli atti relativi a cause nelle quali la parte può stare in giudizio personalmente: in tali ipotesi, magari il soggetto coinvolto nel giudizio, pur essendo lieto di non dover sostenere il costo di un professionista, desidera ricorrere a una macchina che lo aiuti almeno a sottoporre all'autorità un testo predisposto con il rispetto dei crismi formali, che da solo non è in grado di onorare.

Più complesso è però confrontarsi con il successivo art. 2, comma 6, della legge professionale, il quale prevede che «*fuori dei casi in cui ricorrono competenze espressamente individuate relative a specifici settori del diritto e che sono previste dalla legge per gli esercenti altre professioni regolamentate, l'attività professionale di consulenza legale e di assistenza legale stragiudiziale, ove connessa all'attività giurisdizionale, se svolta in modo continuativo, sistematico e organizzato, è di competenza degli avvocati*», precisazione, pure ampliando ulteriormente l'area dell'attività riservata agli avvocati, si palesa di complessa interpretazione e di difficile lettura. Nonostante il testo normativo appaia certamente sibillino, l'attività di consulenza, che pure richiede competenze giuridiche specifiche, non è – di per sé – riservata agli avvocati, ma lo diventa solamente se esercitata secondo determinati crismi, ragione per cui è oggetto di incursione da parte delle società di consulenza e dell'intelligenza artificiale[45].

[43] V. Cass. pen., 7 ottobre 2016, n. 52888, in *Plurisonline.it*. Nei medesimi termini, più di recente, Cass. pen., 16 dicembre 2020, n. 1931, in *Plurisonline.it*.

[44] Trib. Gorizia, 27 luglio 2018, n. 331, in *Dejure.it.*

[45] Nei predetti termini Trib. Bolzano sez. I, 30 giugno 2020, n. 515, in *Dejure*.it; Trib. Cagliari Sez. II, 12 febbraio 2016, in *Plurisonline.it*. Secondo un orientamento consolidato, la prestazione di opere intellettuali, nell'ambito dell'assistenza legale, è riservata agli avvocati iscritti negli albi forensi solo nei limiti della rappresentanza, assistenza e difesa delle parti in giudizio e, comunque, di diretta collaborazione con il giudice nell'ambito del processo. Al di fuori di tali limiti, l'attività

Va, tuttavia, rilevato come, anche laddove la completa sostituzione della macchina al professionista non contrasti con quanto previsto dalla legge, tale soluzione non convinca per ragioni di opportunità, giacché si tratta di attività che – oltre alla conoscenza, alla capacità e velocità nell'elaborare dati delle quali la macchina certamente dispone – postulano talvolta anche una certa capacità creativa, nonché una indubbia ragionevolezza, elementi che paiono estranei ai meccanismi di intelligenza artificiale, i quali – almeno allo stato attuale – agiscono secondo schemi rigidi e scarsamente flessibili.

A tal riguardo, appare evidente come talora la soluzione del caso non si basi meramente su un calcolo matematico, ma postuli la valutazione di altri fattori e il coinvolgimento di altre abilità, quali il saper consigliare, guidare, cogliere le attitudini e le preferenze, giungendo finanche a proporre soluzioni fantasiose. In questo senso, l'intervento umano appare in grado di trovare punti di incontro anche dotati di elasticità, che consentono a tutte le parti di porre fine alla vertenza con un certo senso di soddisfazione.

Tale considerazione appare particolarmente significativa avendo come punto di riferimento le crisi familiari, ove è essenziale rifuggire da soluzioni rigide, applicate senza elasticità, ma promuovere una maggiore flessibilità tra soggetti coinvolti e, nello specifico, tra i genitori, circostanza assai propizia nell'interesse degli eventuali figli minori. In tale senso appare complessa l'opzione di sostituire l'essere umano con la macchina, in quanto il timore è che tale soluzione possa portare a una rigida applicazione del diritto, che non necessariamente conduce a una maggiore soddisfazione per i clienti.

A ciò si aggiunga – elemento che è emerso anche nel dibattito concernente la *cyber* giustizia – che l'applicazione della legge postula il precedente esperimento dell'indagine ermeneutica, con la conseguenza che non è sufficiente possedere una corretta comprensione linguistica per interpretare la legge; ciò non bastasse, numerose sono le clausole generali presenti all'interno del nostro ordinamento, che devono essere riempite di volta in volta di significato, e molteplici sono le situazioni nelle quali il giudice è chiamato a decidere secondo equità; attività nelle quali emergono i limiti della macchina[46].

Ciò non bastasse, a ben vedere è anche maggiore la tutela offerta al soggetto che scelga di avvalersi di un professionista, in quanto questi ha un

di assistenza e consulenza non può considerarsi riservata agli iscritti negli albi professionali, non rientra nella previsione dell'art. 2231 c.c. e dà diritto a compenso a favore di colui che la esercita (così, tra le tante, v. Trib. Padova, Sez. II, 12 gennaio 2015, in *Plurisonline.it*.; Cass. civ. Sez. Un., 3 dicembre 2008, n. 28658, in *Plurisonline.it*). In senso contrario, Cass., pen., 13 gennaio 2017, n. 7630, in *Plurisonline.it*.

[46] Sul punto v. E. Battelli, *Giustizia predittiva, decisione robotica e ruolo del giudice*, in *Giust. civ.*, 2020, p. 281 ss.

obbligo di assicurarsi per la responsabilità civile[47]: in tale ottica, il cliente che ha subito pregiudizi connessi all'inadempimento del contratto avrebbe una tutela rafforzata quanto alla possibilità di vedere soddisfatte le proprie pretese economiche anche in considerazione della delicatezza dell'attività svolta.

Infine, si consideri che chi è iscritto a un determinato albo o elenco è tenuto al rispetto di specifiche regole deontologiche, poste a tutela non solamente del proprio assistito, ma anche della controparte e dei colleghi: a tacer d'altro, l'avvocato deve, esercitare l'attività professionale con indipendenza, lealtà, correttezza, probità, dignità, decoro, diligenza e competenza, tenendo conto del rilievo costituzionale e sociale della difesa, rispettando i principi della corretta e leale concorrenza (art. 9 codice deontologico forense); inoltre, deve mantenere nei confronti dei colleghi un comportamento ispirato alla lealtà e alla correttezza (art. 19 codice deontologico forense).

Diversa è invece, come anticipato, la valutazione circa gli strumenti che si limitano ad adiuvare il professionista, sgravandolo dai compiti più ripetitivi o rendendo più agevole il lavoro di squadra, che consentono allo stesso di concentrarsi sull'attività più stimolante sotto il versante intellettuale o che velocizzano lo svolgimento dell'incarico, poiché in questo caso sussiste sempre la supervisione di chi ha i requisiti per vagliare l'intera attività dell'intelligenza artificiale, la quale comunque – non essendo soggetto – non può essere stimata "ausiliario" del professionista, ma mero strumento di supporto dello stesso.

Sotto tale versante, peraltro, occorre rilevare come la tecnologia, più o meno intelligente, trovi applicazione quotidiana negli studi legali, rivelandosi preziosa per il legale che voglia essere al passo con i tempi ed efficiente, contingenza in grado di comportare sicuri benefici al cliente sotto il versante della riduzione dei costi e dei tempi sottesi allo svolgimento dell'incarico. Anche in questo caso, non si possono tacere le evidenti ripercussioni sul mercato del lavoro se si considera che il titolare dello studio, avvalendosi della stessa, può – in una chiara ottica di contenimento delle spese – non avvertire la necessità di assumere personale per la segreteria o giovani collaboratori, cambiamento con il quale occorre fare i conti[48].

Ciò non bastasse, il fatto che alcuni di tali strumenti siano ormai di uso comune e diffusi negli studi legali, nonché indispensabili per lo svolgimento di taluni compiti, come, a esempio, i programmi per effettuare il deposito telematico di atti o la notifica a mezzo posta elettronica certificata, quelli che consentono di partecipare alle udienze da remoto o di sottoscrivere digital-

[47] V. D.P.R, 7 agosto 2012, n. 137, art. 5.

[48] Così R. Susskind, D. Susskind, *The future of professions*, Oxford University Press, 2015, *passim*.

mente i documenti, incide anche sul contegno in concreto esigibile dal professionista. Non solo l'avvocato desideroso di concorrere con i propri colleghi dovrà necessariamente conoscere quantomeno gli strumenti più elementari e di "uso comune" ed essere in grado di utilizzarli, ma laddove non lo fosse rischierebbe di risultare addirittura inadempiente nei confronti del cliente: pensiamo a chi, non sapendo adoperare un programma per il deposito telematico, ometta di dedurre nei termini perentori previsti i mezzi di prova in una causa nella quale il cliente ricopre il ruolo di attore; non vi è dubbio che tale avvocato sarebbe chiamato a rispondere per inadempimento *ex* art. 1218 c.c.

Il professionista deve, infatti, impiegare una diligenza professionale parametrata al professionista che abbia una preparazione e un'attenzione media e tale soggetto – allo stato attuale – non può ignorare e non saper utilizzare tali strumenti, con ogni conseguenza che ciò comporta in punto responsabilità[49]. In questo senso, l'avvocato per fare fronte a un dovere di competenza e di formazione continua non solo dovrà conoscere la legge, ma dovrà formarsi anche sotto il versante tecnologico, facendo fronte a quello che legittimamente può essere denominato dovere di competenza tecnologica[50]. Di contro, non è certo esigibile il saper ricorrere ai meccanismi di intelligenza artificiale di ultima generazione, che il professionista "medio" è legittimato a non conoscere.

Inoltre, anche l'art. 14 del codice deontologico forense impone un dovere di competenza nel rispetto del quale l'avvocato, al fine di assicurare la qualità delle prestazioni professionali, non deve accettare incarichi che non sia in grado di svolgere con adeguata competenza: ne deriva, a titolo esemplificativo, che il professionista che non è in grado di notificare a mezzo posta elettronica un atto o effettuare un deposito telematico, non dovrebbe neppure assumere l'incarico, salvo incorrere (anche) nelle sanzioni deontologiche previste.

[49] Sul punto, v. le considerazioni riportate all'interno del "CCBE considerations on the legal aspects of artificial intelligence" (https://www.ccbe.eu/fileadmin/speciality_distribution/public/documents/IT_LAW/ITL_Guides_recommendations/EN_ITL_20200220_CCBE-considerations-on-the-Legal-Aspects-of-AI.pdf).

[50] G.C. Hazard e A. Dondi, *Etiche della professione legale*, Il Mulino, 2005, p. 155 ss.; A. Dondi, *Processo civile, new technologies e implicazioni etico-professionali*, in *Riv. trim. dir. proc. civ.*, 2019, p. 874 ss.; nonché Comoglio 2018: 328; Bina 2020: 57 ss. Per uno sguardo negli Stati Uniti, ove l'attenzione al tema è assai più significative, v. tra i tanti: H. Frostestad Kuehl, *Technologically competent: ethical practice for 21st century lawyering*, in *Journal of law, technology & the internet*, 2019, 10, p. 1 ss.; A. Perlman, *The Twenty-First Century Lawyer's Evolving Ethical Duty of Competence*, 22(4), in *The Professional Lawyer (ABA Center for Professional Responsibility)*, 2014, *passim*; B.P. Cooper, *Social Media and the Lawyer's Evolving Duty of Technological Competence*, *Legal Ethics*, 2014, 17:3, 463-466.

5. Conclusioni

La diffusione di sistemi di intelligenza artificiale – chiamati ad adempiere a una prestazione tradizionalmente demandata ai prestatori di opera intellettuale – incontra, dunque, i limiti già da tempo indagati dall'interprete, chiamato a riflettere circa l'estensione dell'attività strettamente riservata ai soggetti in possesso dei requisiti specifici richiesti dalle singole leggi professionali. In questo senso, l'utilizzo di tali meccanismi porta a confrontarsi con problemi che non sono, almeno sotto tale versante, nuovi.

Di là da tale aspetto, certamente il fenomeno oggetto di indagine, ove consentito e legittimo, induce a valutare le innegabili conseguenze che comporta sulla società, sulle modalità di lavoro e sul mercato occupazionale e palesa una "rivoluzione" in un ambito – quello della prestazione di servizi – che a differenza di quello della produzione era ancora prerogativa delle persone.

Come evidenziato, meritano attenzione anche gli strumenti che si limitano a supportare il professionista nello svolgimento dell'incarico, che questi è sempre più chiamato a conoscere in virtù di quello che ben può essere definito un dovere di competenza tecnologica sotto il versante contrattuale nel rapporto con il cliente e sotto il versante deontologico: l'avvocato dovrà, dunque, acquisire una competenza quantomeno con riferimento agli strumenti più elementari (non di ultima generazione), che la tecnologia offre, seguendo il cambiamento, e dovrà essere attento ai rischi della tecnologia, provvedendo a proteggere le informazioni del cliente, con un conseguente e sostanziale mutamento del modo di lavorare e di concepire l'incarico.

In tale ottica, è innegabile che le opportunità e i nuovi problemi sottesi all'evoluzione tecnologica non lasciano indifferente il diritto e portano a ripensare le categorie dello stesso, in un'ottica di tutela dei diritti dei diversi soggetti coinvolti soprattutto alla luce della prevedibile sempre maggiore espansione del ricorso ai citati meccanismi che si diffondono con una velocità e portata senza precedenti.

Più in generale, il breve percorso svolto ha permesso di comprendere come l'attuale panorama costringa l'interprete a un notevole impegno di riflessione per capire di volta in volta quali siano le norme applicabili e attuali nonostante il quadro sia assai mutato rispetto al momento nel quale sono state pensate e introdotte. Va, tuttavia, osservato come – allo stato attuale e perlomeno ragionando delle fatttispecie oggetto di indagine – lo strumento interpretativo delle norme esistenti consenta ancora di fare fronte alla situazione[51],

[51] Del medesimo avviso U. Ruffolo, *Per i fondamenti di un diritto della robotica*, in *Intelligenza artificiale e responsabilità,* cit., p. 3.

anche se un intervento a livello europeo potrebbe – al fine di fornire certezza – individuare *a priori* su chi saranno allocate le conseguenze dell'agire di un sistema di intelligenza artificiale. In questo senso, considerando la portata del fenomeno, che interessa e pone in relazione soggetti che evidentemente si trovano anche oltre i confini nazionali, è fondamentale che – se un intervento vi deve essere – venga effettuato a livello europeo e, dunque, sia uniforme.

Bibliografia

Alcaro F., *Riflessioni critiche intorno alla soggettività giuridica. Significato di un'evoluzione*, Giuffrè, 1976.

Allen T. e Widdison R., *Can computers make contracts?*, in *Harvard journal of law & technology*, 9, I, 1996.

Anastasi A., *Professioni intellettuali – Dir. lavoro*, in *Enc. giur.*, Roma, Ed. Enc. it., XXIV, 1991.

Battelli E., *Giustizia predittiva, decisione robotica e ruolo del giudice*, *Giust. civ.*, 2020, 281-319.

Bianca C.M., *Acontrattualità dei contratti di massa?*, in *Vita not.*, 2001, 1120-1128.

Bina M., *Appunti su deontologia forense, processo civile e nuove tecnologie*, in *L'aula civile*, 2020, 57-60.

Borruso R., *Computer e diritto*, II, Giuffrè, 1988.

Boscarato C., *Who is responsible for Robot's actions? An initial examination of Italian law within a European perspective*, in B. Van Berg e L. Klaming (a cura di), *Technologies on the stand: legal and ethical questions in neuroscience and robotics*, Wolfpublisher, 2011.

Bouhai N. e Saleh I., *Internet of Things: Evolutions and Innovations*, Wiley, 2017.

Bravo F., *Contratto cibernetico*, in *Dir. inform. Informatica*, 2011, 69-211.

Camardi C., *Contratto e rapporto nelle reti telematiche. Un nuovo modello di scambio*, V. Ricciuto e N. Zorzi (a cura di), *Il contratto telematico*, Cedam, 2002.

Caggiano I.A., *Il contratto nel mondo digitale*, in *Nuova giur. civ. comm.*, 2018, II, 1152-1157.

Capilli G., *Responsabilità e robot*, in *Nuova giur. civ. comm.*, 2019, II, 621-631.

Carnelutti F., *Nullità del contratto di patrocinio per difetto del titolo professionale*, in *Riv. dir. proc.*, 1953, I, 313-350.

Castellani G., Smart contracts *e profili di diritto civile*, in *Comp. dir. civ.*, 2019, 1.

Cataudella A., Intuitus personae *e tipo negoziale*, *Studi in onore di F. Santoro Passarelli*, 1972, Esi, 624-646.

Cavaliere G.A. e Iaselli M., *Contratto telematico*, in G. Buffone, I. De Giovanni, A.I. Natale, *Il contratto*, , Cedam, 2013, 1565-1627.

Clarizia R., *I contratti informatici*, in *Trattato dei contratti*, Torino, Giappichelli, 2007.

Comoglio P., *Nuove tecnologie e disponibilità della prova*, Torino, Giappichelli, 2018.

Cooper B.P., *Social Media and the Lawyer's Evolving Duty of Technological Competence*, *Legal Ethics*, 2014, 17:3, 463-466.

Costantini F., *Intelligenza artificiale diritto civile. Verso una "artificial intelligence forensics"?*, in G. Costabile, A. Attanasio e M. Ianulardo (a cura di), *IISFA Memberbook. DIGITAL FORENSICS: Condivisione della conoscenza tra i membri dell'IISFA ITALIAN CHAPTER*, 2017, 17-39.

Costantini F., *Intelligenza artificiale, design tecnologico e futuro del lavoro nell'UE: i presupposti e il contesto*, *Lav. nella giur.*, 2021, 807-814.
Costanza M., *L'Intelligenza Artificiale e gli stilemi della responsabilità civile*, in *Giur. it.*, 2019, 1686-1689.
Cuccuru P., Blockchain *e automazione contrattuale. Riflessioni sugli* smart contracts, in *Nuova giur. civ. comm.*, 2017, II, 107-119.
Di Giovanni F., *Attività contrattuale e Intelligenza Artificiale*, in *Giur. it.*, 2019, 1677-1686.
Di Giovanni F., *Sui contratti delle macchine intelligenti*, in *Intelligenza artificiale. Il diritto, i diritti e l'etica*, Giuffrè, 2020.
Di Rosa G., *Quali regole per i sistemi automatizzati "intelligenti"?*, in *Riv. dir. civ.*, 2021, 823-853.
Di Sabato D., *Gli* smart contracts*: robot che gestiscono il rischio contrattuale*, in *Contr. impr.*, 2017, 378-402.
Dondi A., *Processo civile*, new technologies *e implicazioni etico-professionali*, in *Riv. trim. dir. proc. civ.*, 2019, 863-881.
Dore G., *I doveri di informazione nella rete degli scambi commerciali telematici*, in *Giur. Merito*, 2013,2569-2583.
Finocchiaro G., *I contratti informatici*, in *Tratt. dir. comm. e dir. pub. ec.*, a cura di F. Galgano, Cedam, 1997, 60-78.
Finocchiaro G., *La conclusione del contratto mediante i* "software agents"*: un falso problema giuridico?*, in *Contr. impr.*, 2002, 500-509.
Finocchiaro G., *Il contratto nell'era dell'intelligenza artificiale*, in *Riv. trim. dir. proc. civ.*, 2018, 441-460.
Finocchiaro G., *Intelligenza Artificiale e protezione dei dati personali*, in *Giur. it.*, 2019, p. 1670-1676.
Frostestad Kuehl H., *Technologically competent: ethical practice for 21st century lawyering*, in *Journal of law, technology & the internet*, 2019, 10.
Galgano F., *La categoria del contratto alla soglia del terzo millennio*, in *Contr. impr.*, 2000, 919-929.
Gallo P., *Soggetto di diritto (I agg.) in dig. Civ. agg.*, Utet, 2011.
Giacobbe G., voce *Professioni intellettuali*, in *Enc. dir.*, Giuffrè, 1987, XXXVI.
Goodenough O.R., *Getting to Computational Jurisprudence 3.0*, in A. Santosuosso, O.R. Goodenough, M. Tomasi (a cura di), *The challenge of innovation in law. The Impact of Technology and Science on Legal Studies and Practice*, Pavia University Press, 2015, 3-17.
Greengard S., *Internet delle cose*, Il Mulino, 2017.
Hazard, G.C. e Dondi A., *Etiche della professione legale*, Il Mulino, 2005.
Irti N., *Scambi senza accordo*, in *Riv. trim. dir. proc. civ.*, 1998, 347-364.
Irti N., *«È vero ma...» (Replica a Giorgio Oppo)*, in *Riv. dir. civ.*, I., 1999, 273-278.
Irti N., *Un diritto incalcolabile*, Giappichelli, 2016.
Lega C., *Le libere professioni intellettuali*, Giuffrè, 1974.
Luciani O M., *La decisione giudiziaria robotica*, in *Rivista AIC*, 2018, 872-893.
Mazzariol R., *Attività di psicoanalista e professioni intellettuali «protette»: spunti per una riflessione critica*, in *Nuova giur. civ. comm.*, 2013, II: 423-430.
Monti S., *Il contratto dall'uomo alla macchina e... viceversa?*, in *Contr.*, 2021, p. 459.

Moro, P., *Intelligenza artificiale e professioni legali. La questione del metodo*, in *Journal of Ethics and Legal Technologies*, 2019, I, 24-43.
Musio A., *La storia non finita dell'evoluzione del contratto tra novità tecnologiche e conseguenti esigenze di regolazione*, in *Nuova giur. civ. comm.*, II, 2021, p. 226-237.
Musolino G., *Contratto d'opera professionale. Artt. 2229 – 2238 c.c.*, 2ª ed., Giuffrè, 2016.
Nardi S., *Accordo concluso* online *dal minore di età*, in *Comp. dir. civ.*, 2019, 1.
Nepor S., *Internet e la legge*, Giuffrè, 1999.
Oppo G., *Disumanizzazione del contratto?*, in *Riv. dir. civ.*, I, 1998, 525-546.
Parini G.A., *Riflessioni sul ricorso all'intelligenza artificiale nelle professioni intellettuali*, in *Comp. dir. civ.*, 2019, 1.
Parisi F., *Il contratto concluso mediante computer*, Cedam, 1987.
Parisi F., *Il contratto virtuale. Procedimenti formativi e forme negoziali tra tipicità e atipicità*, Giuffrè, 2005.
Parisi F., *Diritto privato dell'informatica e di internet*, Giuffrè, 2006.
Parola L., Merati P., Gavotti G., Blockchain *e* smart contract*: questioni giuridiche aperte*, in *Contratti*, 2018, 681-688.
Pasquino V., *La vendita attraverso reti telematiche. Profili civilistici*, in *Dir. inform. Informática*, 1990, 697-710.
Pennasilico M., *La conclusione dei contratti on-line tra continuità e innovazione*, in *Dir. inform.*, 2004, 805-834.
Perlingieri P., *La personalità umana nell'ordinamento giuridico*, Esi, 1972.
Perlingieri P., *Metodo, categorie, sistema nel diritto del commercio elettronico*, in *Il diritto dei contratti fra persona e mercato*, Esi, 2003.
Perlman A., The Twenty-First Century Lawyer's Evolving Ethical Duty of Competence, 22(4), in *The Professional Lawyer (ABA Center for Professional Responsibility)*, 2014.
Perulli A., *Il lavoro autonomo*, in *Tratt. Cicu – Messineo*, I, , Giuffrè, 1996.
Rampone F., Smart contrac"*: né* "smart"*, né* "contract", in *Riv. dir. priv.*, 2020, 241-258.
Raskin M., *The law and the legality of smart contracts*, in *Georgetown law Technology Revue*, 2017, 305-334.
Resta G. e Zeno-Zencovich V., *Volontà e consenso nella fruizione dei servizi in rete*, in *Riv. trim. dir. proc. civ.*, 2018, 411-440.
Ricciuto V., Zorzi N., (a cura di), *Il contratto telematico*, in *Tratt. dir. comm. dir. pub. ec.*, Cedam, 2002.
Riva Sanseverino L., *Lavoro autonomo*, in *Del lavoro autonomo*, Bologna, Zanichelli, 1963.
Roppo E., *Behavioural Law and Economics, regolazione del mercato e sistema dei contratti*, in *Riv. dir. priv.*, 2013, 167-186.
Ruffolo U., *Per i fondamenti di un diritto della robotica*, in *Intelligenza artificiale e responsabilità*, Giuffrè, 2018.
Ruffolo U., *Intelligenza Artificiale,* machine learning *e responsabilità da algoritmo*", in *Giur. it.*, 2019, 1689-1704.
Ruscello F., *Istituzioni di diritto civile*, Cacucci, 2017.
Salomone R., *Le libere professioni intellettuali*, Cedam, 2010.
Santoro Passarelli F., voce *Professioni intellettuali*, in *Noviss. dig. it.*, Torino, Utet, XIV, 1968.

Santosuosso A., Boscarato C., Caroleo F., *Robot e diritto: una prima ricostruzione*, in *Nuova giur. civ. comm.*, II, 2012, 492-516.

Santosuosso A., *A general theory of law and technology or a general reconsideration of law?*, in E. Palmerini e E. Stradella (a cura di), *Law and Technology. The Challenge of regulating technological development*, Pisa University Press, 2013.

Santosuosso A., *Technological Innovation in Law: Just an Option or a Strict Necessity?*, in A. Santosuosso, O.R. Goodenough, M. Tomasi (a cura di), *The challenge of innovation in law. The Impact of Technology and Science on Legal Studies and Practice*, Pavia University Press, 2015, 19-34.

Santosuosso A., *Diritto scienza e nuove tecnologie*, Cedam, 2016 (2° ed.).

Sartor G., *Gli agenti software. Nuovi soggetti di ciberdiritto*, in *Contr. impr.*, 2002, 465-499.

Sartor G., *L'informatica giuridica e le tecnologie dell'informatica. Corso d'informatica giuridica*, Giappichelli, 2016.

Sartor G. e Omicini A., *The autonomy of technological systems and responsibilities for their use*, in *Autonomous Weapon Systems. Law, Ethics, Policy*, Cambridge University Press, 2016.

Susskind R. *L'avvocato di domani*, Edizioni Guerini Next, 2019.

Susskind R., Susskind D., *The future of professions*, Oxford University Press, 2015.

Szabo N., *Formalizing and Securing Relationships on Public Networks*, in *First Monday*, 2, 9, 2015.

Teubner G., *Soggetti giuridici digitali? Sullo* status *privatistico degli agenti software autonomi*, a cura di P. Femia, Esi, 2019.

Tosi E., *I contratti di informatica. Tipi contrattuali, formazione e responsabilità*, Giuffrè, 1993.

Tosi E., *Il contratto virtuale. Procedimenti formativi e forme negoziali tra tipicità e atipicità*, Giuffrè, 2005.

Tosi E., *Diritto privato dell'informatica e di internet*, Milano, Giuffrè, 2006.

Turing A., *Computing Machinery and Intelligence*, in *Mind: new series*, 239, 1950, 443–450.

Vaciago G., *Opportunità e cambiamenti dei servizi* legal tech *all'interno degli studi legali all'interno del contesto italiano ed europeo*, *Il diritto dell'internet nell'era digitale*, a cura di G. Cassano e S. Previti, Giuffrè, 2020.

Zatti P., *Persona giuridica e soggettività*, Cedam, 1975.

Decisão Robótica no Direito Italiano

Robotic Decision in Italian Law

Vitulia Ivone*

resumo: Nos últimos anos, a eficácia das normas jurídicas foi muitas vezes condicionada pelo rápido desenvolvimento tecnológico que pôs em causa a sua idoneidade e alcance. Em particular, as inovações decorrentes dos avanços da inteligência artificial exacerbaram a já complicada relação entre direito e tecnologia. É evidente a necessidade de compreender até que ponto as regras existentes podem ser consideradas flexíveis na interpretação e podem responder às necessidades de uma sociedade em evolução.

palavras-chave: inteligência – direito – responsabilidade – robótica – regras – algoritmo

abstract: In recent years, the effectiveness of legal rules has often been conditioned by the rapid technological development that has called into question their suitability and scope. In particular, innovations stemming from advances in artificial intelligence have exacerbated the already complicated relationship between law and technology. There is a clear need to understand the extent to which existing rules can be considered flexible in interpretation and can respond to the needs of an evolving society.

keywords: intelligence – law – responsibility – robotics – rules – algorithm

* Departamento de Ciências jurídicas, Faculdade de Direito Universidade de Salerno, Itália.

SUMÁRIO: 1. Premissa. 2. O tema de linguagens. 3. O sistema ITALGIURE e a justiça preditiva. 4. Regras europeias sobre Inteligência artificial. 5. O caso Yahoo! 6. Os riscos das decisões robóticas. 7. Projetos italianos na área de tecnologia e justiça. 8. Algumas conclusões.

1. Premissa

O aprofundamento do estudo dos sistemas modernos de inteligência artificial e suas implicações do ponto de vista jurídico não pode ocorrer sem uma profunda conscientização das consequências decorrentes da difusão de novas tecnologias em grande escala que oferecem oportunidades sem precedentes para cidadãos, empresas e o público.

Ao mesmo tempo, se tais sistemas não forem adequadamente governados, podem dar origem a riscos para os direitos e liberdades fundamentais[1].

A complexidade do tema é por demais evidente e suscita, de certa forma, novidades preocupantes, cujas reflexões extrapolam a dimensão do fenômeno teórico para o qual o jurista volta sua curiosidade e sua atenção hermenêutica, para projetar e afetar uma mesma ética sobre o comportamento humano e sobre as avaliações e decisões de política jurídica que serão tomadas no futuro.

As novas questões que surgem com o tema da inteligência artificial preocupam e ocupam a atenção dos estudiosos do direito civil, testando tanto as capacidades de resposta do sistema quanto as categorias conceituais tradicionais e atuais, para adaptá-lo à realidade em rápida mudança.

A ideia de que uma máquina, por mais "inteligente" que seja, possa – através da consumação fria e gelada de um algoritmo, mediante uma "decisão robótica" – estabelecer o destino de uma pessoa, esteja ele relacionado apenas ao tamanho de seus bens, ou até mesmo chegando ao ponto de regular o campo do afeto e de relações familiares, do tratamento médico obrigatório e até àqueles sobre as decisões sobre a continuação ou o fim da vida, sem dúvida, suscita preocupação e consternação.

Todas essas questões inevitavelmente afetam o sistema e colocam um questionamento não secundário: o atual conjunto de regras é suficiente para governar um fenômeno tão revolucionário? Se as regras fossem insuficientes, o princípio da segurança jurídica poderia vacilar?

Para tentar dar algumas respostas, é preciso refletir sobre alguns temas.

[1] Erica Palmerini, *Robotica e diritto: suggestioni, intersezioni, sviluppi a margine di una ricerca europea*, in *Resp. civ.*, 2016, p. 1816 ss.

A mudança de época que deriva da utilidade da inteligência artificial trouxe bem-estar, desenvolvimento e novas perspetivas.

No entanto, ao longo deste caminho, surgiu a necessidade de conhecer e regular os vários fenómenos, como o algoritmo e a gestão de dados.

As primeiras aplicações de inteligência artificial no campo do direito levaram à criação do chamado sistemas baseados em conhecimento.

Em outras palavras, o sistema parte de uma "base de conhecimento" e deduz suas consequências, isso é possível inserindo asserções ou declarações expressas em linguagem de computador e operando sobre ela por meio de um motor inferencial.

As inferências permitem "justificar" o resultado, podendo assim proceder-se a uma reconstrução do processo dedutivo seguido, contudo a natureza axiomática da base de conhecimento representa um limite contra o qual o motor inferencial não consegue fazer avaliações.

2. O tema de linguagens

Se pode afirmar que existe uma diferença de linguagens entre as estruturas da mente humana, inteligência artificial e direito[2].

Cada um desses "sujeitos" tem seus próprios meios de comunicação e seus próprios símbolos[3]. E eles não são facilmente trocáveis.

Do ponto de vista prático, são muitas as aplicações da inteligência artificial ao direito. Essas aplicações podem ser classificadas em quatro tipos de sistemas jurídicos baseados em inteligência artificial: 1) sistemas de auxílio à decisão; 2) sistemas baseados em casos; 3) sistemas de recuperação inteligente de informações jurídicas; 4) sistemas de auxílio à redação e produção automática de documentos.

O primeiro: sistemas de auxílio à decisão.

Pretendem reproduzir o raciocínio do jurista chamado a identificar o caso abstrato da norma que deve ser aplicada ao caso concreto e sugerir, com base nas normas identificadas, uma solução para o caso: esse mecanismo de raciocínio não está em realidade unidirecional, mas pauta-se a partir de uma avaliação prudencial que une os dois casos e tenta estabelecer uma relação de ajuste de um ao outro, que leve em consideração as circunstâncias do caso, bem como a experiência do próprio jurista na apreciação dos elementos de direito e de facto relevantes para a resolução do litígio.

[2] Stefano Crisci, *Intelligenza artificiale ed etica dell'algoritmo*, in *F. Amm.*, 2018, p. 1787.

[3] Amedeo Santosuosso – Chiara Boscarato – Franco Caroleo, *Robot e diritto: una prima ricognizione*, in *Nuova g. civ. comm.*, 2012, II, p. 494 ss. e spec. p. 497 ss.; Carolina Perlingieri, *L'incidenza dell'utilizzazione della tecnologia robotica nei rapporti civilistici*, in *Rass. d. civ.*, 2015, p. 1236, nt. 1.

Um exemplo indicativo é oferecido pela reconstrução de áreas consideráveis dos casos da Suprema Corte americana, que assume características topológicas (em particular em termos de hubs semânticos).

O segundo sistema é baseado em casos.

O sistema inteligente baseado em casos serve de suporte em particular à actividade do advogado, pois analisa um número muito elevado de precedentes e permite identificar, nos sistemas de Civil Law, a regra geral e abstracta a aplicar ao novo caso e, em sistemas de Common Law, a decisão a ser tomada para resolver a nova disputa.

O terceiro sistema trata de recuperação inteligente de informações jurídicas.

Hoje essas atividades são certamente mais acessíveis a todos: a referência é a aqueles sistemas inteligentes que auxiliam o jurista na busca de documentos jurídicos, ou seja, documentos que reúnem as fontes do direito formuladas textualmente (leis, sentenças, boatos, máximas etc.) e indexados na base de conhecimento do programa de computador.

3. O sistema ITALGIURE e a justiça preditiva

A inteligência artificial no campo do direito atua desde a busca indexada, passando pela resolução de processos judiciais, até a justiça preditiva e é composta por três principais macro áreas de interesse:

- sistemas de análise jurídica, para a subsunção de uma causa jurídica dentro de uma causa jurídica específica;
- sistemas de planejamento jurídico, que sugere as melhores ações para atingir um determinado resultado;
- sistemas de busca de informação jurídica, que permitem buscar informações de conteúdo conceitual, e não meramente semântica jurídica.

É precisamente esta última categoria que implementa os mais recentes e sofisticados sistemas de aprendizado de máquina e inteligência artificial, enquanto os dois primeiros sistemas ainda se referem aos modelos de sistemas especialistas baseados em conhecimento.

Um exemplo particularmente interessante deste modelo é dado por: thesaurus de ITALGIURE (o sistema de documentação automática do CED do Supremo Tribunal de Cassação), que permite a busca, com base nos termos descritos, indicados pelo usuário no aplicativo endereçado ao programa, de documentos relativos a cada setor e fonte do direito[4].

[4] Claudia Morelli, *Giustizia predittiva: in Francia online la prima piattaforma europea. Uno strumento per garantire la certezza del diritto?*, www.altalex.com, 2017.

Os sistemas de redação assistida nem sempre podem ser definidos como sistemas inteligentes, pois, por vezes, se limitam a agregar apenas as informações necessárias para completar a redação de um documento (componentes de um texto legal, um contrato, uma sentença, etc.) e informações do próprio produtor do documento (repartição legislativa, contratante, juiz) que assim assegura a sua autenticidade[5].

Estes tipos de aplicações, inscritas na vida comum dos envolvidos no direito, indicam que a inteligência artificial nunca se destina a um mero processamento de dados ou execução mecânica de tarefas, mas sim a uma aplicação de novas tecnologias computacionais a fim de obter desempenhos complexos comparáveis àqueles de que as mentes humanas são atualmente capazes.

Juntamente com os sistemas de bancos de dados muito úteis, há o tema da justiça preditiva.

O crescente número de processos tem sufocado o imediatismo da sentença, violando efetivamente o princípio fundamental do devido processo legal, e aumentando os custos para o Estado.

Neste clima difícil, os juristas têm buscado, na inteligência artificial, uma solução concreta capaz de agilizar e otimizar os tempos da justiça.

Se a inteligência artificial é um conjunto de metodologias, técnicas e algoritmos para a análise e previsão automática de fenômenos complexos, implicitamente descritos por bancos de dados históricos que representam experiências "anteriores", compartilhadas e amadurecidas ao longo do tempo[6], o termo justiça preditiva refere-se ao uso de algoritmos e técnicas de inteligência artificial para fazer "previsões" na área jurídica, pretende-se fornecer, de forma simples e sem tecnicismos desnecessários, com particular referência às disciplinas jurídicas, as noções básicas e as potencialidades de aplicação da inteligência artificial[7].

Atualmente, os problemas que dificultam a substituição do juiz pelo robô dizem respeito à compatibilidade quanto aos princípios da imparcialidade, imparcialidade e discricionariedade do juiz, quanto às vastas variáveis dos casos, quanto à ausência de o caso em si.

[5] Edoardo Rulli, *Giustizia predittiva, intelligenza artificiale e modelli probabilistici. Chi ha paura degli algoritmi?*, Bologna, Il Mulino, 2018, p. 532.

[6] Enrico Gabrielli e Ugo Ruffolo (a cura di), *Intelligenza Artificiale e diritto*, in *Giur. It.*, 2019, fasc. 7 (luglio 2019), p.1657 e s.

[7] Ugo Ruffolo, *Per i fondamenti di un diritto della robotica*, in Id. (a cura di), *Intelligenza artificiale e responsabilità*, Milano, 2017, p. 15.

O ponto aleatório é o tema da responsabilidade, pois é sempre necessário poder reassumir a responsabilidade do juiz[8].

O juiz deve necessariamente dar resposta aos pedidos das partes. Esse princípio imutável do direito processual move a análise para o obstáculo de aplicação da justiça preditiva.

O debate sobre o tema da justiça preditiva segue em aberto: ao lado dos que demonizam suas nefastas consequências, há os que destacam seus aspetos positivos para uma justiça mais lenta.

Parece claro que é preciso estabelecer regras éticas que limitem o dano e valorizem seus méritos.

4. Regras europeias sobre Inteligência artificial

A União Europeia iniciou o seu debate sobre Inteligência artificial em 2017 de forma inorgânica, com a resolução do Parlamento Europeu sobre robótica, que previa o nascimento de robôs inteligentes autónomos e evocava a necessidade de atribuir direitos e deveres a estas novas entidades legais. A mesma resolução também exortou a Comissão Europeia a considerar a criação de uma agência de Inteligência artificial e a estabelecer um quadro de políticas globais para mitigar os riscos desta poderosa tecnologia.

Devido à sua atenção quase exclusiva aos riscos da Inteligência artificial, a posição do Parlamento Europeu, embora provoque uma reação muito crítica por parte da comunidade científica, pelo menos colocou a Inteligência artificial no "radar" da política europeia.

A estratégia de Inteligência Artificial da União Europeia atingiu um ponto de viragem em dezembro de 2019 com a chegada da nova Comissão Europeia que, a 19 de fevereiro de 2020, lançou um pacote de regras, contendo as suas ideias e ações para a transformação digital, incluindo um Livro Branco sobre Inteligência Artificial e uma Estratégia Europeia de Dados. O pacote, ao mesmo tempo muito assertivo e abrangente, marca mais um passo na busca europeia por uma Inteligência Artificial "centrada no ser humano". Baseia-se numa visão específica do futuro dos dados e da Inteligência artificial[9].

O Livro Branco tem o duplo objetivo de criar um "ecossistema de excelência" único e um "ecossistema de confiança", baseado principalmente em

[8] Giorgio Comandé, *Intelligenza artificiale e responsabilità tra liability e accountability. Il carattere trasformativo dell'IA e il problema della responsabilità, in Analisi giuridica dell'economia*, 2019, I, p.173; Maria Costanza, *L'Intelligenza Artificiale e gli stilemi della responsabilità civile*, in *Dottrina e attualità giuridiche – Intelligenza Artificiale e diritto*, Giurisprudenza Italiana 2019, p. 1687 ss.

[9] Giuseppe Proietti, *Il Libro Bianco sull'intelligenza artificiale. L'approccio europeo tra diritto ed etica*, in *giustiziacivile.com*, 2020.

uma abordagem "centrada no ser humano". E, por isso, apela à adoção de um quadro regulamentar flexível e ágil, limitado a aplicações de "alto risco", em sectores como a saúde, os transportes, a polícia e a justiça, centrando-se nas disposições relativas à qualidade e rastreabilidade dos dados, transparência e controle humano.

O Livro Branco expressa a necessidade de apoio estratégico para criar um bom grau de confiança do usuário nas tecnologias.

O requisito referido no Livro Branco é baseado no cumprimento de três princípios fundamentais:

o princípio da cognoscibilidade, o princípio da não exclusividade e o princípio da não discriminação.

O primeiro pode ser traduzido em um direito absoluto de acesso, não se limitando ao único conhecimento do tratamento. Na verdade, o interessado deve ser capaz de compreender a lógica da decisão automatizada para poder decifrar seu uso.

A não exclusividade algorítmica lembra a necessidade de que uma decisão sobre uma parte interessada não possa resultar apenas da decisão automatizada.

O tecido regulatório exige que no processo de tomada de decisão haja sempre uma contribuição humana capaz de controlar e validar a decisão algorítmica.

Finalmente, o princípio da não discriminação casa a visão antropocêntrica italiana e europeia sobre os sistemas de inteligência artificial.

Alguns sistemas de IA, como o aprendizado de máquina, usam esquemas de aprendizado que produzem soluções baseadas em recorrências estatísticas e que são mais adequadas a um conceito de probabilidade do que de certeza.

A Comissão europeia apoia fortemente uma abordagem antropocêntrica, com base no cumprimento dos direitos fundamentais, incluindo os direitos à liberdade de expressão e reunião, a dignidade humana, à não discriminação com base no sexo, na raça, na origem étnica, na religião ou crença, na idade ou na orientação sexual (quando aplicável em certas áreas), à proteção dos dados pessoais e da privacidade ou ao direito a um recurso judicial efetivo e a um processo justo, bem como à proteção do consumidor.

Esses riscos podem surgir de defeitos no projeto geral de sistemas de IA ou com o uso de dados sem que sejam corrigidas quaisquer distorções (por exemplo, se um sistema é elaborado usando apenas ou principalmente dados masculinos, levará a resultados abaixo do ideal para as mulheres).

No entanto, à semelhança de outras iniciativas louváveis da União Europeia, mesmo este Livro Branco, se não for apoiado pelas legislações nacionais competentes, corre o risco de permanecer uma mera declaração de princípios.

5. O caso Yahoo!

Interessante aqui tecer algumas breves considerações sobre um caso jurisprudencial italiano muito recente.

Em 8 de fevereiro de 2022, o Tribunal Supremo de Cassação italiano, ao julgar recurso interposto pelo Yahoo!, estabeleceu que o cancelamento da cópia de cache relativa a informações acessíveis por meio de mecanismo de busca exige um equilíbrio entre o direito ao esquecimento do interessado e o direito de divulgar e adquirir informações relativas aos fatos como um todo por meio de palavras-chave também diferentes do nome[10].

O caso é interessante.

Em 2015, uma pessoa transmitiu ao motor de busca Yahoo! um pedido para remover dos resultados de pesquisa na Europa (UERIEL) URLs diferentes e específicas que vinculavam seu nome a uma questão legal que ele considerava não mais relevante para o direito de imprensa.

O motor de busca, por sua vez, tinha decidido não poder responder a este pedido, por entender que não poderia ser qualificado como titular deste tratamento de dados pessoais. Assim, o interessado interpôs recurso junto da Autoridade para a Proteção de Dados Pessoais, com pedidos de remoção dos UERIELES) URLs, bem como a eliminação de cópias em cache de páginas web acessíveis através dos referidos (UERLIES) URLs.

Com disposição de 25 de fevereiro de 2016, a Autoridade para a Proteção de Dados Pessoais[11] acatou parcialmente esses pedidos do interessado, orde-

[10] Cass. Civ., 8 de fevereiro de 2022, n. 3952.

[11] O Garante para a Proteção de Dados Pessoais, também conhecido como Garantidor da Privacidade, é uma autoridade administrativa italiana independente estabelecida pela lei de 31 de dezembro de 1996, n. 675, para garantir a proteção dos direitos e liberdades fundamentais e o respeito à dignidade no tratamento de dados pessoais.
No artigo 26, parágrafo 1º do decreto legislativo n. 196 de 30 de junho de 2003 é claro que "Os dados sensíveis só podem ser tratados com o consentimento por escrito do interessado e com a autorização prévia do Garante, em conformidade com as condições e limites estabelecidos por este código, bem como pelo lei e pelos regulamentos".
No nº 4, novamente do artigo 26, explica-se quando os dados sensíveis podem ser tratados sem consentimento, mas com autorização prévia do garante, se forem recolhidos: "por associações, órgãos ou organizações sem fins lucrativos, ainda que não reconhecidas, de natureza política, filosófica, religiosa ou sindical", "para a proteção da vida ou integridade física de terceiros", para a "realização de investigações defensivas","para cumprir obrigações ou tarefas específicas estabelecidas por lei, por regulamento ou por legislação comunitária para a gestão da relação de trabalho".
O artigo 37 define os casos em que o responsável pelo tratamento deve notificar o Garante, se o tratamento incidir sobre: "dados genéticos, biométricos ou que indiquem a localização geográfica de pessoas ou objectos através de uma rede de comunicações electrónicas", "dados adequados à revelação do estado de saúde e da vida sexual, tratados para efeitos de procriação assistida, prestação

nando ao Yahoo! remover tais URLs e excluir tais cópias em cache, o Tribunal Supremo decidiu em favor da pessoa que havia interposto recurso.

O Yahoo! solicitou ao Tribunal de Milão que cancelasse esta disposição da Autoridade. O juiz de primeira instância, com sentença n. 12.623/2016, negou provimento ao recurso, considerando, em primeiro lugar, que a Autoridade não tinha competência para proferir o dispositivo impugnado. A este respeito, o Tribunal considerou que, nos termos do Regulamento da UE 1215/2012 (designado "Bruxelas I bis"), era necessário considerar o lugar onde ocorreu o facto danoso, pelo que era competente, à escolha do autor, tanto o juiz do lugar do fato gerador do dano, como o juiz do lugar onde ocorreu o próprio fato.

O Tribunal, referindo-se à jurisprudência do TEDH, considerou que os direitos fundamentais do interessado prevalecem sobre o interesse económico do motor de busca, bem como sobre o interesse público em obter informações depois de uma pesquisa relativa ao nome do interessado.

A parte interessada havia solicitado à Autoridade a "desindexação", ou seja, um pedido que consiste em impedir o nome de uma pessoa de aparecer nos resultados de um motor de busca depois de feita uma consulta.

Com a desindexação elimina-se um determinado método de busca de dados, que permanece presente na rede, e que continua acessível, mas com uma busca mais complexa e mais longa.

de serviços de saúde por telemática relativos a bases de dados ou fornecimento de bens, inquéritos epidemiológicos, detecção de doenças mentais, infecciosas e difusas, seropositividade, detecção de órgãos e transplante de tecidos e monitoramento dos gastos com saúde", "dados próprios para revelar a vida sexual ou a esfera psíquica tratados por associações, organismos ou organizações sem fins lucrativos, ainda que não reconhecidos, de natureza política, filosófica, religiosa ou sindical", "dados tratados com o auxílio de ferramentas eletrónicas destinadas à definição do perfil ou personalidade do interessado, ou à análise de hábitos ou escolhas de consumo, ou ao acompanhamento da utilização de serviços de comunicações eletrónicas com exclusão do tratamento tecnicamente indispensável à prestação do serviço mesmos serviços aos usuários" (o chamado perfil do usuário), "dados sensíveis registrados em bancos de dados para fins de seleção de pessoal em nome de terceiros, bem como dados sensíveis utilizados para pesquisas de opinião, pesquisas de mercado e outras pesquisas amostrais", "dados registados em bases de dados especiais geridos com ferramentas electrónicas e relativos ao risco de solvência económica, à situação financeira, ao correcto cumprimento de obrigações, a condutas ilícitas ou fraudulentas"

Pelo artigo 39 existem as obrigações de comunicação ao garante relativamente à transferência de dados pessoais entre entidades públicas e ao tratamento de dados pessoais relativos ao estado de saúde.

O artigo 154, n. 5 define o tempo útil de resposta do Garante aos pedidos de autorizações, "sem prejuízo dos prazos mais curtos estabelecidos na lei, o parecer do Garante é dado nos casos previstos no prazo de quarenta e cinco dias a contar da receção do pedido . Findo o prazo, a administração pode proceder independentemente da aquisição do parecer".

Ou seja, trata-se de uma operação de tratamento de dados pessoais, na aceção do artigo 2º, b) da Diretiva 95/46/CE, bem como confirmado por jurisprudência constante do Tribunal de Justiça da União Europeia.

De acordo com a jurisprudência do Tribunal de Cassação italiano, a desindexação é atribuível ao direito de excluir dados, no âmbito da classificação que o considera-como uma das três possíveis declinações do direito ao esquecimento.

O direito ao esquecimento, segundo a Corte, poderia ser declinado respetivamente:

- como desindexação[12];
- como o direito de não voltar a ver notícias publicadas relativas a acontecimentos legitimamente divulgados no passado, se tiver decorrido um período de tempo adequado entre a primeira e a segunda publicação;
- como uma necessidade de colocar a publicação, que legitimamente ocorreu há muito tempo, no contexto atual.

O que se constata é que é necessário proteger tanto o exercício do direito à liberdade de informação dos internautas – liberdade protegida pelo artigo 11º da Carta dos Direitos Fundamentais da União Europeia – como os direitos do interessado, tal como definidos pelos artigos 7º e 8º da Carta de Nice segundo os quais estes prevalecem sobre os interesses económicos do motor de busca, bem como sobre a liberdade de informação dos utilizadores.

No entanto, esse equilíbrio pode mudar significativamente em circunstâncias particulares, dependendo da natureza das informações, da natureza sensível da vida privada do titular dos dados, do interesse do público em obter essas informações, bem como do papel desempenhado pela pessoa na esfera pública.

[12] O artigo 17º do Regulamento europeu sobre a proteção de dados pessoais regula o direito ao cancelamento dos dados pessoais: "O interessado tem o direito de obter do responsável pelo tratamento o cancelamento dos dados pessoais que lhe digam respeito sem demora injustificada e o o responsável pelo tratamento tem a obrigação de cancelar os dados pessoais sem demora injustificada, se os dados pessoais já não forem necessários para as finalidades para as quais foram recolhidos ou tratados de outra forma" (ponto 1).
Os outros casos em que o interessado tem direito a solicitar o cancelamento dos seus dados pessoais são: revogação do consentimento em que se baseia o tratamento; oposição ao tratamento; dados processados ilegalmente; cancelamento necessário para cumprir uma obrigação legal; obrigação legal ao abrigo da legislação da União Europeia ou do Estado-membro; dados recolhidos relativos à oferta de serviços da sociedade da informação.

6. Os riscos das decisões robóticas

No que diz respeito à abordagem jurídica da tecnologia robótica, os termos da relação relativa podem traduzir-se e resumir-se em ordens específicas de problemas especializados e nas correspondentes técnicas regulatórias; neste sentido, face à ausência de uma disciplina de robótica propriamente dita, é frequente a dúvida se as normas legais vigentes são suficientes para regulamentar a robótica na esfera cível ou, pelo contrário, se é necessário criar as regras[13].

Estas questões, de facto, traduzem tipos de problemas nada estranhos à experiência jurídica e que, por isso, podem em todo o caso ser úteis na configuração do discurso jurídico, partindo antes de mais nada pela necessidade sentida (mesmo nível supranacional) de que lhes sejam garantidos alguns valores éticos como o primado da autonomia humana, o princípio da precaução, o princípio da transparência ou mesmo a justificabilidade (explicabilidade), e o princípio da justiça[14].

Neste sentido, a dimensão científica e ética deve conjugar-se adequadamente com a dimensão jurídica, naquela perspectiva de governança global (nova e global) do progresso tecnocientífico. Um traço evidente disso pode ser encontrado, entre outros, no último documento (em ordem cronológica) da Comissão Europeia, que publicou, em 21 de abril d e 2021, a Proposta de regulamento sobre a abordagem europeia da inteligência artificial, expressão do primeiro quadro da União Europeia neste domínio específico.

Neste documento, os riscos da inteligência artificial são avaliados criteriosamente, com o objetivo de salvaguardar os valores e direitos fundamentais da União Europeia e a segurança dos utilizadores.

Nessa direção entendemos pela previsão de uma série de práticas proibidas de inteligência artificial por serem capazes (em termos absolutamente gerais) de influenciar o comportamento individual ou de grupos específicos de pessoas com correlatos preconceitos que permitem a ocorrência de potenciais danos físicos ou psicológicos; ou no que diz respeito à utilização de sistemas de identificação biométrica remota "em tempo real" em espaços acessíveis ao público para efeitos de aplicação da lei, sem que nenhuma das razões especificamente previstas seja utilizada (nomeadamente para combater a criminalidade), visando assim evitar o perigo extremamente grave do uso de técnicas de vigilância em massa[15].

[13] Giusella Finocchiaro, *Intelligenza artificiale e responsabilità*, in *Contratto e impr.*, 2020, p. 724.

[14] Alessandro Longo – Guido Scorza, *Intelligenza artificiale. L'impatto sulle nostre vite, diritti e libertà*, Milano, 2020.

[15] Salvatore Amato, *Biodiritto 4.0. Intelligenza artificiale e nuove tecnologie*, Torino 2020, p. 100 s.; Arianna Fusaro, *Quale modello di responsabilità per la robotica avanzata? Riflessioni a margine del percorso europeo*, in *Nuova g. civ. comm.*, 2020, p. 1344.

Mas quais são os riscos de uma decisão exclusivamente robótica?

1) o resultado fornecido pelos algoritmos preditivos é necessariamente influenciado pela qualidade dos dados que são colocados como entrada; portanto, é necessário cuidar de sua qualidade, independência da fonte e acessibilidade;
2) o algoritmo pode ter um resultado discriminatório com base em dados pessoais sensíveis, incluindo raça e origem social;
3) a estrutura do algoritmo não é neutra e deve ser verificável; o que é difícil se o algoritmo estiver protegido por direitos de propriedade intelectual; também pode haver erros de projeto; – o algoritmo, quando utilizado como mero suporte à decisão do juiz, requer treinamento de pessoal no âmbito do judiciário;
4) o algoritmo preditivo – a partir de uma elaboração de jurisprudência e casos anteriores – pode indicar não o "resultado" exato de uma determinada disputa, mas seu possível resultado, com vários riscos: o algoritmo não é capaz de "reconhecer" que aquele caso submetido a ele não é um caso semelhante; existem algumas singularidades que um decisor humano talvez notasse e que o levariam a fazer uma distinção;
5) o algoritmo pode favorecer o efeito "manada": correr-se-ia o risco de induzir o juiz menos atento a se conformar com a proposta do algoritmo sem assumir autêntica responsabilidade pela sentença que emite, com o efeito de cristalizar a jurisprudência, que seria menos sensível às mudanças sociais[16].

7. Projetos italianos na área de tecnologia e justiça

Na Itália, nos últimos anos, surgiram vários projetos para a introdução da inteligência artificial no mundo da justiça, com base no princípio segundo o qual uma mecanização dos procedimentos pode fortalecer a garantia do princípio da segurança jurídica, além, é claro, benefícios em termos de economia processual.

A Scuola Superiore Sant'Anna de Pisa criou o laboratório Lider, a primeira plataforma de justiça preditiva para o desenvolvimento de uma metodologia de análise de material jurisprudencial combinando técnicas de aprendizado de máquina e análise de big data. O objetivo é oferecer um melhor suporte e um atendimento mais eficiente aos profissionais, cidadãos e juízes. Nas intenções, profissionais e cidadãos terão maior acesso à jurisprudência, podendo

[16] Guido D'Ippolito, *Profilazione e pubblicità targettizzata* on line. *Real-Time Bidding e behavioural advertising*, Napoli, 2021.

também avaliar autonomamente os possíveis desfechos de um caso concreto, assim como magistrados terão a possibilidade de ter um conhecimento mais aprofundado das sentenças existentes para julgar casos delicados questões como as relativas à pensão de alimentos em caso de separação e divórcio ou à liquidação de danos de saúde. Dessa forma, eles poderiam tomar decisões mais informadas e consistentes com os princípios de justiça e igualdade. Seria também uma plataforma de alto nível, capaz de gerenciar grandes fluxos de dados e suportar sofisticados modelos de inteligência artificial para processamento e compreensão da linguagem jurídica[17].

Outro projeto ambicioso e representativo do desenvolvimento da justiça preditiva na Itália é o do Tribunal de Apelação de Brescia, que começou em 2018 com um acordo de um ano estipulado entre o próprio Tribunal, o Tribunal e a universidade da cidade. O projeto – cujos resultados podem ser consultados no site do Tribunal da Relação – visa fornecer aos utilizadores e advogados dois dados fundamentais para a segurança jurídica e para as relações laborais e sociais: a duração previsível de um processo sobre uma matéria específica, e as orientações jurisprudenciais da Corte e da Corte de Apelações de Brescia. Isso é possível pelo trabalho de estudo e tratamento que é realizado apenas em áreas jurídicas específicas, visando assim uma ampliação progressiva dos dados a serem submetidos à análise dos algoritmos.

Por fim, o Centro de Documentação Eletrônica (C.E.D.) do Tribunal de Cassação e a Escola da Universidade IUSS de Pavia assinaram – em 29 de setembro de 2021 – um novo "Acordo Estrutural" sobre inteligência artificial e justiça preditiva. O acordo tem uma duração inicial de 5 anos, renovável por mais 5 anos através de procedimento específico.

O objetivo é ativar uma colaboração estratégica para pesquisa avançada no setor de ferramentas técnicas para a coleta e organização de material jurídico digital. Em particular, as partes pretendem aumentar a riqueza de conhecimento composta pela jurisprudência e legislação italiana e europeia por meio do uso de análises jurídicas e ferramentas de inteligência artificial[18].

O artigo 3 do acordo prevê três diferentes objetivos de longo prazo: a previsão do resultado dos processos decisórios judiciais, administrativos e políticos, a extração de argumentos jurídicos do corpus de sentenças e decisões Italgiure, a criação automática de máximas por meio de documento automático.

[17] Massimo Luciani, *La decisione giudiziaria robotica*, in *Nuovo Diritto Civile*, III, 2018, p.1; Roberto Bichi, *Intelligenza artificiale tra "calcolabilità" del diritto e tutela dei diritti*, in *Giur. it.*, 2019, p.1772 ss.
[18] Alessandra Carleo (a cura di), *Decisione robotica*, Bologna, 2019.

Para isso, o C.E.D. e a IUSS ativarão formas de troca de informações que deverão ser reguladas por contratos de candidatura posteriores destinados a identificar o tema científico abrangido pelo projeto, os termos e métodos para sua realização, as estruturas e pessoal envolvidos na iniciativa e os recursos necessários à sua implementação, incluindo as regras para a divisão dos mesmos entre os órgãos envolvidos.

8. Algumas conclusões

Itália há muito se junta à trilha da digitalização internacional iniciada em vários países, como Estados Unidos, França e Holanda.

Aliás, costumamos falar da hipótese de que, um dia, será possível prever o desfecho de um processo com um grau de precisão quase infalível, a começar pelas pequenas causas. Já existem vários projetos nesse sentido, principalmente no exterior. No entanto, é um horizonte bastante temido pela maioria dos insiders, para quem o termo "robô juiz", que há algum tempo se fala e se escreve, deve ser considerado negativo. De fato, a possibilidade de confiar o destino e os interesses das pessoas às máquinas não é bem-vinda, pelo contrário, até hoje o aspeto humano dos órgãos de julgamento ainda é considerado insubstituível.

Além deste último caso polêmico, que ainda está longe de ser uma realidade do sistema judicial italiano, os desenvolvimentos recentes são de grande interesse, pois podem trazer benefícios significativos para o trabalho de operadores e profissionais, tanto em termos de qualidade quanto de tempo[19].

A condição é naturalmente a comum ao uso de qualquer tecnologia, ou seja, a adoção de todas as medidas técnicas necessárias para proteger os direitos dos sujeitos envolvidos e evitar todas aquelas formas de discriminação que já se materializaram em diversos casos de aplicação de algoritmos para bases de dados legais.

Como vimos, a perspetiva jurídica coloca um tema de correção na relação entre direito e técnica em relação ao qual, a questão da decisão robótica assume características de extrema complexidade, como "uma questão difícil e séria, que o jurista enfrenta com medo ou desconfiança, quase como se estivesse em jogo a relação integral entre o homem e a tecnologia, para além do próprio trabalho e dos estudos"[20].

[19] Vittorio Frosini, *Informatica diritto e società*, Milano, 1988, P.70 ss.; Id., *Cibernetica, diritto e società*, Milano, 1973, 104 ss.; Paolo Grossi, *Sulla odierna 'incertezza' del diritto*, in *Giustizia Civile: rivista giuridica trimestrale*, n. 4, 2014, p.921-955; Amedeo Santosuosso, *Intelligenza artificiale e diritto*, Milano, 2020, P.101-120.

[20] Natalino Irti, *Il tessitore di Goethe (per la decisione robotica)*, in *R. d. proc.*, 2018, p. 1180 s.; Ugo Ruffolo (a cura di), *Intelligenza artificiale. Il diritto, i diritti, l'etica*, Milano, 2020; Guido Alpa

Surge claramente o pressuposto de que a certeza jurídica e as regras estabelecidas não possuem aquela agilidade e elasticidade que os sistemas tecnológicos em constante mudança exigem[21].

Isso não significa abdicar de princípios, reduzindo o alcance dos valores fundamentais. Pelo contrário, envolve a comunidade académica mais sensível numa formação mais ampla, mais atenta, aberta, mas, ainda assim, prudente.

Bibliografia

Alpa, Guido (a cura di), *Diritto e intelligenza artificiale. Profili generali, soggetti, contratti, responsabilità civile, diritto bancario e finanziario, processo civile*, Pisa, 2020.

Amato, Salvatore, *Biodiritto 4.0. Intelligenza artificiale e nuove tecnologie*, Torino 2020, p. 100 s.

Bichi, Roberto, *Intelligenza artificiale tra "calcolabilità" del diritto e tutela dei diritti*, in *Giur. it.*, 2019, p.1772 ss.

Carleo, Alessandra (a cura di), *Decisione robotica*, Bologna, 2019.

Comandé, Giorgio, *Intelligenza artificiale e responsabilità tra liability e accountability. Il carattere trasformativo dell'IA e il problema della responsabilità, in Analisi giuridica dell'economia*, 2019, I, p.173;

Costanza, Maria, *L'Intelligenza Artificiale e gli stilemi della responsabilità civile*, in *Dottrina e attualità giuridiche – Intelligenza Artificiale e diritto*, Giurisprudenza Italiana 2019, p. 1687 ss.

Crisci, Stefano, *Intelligenza artificiale ed etica dell'algoritmo*, in *F. Amm.*, 2018, p. 1787.

D'avack, Lorenzo, *La rivoluzione tecnologica e la nuova era digitale: problemi etici*, in *Intelligenza artificiale. Il diritto, i diritti, l'etica*, a cura di U. Ruffolo, con prefazione di G. Alpa-A. Barbera, Milano 2020, p. 26 s.

D'ippolito, Guido, *Profilazione e pubblicità targettizzata* on line. *Real-Time Bidding e behavioural advertising*, Napoli, 2021.

Finocchiaro, Giusella, *Intelligenza artificiale e responsabilità*, in *Contratto e impr.*, 2020, p. 724.

Frosini, Vittorio, *Informatica diritto e società*, Milano, 1988, 70 ss.; Id., *Cibernetica, diritto e società*, Milano, 1973, 104 ss.

Fusaro, Arianna, *Quale modello di responsabilità per la robotica avanzata? Riflessioni a margine del percorso europeo*, in *Nuova g. civ. comm.*, 2020, p. 1344.

Gabrielli, Enrico – Ruffolo, Ugo (a cura di), *Intelligenza Artificiale e diritto*, in *Giur. It.*, 2019, fasc. 7 (luglio 2019), p.1657 e s.

Grossi, Paolo, *Sulla odierna 'incertezza' del diritto*, in *Giustizia Civile: rivista giuridica trimestrale*, n. 4, 2014, p.921-955.

(a cura di), *Diritto e intelligenza artificiale. Profili generali, soggetti, contratti, responsabilità civile, diritto bancario e finanziario, processo civile*, Pisa, 2020.

[21] Lorenzo D'avack, *La rivoluzione tecnologica e la nuova era digitale: problemi etici*, in *Intelligenza artificiale. Il diritto, i diritti, l'etica*, a cura di U. Ruffolo, con prefazione di G. Alpa-A. Barbera, Milano 2020, p. 26 s.

Irti, Natalino, *Il tessitore di Goethe (per la decisione robotica)*, in *R. d. proc.*, 2018, p. 1180 s.
Longo, Alessandro – SCORZA, Guido, *Intelligenza artificiale. L'impatto sulle nostre vite, diritti e libertà*, Milano, 2020.
Luciani, Massimo, *La decisione giudiziaria robotica*, in *Nuovo Diritto Civile*, III, 2018, p.1.
Morelli, Claudia, *Giustizia predittiva: in Francia online la prima piattaforma europea. Uno strumento per garantire la certezza del diritto?*, www.altalex.com, 2017.
Palmerini, Erica, *Robotica e diritto: suggestioni, intersezioni, sviluppi a margine di una ricerca europea*, in *Resp. civ.*, 2016, p. 1816 ss.
Perlingieri, Carolina, *L'incidenza dell'utilizzazione della tecnologia robotica nei rapporti civilistici*, in *Rass. d. civ.*, 2015, p. 1236, nt. 1.
Proietti, Giuseppe, *Il Libro Bianco sull'intelligenza artificiale. L'approccio europeo tra diritto ed etica*, in *giustiziacivile.com*, 2020.
Ruffolo, Ugo, *Per i fondamenti di un diritto della robotica*, in Id. (a cura di), *Intelligenza artificiale e responsabilità*, Milano, 2017, p. 15.
Ruffolo, Ugo, (a cura di), *Intelligenza artificiale. Il diritto, i diritti, l'etica*, Milano, 2020.
Rulli, Edoardo, *Giustizia predittiva, intelligenza artificiale e modelli probabilistici. Chi ha paura degli algoritmi?*, Bologna, Il Mulino, 2018, p. 532.
Santosuosso, Amedeo, *Intelligenza artificiale e diritto*, Milano, 2020, p101-120.
Santosuosso, Amedeo – Boscarato Chiara – Caroleo Franco, *Robot e diritto: una prima ricognizione*, in *Nuova g. civ. comm.*, 2012, II, p. 494 ss. e spec. p. 497 ss.;

III
PERSONALIDADE E PESSOA NUMA SOCIEDADE DIGITAL

Frank e o Robô: a robótica e a hipervulnerabilidade do idoso

Robot and Frank: robotics and elderly hypervulnerability

Cristina Stringari Pasqual*

RESUMO: O uso da tecnologia, e em especial da robótica, traz indagações importantes ao mundo jurídico. Uma destas indagações é sua utilização frente aos chamados hipervulneráveis, afinal são sujeitos que por determinadas características pessoais que apresentam, exigem um tratamento diferenciado, uma proteção mais adequada frente à sua debilidade. Incluem-se neste grupo, os idosos, que em decorrência de fatores biológicos têm reconhecido pelo ordenamento jurídico brasileiro diversos direitos. Uma questão importante que se põe, entretanto, é se a disciplina jurídica existente no Brasil é suficiente para tutelar o idoso frente às novas tecnologias ou é necessário um novo marco legal. Esta é a problemática trazida por este pequeno ensaio.

PALAVRAS-CHAVE: Tecnologia, idoso, hipervulnerabilidade, marco legal.

ABSTRACT: Technology, especially robotics, raises important questions to the legal world. One of these questions is the use of technology regarding the so called hypervulnerable, considering they have some specific characteristics that demands a differentiated treatment and a more suitable protection compared to their weaknesses. This group includes the elderly, whose limiting biological factors have been recognized by the Brazilian legal system, generating several specific rights. Nonetheless, an important question is if the current legal discipline

* Fundação Escola Superior do Ministério Público (FMP), Brasil

can fully protect the elderly facing new technologies or do we need a new legal framework to protect them. This is the problematic brought by this short essay.

KEYWORDS: Tecnology, elderly, hipervulnerability, legal milestone.

SUMÁRIO: 1. Introdução. 2. A hipervulnerabilidade do idoso. 3. A robótica e a tutela jurídica do idoso: primeiras reflexões. 4. Conclusão.

1. Introdução[1]

O filme Frank e o robô se desenvolve em um futuro próximo para sua época, descrevendo a relação que se forma entre um idoso e um robô.[2] Talvez hoje, passados aproximadamente seis anos de seu lançamento, possa-se dizer que o filme é capaz de retratar uma realidade, pois não é mais ficção a utilização de robôs como cuidadores.[3]

O filme relata a vida de Frank, um aposentado, divorciado, pai de dois filhos adultos, que vive só, e que manifesta sinais de demência, atrapalhando-se com pequenas tarefas do dia a dia e esquecendo-se de fatos ocorridos.

A fragilidade de Frank preocupa seu filho, que acredita que não tem o pai condições de ficar sem companhia. Assim, a fim de evitar a internação de Frank em uma clínica geriátrica, opta por comprar um robô que, apesar de não ter características humanoides, anda, fala e é programado para cuidar da casa e da saúde física e mental de seu dono.

Frank em um primeiro momento rejeita a nova companhia, mas com o tempo acaba se apegando ao robô-cuidador, desenvolvendo-se uma relação de "amizade" e "cumplicidade" entre ambos. Surge uma impensada relação

[1] Artigo publicado em *Coletânea do V Seminário Nacional Tutelas à Efetivação de Direitos Indisponíveis*. Porto Alegre: FMP, 2019.

[2] Filme norte americano, dirigido por Jake Schreier, lançado em 12 de agosto de 2012.

[3] A empresa Soft Banck Robotics criou um robô humanoide que já é utilizado em casas de repouso. O robô se chama Pepper e consegue reagir e conversar segundo a interpretação que faz sobre o paciente. Como mostra a reportagem publicada em https://institutomongeralaegon.org/tecnologia/robo-cuidador-de-idosos. Também há na Holanda a robô Tessa criada pela empresa Tinybots que usa tecnologia de voz e estímulos musicais para ajudar idosos com deficiência. Veja em https://revistapegn.globo.com/Startups/noticia/2018/01/empresa-da-vida-um-robo-cuidador-que-ajuda-idosos.html. Sobre as tecnologias de assistência para idosos, com informações sobre diversos robôs já desenvolvidos e utilizados como cuidadores ver por todos Elisangela Gisele do Carmo; Maria Silvana Zazzetta; José Luiz Riani Costa. Robótica na assistência ao idoso com doença de Alzheimer: as vantagens e os desafios dessa intervenção. *Estudo Interdisciplinar do envelhecimento* vol. 21, n.2. Porto Alegre, 2016, p. 59-66.

entre homem e máquina, e com ela o personagem principal revive seu passado transgressor, trazendo-lhe novo ânimo, planejando um furto, o que para o robô não é recebido como uma ideia inviável, afinal não tem como avaliar a existência de uma conduta contrária à lei, pois isso não faz parte de sua programação, a qual foi composta pelos dados necessários para manter seu dono ocupado e saudável.

O pano de fundo do filme, e que o torna instigante, é a repercussão que pode surgir da relação entre um idoso e um robô, o qual mesmo não tendo uma aparência de humano, ao falar e se deslocar facilmente, acaba por criar em Frank uma sensação de cuidado, proximidade, fazendo surgir um vínculo forte entre ele e o robô, emoções consideradas comuns entre pessoas humanas, pois têm sentimentos, mas evidentemente sem reciprocidade.

A fragilidade resultante da debilidade física e emocional de Frank, faz com que ele, a partir de um certo momento, em virtude da empatia que sente pela máquina, não seja capaz de identificar que um robô não sente, não cria afetos, mas simplesmente segue sua programação, podendo atualmente, ao máximo, em virtude dos mecanismos que compõem a inteligência artificial, ter respostas e atitudes inesperadas, pois são capazes de reagir de acordo com o ambiente.

O filme traz à tona questões que merecem ser debatidas nas mais diversas áreas, inclusive, no âmbito jurídico, pois não pode o Direito afastar-se da realidade social, mas deve sim proporcionar respostas as novas situações postas em sociedade. Como a utilização de robôs cuidadores de idosos não é mais uma ficção futurista, mas uma realidade que traz questionamentos valorativos, não pode o Direito deixar de intervir. A questão é identificar como isso dever ocorrer e se no Brasil o arcabouço legislativo é suficiente para a proteção do idoso frente as inovações tecnológicas e em especial a robótica e a inteligência artificial. Esta é a reflexão que pretendemos aqui desenvolver.

2. A hipervulnerabilidade do idoso

A finitude humana é algo inevitável. O homem nasce, se desenvolve e morre. O que claramente hoje tem se identificado, todavia, é que apesar das desigualdades sociais continuarem existindo, a expectativa de vida humana tem aumentado. O meio social em que a pessoa vive, a situação econômica, assim como dos avanços da medicina, são a consequência deste aumento gradativo da população idosa.[4] E a maior longevidade faz com que cresça ainda mais a visibilidade dos idosos e legitimem-se suas demandas.

[4] Segundo a OMS entre os anos de 2000 e 2015 a expectativa de vida aumentou cinco anos globalmente, sendo tal evolução a mais rápida desde a década de 1960. Sobre o tema ver: https://

Como consequência de tal realidade foi sendo observado mundialmente uma crescente preocupação com a tutela do idoso, reconhecendo-se a necessidade de uma tutela especial em face da fragilidade inerente ao avanço da idade, resultante dos efeitos biológicos e psíquicos do gradativo envelhecimento.

O marco mundial ocorreu em 1978 quando por meio da Resolução 33/52 da Assembleia Geral da ONU, de 14 de dezembro de 1978, decidiu-se pela convocação, para 1982, de uma assembleia mundial sobre o envelhecimento, objetivando dar início a um plano internacional para garantir a segurança econômica e social das pessoas de idade avançada, sendo então aprovado o texto da Recomendação nº 18, do Plano de Ação de Viena sobre Envelhecimento, no qual constaram diversas garantias, incentivos, facilitações que os países deveriam promover aos idosos.[5]

No Brasil, nem sempre foi reconhecida a necessária proteção especial aos idosos. Foi com a Constituição Federal de 1988 que se identificou uma efetiva preocupação em tutelar este grupo, estabelecendo o artigo 230 da Carta Magna que "o Estado tem o dever de amparar as pessoas idosas, assegurando sua participação na comunidade, defendendo sua dignidade e bem-estar e garantindo-lhes à vida". Foi com o reconhecimento constitucional da fragilidade do idoso, de sua vulnerabilidade[6], que se materializou a imposição de uma proteção especial a este ser humano portador de uma debilidade física e mental que vai surgindo em decorrência do tempo.

A partir da previsão constitucional brasileira foram surgindo leis no âmbito federal[7] estabelecendo direitos especiais ao idoso no intuito de proporcio-

nacoesunidas.org. A previsão para 2050 é de que haverá dois bilhões de pessoas no mundo com 60 anos, sendo que o Brasil será o sexto país do mundo em idosos até 2025. Sobre o tema ver Elisangela Gisele do Carmo; Maria Silvana Zazzetta; José Luiz Riani Costa. Robótica na assistência ao idoso com doença de Alzheimer: as vantagens e os desafios dessa intervenção. *Estudo Interdisciplinar do envelhecimento* vol. 21, n.2. Porto Alegre, 2016, p. 48.

[5] Ver por todos Cristiano Heineck Schmitt. Consumidores Hipervulneráveis: a proteção do idoso no mercado de consumo. São Paulo: Atlas, 2014, p. 218.

[6] Fernando Vasconcelos; Maurilio Casas Maia. A tutela do melhor interesse do vulnerável: uma visão a partir dos julgados relatados pelo Min. Herman Benjamin (STJ). *Revista de Direito do Consumidor* vol. 103, jan.- fev. 2016, p. 243-271.

[7] Como a Lei nº 8.742/1993, Lei Orgânica da Assistência Social, a Lei nº 8.842/1994, que estabeleceu uma política nacional para a tutela do idoso, objetivando assegurar seus direitos sociais no intuito de promover-lhes "autonomia, integração e efetiva participação na sociedade"; a Lei nº 10.048/2000, que estabeleceu a prioridade de atendimento de idosos em repartições públicas e concessionárias de serviços públicos; e, a Lei nº 10.173/2001, impondo a preferência na tramitação de processos judiciais.

nar-lhe um efetivo tratamento isonômico. Mas foi em 2003 que se aprovou no Brasil a legislação que consagrou ao idoso o status de hipervulnerável[8], viabilizando plenamente o já instituído constitucionalmente, ou seja, uma proteção especial da intrínseca debilidade fundada na faixa etária do sujeito considerado idoso.[9]

O Estatuto do Idoso, Lei 10.741/2003, fixou o critério etário para o reconhecimento do que juridicamente seria reconhecido como idoso, considerando-o aquele com idade igual ou superior a sessenta anos (art. 1º). Destacou que além de ser titular de todos os direitos fundamentais constitucionalmente reconhecidos, deve ter assegurado "todas as oportunidades e facilidades, para preservação de sua saúde física e mental e seu aperfeiçoamento moral, intelectual, espiritual e social, em condições de liberdade e dignidade" (art. 2º). Também determinou que se garanta com total prioridade "a efetivação do direito à vida, à saúde, à alimentação, à educação, à cultura, ao esporte, ao lazer, ao trabalho, à cidadania, à liberdade, à dignidade, ao respeito e à convivência familiar e comunitária" (art. 3º).

A Lei nº 10.741/2003 instituiu um microssistema legislativo o qual trouxe não só normas de cunho geral, enumerando seus direitos fundamentais e especiais, mas também medidas de proteção, fiscalização com previsão de punições, assim como também regras de acesso à justiça, de tutela de interesses coletivos e tipificação de crimes com suas respectivas sanções.

Reconheceu o Estatuto a vulnerabilidade agravada[10] do idoso com base em critérios biológicos[11], considerando que a idade lhe coloca em uma posi-

[8] Expressão consolidada na doutrina e na jurisprudência. No ano de 2009 foi proferida no STJ a primeira decisão referindo a expressão hipervulnerabilidade. A discussão estabelecida em juízo dizia respeito a interpretação a ser dada a norma na tutela dos indígenas. REsp n. 1.064.009 – SC. Segunda Turma, Relator Ministro Herman Benjamin, DJ 27.04.2011.

[9] Assim, Fabiana Rodrigues Barletta. O direito à saúde da pessoa idosa. São Paulo: Saraiva, 2010, p. 118.

[10] Segundo Adalberto Pasqualotto e Flaviana Rampazzo Soares, "a hipervulnerabilidade representa a vulnerabilidade agravada e essa intensificação da suscetibilidade ao dano pode provir de distintas fontes, decorrentes de fatores de duração permanente ou temporária, a considerar condições individuais ou coletivas, com *potencialidade* de gerar a hipervulnerabilidade". Consumidor Hipervulnerável: análise crítica, substrato axiológico, contornos e abrangência. *Revista de Direito do Consumidor* vol.113, set.- out. 2017, p. 81-10.

[11] Importa, entretanto, ressaltar o referido por Cláudia Mara de Almeida Rabelo Viegas e Marília Ferreira de Barros ao afirmarem ser difícil determinar ao certo o idoso por um critério etário exclusivamente, pois o envelhecer é uma característica individual de cada pessoal, estando suas condições não só ligadas à idade cronológica, notando-se na sociedade diferenças significativas relacionadas à saúde e fatores sociais e econômicos. Cláudia Mara de Almeida Rabelo Viegas. Abandono Afetivo inverso: o abandono do idoso e a violação do dever de cuidado por

ção deficitária, sendo por isso merecedor de uma proteção especial. A disciplina legislativa trazida pelo microssistema visa equilibrar as relações jurídicas estabelecidas com idosos, a fim de garantir-lhes dignidade, solidariedade, igualdade real, justiça e liberdade.[12] Assim, em toda e qualquer relação jurídica que tenha por partícipe um idoso, necessário se faz um olhar jurídico a partir não só da Constituição Federal, mas também da legislação protetiva deste grupo vulnerável (Estatuto do Idoso), em conjunto com outras leis que tenham também aplicação ao objeto da relação *in concreto (v.g.* o Código de Defesa do Consumidor)[13].

Podemos dizer que o idoso pode ter uma vulnerabilidade biológica, a qual pode ser agravada por seu "ciclo vital"[14], mas que também pode ser intensificada por questões psicológicas, culturais, sociais, ambientais e até mesmo afetivas, sendo possível assim falar-se na hipervulnerabilidade decorrente de um processo "biopsicossocial"[15].

Frente a tal realidade, imprescindível um olhar cauteloso a toda e qualquer relação jurídica entabulada por um idoso, e por isso o uso das novas tecnologias postas em sociedade traz à tona o questionamento de se o que hoje há no Brasil em matéria legislativa é suficiente para garantir a proteção efetiva da hipervulnerabilidade decorrente do envelhecimento, e em especial em face da utilização de robôs como instrumentos ou ferramentas para o acompanhamento ou tratamento de necessidades especiais demandadas pelos idosos.

3. A robótica e a tutela jurídica do idoso: primeiras reflexões

O desenvolvimento da tecnologia trouxe consigo uma revolução social inquestionável. Permitiu o acesso a novos conhecimentos, a agilidade na produção e um maior crescimento econômico.

parte da prole. *Cadernos do Programa de Pós-Graduação de Direito da UFGRS*, edição digital, n. 03, 2016. Porto Alegre, p. 168 – 201.

[12] Fernando Vasconcelos; Maurilio Casas Maia. A tutela do melhor interesse do vulnerável: uma visão a partir dos julgados relatados pelo Min. Herman Benjamin (STJ). *Revista de Direito do Consumidor* vol. 103, jan.- fev. 2016, p. 172.

[13] Como destaca Cristiano Heineck Schimitt. Consumidores Hipervulneráveis: a proteção do idoso no mercado de consumo. São Paulo: Atlas, 2014.

[14] Expressão utilizada por Michelle Rinco; Andrea Lopes; Marisa Accioly Domingues. Envelhecimento e Vulnerabilidade Social: discussão conceitual à luz das políticas públicas e suporte social. *Revista Temática Kairós Gerontologia* n. 15, online, p.87.

[15] Michelle Rinco; Andrea Lopes; Marisa Accioly Domingues. Envelhecimento e Vulnerabilidade Social: discussão conceitual à luz das políticas públicas e suporte social. *Revista Temática Kairós Gerontologia* n. 15, online, p. 95.

Durante muito tempo a *internet* foi considerada um grande marco no desenvolvimento tecnológico, afinal ela proporcionou a redução das distâncias, a comunicação instantânea e a obtenção de dados e informações a um custo menor e em menos tempo. Atualmente, em matéria de inovação tecnológica, o tema de destaque diz respeito às tecnologias emergentes, estando entre elas a robótica e a chamada inteligência artificial.

Os robôs já há muito tempo vêm ocupando os mais diversos setores da economia. A Coréia do Sul, Singapura e Japão ocupam as primeiras posições em número de robôs por habitante, mas países da Europa e os Estados Unidos também apresentam um número significativo.[16] Foram os robôs inicialmente concebidos para serem máquinas industriais, visando auxiliar o trabalho humano, proporcionando maior produtividade com redução de custos, executando tarefas absolutamente mecanizadas. Atualmente, entretanto, já estão chegando no setor de serviços[17] e comércio, como caso do xrobô[18], e até mesmo na vida doméstica das pessoas.

Com o desenvolvimento da inteligência artificial para a robótica, a execução de atividades que outrora eram impensáveis, vistas como passíveis de serem executadas somente por humanos, tornaram-se realidade.[19] No caso de robôs criados para cuidados pessoais, há até mesmo aqueles que exercem a função de "companheiro emocional do proprietário", podendo falar com ele, como também entender e expressar as emoções humanas.[20]

O desenvolvimento da robótica tem atingido tamanha velocidade que hoje em dia muitos robôs apresentam aparência semelhante a de um humano, como o caso da Sophia, criado pela Hanson Robotics[21], com comportamento flexível, capazes de manter um diálogo, de apresentar opiniões e decisões.[22]

[16] Veja reportagem publicada na revista Veja publicou reportagem acerca dos dez países mais robotizados do mundo. https://veja.abril.com.br/tecnologia/conheca-os-10-paises-mais-robotizados-do-mundo/.

[17] Iván Mateo Borge. La robótica y la inteligencia artificial en prestación de servicios jurídicos. *Inteligencia artificial, Tecnología, Derecho.* Susana Navas Navarro (Dir.). Valencia: Tirant lo Blanch, 2017, p. 123.

[18] Ver sobre o tema, http://xrobo.com.br/

[19] Erica Palmerini. Robótica y derecho: sugerencias, confluencias, evoluciones en el marco de una investigación europea. *Revista de Derecho Privado* n. 32, enero – junio de 2017. Colombia: Bogotá, p. 02.

[20] Nesse sentido Moisés Barrio Andrés. Robótica, inteligencia artificial y Derecho. *Cyber Elcano* n. 36, set. 2018. Real Instituto Elcano, p. 2.

[21] A cerca do robô Sophia ver https://www.hansonrobotics.com/wp-content/uploads/2018/09/gds_Sophia_D058.jpg.

[22] Rafael de Asis. Sobre ética y tecnologías emergentes. *Papeles el tiempo de los derechos* n. 7, 2013, p. 8.

Dotados de inteligência artificial tornam-se autômatos, executam atividades que vão além o programado por um computador, tendo capacidade de escolhas e de aplica-las independentemente de controle ou influência externa, mas a partir dos dados que capta.[23] Estão os robôs superando os seres humanos em muitas dimensões e aspectos.[24]

Estas máquinas dotadas de uma tecnologia crescente, saíram da ficção, dos filmes, dos livros, e passaram a ocupar um espaço na realidade das pessoas, e esta nova realidade traz uma série de questionamentos e nas mais diversas áreas da ciência, como a economia, filosofia, psicologia, ética e o direito, cada qual destacando a complexidade dos efeitos que podem surgir em cada um dos seus setores. Fala-se agora em "segunda era das máquinas" ou em "indústria 4.0" para descrever esta nova fase mundial.[25]

No que diz respeito à assistência as pessoas, como se dá no caso dos robôs para cuidados domésticos e pessoais, a análise ético-jurídica é fundamental, em especial quando o sujeito da relação é um hipervulnerável como se identifica no caso do idoso, impondo-se uma cautela redobrada para que não sejam lesados direitos de personalidade do mesmo, para que sua dignidade seja preservada.

Não se pode admitir que o idoso seja excluído na vida social e familiar, afastando-o do convício humano, sob o pretexto de que existem tecnologias que permitem um cuidado mais qualificado para a saúde do idoso.[26] Mais, a imposição de um convívio contínuo entre máquina e idoso pode indiscutivelmente ensejar a criação de uma relação afetiva entre homem e máquina, como claramente demonstrado na ficção Frank e o Robô, mas é indiscutível que mesmo que dotados de inteligência artificial, jamais serão os robôs capazes de desenvolver sensibilidade ou afeto, pois só o ser humano é detentor de tal capacidade.

[23] Assim, Bruno Farage da Costa Felipe. Direitos dos robôs, tomadas de decisões e escolhas morais: algumas considerações acerca da necessidade de regulamentação ética e jurídica da inteligência artificial. *Revista Juris Poiesis* n. 22, abr. 2017, Rio de Janeiro, p. 155.

[24] Luis Daniel Crovi. Los animales y los robots frente al Derecho. *Revista Venezolana de Legislación y Jurisprudencia* n. 10, 2018. Caracas, p. 139.

[25] Assim, Luis Daniel Crovi. Los animales y los robots frente al Derecho. *Revista Venezolana de Legislación y Jurisprudencia* n. 10, 2018. Caracas, p. 139.

[26] Em pesquisa realizada envolvendo idosos, gerontologistas e cuidadores, concluiu-se que os robôs cuidadores poderiam ter um papel positivo em três setores: segurança (auxiliar a andar, a evitar e detectar quedas, monitorar aspectos incomuns de comportamentos e lembrar de tomar medicamentos), socialização (lembrando de compromissos médios e familiares, sinalizando se a pessoa ficou muitos dias sem falar com ninguém, propondo chamadas, estimulando a realização de atividades cognitivas) e tarefas diárias. Em: https://institutomongeralaegon.org/saude-e-bem-estar/autonomia/o-papel-da-robotica-na-sociedade-em envelhecimento.

Reflexões e debates importantes sobre a robótica e direito já tem sido identificadas nos países mais desenvolvidos. Um grande exemplos disso é o RoboLaw[27], projeto europeu de pesquisa sobre robótica e suas repercussões legais, éticas, sociais e técnicas, o qual objetivou discutir se os marcos jurídicos existentes e vigentes são adequados para atender a acelerada evolução tecnológica, criando como resultado recomendações que foram remetidas à Comissão Europeia.[28] Outro exemplo de grande importância é a Resolução do Parlamento Europeu de 16 de fevereiro de 2017, que contém recomendações à comissão sobre disposições de Direito Civil sobre Robótica.[29]

Da referida Resolução identifica-se claramente a preocupação com os reflexos da robótica perante os idosos, sendo destacado o crescimento consideravelmente rápido da população de idade superior a 80 anos e que a robótica e a inteligência artificial devem ser concebidas de forma a preservar a dignidade, a autonomia e a autodeterminação do indivíduo, especialmente na seara dos cuidados e da companhia dos humanos.[30]

No que diz respeito à reflexão ética, em específico, muitas discussões já foram entabuladas, chegando-se até mesmo a criar-se o termo roboética, a partir da reflexão de filósofos, juristas, cientistas da robótica, antropólogos e sociólogos, na busca de fixar bases éticas no desenho, desenvolvimento e emprego dos robôs. Considera-se roboética "um conjunto de critérios ou teorias com as quais se pretende das respostas a todos os problemas éticos que

[27] RoboLaw-Regulating emerging robotic Technologies: Robotics facing law and ethics, disponível em www.robolaw.eu.

[28] Verónica E. Melo, El derecho ante la inteligência artificial y la robótica. *El Derecho: Diario de doctrina y jurisprudencia* n. 14.343, feb. 2018.

[29] Em http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+TA+P8-TA-2017 0051+0+DOC+XML+V0//PT.

[30] Na introdução da Resolução (letra F) consta expressamente: "Considerando que o envelhecimento da população se deve a um aumento da esperança de vida em consequência da melhoria das condições de vida e do progresso na medicina moderna, constituindo um dos principais desafios políticos, sociais e económicos do século XXI com que as sociedades europeias se deparam; que, em 2025, mais de 20% dos europeus terão uma idade igual ou superior a 65 anos, assistindo-se a um aumento particularmente rápido do número de pessoas de 80 anos ou mais, pelo que o equilíbrio entre gerações nas nossas sociedades será fundamentalmente diferente, e que é do interesse da sociedade que os idosos gozem de boa saúde e permaneçam ativos o máximo de tempo possível". Da mesma forma na letra O: "Considerando que os desenvolvimentos na robótica e inteligência artificial podem e devem ser concebidos de tal forma que preservem a dignidade, a autonomia e a autodeterminação do indivíduo, especialmente nos domínios dos cuidados e da companhia dos humanos, e no contexto dos dispositivos médicos, da «reparação» ou melhoria dos seres humanos."

surgem na criação e no uso de robôs, e que se projetam em seus fabricantes e usuários, e inclusive nos próprios robôs"[31].

Frente a esta realidade identifica-se claramente a necessidade de um marco jurídico capaz de trazer as respostas necessárias, a segurança possível às relações que passam a surgir no âmbito da robótica e da inteligência artificial e este cuidado deve ser tomado no Brasil.

Em nosso país o desenvolvimento da robótica não se apresenta tão célere como nos países desenvolvidos, mas já se tem observar um crescimento no setor, o que deve despertar nos operadores do Direito uma atenção especial, pois há que se verificar se o manancial legislativo hoje existente é capaz de proporcionar segurança à este grupo social.

E a discussão não pode tardar, pois não demorará para no Brasil surgir a necessidade real de respostas. Prova disso é que está sendo desenvolvido, desde 2013, no Centro Universitário um robô 100% brasileiro, o qual foi batizado de Judith, e que foi concebido para ser aplicado no auxílio de idosos ou pessoas com dificuldade de locomoção em atividades domésticas. Trata-se de um protótipo autômato que consegue se locomover pela casa, reconhecer pessoas caídas no chão e identificar o estado de espírito delas por meio das expressões faciais. Segundo os criadores de Judith, espera-se que ela tenha um bom desenvolvimento de interação com os humanos e a inteligência artificial, para que não configure um modelo invasivo e autoritário.[32]

Verifica-se assim que apesar de no Brasil existirem legislações bastante avançadas que visam proteger os idosos em suas relações sociais, elas indiscutivelmente não foram construídas pensando na realidade da robótica e da inteligência artificial e por isso uma cautelosa reflexão sobre o tema é fundamental para a segurança da população.

[31] Rafael de Asis. Sobre ética y tecnologías emergentes. *Papeles el tiempo de los derechos*, n. 7, 2013, p. 9. Destaca o autor que o termo roboética foi proposto oficialmente durante o Primeiro Simposio Internacional de roboética em San Remo, no ano de 2004. Ressalta também que no mesmo ano no Japão, foi firmado por cientistas e representantes da indústria robótica japonesa, na International Robot Fair de Fukuoka, A World Robot Declaration, que equivale a uma versão robótica de juramento hipocrático, sendo nela afirmado que "a próxima geração de robôs coexistirá com os humanos. Assistirão os humanos física e psicologicamente. Contribuirão para a realização de uma sociedade segura y pacífica. Com o objetivo de que a sociedade aceite a acolha os robôs, será necessário definir e aplicar determinados standards, modificar os ambientes de vida e trabalho, y as instituições púbicas promoverão a introdução de robôs. Importa também destacar que da Resolução do Parlamento Europeu sobre Robótica antes referida, também surgiu um Código de Ética para engenheiros de robótica, o que mais uma vez destaca a preocupação com o tema.

[32] Em https://startse.com/noticia/conheca-judith-o-primeiro-robo-100-brasileiro.

4. Conclusão

Como demonstra a cena final do filme Frank e o Robô, quando Frank, já residindo em um lar para idosos identifica ao seu redor, acompanhando outros idosos, um número significativo de robôs idênticos aquele com quem ele conviveu em sua casa por um período, o crescimento da utilização de robôs tem se demonstrado uma realidade.

A robótica ocupa um mercado em célere expansão, um mercado estratégico economicamente e de grande impacto social, e que por isso exige uma indiscutível intervenção por parte do Direito, a fim de garantir uma maior segurança às relações, fixando bases éticas claras, direitos e deveres capazes de garantir a necessária tutela aos envolvidos nas relações jurídicas que se constituírem. Um marco regulatório capaz de proteger a parte mais débil na relação parece viável, mas ao mesmo tempo há que se ter a cautela para que não surja uma disciplina que impeça o desenvolvimento tecnológico necessário para o crescimento do setor.

Não pode o Direito fechar os olhos à tutela de direitos de personalidade inerentes ao ser humano, mas não pode ele servir de entrave às novas criações tecnológicas. A difícil tarefa que surge, portanto, é buscar um equilíbrio. Não se pode permitir que valores fundamentais, como a proteção da dignidade humana venha a ser desrespeitada.

Referências

Andrés, Moisés Barrio. Robótica, inteligencia artificial y Derecho. *Cyber Elcano*, n. 36, set. 2018. Real Instituto Elcano, p. 1-7.

Asis, Rafael de. Sobre ética y tecnologías emergentes. *Papeles el tiempo de los derechos*, n. 7, 2013, p. 1 – 14.

Barletta, Fabiana Rodrigues. O direito à saúde da pessoa idosa. São Paulo: Saraiva, 2010.

Borge, Iván Mateo. La robótica y la inteligencia artificial en prestación de servicios jurídicos. *Inteligencia artificial, Tecnología, Derecho*. NAVARRO, Susana Navas (Dir). Valencia: Tirant lo Blanch, 2017, p. 123 – 150.

Carmo, Elisangela Gisele do; Zazzetta, Maria Silvana; Costa, José Luiz Riani. Robótica na assistência ao idoso com doença de Alzheimer: as vantagens e os desafios dessa intervenção. *Estudo Interdisciplinar do envelhecimento* vol. 21, n. 2. Porto Alegre, 2016, p. 47-74.

Costa Felipe, Bruno Farage da. Direitos dos robôs, tomadas de decisões e escolhas morais: algumas considerações acerca da necessidade de regulamentação ética e jurídica da inteligência artificial. *Revista Juris Poiesis*, n°22, abr. 2017, Rio de Janeiro, p. 150-169.

Crovi, Luis Daniel. Los animales y los robots frente al Derecho. *Revista Venezolana de Legislación y Jurisprudencia*, n. 10, 2018. Caracas, p. 133-144.

Melo, Verónica E. El derecho ante la inteligencia artificial y la robótica. *El Derecho: Diario de doctrina y jurisprudencia* n. 14.343, feb. 2018.

Palmerini, Erica. Robótica y derecho: sugerencias, confluencias, evoluciones en el marco de una investigación europea. *Revista de Derecho Privado*, n. 32, enero – junio de 2017. Bogotá, p. 53 a 97.

Pasqualotto, Adalberto; Soares, Flaviana Rampazzo. Consumidor Hipervulnerável: análise crítica, substrato axiológico, contornos e abrangência. *Revista de Direito do Consumidor* vol.113, set.- out. 2017, p. 81-10.

Resolução do Parlamento Europeu, 16 de fevereiro de 2017. Disposições de Direito Civil sobre Robótica. http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+TA+P8-TA-2017-0051+0+DOC+XML+V0//PT.

Rinco, Michelle; Lopes, Andrea; Domingues, Marisa Accioly. Envelhecimento e Vulnerabilidade Social: discussão conceitual à luz das políticas públicas e suporte social. *Revista Temática Kairós Gerontologia* n. 15, online, p.79-95.

Schmitt, Cristiano Heineck. Consumidores Hipervulneráveis: a proteção do idoso no mercado de consumo. São Paulo: Atlas, 2014.

Viegas, Cláudia Mara De Almeida Rabelo. Abandono Afetivo inverso: o abandono do idoso e a violação do dever de cuidado por parte da prole. *Cadernos do Programa de Pós-Graduação de Direito da UFGRS*, edição digital, n. 03, 2016. Porto Alegre, p. 168 – 201.

Vasconcelos, Fernando; Maia, Maurilio Casas. A tutela do melhor interesse do vulnerável: uma visão a partir dos julgados relatados pelo Min. Herman Benjamin (STJ). *Revista de Direito do Consumidor* vol. 103, jan.- fev. 2016, p. 243 – 271.

Una clasificación de la Inteligencia Artificial jurídica desde la perspectiva de la Filosofía del Derecho*

A classification of Artificial legal Intelligence from the perspective of Legal Philosophy

Jorge Crego**

RESUMEN: Hace ya más de 40 años desde que se comenzase a explorar las posibilidades de emplear inteligencia artificial en el campo del derecho. Existen numerosos usos y propuestas en el marco de la IA y el derecho, desde la automatización de la búsqueda de la información jurídica hasta la personalización del derecho. Estas propuestas plantean diversas cuestiones relacionadas con la filosofía del derecho. La proliferación de propuestas, en ocasiones, dificulta identificar la relación existente entre cada una de ellas con debates específicos de la filosofía del derecho. Este trabajo ofrece una clasificación de usos de la IA jurídica que facilita la identificación de los temas de filosofía de derecho especialmente afectados por cada uso. Se propone distinguir tres grandes categorías: derecho formalizado, práctica jurídica algorítmica, y derecho personalizado. Esta propuesta incorpora los avances recientes en *machine learning* y propuestas novedosas como la personalización del derecho. Con esta clasificación resulta más sencillo identificar las premisas iusfilosóficas en que se asientan o sobre las que deben reflexionar los diversos proyectos de empleo de IA en el derecho.

PALABRAS CLAVE: derecho algorítmico; derecho personalizado; imperio de la ley; machine learning; teoría del derecho; legal tech.

* Este trabajo se ha realizado en el marco del Proyecto de Investigación "Inteligencia artificial jurídica y Estado de Derecho" (PID2022-139773OB-100), perteneciente a la convocatoria de Proyectos de Generación de Generación de Conocimiento 2022, del Programa Estatal para Impulsar la Investigación Científico-Técnica y su Transferencia.

** Profesor Ayudante Doctor. ORCID: 0000-0001-7072-6569. jorge.crego@udc.es. Universidade da Coruña

ABSTRACT: Over 40 years ago, the possibility of using artificial intelligence in law began to be explored. There exist many uses and proposals in AI and law, ranging from the automation of legal information retrieval to personalized law. These proposals raise many questions related to legal philosophy. Proliferation of proposals sometimes complicates the identification of the relationship between each proposal and specific topics of legal philosophy. This study offers a classification of the uses of legal AI that contributes to identifying the topics of legal philosophy especially affected by each use. Three different categories are proposed: formalised law, algorithmic legal practise, and personalized law. This proposal incorporates recent developments in machine learning and novel proposals such as the personalization of law. The classification helps in the identification of the underlying jurisprudential assumptions and topics which require consideration that are associated with each of the AI and law proposals.

KEYWORDS: algorithmic law; personalized law; rule of law; machine learning; legal theory; legal tech.

1. Introducción

El campo de la inteligencia artificial (IA) y el derecho tiene ya más de cuarenta años de historia desde sus inicios en los años ochenta[1]. Durante los últimos

[1] Trevor Bench-Capon, "Thirty years of Artificial Intelligence and Law: Editor's Introduction", en *Artificial Intelligence and Law,* 2022, v. 30, nº 4, pp. 475-479. La relación entre derecho e "inteligencia artificial", entendida esta como una idea o una hipótesis teórica es antigua. Es habitual referirse a Leibniz como origen de la idea de computar el derecho; *vid. infra,* nota 74. No es extraño encontrar en los trabajos de autores del ámbito de la filosofía del derecho referencias a una "inteligencia artificial" jurídica; *vid., v.gr.,* Ronald Dworkin, *Law's Empire*, Cambridge,

años, la disciplina de la IA ha crecido al calor de los avances en la capacidad de computación, la disponibilidad de grandes cantidades de datos y el desarrollo de nuevos métodos como el *machine learning*[2]. El uso de sistemas de IA está cada vez más extendido y no es de extrañar que este florecimiento contemporáneo de la IA también se presencie en el ámbito de la IA jurídica[3].

Los nuevos enfoques basados en *machine learning* y nuevas propuestas hacen obsoletas viejas consideraciones sobre la IA jurídica desde una perspectiva iusfilosófica. En el pasado, autores como Susskind han tratado de identificar cuestiones de filosofía del derecho asociables al uso de IA en el ámbito jurídico[4]. Sin embargo, estas reflexiones se basan en el enfoque de los sistemas jurídicos expertos, que actualmente ha sido desplazado por el enfoque del *machine learning*. Además, las nuevas posibilidades derivadas de los avances en el campo de la IA han permitido la aparición de propuestas radicalmente diferentes a las preexistentes como es el caso del "derecho personalizado"[5]. Aunque trabajos como el de Susskind tienen todavía un valor innegable, resulta preciso completarlos a la luz de los nuevos desarrollos de la IA.

El propósito de este trabajo es ofrecer una clasificación de las propuestas actuales sobre el uso de la IA en el derecho que pueda resultar adecuada para la filosofía del derecho. En ocasiones se ha tratado la IA jurídica como un todo unívoco y se ha achacado a dicho concepto abstracto una serie de consecuencias que se podrían asociar a usos específicos de la IA jurídica pero no a todos los usos propuestos. Una adecuada clasificación de los usos de la IA

Harvard University Press, 1986, p. 412; Norberto Bobbio, *El positivismo jurídico. Lecciones de Filosofía del Derecho reunidas por el doctor Nello Morra*, Madrid, Debate, 1993, p. 143; Richard A. Posner, *How Judges Think*, Cambridge, Harvard University Press, 2008.

[2] Para una breve historia moderna de la inteligencia artificial, *vid*. Stuart J. Russell, Peter Norvig, *Artificial Intelligence: A Modern Approach*, 4ª ed., Harlow, Pearson Education, 2022, pp. 35-48.

[3] Michael Chui et al., "The state of AI in 2021," (McKinsey Analytics, 2021). En España, son varios los autores que han estudiado el fenómeno de la IA jurídica; *vid., v.gr.,* Daniéle Bourcier, Pompeu Casanovas, *Inteligencia artificial y derecho*, Barcelona, UOC, 2003; Pompeu Casanovas, "Inteligencia Artificial y Derecho: a vuelapluma", en *Teoría y Derecho. Revista de Pensamiento Jurídico,* v. 7, 2010; José Ignacio Solar Cayón, *La inteligencia artificial jurídica. El impacto de la innovación tecnológica en la práctica del Derecho y el mercado de servicios jurídicos*, Cizur Menor, Thomson Reuters, Aranzadi, 2019.

[4] Richard Susskind, *Transforming the Law: Essays on Technology, Justice and the Legal Marketplace*, Oxford, Oxford University Press, 2003, pp. 177-206.

[5] *V.gr.,* Omri Ben-Shahar, Ariel Porat, *Personalized Law: Different Rules for Different People*, Nueva York, Oxford University Press, 2021; Christoph Busch, Alberto De Franceschi, *Algorithmic Regulation and Personalized Law: A Handbook*, Londres, Bloomsbury, 2021.

jurídica ayudará a afinar en las críticas y a la asociación de ciertos problemas de la filosofía del derecho a usos concretos de la IA jurídica.

El concepto de filosofía del derecho se entiende aquí en un sentido amplio, incluyendo tanto lo que se suele denominar filosofía del derecho en sentido estricto como la teoría del derecho[6]. Además, este trabajo parte de una descripción simplificada del derecho que servirá para justificar la clasificación. Esta simplificación descansa en dos premisas. Primero, el derecho contemporáneo opera fundamentalmente a través normas generales, dirigidas a tipos y no a casos particulares. Segundo, la tarea esencial del razonamiento jurídico es determinar si cierta regla o grupo de reglas se aplica a un caso particular. Estas dos ideas permiten identificar de forma clara los cambios más significativos asociados a las categorías de IA jurídica propuestas en este trabajo. Por último, resulta necesario señalar que este trabajo no pretende ofrecer una enumeración exhaustiva de las cuestiones de filosofía del derecho planteadas por el uso de la IA jurídica y su correspondiente clasificación atendiendo a las categorías propuestas. El objetivo del presente trabajo es más humilde: consiste simplemente en ofrecer un marco razonablemente abstracto de clasificación de dichos usos e ilustrar cómo permite conectar cada categoría con una serie de problemas de filosofía del derecho. La identificación de subcategorías y la clasificación de más temas de filosofía del derecho podría resultar adecuada, pero tal tarea debería ser objeto de un estudio posterior.

La propuesta de clasificación está formada por tres grandes categorías: derecho formalizado, práctica jurídica algorítmica, y derecho personalizado. El derecho formalizado se refiere a la representación del derecho existente en lenguajes computables para el posterior perfeccionamiento del sistema jurídico. La práctica jurídica algorítmica se refiere a todo uso de IA jurídica dirigido a automatizar, en mayor o menor grado, las tareas típicas de los juristas. El derecho personalizado es una propuesta reciente dirigida a sustituir las normas generales típicas de los sistemas jurídicos contemporáneos por conjuntos complejos de normas personalizadas, adaptadas a las particularidades de cada sujeto jurídico y las circunstancias del caso particular. Esta clasificación del campo de la IA y el derecho, como se ha señalado, clarifica la asociación de ciertas cuestiones fundamentales de la filosofía del derecho a cada una de las categorías de la IA jurídica.

Tras esta breve introducción el trabajo se divide en cuatro apartados. Primero, se presenta la simplificación del derecho que servirá de base para la justificación de la clasificación propuesta. Seguidamente, se identifica el enfo-

[6] Pedro Serna, "Teoría del derecho y filosofía del derecho", en *Persona y Derecho,* v. 32, 1995.

que actualmente dominante en el ámbito de la IA y el derecho, a través de la distinción entre el enfoque de codificación manual y el enfoque de *machine learning*. En tercer lugar, se presenta las tres categorías de IA jurídica propuestas y se describe la naturaleza de esta clasificación. En cuarto lugar, se relaciona cada una de las categorías con una serie de cuestiones iusfilosóficas, tratando de mostrar cómo la clasificación ayuda a identificar las cuestiones de filosofía del derecho que se ven especialmente afectadas por cada una de las categorías. Un último apartado recoge las principales conclusiones alcanzadas en el trabajo.

2. El derecho como conjunto de reglas y el razonamiento jurídico como subsunción

Con el objetivo de presentar el desarrollo de la IA en el derecho, este trabajo parte de una caracterización del derecho actual como una práctica basada en reglas generales. Consecuentemente, presenta el razonamiento jurídico como un proceso dirigido a enmarcar un caso particular en un conjunto más o menos complejo de reglas generales. Por tanto, en este modelo, la operación central del razonamiento jurídico es la subsunción.

Más allá de la discusión relativa a si el concepto de derecho incluye como condición necesaria la regulación de la conducta humana a través de reglas o, al menos, la existencia de algunas normas generales[7], parece razonable afirmar que, *de facto*, en aquellos sistemas que actualmente denominamos "derecho", las reglas tienen un papel dominante. La idea contemporánea de sistema jurídico encaja en la descripción de Fuller: se trata de sistemas "para someter la conducta humana al gobierno de reglas", entendidas estas últimas como normas generales[8]. El derecho contemporáneo opera a través de reglas y la idea de reglas remite a la idea de "generalizaciones"[9]. Como afirma Schauer, las reglas se dirigen "a *tipos* y no a casos particulares"[10]. En palabras de Kramer, la idea del gobierno a través de reglas se basa en normas que, por ser generales, "se aplican a un tipo de conducta en lugar de solamente a algunas instancias particulares de conducta y la mayoría de este tipo de normas se dirigen a cate-

[7] *Vid.*, por ejemplo, Georg Henrik von Wright, *Norma y acción. Una investigación lógica*, Santiago de Chile, Ediciones Olejnik, 2019, pp. 70-72; Frederick Schauer, *Las reglas en juego. Un examen filosófico de la toma de decisiones basada en reglas en el derecho y en la vuda cotidiana*, Madrid, Marcial Pons, 2004, p. 229 y ss.

[8] Lon L. Fuller, *The Morality of Law*, 2ª ed., New Haven, Yale University Press, 1969, p. 46.

[9] Frederick Schauer, *Las reglas en juego, cit.*, pp. 75-96.

[10] *Ibid.*, p. 76.

gorías generales de personas en lugar de a individuos designados"[11]. Podría cuestionarse la idea de que los sistemas jurídicos contemporáneos se basan fundamentalmente en reglas a partir de la defensa de la importancia de los denominados "principios" en el derecho actual[12]. Sin embargo, al describir aquí las reglas por su carácter general, la idea de principios puede considerarse un subtipo de la categoría amplia de reglas. De este modo, su presencia en los sistemas jurídicos contemporáneos no cuestiona la idea de que dichos sistemas funcionan predominantemente a través de normas generales.

Si se parte de la consideración de que los sistemas jurídicos contemporáneos están constituidos principalmente por reglas, entendidas como normas generales, puede aceptarse también que la actividad central del razonamiento jurídico es la subsunción de casos particulares en reglas generales[13]. La subsunción se refiere a la aplicación de la regla al caso particular y tiene como objetivo resolver lo que MacCormick denomina "el problema de clasificación"[14]. Este problema se refiere a la cuestión de si un caso al que nos enfrentamos puede considerarse una instancia del supuesto de hecho de la regla general. Por ejemplo, si tomamos una regla que prohíbe comprar bebidas alcohólicas a las personas menores de 18 años, su aplicación exige dilucidar si, en un caso concreto, la bebida que pretende comprar una persona es una bebida alcohólica y si la persona que pretende comprarla es mayor o menor de 18 años. Por tanto, el proceso de subsunción implica la identificación del supuesto de hecho de una regla general con un caso particular. Esta operación consiste entonces en la individualización de la regla general, identificando una norma particular que rige el caso concreto. Por ejemplo, la regla "prohibido comprar bebidas alcohólicas a las personas menores de 18 años" se transforma en un caso concreto en "dado que esta cerveza es una instancia de bebida alcohólica y esta persona particular es menor de 18 años, está prohibido que esta persona compre esta cerveza".

[11] Matthew Kramer, *Objectivity and the Rule of Law*, Cambridge, Cambridge University Press, 2007, p. 110.

[12] Ronald Dworkin, *Los derechos en serio*, Barcelona, Ariel, 1984, pp. 61-101; Robert Alexy, *El concepto y la validez del derecho*, Barcelona, Gedisa, 1993, pp. 161-174.

[13] La afirmación de que la principal tarea del razonamiento jurídico es la subsunción no quiere decir que sea la única ni que tal subsunción sea una operación formal simple; *vid*. Andrei Marmor, *Interpretation and Legal Theory*, 2ª ed., Oxford, Portland, Hart Publishing, 2005, pp. 95-98. Una descripción más detallada debería hacer mención a otro tipo de razonamientos jurídicos, como la analogía, la ponderación, el juicio de proporcionalidad, o la cuestión de la interpretación, así como estudiar si esos razonamientos pueden concebirse como tipos de subsunción o medios relacionados con la subsunción.

[14] Neil MacCormick, *Legal Reasoning and Legal Theory*, Oxford, Clarendon Press, 1994, pp. 93-97.

Esta caracterización de los sistemas jurídicos contemporáneos y del razonamiento jurídico asociado a dichos sistemas domina de algún modo el campo de la IA y el derecho. En ocasiones se ha identificado la "personalización masiva" [*mass customization*] como el objetivo último del campo de la IA jurídica[15]. Ashley presenta esta concepción de la tarea de la IA en el derecho a través de la siguiente pregunta: "si la ingeniería de procesos de servicios jurídicos implica repensar cómo ofrecer 'soluciones muy baratas y de mucha calidad', ¿quién o qué será responsable de personalizar [*tailoring*] aquellas soluciones al problema particular del cliente?"[16]. Esta pregunta ya definía el objetivo de la IA jurídica en la época del dominio de los sistemas jurídicos expertos (*legal expert systems*)[17]. Susskind consideraba ya en dicha época (fundamentalmente, los años 80) que la implementación de la interpretación y la aplicación de normas jurídicas a problemas jurídicos concretos es el "sello distintivo del conocimiento jurídico", y la IA jurídica debe dirigirse a implementar esa interpretación y aplicación de forma automatizada[18]. En la actualidad, los modelos de *machine learning* que dominan en el ámbito de la IA jurídica mantienen la idea de la personalización como ideal regulador.

En conclusión, para los objetivos de este trabajo, los sistemas jurídicos pueden caracterizarse como sistemas normativos basados principalmente en reglas generales y el razonamiento jurídico puede concebirse como dirigido a la subsunción de un caso particular en una o varias reglas generales del sistema jurídico. Estas dos ideas servirán para identificar las cuestiones centrales de la filosofía del derecho afectada por la IA. La descripción del derecho actual a partir de estas dos ideas es, sin lugar a duda, una simplificación. Sin embargo, se trata de una simplificación razonable que propicia una mejor comprensión de los efectos de la IA en el derecho.

3. La inteligencia artificial jurídica contemporánea y los modelos de *machine learning*

En el ámbito de la IA y, consecuentemente, en la IA jurídica es común distinguir entre dos enfoques diferentes: el enfoque de codificación manual y el

[15] Este término se introdujo por primera vez en la obra de Davis y se ha identificado como el propósito clave, o al menos la tendencia dominante, del uso de la IA en la actualidad; Davis Stanley, *Future perfect*, Reading, Addison-Wesley, 1987; Karen Yeung, "Five Fears about Mass Predictive Personalisation in an Age of Surveillance Capitalism", en *International Data Privacy Law,* v. 8, n. 3, 2018.

[16] Kevin D. Ashley, *Artificial Intelligence and Legal Analytics: New Tools for Law Practice in the Digital Age*, Cambridge, Cambridge University Press, 2017, p. 355.

[17] *Vid., infra,* apartado 3.

[18] Richard Susskind, *Transforming the Law, cit.*, p. 209.

enfoque del *machine learning*[19]. En el enfoque de la codificación manual, una persona codifica en un lenguaje computable las reglas que quien tiene experiencia en un campo concreto le transmite. Este enfoque dominó el campo de la IA hasta mediados de los años 80 y cuajó en los denominados "sistemas jurídicos expertos (*legal expert systems*)[20]. El enfoque del *machine learning* se refiere a algoritmos que identifican patrones en datos para resolver diversos tipos de tarea y que, además, mejoran automáticamente su desempeño[21]. Este enfoque ha dominado el campo de la IA desde finales de la década de los 80[22], pero ha ganado importancia en el campo de la IA jurídica a partir de la década de 2010[23]. El enfoque de los *legal expert systems* perdió fuerza por varios motivos, entre ellos la constatación de las dificultades para transformar el derecho en un conjunto de reglas perfectamente computables y el cuello de botella derivado de los costes y el tiempo necesario para la codificación manual[24].

Los modelos de *machine learning* ofrecen varias ventajas que son relevantes en el campo de la IA jurídica. Enfrentados a fenómenos complejos, los sistemas de *machine learning* infieren patrones a partir de los datos existentes. Por tanto, en lugar de que programadores humanos establezcan reglas explícitas, estos programadores preparan un algoritmo para que, siguiendo una serie de instrucciones, identifique el patrón que más se ajusta a los datos recibi-

[19] Harry Surden, "Artificial intelligence and law: An overview", en *Georgia State University Law Review,* v. 35, 2019, pp. 1310-1321. Distinciones similares se pueden encontrar en Stuart J. Russell, Peter Norvig, *Artificial Intelligence: A Modern Approach, cit.*, p. p. 53 (señalando que la IA ha pasado de la programación manual del conocimiento al *machine learning* a partir de datos); Mireille Hildebrandt, "Algorithmic Regulation and the Rule of Law", en *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences,* v. 376, n. 2128, 2018; Monika Zalnieriute, Lyria Bennett Moses, George Williams, "The rule of law and automation of government decision-making", en *The Modern Law Review,* v. 82, n. 3, 2019, pp. 432-435; Mireille Hildebrandt, "Code-Driven Law: Freezing the Future and Scaling the Past", en *Is Law Computable?: Critical Perspectives on Law and Artificial Intelligence*, Markou, Christopher; Deakin, Simon (ed.), Oxford, Hart Publishing, 2020.

[20] Stuart J. Russell, Peter Norvig, *Artificial Intelligence: A Modern Approach, cit.*, pp. 40-42.

[21] Harry Surden, "Machine learning and law: an overview", en *Research Handbook on Big Data Law*, Vogl, Roland (ed.), Cheltenham, Northampton, Edward Elgar Publishing, 2021, p. 171. Las bases del *machine learning* están explicadas de forma comprensible para legos en la materia en David Lehr, Paul Ohm, "Playing with the data: what legal scholars should learn about machine learning", en *University of California, Davis Law Review,* v. 51, 2017.

[22] Stuart J. Russell, Peter Norvig, *Artificial Intelligence: A Modern Approach, cit.*, pp. 42-45.

[23] Trevor Bench-Capon, "Thirty years of Artificial Intelligence and Law: Editor's Introduction", *cit.*; Serena Villata et al., "Thirty years of artificial intelligence and law: the third decade", *ibid.*

[24] Kevin D. Ashley, *Artificial Intelligence and Legal Analytics, cit.*, p. 11; Richard Susskind, *Transforming the Law, cit.*, pp. 161-176.

dos[25], correlacionando una serie de datos conocidos (*input*) con un dato que se trata de descubrir (*output*). De este modo, se puede hacer frente a problemas más complicados de los que se podrían resolver a través de reglas codificadas manualmente[26], ya que se pueden identificar correlaciones complejas entre una gran cantidad de *inputs* y un *output* específico. Los *legal expert systems* y, en general, lo sistemas de codificación manual tienen un límite en cuanto a la complejidad que pueden representar. Al estar codificados por seres humanos, no pueden representar un conocimiento más complejo que el que los seres humanos pueden explicitar[27]. Es importante tener en cuenta que el uso de *machine learning* no implica necesariamente la formulación de patrones complejos. Los sistemas de *machine learning* pueden emplearse para identificar automáticamente patrones simples que podrían considerarse reglas y que, en ciertos casos, un ser humano podría codificar[28]. Sin embargo, se trata de un enfoque especialmente valioso para la identificación de patrones complejos.

En resumen, en la actualidad conviven un enfoque de codificación manual y un enfoque de *machine learning*. El primero supone la estructuración manual de un conjunto de reglas y el segundo implica la identificación automática de patrones, generalmente complejos, a partir de una serie de datos. El segundo domina el campo de la IA jurídica contemporánea aunque la codificación manual es todavía relevante.

4. Tres categorías de la inteligencia artificial jurídica

Atendiendo a la caracterización del derecho y el razonamiento jurídico presentada en la anterior sección, es posible estructurar el campo de la IA jurídica en tres categorías generales: el derecho formalizado, la práctica jurídica

[25] Esta explicación es excesivamente simple y esconde excesivamente el rol del ser humano en el proceso. Lehr y Ohm ofrecen una explicación más detallada que permite entender el funcionamiento de estos algoritmos; David Lehr, Paul Ohm, "Playing with the data", *cit.*, *passim*.

[26] Harry Surden, "Machine learning and law", en *Washington Law Review,* v. 89, n. 1, 2014, p. 94.

[27] También cabría añadir el obstáculo de que el derecho no puede representarse en modelos de lógica formal. A esto, varios autores han señalado que avances contemporáneos en computación y modelos no-clásicos de lógica facilitan la representación formal del conocimiento jurídico: L. Wolfgang Bibel, "AI and the conquest of complexity in law", en *Artificial Intelligence and Law,* v. 12, n. 3, 2004; Kevin D. Ashley, *Artificial Intelligence and Legal Analytics, cit.*, pp. 127-168.

[28] *Vid.*, a modo de ejemplo, Ian H. Witten et al., *Data Mining. Practical Machine Learning Tools and Techniques*, 4ª ed., Cambridge, Morgan Kaufmann, 2016, pp. 219 y ss. (el capítulo 6 de este manual se dedica, precisamente, a explicar el aprendizaje de reglas y árboles de decisión); Monika Zalnieriute, Lyria Bennett Moses, George Williams, "The rule of law and automation of government decision-making", *cit.*, p. 434; Mireille Hildebrandt, "Code-Driven Law", *cit.*, p. 67.

algorítmica, y el derecho personalizado. Cada una de estas categorías tiene un impacto particular en la filosofía del derecho que permite diferenciarlas entre sí, aunque existan similitudes y conexiones entre ellas. En lo que resta de sección se presentan las tres categorías. La sección finaliza con una breve explicación sobre la naturaleza de la clasificación propuesta y las relaciones entre cada categoría.

4.1. El derecho formalizado

Una primera categoría de la IA jurídica es el derecho formalizado. Esta categoría se refiere a los usos de IA para la reformulación, clarificación, compleción, o, en general, el perfeccionamiento de los textos jurídicos[29]. Su objeto es, por tanto, el conjunto de normas preexistentes en el sistema jurídico, que se transforma de algún modo a través de las posibilidades que ofrece la formalización del derecho.

La base del derecho formalizado reside en tareas como la representación de conjuntos de normas jurídicas en lenguajes formales o computables[30], la creación de redes jurídicas (*legal networks*)[31], la creación de ontologías jurídicas[32], etc. En general, el propósito de la formalización es reformular el modo en que se presentan las normas en los textos jurídicos (tanto los cuerpos jurídicos como las sentencias) de manera tal que se obtenga una representación del texto inicial que incorpore cierto conocimiento jurídico. Sin embargo, la idea de derecho formalizado no se refiere a la simple representación, sino al uso de IA jurídica para alterar el contenido del sistema jurídico con la intención de perfeccionarlo. De este modo, el resultado de las tareas mencionadas serviría para perfeccionar el sistema jurídico con posterioridad. Un ejemplo temprano de las potencialidades del derecho formalizado es la representación de cuerpos jurídicos en lenguaje formal con el propósito de identificar ambigüedades sintácticas para la posterior clarificación del significado de ciertos conectores[33]. Por otro lado, es necesario destacar que el derecho formalizado no se refiere necesariamente a un sistema jurídico que se presenta a

[29] Trevor Bench-Capon, Henry Prakken, "Introducing the logic and law corner", en *Journal of logic and computation*, v. 18, n. 1, 2008, pp. 2-4.

[30] L. Wolfgang Bibel, "AI and the conquest of complexity in law", *cit.*, pp. 166-171.

[31] Kevin D. Ashley, *Artificial Intelligence and Legal Analytics, cit.*, pp. 70-72.

[32] *Ibid.*, 171-209.

[33] Vid., *v.gr.*, Layman E. Allen, "Symbolic logic: A razor-edged tool for drafting and interpreting legal documents", en *Yale Law Journal*, v. 66, n. 6, 1957. Branting enumera varios beneficios derivables del análisis de redes; *vid.* Branting, L. Karl, "Data-centric and logic-based models for automated legal problem solving", en *Artificial Intelligence and Law*, v. 25, n. 1, 2017, pp. 16-17.

los sujetos jurídicos a través de un lenguaje formal. Lo fundamental es que el sistema jurídico se formalice en algún momento o se establezcan las relaciones entre conceptos o textos jurídicos, aprovechando tales operaciones para pulir el texto jurídico[34]. Cuestión distinta es que, *a posteriori*, el texto definitivo se publique a través de lenguaje natural o de lenguaje formal.

El derecho formalizado no se limita a pulir los cuerpos jurídicos identificando y resolviendo contradicciones o lagunas. Este uso de la IA jurídica también permite integrar la jurisprudencia en los propios cuerpos jurídicos, incorporando las precisiones interpretativas elaboradas por los tribunales. Ashley señala, siguiendo a Levi, que "las reglas cambian a medida que las reglas se aplican"[35]. Esto sucede claramente al menos en el caso de conceptos vagos pues, como afirma el propio Ashley, "un concepto jurídico se expande o contrae a medida que los tribunales decide que se aplican o no a nuevos casos". Por tanto, un potencial modo de perfeccionar los sistemas jurídicos sería a través de la incorporación del conocimiento jurídico presente en la jurisprudencia a los cuerpos jurídicos[36].

El derecho formalizado se distingue de la práctica jurídica algorítmica por el objeto de transformación. Mientras que la práctica jurídica algorítmica se dirige a la automatización de tareas propias del jurista, el derecho formalizado no altera directamente dichas tareas, sino que se dirige a la reformulación del sistema jurídico entendido como conjunto de normas para alcanzar un mayor grado de precisión y plenitud.

De acuerdo con Bench-Capon y Prakken, los primeros intentos dirigidos a formalizar el derecho mostraron la existencia ciertas limitaciones de la lógica clásica en el ámbito jurídico[37]. Pese a los intentos por adaptar los modelos lógicos a la realidad jurídica, la perspectiva dominante en el campo de la IA y el derecho pasó de la representación formal a la automatización del proceso de razonamiento jurídico.

[34] Sarah B. Lawsky, "Formalizing the code", en *Tax Law Review*, v. 70, n. 2, 2017, p. 395.

[35] Kevin D. Ashley, *Artificial Intelligence and Legal Analytics, cit.*, pp. 74-77; Edward H. Levi, *An introduction to legal reasoning*, Chicago, Londres, University of Chicago Press, 2013, pp. 3-4.

[36] Branting, L. Karl, "Data-centric and logic-based models for automated legal problem solving", *cit.*, pp. 10-11. Sobre el modelado del razonamiento jurídico presente en sentencias, *vid.* Kevin D. Ashley, *Artificial Intelligence and Legal Analytics, cit.*, pp. 73-106. Para una crítica del derecho formalizado, *vid.* Mireille Hildebrandt, *Smart Technologies and the End(s) of Law: Novel Entanglements of Law and Technology*, Cheltenham, Northampton, Edward Elgar Publishing, 2015, pp. 133-185; Mireille Hildebrandt, "Code-Driven Law", *cit. passim.* Un argumento relevante de Hildebrandt es precisamente que al fijar el significado de los conceptos jurídicos, el derecho formalizado impide la discusión del significado de dichos conceptos y limita la protección de los ciudadanos frente al ejercicio del poder.

[37] Trevor Bench-Capon, Henry Prakken, "Introducing the logic and law corner", *cit.*, p. 4.

4.2. La práctica jurídica algorítmica

La práctica jurídica algorítmica es una segunda categoría en la que se incluyen los usos de sistemas de IA dirigidos a automatizar el razonamiento jurídico, es decir, las tareas cuyo objetivo final es la subsunción de un caso particular en una regla general. Una característica común a todas las funcionalidades que forman parte de esta categoría es que, al menos de manera directa, no transforman el sistema jurídico entendido como sistema de normas. Por el contrario, suponen la automatización de actividades que los juristas llevan a cabo con dichas normas.

La práctica jurídica algorítmica incluye la automatización de la recuperación de información, los algoritmos predictivos, y la interpretación y aplicación algorítmica del derecho. La recuperación automatizada de información jurídica es quizá el primer ámbito en que se empleó IA jurídica y el más desarrollado[38]. Actualmente, el uso de sistemas de IA en la búsqueda jurídica es común. Los esfuerzos actuales se dirigen a mejorar las prestaciones de los buscadores a través del paso de la búsqueda basada en datos a la búsqueda basada en conocimiento[39].

Otros sistemas pertenecientes a la práctica jurídica algorítmica son los algoritmos predictivos, que algunos autores ven como una evolución de la búsqueda de información[40]. La idea de estas propuestas es que, gracias al creciente almacenamiento de datos jurídicos, es posible predecir de forma automática el resultado de un litigio, en ocasiones con más precisión que cualquier jurista humano[41]. Igual que sucede con la búsqueda automatizada, algunos autores han explorado la posibilidad de mejorar los sistemas predictivos actuales, pasando de modelos basados en datos a modelos basados en conocimiento jurídico explícito[42].

[38] Trevor Bench-Capon et al., "A history of AI and Law in 50 papers: 25 years of the international conference on AI and Law", en *Artificial Intelligence and Law,* v. 20, n. 3, 2012, pp. 218-219; Luigi Lombardi Vallauri, "Verso un Sistema Esperto Giuridico Integrale", en *Persona y Derecho,* n. 31, 1994, p. 172; Kevin D. Ashley, *Artificial Intelligence and Legal Analytics, cit.*, pp. 210-233.

[39] Richard Susskind, *Transforming the Law, cit.*, pp. 178-180, 185-189; Kevin D. Ashley, *Artificial Intelligence and Legal Analytics, cit.*, pp. 248-254.

[40] Daniel Martin Katz, "Quantitative Legal Prediction – Or – How I Learned to Stop Worrying and Start Preparing for the Data-Driven Future of the Legal Services Industry", en *Emory Law Journal,* v. 62, n. 4, 2013, p. 947.

[41] *Vid., v.gr.,* Nikolaos Aletras et al., "Predicting Judicial Decisions of the European Court of Human Rights: A Natural Language Processing Perspective", en *PeerJ Computer Science,* v. 2, n. e93, 2016. Este trabajo ha sido objeto de fuertes críticas; *vid., v.gr.,* Frank Pasquale, Glyn Cashwell, "Prediction, persuasion, and the jurisprudence of behaviourism", en *University of Toronto Law Journal,* v. 68, n. Supplement 1, 2018.

[42] Kevin D. Ashley, *Artificial Intelligence and Legal Analytics, cit.*, pp. 107-126.

Por último, la interpretación y aplicación algorítmica del derecho también es una subcategoría de la práctica jurídica algorítmica[43]. La predicción de la solución de un caso y la identificación de textos jurídicos relevantes llevada a cabo por la recuperación automatizada de información jurídica no agotan las actividades de la práctica jurídica. Aun habiendo identificado los textos relevantes para un caso concreto y previendo la solución de dicho caso, los juristas deben interpretar el significado de esos textos y argumentar por qué el caso particular es una instancia del supuesto de hecho de las normas contenidas en los textos, o de la interpretación concreta de las normas sostenida por los tribunales. Estas complejas tareas requerirían el desarrollo de sistemas basados en conocimiento jurídico y la superación de una serie de dificultades todavía no resueltas[44]. Dentro de esta subcategoría puede distinguirse entre el razonamiento basado en reglas y el razonamiento basado en casos[45], así como los modelos híbridos, como los modelos computacionales de argumentos jurídicos[46].

En resumen, la práctica jurídica algorítmica automatiza las tareas de identificación de normas relevantes, la interpretación y la aplicación de dichas normas a la luz de las consideraciones de los tribunales y la predicción de la solución del caso. En otras palabras, esta categoría se refiere a todos los usos de la IA dirigidos a automatizar diversas tareas del trabajo humano de subsunción de los casos particulares en las reglas generales.

Existen otros usos de IA jurídica que pueden concebirse como instancias de uno de los tres usos ya mencionados o como combinaciones de dichos usos. Un ejemplo es el caso de la respuesta a preguntas jurídicas[47]. Este tipo de sistemas ofrecen una respuesta a cuestiones jurídicas transmitidas por individuos a través de lenguaje natural. La respuesta a este tipo de preguntas puede concebirse como el resultado de un ensamblaje de un sistema de

[43] Se entiende la interpretación jurídica como la comprensión o explicación del significado de un texto jurídico; Andrei Marmor, *Interpretation and Legal Theory, cit.*, p. 10. La aplicación es entonces la actividad dirigida a determinar que un caso particular es una instancia del supuesto de hecho de una norma jurídica; Neil MacCormick, *Legal Reasoning and Legal Theory, cit.*, pp. 93-97.

[44] Richard Susskind, *Transforming the Law, cit.*, p. 197; Luigi Lombardi Vallauri, "Verso un Sistema Esperto Giuridico Integrale", *cit.*, pp. 167-172; Kevin D. Ashley, *Artificial Intelligence and Legal Analytics, cit.*, pp. 38-56.

[45] Kevin D. Ashley, *Artificial Intelligence and Legal Analytics, cit.*, pp. 38-106.

[46] *Ibid.*, pp. 127-168.

[47] Albert H. Yoon, "The Post-Modern Lawyer: Technology and the Democratization of Legal Representation", en *University of Toronto Law Journal*, v. 66, n. 4, 2016, pp. 467-468; Richard Susskind, *Tomorrow's Lawyers: An Introduction to Your Future*, 2ª ed., Oxford, Oxford University Press, 2017, pp. 54-55.

recuperación de información jurídica (la identificación de los textos jurídicos para resolver la pregunta) y un sistema de interpretación y aplicación que reformula la información jurídica recuperada para transformarla en una respuesta directa al caso concreto. Algo similar sucedería con las propuestas de sustituir a jueces por sistemas de IA jurídica, ya que en este caso se trataría de sistemas de interpretación y aplicación dotados de autoridad[48].

4.3. El derecho personalizado

La última categoría de la IA jurídica es el llamado "derecho personalizado". En los últimos años, diversos autores han defendido la posibilidad de trascender la idea de un derecho basado en reglas generales e implantar un sistema normativo basado en normas personalizadas[49]. La premisa común es que los avances en métodos de *machine learning* y el desarrollo del *big data* permite identificar las normas particulares que avanzan un determinado propósito normativo[50].

Casey y Niblett ofrecen una explicación de cómo implementar el derecho personalizado que puede servir de modelo general[51]. El proceso comienza con la identificación de un objetivo político por parte del poder legislativo. Se trataría de objetivos formulados a modo de estándares cuya calibración se produciría en el momento de aplicación, adaptándose a las particularidades del caso, al estilo de los principios en el derecho[52]. Estos objetivos serían transformados en "microdirectivas" a través del uso de "tecnología predictiva". En este sentido, un sistema de *machine learning* "analiza una cantidad masiva de datos de manera instantánea para predecir qué reglas pueden alcanzar con precisión el objetivo político"[53]. El algoritmo aprende qué combinación

[48] Volokh describe las "IA juezas" como capaces de aplicar reglas y estándares vagos, realizar inferencias de hechos y presentar escritos lo suficientemente persuasivos; Eugene Volokh, "Chief Justice Robots", en *Duke Law Journal,* v. 68, 2018. De esta descripción puede extraerse que las "IA juezas" hacen algo más que interpretar y aplicar el derecho: también determinan los hechos probados.

[49] John O. McGinnis, Steven Wasick, "Law's algorithm", en *Florida Law Review,* v. 66, 2014; Benjamin Alarie, "The Path of the Law: Towards Legal Singularity", en *University of Toronto Law Journal,* v. 66, n. 4, 2016; Anthony J. Casey, Anthony Niblett, "The Death of Rules and Standards", en *Indiana Law Journal,* v. 92, n. 4, 2017; Anthony J. Casey, Anthony Niblett, "Framework for the New Personalization of Law", en *The University of Chicago Law Review,* v. 86, n. 2, 2019; Omri Ben-Shahar, Ariel Porat, *Personalized Law, cit., passim.*

[50] Benjamin Alarie, "The Path of the Law", *cit.*, p. 445.

[51] Anthony J. Casey, Anthony Niblett, "The Death of Rules and Standards", *cit.*, pp. 1410-1412.

[52] Ronald Dworkin, *Los derechos en serio, cit.*, pp. 72-80.

[53] Anthony J. Casey, Anthony Niblett, "The Death of Rules and Standards", *cit.*, p. 1410.

de normas singulares adaptadas a las particularidades de cada caso es la que optimiza el objetivo político. Esta propuesta altera por completo la naturaleza del sistema jurídico, puesto que ahora estará compuesto por objetivos amplios y por un vasto conjunto de normas personalizadas adaptadas a las particularidades de cada caso. En cierto modo, ya no existirían reglas generales sino normas singulares para cada problema jurídico.

Casey y Niblett asumen que un ser humano no podrá guiarse por este sistema complejo de normas, por lo que es necesaria una "tecnología comunicativa" que reciba datos sobre la situación particular de un individuo o sobre el problema jurídico concreto e identifique la norma particular que se aplica a su situación. Por tanto, "la norma que controla un escenario particular debe tener en cuenta cientos o miles de factores, pero el individuo recibirá un mandato simple como una luz verde o roja"[54]. En definitiva, el sistema de *machine learning* identifica qué norma optimiza el objetivo político a la luz de las particularidades del caso concreto.

Según sus defensores, el derecho personalizado es más conveniente que un sistema de reglas generales por tres motivos: el beneficio de la precisión, la consecución de una seguridad jurídica perfecta, y la optimización de la justicia del sistema. El beneficio de la precisión se refiere al hecho de que las "reglas personalizadas tienen el potencial de lograr los fines subyacentes de cualquier ley de manera más efectiva"[55]. A modo de ejemplo, si el propósito del código de circulación es asegurar una circulación fluida reduciendo los accidentes a una determinada tasa y minimizando razonablemente la contaminación, un sistema de *machine learning* podría identificar el conjunto de reglas personalizadas que maximizase dichos objetivos[56]. Además, Alarie afirma que el derecho personalizado resolvería el problema de la falta de la seguridad jurídica[57], al ofrecer una norma singular para cada caso concreto. De este modo, no existirían problemas de interpretación o aplicación que hiciesen al sujeto jurídico dudar acerca del modo en que el derecho regulará su conducta: la subsunción no sería necesaria y no habría discusión posible sobre si la norma se aplica al caso concreto. Por último, la personalización

[54] *Ibid*., 1411.

[55] Omri Ben-Shahar, Ariel Porat, *Personalized Law, cit*., p. 39.

[56] Sobre la dificultad existente para fijar estos objetivos políticos y las alternativas existentes, *vid*. Anthony J. Casey, Anthony Niblett, "Framework for the New Personalization of Law", *cit*., pp. 338-345.

[57] Benjamin Alarie, "The Path of the Law", *cit*., pp. 445-446; en un sentido similar, *vid*. Anthony J. Casey, Anthony Niblett, "The Death of Rules and Standards", *cit*., p. 1405.

del derecho también traería consigo mejoras en la justicia del sistema[58]. Por emplear el argumento de Ben-Shahar y Porat, "si las personas varían en referencia a lo que merecen o necesitan, y si el mérito y la necesidad dependen de numerosos atributos, un sistema justo tendría que tratar a las personas *de manera diferente*, en consonancia con el modo en que sus atributos relevantes difieren"[59]. El derecho personalizado se adapta a esos atributos, mientras que las reglas generales no[60].

En definitiva, el derecho personalizado se apoya en los avances en *machine learning* para sustituir las reglas generales en las que se basa el derecho contemporáneo por reglas personalizadas, adaptadas a las particularidades de cada situación.

4.4. Anotaciones sobre la naturaleza de la clasificación propuesta

La clasificación que se ha presentado no es ni excluyente ni exhaustiva, pero permite identificar las tres líneas generales del uso de IA en derecho y tiene la virtud de estructurarse alrededor de una de las características esenciales de los sistemas sociotécnicos contemporáneos basados en IA. Además, las categorías están íntimamente relacionadas, como han señalado diversos autores.

Esta clasificación no pretende ser excluyente porque es compatible con otras clasificaciones de los sistemas de IA jurídica. De hecho, en el primer apartado ya se ha dejado constancia de una clasificación diferente de naturaleza técnica: la distinción entre modelos de codificación manual y modelos de *machine learning*. Tampoco se trata de una clasificación necesariamente exhaustiva, aunque sí identifica los tres grandes ámbitos en los que se ha desarrollado el uso de IA en el derecho. Podría preguntarse, por ejemplo, en qué lugar quedan usos de IA jurídica como la elaboración automática de documentos jurídicos, las tareas de averiguación procesal de hechos relevantes, o la resolución automática de disputas[61]. Pese a todo, el objetivo es presentar

[58] Para Casey y Niblett, se evitaría el carácter sobre- e infra-incluyente de las reglas generales, que implican necesariamente la prohibición de conductas que deberían permitirse y la permisión de conductas que deberían prohibirse; Anthony J. Casey, Anthony Niblett, "The Death of Rules and Standards", *cit.*, p. 1407. La idea del beneficio de la precisión ya mencionada también mejoraría la justicia, dado que el derecho personalizado crea normas que "se ajustan mejor a cada caso"; Omri Ben-Shahar, Ariel Porat, *Personalized Law, cit.*, p. 41.

[59] Omri Ben-Shahar, Ariel Porat, *Personalized Law, cit.*, p. 122.

[60] Este argumento relativo a la comparación entre reglas y normas singulares se encuentra ya en Platón; Platón, "Político", en *Diálogos*, traducción de Rico Gómez, María, Madrid, Centro de Estudios Políticos y Constitucionales, 2007, pp. 294a-295b.

[61] José Ignacio Solar Cayón, *La inteligencia artificial jurídica, cit.*, pp. pp. 137-185. Pese a las diferencias particularidades de estas tareas, parece que sería posible relacionarlas con los usos de

una clasificación conveniente para la identificación de cuestiones de filosofía jurídica, de ahí que se parta de una caracterización simplificada del derecho y el razonamiento jurídico como criterios dirigidos a estructurar la clasificación. En el fondo, la práctica jurídica algorítmica tiene como propósito rector la automatización del razonamiento jurídico y el derecho personalizado la transformación de un derecho basado en reglas en un sistema normativo basado en normas particulares. En el caso del derecho formalizado, el objetivo general es refinar el modelo de derecho existente para que se ajuste mejor al ideal de un sistema completo, coherente y predecible.

Por último, resulta fundamental destacar que las tres categorías no son compartimentos estanco. Existen importantes conexiones entre las tres[62]. Por ejemplo, la práctica jurídica algorítmica y el derecho formalizado están íntimamente relacionados. Los avances en la formalización del derecho favorecerían la algoritmización de la práctica jurídica, ya que al reformular el derecho en términos computables la automatización del razonamiento jurídico sería más sencilla[63]. En el plano teórico también existe una importante vinculación. Por ejemplo, como ya se ha señalado, Bench-Capon y Prakken destacan que el interés por el razonamiento jurídico en el ámbito de la IA surgió al constatar que la representación formal del derecho se enfrentaba a limita-

IA jurídica recogidas en este trabajo. Por ejemplo, la averiguación procesal de hechos relevantes está íntimamente vinculada con la actividad de aplicación del derecho y la resolución automática de disputas se relaciona con la interpretación y aplicación del derecho o la predicción de decisiones judiciales.

[62] También se han señalado conexiones entre usos específicos dentro de cada categoría; *vid., v.gr.,* Daniel Martin Katz, "Quantitative Legal Prediction", *cit.*, p. 947 (sobre la relación entre los buscadores jurídicos y la predicción); JC Smith, "Machine Intelligence and Legal Reasoning", en *Chicago-Kent Law Review,* v. 73, n. 1, 1997, p. 333 (sobre la relación entre el razonamiento jurídico y los buscadores jurídicos); Frank Pasquale, "A Rule of Persons, Not Machines: The Limits of Legal Automation", en *George Washington Law Review,* v. 87, n. 1, 2019, p. 51 (sobre cómo la predicción algorítmica presiona para la formalización y la automatización de la interpretación).

[63] A modo de ejemplo, *vid.* L. Wolfgang Bibel, "AI and the conquest of complexity in law", *cit.*, pp. 168-171 (sobre la relación entre la formalización del derecho y la mejora en la búsqueda de información jurídica o en la interpretación y aplicación automatizada); Kevin D. Ashley, *Artificial Intelligence and Legal Analytics, cit.*, pp. 45-47 (sobre cómo la automatización del razonamiento jurídico basado en reglas se vería facilitada por la formalización del sistema jurídico); John O. McGinnis, Steven Wasick, "Law's algorithm", *cit.*, p. passim (sobre la relación entre búsqueda de información y personalización); Anthony J. Casey, Anthony Niblett, "Self-Driving Laws", en *University of Toronto Law Journal,* v. 66, n. 4, 2016, p. 435 (señalando la relación entre predicción y personalización: "la predición se convierte en la ley"); Anthony J. Casey, Anthony Niblett, "Framework for the New Personalization of Law", *cit.*, pp. 338-345 (sobre la relación entre la personalización del derecho y, por ejemplo, el uso de "IA juezas" o la formalización de principios detallados en sentencias judiciales).

ciones relevantes derivadas de la naturaleza del razonamiento jurídico[64]. En cierto modo, siguiendo a estos autores, podría decirse que la práctica jurídica algorítmica se desarrolló por las insuficiencias del derecho formalizado. Estas conexiones no son un problema para la clasificación, sino más bien una virtud, ya que permiten identificar las relaciones entre los diversos usos de la IA jurídica e incluso identificar posibles sinergias y tendencias en la transformación del derecho[65].

5. Las categorías de la inteligencia artificial jurídica como marco de identificación de las cuestiones de filosofía del derecho

Una vez expuestas las tres grandes categorías de la IA jurídica es posible mostrar de qué modo esta clasificación ayuda a relacionar los usos de IA jurídica con las diferentes cuestiones clásicas de la filosofía del derecho. En este apartado se presentan algunas relaciones a modo de ilustración, justificando por qué cada una de las categorías plantea específicamente algunos temas destacados de la disciplina. En general, cada categoría se asienta en una serie de premisas o concepciones particulares del derecho que se relacionan directamente con diferentes temas de la filosofía jurídica. En este trabajo no es posible ofrecer una relación exhaustiva de cuestiones de la filosofía del derecho vinculadas a cada una de las categorías. Bastará con mostrar que la clasificación es adecuada para tal empresa.

5.1. El derecho formalizado y el problema de la coherencia y la plenitud del sistema jurídico

La idea del derecho formalizado es una plasmación de dos ideales tradicionalmente asociados a la noción de sistema jurídico: la coherencia y, quizá en menor medida, la plenitud[66]. Como se ha señalado, el derecho formalizado consiste en la representación de las normas del sistema de un modo tal que cierto conocimiento jurídico se incorpore a dicha representación. Un ejemplo es la identificación de ambigüedades o contradicciones en el sistema jurídico.

[64] Trevor Bench-Capon, Henry Prakken, "Introducing the logic and law corner", *cit.*, p. 4.

[65] De hecho, de lo expuesto hasta ahora puede entreverse cierta tendencia hacia la personalización del derecho. Dada la caracterización del razonamiento jurídico ofrecida en este trabajo, todo uso de la IA jurídica parece dirigirse a automatizar o facilitar el proceso de transformación de reglas generales en normas singulares. Cuestión distinta es la pregunta normativa acerca de si tal tendencia es deseable. En este trabajo, sin embargo, no es posible detallar ni los motivos que justifican la identificación de tal tendencia ni los argumentos a favor o en contra de seguir dicha tendencia.

[66] Norberto Bobbio, *El positivismo jurídico*, *cit.*, pp. 201-213.

Según Alarie, la IA jurídica serviría para identificar lagunas y reformular el sistema jurídico de un modo más completo[67]. Así es como se plasmarían en la realidad los ideales de plenitud y coherencia del sistema jurídico defendidos por ciertos autores.

Sin embargo, hay quien cuestiona tanto la posibilidad como la conveniencia de alcanzar los ideales de coherencia y plenitud, entendidos como características necesarias de un sistema jurídico que permite ofrece una respuesta incontrovertible para cada problema jurídico. Autores como Lombardi Vallauri señalan que no es posible formalizar el lenguaje legislativo para evitar ambigüedades semánticas[68]. Para Lombardi Vallauri, la plenitud y la coherencia no son completamente alcanzables no solo por las "lagunas estáticas" existentes en el derecho, es decir, las derivadas de la imposibilidad de fijar en textos jurídicos el significado preciso de todas y cada una de las palabras empleadas. También existen las que él denomina "lagunas dinámicas", lagunas existentes cuando no se ha previsto un caso futuro (incógnitas del devenir) o cuando es discutible en qué regla concreta se subsume un caso particular (incógnitas de lo individual)[69]. El autor italiano afirma que, aunque se intentasen colmar las lagunas dinámicas con nuevas normas adaptadas a nuevos casos futuros, el aumento del número de normas aumenta las posibilidades de que se generen contradicciones y ambigüedades. Por tanto, los intentos de reducir las lagunas dinámicas aumentan la posibilidad de generar nuevas lagunas estáticas. Si tratamos de especificar el significado de las normas existentes extendiendo los textos jurídicos con más formulaciones normativas o con formulaciones normativas más detalladas, abriremos la puerta a una cantidad mayor de lagunas estáticas. Curiosamente, la única solución sería "crear un mapa tan grande como el territorio a describir"[70], una idea que se puede relacionar con el derecho personalizado.

También se puede cuestionar la conveniencia de un derecho sin lagunas. Primero, es común señalar que ciertas ambigüedades existentes en los ordenamientos jurídicos son intencionales y cumplen una función que se podría considerar valiosa, como por ejemplo ofrecer mayor discrecionalidad al intérprete del derecho, o mayor capacidad para adaptar la norma al caso particular[71].

[67] Benjamin Alarie, "The Path of the Law", *cit.*, pp. 451-453.

[68] Luigi Lombardi Vallauri, *Corso di filosofia del diritto*, Milán, CEDAM, 1981, p. 32. *Cfr.*, nota 27.

[69] *Ibid*., 35-39.

[70] *Ibid*., 39.

[71] *Ibid*., 33; Kevin D. Ashley, *Artificial Intelligence and Legal Analytics, cit.*, pp. 39-40.

Por otro lado, Hildebrandt destaca que la plenitud y coherencia algorítmica del derecho no es un ideal deseable por socavar algunas protecciones valiosas existentes en los sistemas jurídicos contemporáneos. Para ella, un derecho sin lagunas que permita adaptarse a nuevas situaciones o que permita discutir si un caso particular se ha de subsumir en una regla general es un derecho "legalista" y "congelado" que desvirtúa el imperio de la ley o *rule of law*[72]. Por ejemplo, Hildebrandt afirma que un derecho sin lagunas, en el que las leyes ofrecen un criterio preciso e incontrovertible para resolver un caso concreto, colapsa la figura del juez en la del legislador y elimina los controles y contrapesos típicos de la división de poderes[73], ya que los sujetos jurídicos no pueden disputar el significado de las normas y su aplicación al caso concreto.

En definitiva, el derecho formalizado puede concebirse como una aspiración a la plasmación de los ideales de plenitud y coherencia del sistema jurídico a través de la IA. Por tanto, esta categoría y los usos de IA jurídica contenidos en ella se pueden relacionar con las discusiones relativas a la naturaleza del sistema jurídico y a la posibilidad y conveniencia de un derecho sin lagunas.

5.2. La práctica jurídica algorítmica y la aspiración a computar el razonamiento jurídico

La práctica jurídica algorítmica, como ya se ha señalado, está especialmente vinculada a la automatización del razonamiento jurídico. La discusión acerca de la posibilidad de automatizar el razonamiento jurídico tiene raíces profundas en la filosofía del derecho y se relaciona con la discusión sobre la naturaleza del razonamiento en el derecho[74]. Esto no impide señalar que ciertas aplicaciones de la práctica jurídica, como puede ser la búsqueda automatizada de textos jurídicos, no plantean excesivos problemas de filosofía del derecho y, actualmente, están incorporadas a la práctica habitual de los juristas.

[72] Mireille Hildebrandt, "Code-Driven Law", *cit.*, *passim.*; Mireille Hildebrandt, "Law as computation in the era of artificial legal intelligence: Speaking law to the power of statistics", en *University of Toronto Law Journal*, v. 68, n. 1, 2018.

[73] Mireille Hildebrandt, "Algorithmic Regulation and the Rule of Law", *cit.*, *passim*. Lombardi Vallauri, al explorar las posibles soluciones a las lagunas dinámicas, menciona precisamente cómo la idea de que el legislador resolviese todas las lagunas terminaría por convertir al legislador en juez o incluso notario; *vid*. Luigi Lombardi Vallauri, *Corso di filosofia del diritto, cit.*, p. 39.

[74] En la literatura sobre IA y derecho es común encontrar referencias a Leibniz como primer autor en señalar la posibilidad de calcular cuestiones morales o jurídicas; *vid., v.gr.,* JC Smith, "Machine Intelligence and Legal Reasoning", *cit.*, p. 279; Simon Deakin; Christopher Markou, "From Rule of Law to Legal Singularity", en *Is Law Computable? Critical Perspectives on Law and Artificial Intelligence*, Deakin, Simon; Markou, Christopher (eds.), Oxford, Nueva York, Hart Publishing, 2020, pp. 9-14.

Respecto del problema de la automatización de la interpretación y la aplicación, quienes consideran que tal objetivo es inalcanzable suelen afirmar que sus oponentes parten de una concepción "formalista" del razonamiento jurídico, generalmente con cierto tono despectivo[75]. Evidentemente, este modo de caracterizar los intentos de automatizar la interpretación y la aplicación asocia a los defensores de la automatización del razonamiento jurídico al denominado "formalismo jurídico"[76]. Por el contrario, quienes defienden que el derecho no es computable suelen traer a colación concepciones del razonamiento jurídico de naturaleza hermenéutica o interpretativista[77].

La práctica jurídica algorítmica también guarda relación con el realismo jurídico[78]. La influencia de este movimiento intelectual se puede encontrar tanto en la relevancia que se otorga a la predicción de las decisiones de los tribunales como al trabajo dirigido a crear modelos computacionales del razonamiento basado en casos. Holmes, predecesor del realismo jurídico, es una referencia habitual en los trabajos relacionados con la práctica jurídica algorítmica[79]. En ocasiones, la concepción del derecho dominante entre los proponentes de la práctica jurídica algorítmica más compleja se resume en una afirmación de Holmes: "Las profecías de lo que de hecho harán los tribunales, y nada más pretencioso, es lo que entiendo como derecho"[80]. Si el derecho se reduce a predecir qué harán los tribunales, los algoritmos predictivos de tales decisiones se convierten en el paradigma de la inteligencia artificial jurídica y el modelado del razonamiento a través de casos una herramienta fundamental. Tampoco es de extrañar que, en el marco del derecho personalizado, las ideas de Holmes y el realismo jurídico también sean especialmente relevantes[81].

[75] *Vid., v.gr.,* Frank Pasquale, "A Rule of Persons, Not Machines", *cit.*, p. 36.

[76] Ernest J. Weinrib, "Legal formalism", en *A Companion to Philosophy of Law and Legal Theory*, Patterson, Dennis (ed.), Malden, Blackwell, 1996, pp. 327-338; Norberto Bobbio, *Iusnaturalismo y positivismo jurídico*, Madrid, Trotta, 2015, pp. 95-98.

[77] *Vid., v.gr.* Mireille Hildebrandt, "The Meaning and the Mining of Legal Texts", en *Understanding Digital Humanities*, Berry, David M. (ed.), Nueva York, Palgrave Macmillan, 2012, pp. 145-160.

[78] Brian Leiter, "Legal realism", en *A Companion to Philosophy of Law and Legal Theory*, Patterson, Dennis (ed.), Malden, Blackwell, 1996.

[79] *Vid., v.gr.,* Daniel Martin Katz, "Quantitative Legal Prediction", *cit.*, p. 936; Mireille Hildebrandt, "The Meaning and the Mining of Legal Texts", *cit.*, p. 149; Richard Susskind, *Tomorrow's Lawyers, cit.*, pp. 53-54.

[80] Oliver Wendell Holmes, "The Path of the Law", en *Harvard Law Review,* v. 110, n. 5, 1997, p. 994.

[81] Anthony J. Casey, Anthony Niblett, "The Death of Rules and Standards", *cit.*, p. 1422; John O. McGinnis, Steven Wasick, "Law's algorithm", *cit.*, p. 1023; Benjamin Alarie,

En definitiva, la práctica jurídica algorítmica se asienta en una serie de preconcepciones de la naturaleza del razonamiento jurídico que han sido objeto central de discusión en la filosofía del derecho. El éxito o el fracaso de la automatización del razonamiento jurídico está, por tanto, íntimamente relacionado con el éxito o fracaso de las diferentes concepciones acerca de la naturaleza del derecho y de la interpretación jurídica.

5.3. El derecho personalizado y la transformación del paradigma del *rule of law*

El derecho personalizado ataca ciertos conceptos fundamentales de la concepción contemporánea del derecho, en concreto la idea del *rule of law* o imperio de la ley y, en cierto modo, la noción dominante de norma jurídica[82]. Esta categoría de uso de la IA jurídica también afecta de un modo radical a cuestiones de filosofía del derecho ya tratadas. Por ejemplo, el derecho personalizado se plantea como solución radical al problema de la plenitud y coherencia del ordenamiento, o incluso como disolución del problema. En cierto modo, este problema deriva de las dificultades de identificar una norma general incontrovertible para cada caso particular. Según sus defensores, el derecho personalizado anularía tal problema al especificar una norma singular para cada caso particular[83]: el derecho es pleno porque cada caso particular tiene su propia norma singular y no existen dificultades para identificar tal norma. Empleando la metáfora de Lombardi Vallauri, se conseguiría la creación de un mapa tan grande como el territorio a representar.

El derecho personalizado también rompe con el ideal del *rule of law* o imperio del derecho, entendido como ideal de sometimiento de la conducta humana al gobierno de reglas[84]. Sin reglas generales, el gobierno a través de reglas desaparece. En su lugar se establece lo que, por oposición, se podría llamar *rule of algorithms* o imperio de los algoritmos, en el que en lugar de reglas generales la conducta humana se regularía a través de microdirectivas o normas singulares[85].

El paso del derecho basado en reglas al derecho basado en normas singulares afecta especialmente a la discusión acerca del concepto de derecho y a

"The Path of the Law", *cit.*, *passim*. La influencia es tal que Alarie emplea el título del trabajo más destacado de Holmes para su propio artículo.

[82] Se emplea aquí una concepción formal del *rule of law*; *vid.* Paul Craig, "Formal and Substantive Conceptions of the Rule of Law: An Analytical Framework", en *Public Law,* v. 3, 1997.

[83] O, al menos, identificar tal norma de forma automática y prácticamente instantánea.

[84] Lon L. Fuller, *The Morality of Law, cit.*, p. 46.

[85] Anthony J. Casey, Anthony Niblett, "The Death of Rules and Standards", *cit.*, p. 1404.

la teoría de la norma jurídica. Numerosos autores han señalado la conveniencia de cuestionarse si la generalidad de las normas jurídicas es una condición necesaria del concepto de derecho[86]. En línea con la discusión anterior, también se debate si una norma singular, entendida como una norma dirigida a un individuo singular en unas circunstancias singulares, puede considerarse una norma jurídica[87].

Otra cuestión de teoría de la norma jurídica se refiere a los elementos o estructura de la norma jurídica. Generalmente, se identifican dos elementos de las normas prescriptivas: supuesto de hecho y consecuencia jurídica[88]. El supuesto de hecho describe un comportamiento humano; la consecuencia jurídica, también llamada carácter, se refiere a la calificación del comportamiento como prohibido, obligatorio o permitido[89]. En el supuesto de hecho se puede distinguir entre el agente, el comportamiento o acto en sí, y una serie de condiciones[90]. En la norma "el profesorado está obligado a fijar tutorías de septiembre a julio", la consecuencia jurídica sería la obligación, el sujeto "el profesorado", el comportamiento en sí "fijar tutorías" y la condición "de septiembre a julio". Guastini habla de la "estructura condicional" de la norma jurídica porque, según él, las normas jurídicas suelen estar formadas por supuestos de hecho que incorporan condiciones de aplicación[91]: *si* estamos "entre septiembre y julio" *entonces* los profesores deben fijar tutorías.

La cuestión que se plantea es si las microdirectivas encajan en esta estructura y si las peculiaridades de la estructura de las microdirectivas requieren una reflexión sobre la naturaleza de las normas jurídicas. Las microdirec-

[86] *Vid,* por ejemplo, Georg Henrik von Wright, *Norma y acción. Una investigación lógica, cit*., p. 72; Frederick Schauer, *Las reglas en juego, cit*., pp. 67-69; Frederick Schauer, "Rules and the Rule of Law", en *Harvard Journal of Law & Public Policy,* v. 14, n. 3, 1991.

[87] Norberto Bobbio, *Contribución a la teoría del derecho*, Madrid, Debate, 1990, pp. 283-295; Georg Henrik von Wright, *Norma y acción. Una investigación lógica, cit*., pp. 71-72.

[88] José Juan Moreso, Josep María Vilajosana, *Introducción a la teoría del derecho,* Madrid, Marcial Pons, 2004, pp. pp. 72-73. Los autores distinguen entre normas prescriptivas y constitutivas y diferencian los tipos de consecuencias jurídicas de cada clase de norma. Tales distinciones no son relevantes para el objeto de este trabajo.

[89] También se ha considerado la consecuencia jurídica como equivalente a la sanción. Sin embargo, para el objeto de este trabajo, resulta más ilustrativa la estructura propuesta en el texto.

[90] Joseph Raz, *The Concept of a Legal System. An Introduction to the Theory of Legal System*, 2ª ed., Oxford, Clarendon Press, 1980, pp. 52-53. Para un análisis más detallado de los componentes de las normas prescriptivas, *vid*. Georg Henrik von Wright, *Norma y acción. Una investigación lógica, cit*., pp. 63-68.

[91] Riccardo Guastini, *Filosofía del Diritto positivo: lezioni*, Turín, G. Giappichelli Editore, 2017, p. 34.

tivas son normas que cuentan con una cantidad ingente de condiciones en su supuesto de hecho. En cierto modo, el elevado número de condiciones de las microdirectivas es el motivo por el que no tiene sentido concebirlas como reglas generales. Además, debido al funcionamiento de la tecnología comunicativa, estas condiciones no son transmitidas al sujeto jurídico, que solamente recibe información acerca del acto prohibido y la consecuencia jurídica o carácter. Incluso aunque fuese posible transmitir al sujeto jurídico todas las circunstancias del supuesto de hecho, algo cuestionable debido a la opacidad de los algoritmos de *deep learning*[92], la gran cantidad de circunstancias que determinan el acto prohibido le sería inaprehensible. Este modo de presentación de las microdirectivas choca con la idea de las normas jurídicas como normas con estructura hipotética, porque el sujeto jurídico no recibe las circunstancias incorporadas en el supuesto de hecho que le indicarían qué condiciones han de darse para que la conducta sea exigida. Tal modo de presentar las normas afecta a ciertos rasgos del derecho, como su carácter autoaplicativo[93], especialmente vinculado al *rule of law*.

En resumen, dado que el derecho personalizado opera a través de normas singulares o que incorporan un número ingente de condiciones, en lugar de a través de reglas generales, esta categoría exige reflexionar acerca de numerosas cuestiones centrales de la filosofía del derecho, fundamentalmente aquellas relacionadas con el *rule of law* y la teoría de la norma jurídica.

6. Conclusión

Este trabajo ha tratado de presentar una clasificación general de los usos de la IA jurídica que facilite la asociación de las cuestiones clásicas de la filosofía del derecho a aquellos usos que afecten especialmente a tales cuestiones. Con ello se ha pretendido reducir razonablemente el nivel de abstracción que en ocasiones existe al evaluar el uso de IA en el derecho.

Se han identificado tres categorías: el derecho formalizado, la práctica jurídica algorítmica y el derecho personalizado. El derecho formalizado consiste en la representación de un conjunto de normas de modo que se incorpore cierto conocimiento jurídico acerca de ellas con el fin de perfeccionar dicho conjunto de normas, por ejemplo, evitando ambigüedades o contradicciones. La práctica jurídica algorítmica se refiere a los usos de IA jurídica para la automatización de ciertas tareas típicas de los juristas, como la búsqueda

[92] Jenna Burrell, "How the Machine 'Thinks': Understanding Opacity in Machine Learning Algorithms", en *Big Data & Society*, v. 3, n. 1, 2016.

[93] Henry Melvin Hart, Albert Martin Sacks, *The Legal Process: Basic Problems in the Making and Application of Law*, Westbury, Nueva York, Foundation Press, 1994, pp. 120-121.

de información jurídica o todo o parte del razonamiento jurídico. El derecho personalizado se refiere a la sustitución del derecho basado en reglas por un sistema basado en normas singulares, adaptadas con detalle a las particularidades de un caso.

La clasificación de los usos de la IA jurídica en estas tres categorías facilita la asociación de cada uso concreto a una serie de discusiones clásicas de la filosofía del derecho. A modo de ejemplo, se ha señalado la relación entre el derecho formalizado y el problema de la plenitud y la coherencia del sistema jurídico, la relación entre la práctica jurídica algorítmica y la naturaleza del razonamiento jurídico, y la relación entre el derecho personalizado y la teoría de la norma jurídica. Parafraseando a Susskind, podría decirse que toda propuesta de uso de IA jurídica necesariamente realiza asunciones sobre la naturaleza del derecho y el razonamiento jurídico, y es importante conectar dichas asunciones con cada uso específico de la IA jurídica[94].

La clasificación ofrecida en este trabajo resulta de utilidad para futuras investigaciones en el ámbito de la IA y el derecho. Primero, puede servir como punto de inicio para una clasificación más detallada, en la que se incorporen subcategorías relacionadas con las tres categorías propuestas o incluso se explore la necesidad de contar con nuevas categorías generales. Segundo, puede ayudar a navegar entre los diferentes usos de IA jurídica sin perder el rumbo, identificando siempre el tipo de uso que se está proponiendo y asociándolo de este modo a unas asunciones o premisas específicas que lo relacionan con una serie de cuestiones específicas de la filosofía del derecho.

Bibliografía

Alarie, Benjamin, "The Path of the Law: Towards Legal Singularity", en *University of Toronto Law Journal,* v. 66, n. 4, 2016, pp. 443-455.

Aletras, Nikolaos; Tsarapatsanis, Dimitrios; Preoţiuc-Pietro, Daniel; Lampos, Vasileios, "Predicting Judicial Decisions of the European Court of Human Rights: A Natural Language Processing Perspective", en *PeerJ Computer Science,* v. 2, n. e93, 2016, pp. 1-19.

Alexy, Robert, *El concepto y la validez del derecho*, Barcelona, Gedisa, 1993.

Allen, Layman E., "Symbolic logic: A razor-edged tool for drafting and interpreting legal documents", en *Yale Law Journal,* Layman: v. 66, n. 6, 1957, pp. 833-879.

Ashley, Kevin D., *Artificial Intelligence and Legal Analytics: New Tools for Law Practice in the Digital Age*, Cambridge, Cambridge University Press, 2017.

Ben-Shahar, Omri; Porat, Ariel, *Personalized Law: Different Rules for Different People*, Nueva York, Oxford University Press, 2021.

[94] Richard Susskind, *Transforming the Law, cit*., p. 194.

Bench-Capon, Trevor, "Thirty years of Artificial Intelligence and Law: Editor's Introduction", en *Artificial Intelligence and Law,* 2022, v. 30, n. 4, pp. 475-479.

Bench-Capon, Trevor; Araszkiewicz, Michał; Ashley, Kevin; Atkinson, Katie; Bex, Floris; Borges, Filipe; Bourcier, Daniele, *et al.*, "A history of AI and Law in 50 papers: 25 years of the international conference on AI and Law", en *Artificial Intelligence and Law,* v. 20, n. 3, 2012, pp. 215-319.

Bench-Capon, Trevor; Prakken, Henry, "Introducing the logic and law corner", en *Journal of logic and computation,* v. 18, n. 1, 2008, pp. 1-12.

Bibel, L. Wolfgang, "AI and the conquest of complexity in law", en *Artificial Intelligence and Law,* v. 12, n. 3, 2004, pp. 159-180.

Bobbio, Norberto, *Contribución a la teoría del derecho*, Madrid, Debate, 1990.

—, *El positivismo jurídico. Lecciones de Filosofía del Derecho reunidas por el doctor Nello Morra*, Madrid, Debate, 1993.

—, *Iusnaturalismo y positivismo jurídico*, Madrid, Trotta, 2015.

Bourcier, Daniéle; Casanovas, Pompeu, *Inteligencia artificial y derecho*, Barcelona, UOC, 2003.

Branting, L. Karl, "Data-centric and logic-based models for automated legal problem solving", en *Artificial Intelligence and Law,* v. 25, n. 1, 2017, pp. 5-27.

Burrell, Jenna, "How the Machine 'Thinks': Understanding Opacity in Machine Learning Algorithms", en *Big Data & Society,* v. 3, n. 1, 2016, pp. 1-12.

Busch, Christoph; De Franceschi, Alberto, *Algorithmic Regulation and Personalized Law: A Handbook*, Londres, Bloomsbury, 2021.

Casanovas, Pompeu, "Inteligencia Artificial y Derecho: a vuelapluma", en *Teoría y Derecho. Revista de Pensamiento Jurídico,* v. 7, 2010, pp. 203-221.

Casey, Anthony J.; Niblett, Anthony, "Self-Driving Laws", en *University of Toronto Law Journal,* v. 66, n. 4, 2016, pp. 429-442.

—, "The Death of Rules and Standards", en *Indiana Law Journal,* v. 92, n. 4, 2017, pp. 1401-1447.

—, "Framework for the New Personalization of Law", en *The University of Chicago Law Review,* v. 86, n. 2, 2019, pp. 333-358.

Chui, Michael; Hall, Bryce; Singla, Alex; Sukharevsky, Alex. "The state of AI in 2021." McKinsey Analytics, 2021.

Craig, Paul, "Formal and Substantive Conceptions of the Rule of Law: An Analytical Framework", en *Public Law,* v. 3, 1997, pp. 467-487.

Deakin, Simon; Markou, Christopher, "From Rule of Law to Legal Singularity", en *Is Law Computable? Critical Perspectives on Law and Artificial Intelligence*, Oxford, Nueva York, Hart Publishing, 2020, pp. 1-29.

Dworkin, Ronald, *Los derechos en serio*, Barcelona, Ariel, 1984.

Dworkin, Ronald, *Law's Empire*, Cambridge, Harvard University Press, 1986.

Fuller, Lon L., *The Morality of Law*, 2ª ed., New Haven, Yale University Press, 1969.

Guastini, Riccardo, *Filosofía del Diritto positivo: lezioni*, Turín, G. Giappichelli Editore, 2017.

Hart, Henry Melvin; Sacks, Albert Martin, *The Legal Process: Basic Problems in the Making and Application of Law*, Westbury, Nueva York, Foundation Press, 1994.

HILDEBRANDT, MIREILLE, "The Meaning and the Mining of Legal Texts", en *Understanding Digital Humanities*, David M. Berry (ed.), Nueva York, Palgrave Macmillan, 2012, pp. 145-160.

—, *Smart Technologies and the End(s) of Law: Novel Entanglements of Law and Technology*, Cheltenham, Northampton, Edward Elgar Publishing, 2015.

—, "Algorithmic Regulation and the Rule of Law", en *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences,* v. 376, n. 2128, 2018, pp. 1-11.

—, "Law as computation in the era of artificial legal intelligence: Speaking law to the power of statistics", en *University of Toronto Law Journal,* v. 68, n. 1, 2018, pp. 12-35.

—, "Code-Driven Law: Freezing the Future and Scaling the Past", en *Is Law Computable?: Critical Perspectives on Law and Artificial Intelligence*, Christopher Markou and Simon Deakin (ed.), Oxford, Hart Publishing, 2020, pp. 67-83.

HOLMES, OLIVER WENDELL, "The Path of the Law", en *Harvard Law Review,* v. 110, n. 5, 1997, pp. 991-1009.

KATZ, DANIEL MARTIN, "Quantitative Legal Prediction – Or – How I Learned to Stop Worrying and Start Preparing for the Data-Driven Future of the Legal Services Industry", en *Emory Law Journal,* v. 62, n. 4, 2013, pp. 909-966.

KRAMER, MATTHEW, *Objectivity and the Rule of Law*, Cambridge, Cambridge University Press, 2007.

LAWSKY, SARAH B., "Formalizing the code", en *Tax Law Review,* v. 70, n. 2, 2017, pp. 377-408.

LEHR, DAVID; OHM, PAUL, "Playing with the data: what legal scholars should learn about machine learning", en *University of California, Davis Law Review,* v. 51, 2017, pp. 653-717.

LEITER, BRIAN, "Legal realism", en *A Companion to Philosophy of Law and Legal Theory*, Dennis Patterson (ed.), Malden, Blackwell, 1996, pp. 261-279.

LEVI, EDWARD H., *An introduction to legal reasoning,* Chicago, Londres, University of Chicago Press, 2013.

LOMBARDI VALLAURI, LUIGI, *Corso di filosofia del diritto*, Milán, CEDAM, 1981.

—, "Verso un Sistema Esperto Giuridico Integrale", en *Persona y Derecho,* n. 31, 1994, pp. 157-182.

MACCORMICK, NEIL, *Legal Reasoning and Legal Theory*, Oxford, Clarendon Press, 1994.

MARMOR, ANDREI, *Interpretation and Legal Theory*, 2ª ed., Oxford, Portland, Hart Publishing, 2005.

MCGINNIS, JOHN O.; WASICK, STEVEN, "Law's algorithm", en *Florida Law Review,* v. 66, 2014, pp. 991-1050.

MORESO, JOSÉ JUAN; VILAJOSANA, JOSEP MARÍA, *Introducción a la teoría del derecho*, Madrid, Marcial Pons, 2004.

PASQUALE, FRANK, "A Rule of Persons, Not Machines: The Limits of Legal Automation", en *George Washington Law Review,* v. 87, n. 1, 2019, pp. 1-55.

PASQUALE, FRANK; CASHWELL, GLYN, "Prediction, persuasion, and the jurisprudence of behaviourism", en *University of Toronto Law Journal,* v. 68, n. Supplement 1, 2018, pp. 63-81.

PLATÓN, "Político", en *Diálogos*, traducción de María Rico Gómez, Madrid, Centro de Estudios Políticos y Constitucionales, 2007.

POSNER, RICHARD A., *How Judges Think*, Cambridge, Harvard University Press, 2008.

Raz, Joseph, *The Concept of a Legal System. An Introduction to the Theory of Legal System*, 2ª ed., Oxford, Clarendon Press, 1980.

Russell, Stuart J.; Norvig, Peter, *Artificial Intelligence: A Modern Approach*, 4ª ed., Harlow, Pearson Education, 2022.

Schauer, Frederick, "Rules and the Rule of Law", en *Harvard Journal of Law & Public Policy*, v. 14, n. 3, 1991, pp. 645-694.

—, *Las reglas en juego. Un examen filosófico de la toma de decisiones basada en reglas en el derecho y en la vuda cotidiana*, Madrid, Marcial Pons, 2004.

Serna, Pedro, "Teoría del derecho y filosofía del derecho", en *Persona y Derecho,* v. 32, 1995, pp. 267-298.

Smith, JC, "Machine Intelligence and Legal Reasoning", en *Chicago-Kent Law Review,* v. 73, n. 1, 1997, pp. 277-347.

Solar Cayón, José Ignacio, *La inteligencia artificial jurídica. El impacto de la innovación tecnológica en la práctica del Derecho y el mercado de servicios jurídicos*, Cizur Menor, Thomson Reuters, Aranzadi, 2019.

Stanley, Davis, *Future perfect*, Reading, Addison-Wesley, 1987.

Surden, Harry, "Machine learning and law", en *Washington Law Review,* v. 89, n. 1, 2014, pp. 87-115.

—, "Artificial intelligence and law: An overview", en *Georgia State University Law Review,* v. 35, 2019, pp. 19-22.

—, "Machine learning and law: an overview", en *Research Handbook on Big Data Law*, Roland Vogl (ed.), Cheltenham, Northampton, Edward Elgar Publishing, 2021.

Susskind, Richard, *Transforming the Law: Essays on Technology, Justice and the Legal Marketplace*, Oxford, Oxford University Press, 2003.

—, *Tomorrow's Lawyers: An Introduction to Your Future*, 2ª ed., Oxford, Oxford University Press, 2017.

Villata, Serena; Araszkiewicz, Michal; Ashley, Kevin; Bench-Capon, Trevor; Branting, L. Karl; Conrad, Jack G.; Wyner, Adam, "Thirty years of artificial intelligence and law: the third decade", en *Artificial Intelligence and Law,* 2022, v. 30, n. 4, pp. 561-591.

Volokh, Eugene, "Chief Justice Robots", en *Duke Law Journal,* v. 68, 2018, pp. 1135-1192.

von Wright, Georg Henrik, *Norma y acción. Una investigación lógica*, Santiago de Chile, Ediciones Olejnik, 2019.

Weinrib, Ernest J., "Legal formalism", en *A Companion to Philosophy of Law and Legal Theory*, Dennis Patterson (ed.), Malden, Blackwell, 1996, pp. 327-338.

Witten, Ian H.; Frank, Eibe; Hall, Mark A.; Pal, Christopher, J., *Data Mining. Practical Machine Learning Tools and Techniques*, 4ª ed., Cambridge, Morgan Kaufmann, 2016.

Yeung, Karen, "Five Fears about Mass Predictive Personalisation in an Age of Surveillance Capitalism", en *International Data Privacy Law,* v. 8, n. 3, 2018, pp. 258-269.

Yoon, Albert H., "The Post-Modern Lawyer: Technology and the Democratization of Legal Representation", en *University of Toronto Law Journal,* v. 66, n. 4, 2016, pp. 456-471.

Zalnieriute, Monika; Moses, Lyria Bennett; Williams, George, "The rule of law and automation of government decision-making", en *The Modern Law Review,* v. 82, n. 3, 2019, pp. 425-455.

IA, *Profiling* e direitos de personalidade*

AI, Profiling and personality rights

Maria Raquel Guimarães**

resumo: A grande capacidade computacional hoje alcançada, aliada à utilização de inteligência artificial, permite o tratamento de um enorme volume de dados muito diversificados, a uma velocidade sem precedentes, desafiando os instrumentos tradicionais de protecção da pessoa. A cláusula geral de protecção da personalidade prevista no Código Civil português, bem como os direitos especiais de personalidade que dela emanam, coadjuvados pelo Regulamento geral de protecção de dados e pela Carta portuguesa de direitos humanos na era digital, são postos à prova pela intensificação de actividades de *profiling* que analisam, avaliam e catalogam os nossos comportamentos *online*, de uma forma pouco transparente ou até mesmo subterrânea. À vulgarização destes tratamentos de dados soma-se o seu potencial cruzamento com dados recolhidos, entre outros dispositivos, por câmaras de segurança, *drones*, ou dispositivos inteligentes, aproximando perigosamente o nosso quotidiano do que pensávamos serem argumentos improváveis de ficção científica ou realidades política e geograficamente distantes.

É sobre estas questões que procuramos reflectir neste texto, olhando criticamente para algumas iniciativas legislativas recentes.

palavras-chave: IA; *Profiling*; elaboração de perfis; direito ao carácter; protecção de dados.

* Este texto encontra-se publicado, com o mesmo título, em *Inteligência artificial e robótica. Desafios para o direito do século XXI,* (coord. Eva Sónia Moreira, Pedro Freitas), Coimbra, Gestlegal, 2022, pp. 187-211, introduzindo-se agora actualizações pontuais à luz das mais recentes novidades legislativas.

** Professora Associada da Faculdade de Direito da Universidade do Porto. Investigadora do CIJ.

ABSTRACT: The great computational capacity achieved in the last years, combined with the use of artificial intelligence, allows the processing of an extraordinary volume of diversified data, at an unprecedented speed, challenging the traditional instruments for the protection of the person. The "general personality right", a foundational right where all the special rights of personality may be involved, based on Article 70 of the Portuguese Civil Code, as well as the General Data Protection Regulation and the Portuguese Charter of Human Rights in the digital era, are put to the test by the intensification of profiling activities that analyse, evaluate, and catalogue our online behaviour, in a non-transparent manner. These data processing activities combined with data collected, among other devices, by security cameras, drones, or smart devices, may bring our daily lives dangerously closer to what we thought were improbable science fiction scenarios or politically and geographically distant realities.

This aim of this text is to reflect critically on these issues, taking into account some recent legislative initiatives.

KEYWORDS: AI; Profiling; Right to one's character; Data protection.

1. IA e *Profiling*: da ficção à realidade

A utilização de inteligência artificial (IA) na definição de perfis[1] e na monitorização dos comportamentos dos indivíduos há já alguns anos que extrapolou

[1] Utilizamos aqui a expressão "definição de perfis" no sentido do artigo 4º, nº 4, do Regulamento Geral de Protecção de Dados (RGPD), enquanto "qualquer forma de tratamento automatizado de dados pessoais que consista em utilizar esses dados pessoais para avaliar certos aspetos pessoais de uma pessoa singular, nomeadamente para analisar ou prever aspetos relacionados com o seu desempenho profissional, a sua situação económica, saúde, preferências pessoais, interesses, fiabilidade, comportamento, localização ou deslocações" [Regulamento (UE) 2016/679 do Parlamento Europeu e do Conselho, de 27 de abril de 2016, relativo à proteção das pessoas singulares no que diz respeito ao tratamento de dados pessoais e à livre circulação desses dados e que revoga a Diretiva 95/46/CE (Regulamento Geral sobre a Proteção de Dados), *in JO* L 119 de 4.5.2016, pp. 1-88]. Para mais desenvolvimentos sobre esta noção, v. MAFALDA MIRANDA BARBOSA, *Inteligência artificial. Entre a utopia e a distopia, alguns problemas jurídicos*, Coimbra, Gestlegal, 2021, p. 137 ss. Para um conceito de inteligência artificial e para uma distinção entre inteligência artificial

o mundo da ciência-ficção dos argumentos de séries televisivas como a *Black Mirror*[2] e integrou a realidade diária de um número extraordinário de pessoas, desde logo aquelas abrangidas pelo sistema de "crédito social" na China[3]. Este sistema de "crédito social" – que, na verdade, se compõe, essencialmente, de três sistemas distintos, não consolidados[4] – procura classificar os visados como "bons" ou "maus cidadãos" cruzando enormes quantidades de dados, nomeadamente bancários e judiciais, dados de redes sociais, de aplicações móveis e plataformas digitais, como a *Alibaba*, com dados obtidos através do sistema de vigilância em massa *Skynet*, baseado no que se estima estarem a caminho de ser mais de mil milhões de câmaras[5]. Os dados são tratados em várias plataformas que cruzam a informação recolhida[6], cruzamento facilitado pela identificação dos cidadãos através de um código de crédito social único, um sistema de identificador também único de acesso às plataformas digitais e de um número perpétuo de identificação nacional, ao que acresce a utilização de inteligência artificial no reconhecimento facial[7].

débil e forte, v. José Ignacio Solar Cayón, *La inteligencia artificial jurídica, El impacto de la innovación tecnológica en la práctica del derecho y el mercado de servicios jurídicos*, Pamplona, Aranzadi, 2019, pp. 21 ss., 23-25.

[2] Cfr. *Black Mirror*, Episódio "Nosedive" ("Em queda livre"), temporada 3, Netflix, 2016, em que a personagem principal, Lacie (Bryce Dallas Howard), procura desesperadamente alcançar a pontuação de 4.5, mediante a votação daqueles com quem interage no dia a dia, necessária para comprar uma casa num condomínio exclusivo.

[3] Sobre os paralelismos e as diferenças da sociedade retratada no referido episódio da série *Black Mirror* e o sistema de "crédito social" chinês, v. Daithí Mac Síthigh; Mathias Siems, "The chinese social credit system: a model for other countries?", European University Institute, Department of Law, Working Paper LAW 2019/01, disponível em https://cadmus.eui.eu/bitstream/handle/1814/60424/LAW_2019_01.pdf [consultado em 12/06/2022], p. 29, e Lizzy Rettinger, "The Human Rights Implications of China's Social Credit System", *Journal of High Technology Law*, 21, no. 1, 2021, pp. 1-2, bem como a bibliografia aí referida.

[4] O sistema é composto de listas negras nacionais, listas de crédito social em cidades-piloto e listas de crédito social de instituições financeiras. Assim, Daithí Mac Síthigh; Mathias Siems, "The chinese social credit system: a model for other countries?", cit., pp. 12, 16, Liav Orgad; Wessel Reijers, "How to Make the Perfect Citizen? Lessons from China's Social Credit System", *Vanderbilt Journal of Transnational Law*, 54, no. 5, November 2021, pp. 1092-1093, e Lizzy Rettinger, "The Human Rights Implications of China's Social Credit System", cit., pp. 6, 11-12.

[5] John Frank Weaver, "Everything Is Not *Terminator*. Is China's Social Credit System the Future?", *RAIL: The Journal of Robotics, Artificial Intelligence & Law*, 2, no. 6 (November-December), 2019, p. 445. Lizzy Rettinger, "The Human Rights Implications of China's Social Credit System", cit., p. 11, refere o número de 626 milhões de câmaras, reportado ao ano de 2020.

[6] John Frank Weaver, "Everything Is Not *Terminator*. Is China's Social Credit System the Future?", cit., p. 445.

[7] Lizzy Rettinger, "The Human Rights Implications of China's Social Credit System", cit., pp. 7-8, 10-11.

O sistema de "crédito social", implementado a partir de 2014 em várias cidades-piloto, não estará ainda consolidado a nível nacional mas permite já graduar uma parte importante da população com base na pontuação obtida, bem como criar "listas vermelhas" e "listas negras" de cidadãos "bem" e "mal comportados", em função não só de dívidas fiscais e condenações judiciais mas também de infracções de trânsito ou outras "violações", como atravessar as ruas fora das passadeiras, devolver com atraso um livro numa biblioteca ou comprar álcool e *fastfood* num supermercado[8]. A pontuação obtida dita o acesso aos transportes, a férias e viagens ao estrangeiro, a escolas de maior qualidade, crédito bancário, seguros, subsídios, à compra de uma casa ou de um carro, ou até a possibilidade de ter um animal de estimação, sendo as listas negras publicadas *online*[9] e, em alguns casos, em *placards* electrónicos instalados em espaços públicos, com a identificação dos visados, ou sendo até acessíveis através de *apps* que denunciam a aproximação de alguém assim classificado[10].

Não obstante diferenças significativas, nos países democráticos ocidentais a avaliação de indivíduos e a atribuição de pontuação em função da sua actuação social e da interacção com os demais não é uma realidade longínqua. O sistema de *ratings* é utilizado em plataformas da chamada economia colaborativa – como a *Uber*, a *Airbnb*, a *eBay*, a *Couchsurfing*, entre muitas outras – com vista a superar os défices de conhecimento e de confiança entre os diferentes utilizadores, muitas vezes em associação com redes sociais, como o *Facebook*[11].

[8] Sobre as origens e o desenvolvimento deste sistema, e a sua implementação em cidades-piloto, cfr. Daithí Mac Síthigh; Mathias Siems, "The chinese social credit system: a model for other countries?", cit., pp. 12-14, e Lizzy Rettinger, "The Human Rights Implications of China's Social Credit System", cit., pp. 3-6, 11-12. V., também, Liav Orgad; Wessel Reijers, "How to Make the Perfect Citizen? Lessons from China's Social Credit System, cit., pp. 1088-1089.

[9] Cfr. o *site* https://www.creditchina.gov.cn/ [consultado em 12/06/2022], indicado por Daithí Mac Síthigh; Mathias Siems, "The chinese social credit system: a model for other countries?", cit., p. 13. Sobre as consequências que decorrem da pontuação obtida, v. Lizzy Rettinger, "The Human Rights I mplications of China's Social Credit System", cit., pp. 5-6, 9-10, 12. V., também, Daithí Mac Síthigh; Mathias Siems, "The chinese social credit system: a model for other countries?", cit., pp. 1, 13.

[10] Assim, Liav Orgad; Wessel Reijers, "How to Make the Perfect Citizen? Lessons from China's Social Credit System", cit., p. 1094.

[11] Estes sistemas de *ratings* são chamados à colação neste contexto por Daithí Mac Síthigh; Mathias Siems, "The chinese social credit system: a model for other countries?", cit., pp. 5-7, Liav Orgad; Wessel Reijers, "How to Make the Perfect Citizen? Lessons from China's Social Credit System", cit., pp. 1101-1102, e Lizzy Rettinger, "The Human Rights Implications of China's Social Credit System", cit., pp. 28-29. Sobre a função dos *ratings* na economia colaborativa, v. Diogo Rodrigues da Silva, "Consequences of ratings/reviews on sharing economy platforms",

Os utilizadores destas plataformas – prestadores de serviços profissionais ou não profissionais e, em muitos casos, os beneficiários dos serviços, consumidores ou não consumidores, num sistema de avaliação recíproca – são pontuados com base na sua simpatia, disponibilidade, prestabilidade, educação, *hobbies*, maior ou menor identificação com o "espírito" da plataforma, e não só tendo em conta o cumprimento cabal das obrigações contratuais a que estão adstritos.

Ainda assim, existem diferenças determinantes no que respeita às finalidades destas avaliações face ao sistema de crédito social, uma vez que as primeiras são sempre avaliações parcelares, no contexto limitado de um serviço específico prestado ou de um outro contrato celebrado entre os utilizadores da plataforma, e não visam classificar a pessoa enquanto tal, a pessoa íntima ou o cidadão, de acordo com um modelo de pessoa ideal ou perfeita[12].

Por outro lado, verificam-se também diferenças decisivas no que respeita às entidades que levam a cabo as avaliações e aos regimes políticos que servem de pano de fundo aos diferentes sistemas[13]. Na China, os sistemas de avaliações são organizados pelo próprio Estado, socorrendo-se também das bases de informações de companhias comerciais, Estado que depois impõe autoritariamente sanções com repercussão nos direitos dos cidadãos, sem que estejam devidamente assegurados mecanismos que permitam o contraditório e nem sequer sendo clara a distinção entre comportamentos antijurídicos, imorais ou antissociais[14]. Os limites que a lei impõe aos particulares e ao Estado no tratamento de dados pessoais nos países da União Europeia e, de uma forma geral, o primado da lei nos países democráticos permitem perspectivar de forma diferente as definições de perfis, relativizando os perigos associados de abusos, repressão e discriminação.

Ainda assim, os sistemas de avaliações no mundo ocidental não são exclusivos do sector privado. Também o sector público se socorre destes métodos ainda que para fins específicos e determinados. Entre nós, o Banco de Portugal

in M. Regina Redinha / M. Raquel Guimarães / F. Liberal Fernandes (coords.), *Sharing Economy: Legal Problems of a Permutations and Combinations Society*, Newcastle upon Tyne, Cambridge Scholars, 2019, pp. 382-384.

[12] Assinalam estas diferenças Liav Orgad; Wessel Reijers, "How to Make the Perfect Citizen? Lessons from China's Social Credit System", cit., p. 1104.

[13] *Idem*, p. 1105.

[14] No sentido de que o sistema de crédito social chinês é, essencialmente, um "sistema de virtude social", que promove a moral pessoal e pública, legalmente reconhecida ou não, v. Liav Orgad; Wessel Reijers, "How to Make the Perfect Citizen? Lessons from China's Social Credit System, cit., p. 1107 ss.

gere a Central de Responsabilidades de Crédito, nos termos do Decreto-Lei nº 204/2008, de 14 de Outubro, que agrega informações de crédito de pessoas singulares e colectivas com vista à sua partilha entre as entidades participantes e a permitir a avaliação dos riscos dos devedores. Este tratamento de dados permite a organização de "listas negras" de devedores que são consideradas, até por imperativo legal, na decisão de concessão de crédito[15].

Estas avaliações de responsabilidades de crédito são comuns em estados liberais e são as mesmas que estiveram na base do sistema de crédito social chinês[16]. A questão é que, como salientam Daithí Mac Síthigh e Mathias Siems no seu texto "The chinese social credit system: a model for other countries?", estamos hoje longe de poder assegurar que "*what happens in China, stays in China*"[17]...

2. IA: novidades legislativas

A existência de modelos de crédito social como os adoptados na China, ainda que possa parecer uma realidade distante e dificilmente concretizável entre nós, permite-nos perceber as mudanças radicais que a capacidade computacional tem introduzido no mundo e que afectam a vida em sociedade e a forma

[15] O Decreto-Lei nº 133/2009, de 2 de Junho, e o Decreto-Lei nº 74-A/2017, de 23 de Junho, artigos 10º e 16º, respectivamente, impõem deveres de avaliação da solvabilidade dos consumidores na celebração de contratos de crédito ao consumo e crédito imobiliário, e prevêem, para o efeito, a consulta das bases de dados de responsabilidades de crédito. Por sua vez, a Proposta de Diretiva do Parlamento Europeu e do Conselho relativa aos créditos aos consumidores [COM(2021) 347 final, Bruxelas, 30.06.2021], logo na p. 1 da sua "exposição de motivos", chama a atenção para o facto de a digitalização ter introduzido "novas formas de divulgar informações digitalmente e de avaliar a solvabilidade dos consumidores através de sistemas automatizados de decisão e de dados não tradicionais", mostrando-se preocupada com a utilização de dados "não tradicionais" para esta avaliação. Assim, propõe no seu "considerando" 47 que "[o]s dados pessoais, tais como os dados pessoais existentes nas plataformas de redes sociais ou dados de saúde, incluindo dados sobre o cancro, não dev[am] ser utilizados ao efetuar uma avaliação da solvabilidade". Cfr., também, o artigo 18º da Proposta.

[16] Assim, Daithí Mac Síthigh; Mathias Siems, "The chinese social credit system: a model for other countries?", cit., p. 20, John Frank Weaver, "Everything Is Not *Terminator*. Is China's Social Credit System the Future?", cit., p. 446, Liav Orgad; Wessel Reijers, "How to Make the Perfect Citizen? Lessons from China's Social Credit System, cit., p. 1106, e Lizzy Rettinger, "The Human Rights Implications of China's Social Credit System", cit., pp. 3, 27-28. Para um breve enquadramento histórico dos sistemas de avaliação de crédito, v., ainda, Daithí Mac Síthigh; Mathias Siems, "The chinese social credit system: a model for other countries?", cit., pp. 2-5. V., também, Alba Soriano Arnanz, *Data protection for the prevention of algorithmic discrimination*, Cizur Menor, Aranzadi, 2021, pp. 47-49.

[17] Daithí Mac Síthigh; Mathias Siems, "The chinese social credit system: a model for other countries?", cit., p. 1.

como nos relacionamos com os outros[18]. E podemos perguntar-nos mesmo, com John Frank Weaver, se a utilização de inteligência artificial não conduzirá inevitavelmente a este tipo de resultados[19], sobretudo numa época em que as pessoas estão mais receptivas à monitorização, ao rastreamento de movimentos e até à exposição dos "indesejáveis" – pense-se nas aplicações móveis destinadas à identificação de contactos de risco no contexto situação epidemiológica provocada pelo coronavírus SARS-CoV-2 e pela doença Covid-19[20].

Todas as alterações geradas pelo aumento da capacidade computacional e pela sua aplicação a diferentes sectores da vida social convocaram a intervenção do legislador europeu e do legislador nacional. O ano de 2021 foi particularmente intenso no que toca a iniciativas legislativas motivadas pelo advento da utilização de meios electrónicos e, em particular, da IA. No âmbito nacional há que referir a publicação do Decreto-Lei nº 12/2021, de 9 de Fevereiro, que assegura a execução na ordem jurídica interna do Regulamento (UE) 910/2014, relativo à identificação eletrónica e aos serviços de confiança para as transações eletrónicas no mercado interno[21]. No seguimento deste diploma, a 11 de Março, foi publicado o Despacho nº 2705/2021, do Gabinete Nacional de Segurança, que veio definir requisitos e instruções relativamente à possibilidade de os prestadores qualificados de serviços de confiança adotarem formas de identificação de pessoas físicas através de procedimentos de identificação à distância com recurso a sistemas biométricos automáticos de reconhecimento facial. E, sobretudo, em Maio foi aprovada a Carta portuguesa de direitos humanos na era digital, pela Lei nº 27/2021, de 17 de Maio, prevendo que "[a] utilização da inteligência artificial deve ser orientada pelo respeito dos direitos fundamentais, garantindo um justo equilíbrio entre os princípios

[18] Também Liav Orgad e Wessel Reijers, "How to Make the Perfect Citizen? Lessons from China's Social Credit System", cit., p. 1090, constatam que o sistema de créditos sociais chinês é representativo de quão rapidamente o mundo está a mudar.

[19] John Frank Weaver, "Everything Is Not *Terminator*. Is China's Social Credit System the Future?", cit., p. 445.

[20] Sobre estas aplicações, v. Giorgio Resta, "La protezione dei dati personali nel diritto dell'emergenza Covid-19", *Giustizia civile.com*, 5, 2020, disponível em http://giustiziacivile.com/pdfpage/2262 [consultado em 12/06/2022]; Laura Bradford, Mateo Aboy e Kathleen Liddell, "COVID-19 Contact Tracing Apps: A Stress Test for Privacy, the GDPR and Data Protection Regimes", *Journal of Law and the Biosciences*, volume 7, 1, January-June 2020, disponível em https://doi.org/10.1093/jlb/lsaa034 [consultado em 12/06/2022]; e Hyunghoon Cho, Daphne Ippolito e Yun William Yu, "Contact Tracing Mobile Apps for COVID-19: Privacy Considerations and Related Trade-offs", *Cornell University* arXiv:2003.11511v2 [cs.CR], 2020, disponível em https://arxiv.org/abs/2003.11511v2 [consultado em 12/06/2022].

[21] *In JO* L 257, pp. 73-114.

da explicabilidade, da segurança, da transparência e da responsabilidade, que atenda às circunstâncias de cada caso concreto e estabeleça processos destinados a evitar quaisquer preconceitos e formas de discriminação"[22].

Entretanto, em 21 de Abril, o legislador europeu publicou uma Proposta de Regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial (Regulamento Inteligência Artificial)[23]. E posteriormente, já a 7 de Setembro, foi publicada entre nós a Proposta de Lei nº 111/XIV/2ª, da iniciativa do Governo, que veio a culminar na Lei nº 95/2021, de 29 de Dezembro, que regula a utilização e o acesso pelas forças e serviços de segurança a sistemas de videovigilância para captação, gravação e tratamento de imagem e som[24].

Esta última Proposta nacional, que acabou por ser bastante alterada e amputada, previa o alargamento da utilização de sistemas de videovigilância sem a necessária definição de "um regime legal bem densificado que prev[isse] as condições da utilização de cada tipo de meio utilizado para captar e gravar imagens e som, e as respetivas salvaguardas, tendo em conta os específicos riscos ou impactos que cada um deles implica sobre os direitos fundamentais dos cidadãos"[25]. Como teve necessidade de lembrar a Comissão Nacional de Protecção de Dados (CNPD), na sua apreciação da Proposta,

> "num Estado de Direito democrático não é admissível a mera previsão genérica de utilização de sistemas de videovigilância, em especial com recurso a tecnologias que potenciam os seus efeitos, sem a especificação de condições, limites e critérios necessários a garantir a sua idoneidade para prossecução de finalidades de interesse público, mas também imprescindíveis para assegurar que a afetação dos direitos fundamentais ocorra na medida do estritamente indispensável e sem excesso"[26].

Sobre esta Proposta e as críticas de que foi alvo iremos ainda fazer alguns comentários *infra*. No que toca à Proposta de um "Regulamento Inteligência Artificial", saliente-se a definição, prevista no nº 1 do artigo 3º, do que é considerado, para o legislador europeu, um sistema de inteligência artificial,

[22] Cfr. o artigo 9º, nº 1, da Carta portuguesa de direitos humanos na era digital.

[23] COM(2021) 206 final, Bruxelas, 21.04.2021.

[24] Este processo legislativo pode ser conferido em https://www.parlamento.pt/ActividadeParlamentar/Paginas/DetalheIniciativa.aspx?BID=121083 [consultado em 31/05/2022].

[25] Neste sentido, cfr. o Parecer da Comissão Nacional de Protecção de Dados (CNPD) 2021/143, de 4 de Novembro de 2021, sobre a Proposta de Lei nº 111/XIV/2.a., p. 1v, nº 5.

[26] *Idem*, *ibidem*, nº 7.

concretamente um programa informático "capaz de, tendo em vista um determinado conjunto de objetivos definidos por seres humanos, criar resultados, tais como conteúdos previsões, recomendações ou decisões que influenciam os ambientes com os quais interage". Note-se ainda a definição de dados biométricos, consentânea com aquela prevista no RGPD, enquanto categoria de dados pessoais que resultam do tratamento de características físicas, fisiológicas mas também comportamentais de um indivíduo que o permitem identificar, nomeadamente através do reconhecimento facial[27], eventualmente à distância, mediante a comparação dos dados biométricos de uma pessoa com os dados biométricos contidos numa base de dados[28], "em tempo real" ou "em diferido", consoante a recolha de dados biométricos, a comparação e a identificação ocorrem sem ou com "atraso significativo"[29]. E, sobretudo, a Proposta veio categorizar os sistemas e aplicações de IA consoante os riscos que envolvem, prevendo quatro categorias de riscos, concretamente risco inaceitável, elevado, limitado e mínimo, e, portanto, também quatro categorias de sistemas e aplicações de inteligência artificial[30].

As práticas de IA de risco inaceitável serão, de acordo com a Proposta, proibidas e abrangem sistemas ou aplicações de inteligência artificial que manipulam o comportamento humano para iludir o livre arbítrio dos utilizadores, sistemas que permitem uma "classificação social" por autoridades públicas e, regra geral, a utilização de sistemas de identificação biométrica à distância em "tempo real" em espaços acessíveis ao público para efeitos de manutenção da ordem pública[31].

São consideradas práticas de IA de risco elevado todos os sistemas de identificação biométrica à distância, impondo-lhes a Proposta o cumprimento de requisitos rigorosos antes de poderem ser colocados no mercado. Nomeadamente, têm que assegurar meios adequados de avaliação dos riscos, a elevada qualidade dos dados a tratar, o registo da actividade, de modo a assegurar a sua rastreabilidade, a transparência e prestação de informações aos utilizadores e uma supervisão humana adequada[32].

[27] Cfr. a Proposta do "Regulamento Inteligência Artificial", artigo 3º, nº 33, e o artigo 4º, nº 14, do RGPD.

[28] Cfr. a Proposta do "Regulamento Inteligência Artificial", artigo 3º, nº 36.

[29] Cfr. a Proposta do "Regulamento Inteligência Artificial", artigo 3º, nºs 37 e 38.

[30] Cfr. a Proposta do "Regulamento Inteligência Artificial", p. 14. Sobre o tema, v., entre nós, Mafalda Miranda Barbosa, *Inteligência artificial...*, cit., pp. 150-152.

[31] Cfr. o artigo 5º da Proposta do "Regulamento Inteligência Artificial".

[32] Cfr. o "considerando" 27 ss. e os artigos 6º-15º da Proposta do "Regulamento Inteligência Artificial".

Os sistemas de IA concebidos para interagir com pessoas, como os robôs de conversação, ou para criar conteúdos, comportam, de acordo com o mesmo documento, um risco limitado e implicam obrigações de transparência específicas, como assegurar que os utilizadores sabem que estão a interagir com uma máquina, ou, se o sistema puder ser utilizado para gerar ou manipular conteúdos de imagem, áudio ou vídeo consideravelmente semelhantes a conteúdos autênticos – as chamadas *deepfakes* ou "falsificações profundas" –, deverá ser obrigatório tornar claro que os conteúdos são gerados por meios automatizados[33].

Os sistemas de risco mínimo, como aplicações de jogos de vídeo ou filtros de *spam* baseados em IA, não são regulados pela Proposta de regulamento.

Já em 2022, respectivamente em 14 de Setembro e 19 de Outubro, foram aprovados os Regulamentos (UE) 2022/1925 ("Regulamento dos Mercados Digitais") e 2022/2065 ("Regulamento dos Serviços Digitais")[34]. Em concreto, este último diploma veio impor regras às plataformas em linha quanto às informações que deverão ser fornecidas aos destinatários das mensagens publicitárias exibidas, quando estas se baseiam na definição de perfis, em concreto quais os principais critérios de definição de perfis utilizados e quais os parâmetros que estiveram na base da exibição de um anúncio publicitário específico[35]. Acresce que estas plataformas não poderão exibir anúncios publicitários que tenham por base uma definição de perfis que recorra às categorias especiais de dados pessoais do artigo 9º do RGPD, nomeadamente dados que revelem a origem racial ou étnica, as opções políticas, religiosas ou filosóficas, dados médicos ou dados relativos à orientação sexual[36].

[33] Cfr. a Proposta do "Regulamento Inteligência Artificial", p. 16, e o artigo 52º. Sobre as *deepfakes*, v. Vítor Palmela Fidalgo, "§ 11. Inteligência artificial e direitos de imagem", *in* Manuel Lopes Rocha e Rui Soares Pereira (coords.), *Inteligência artificial & Direito*, Coimbra, Almedina, 2020, p. 140.

[34] Regulamento (UE) 2022/1925 do Parlamento Europeu e do Conselho, de 14 de setembro de 2022, relativo à disputabilidade e equidade dos mercados no setor digital e que altera as Diretivas (UE) 2019/1937 e (UE) 2020/1828 (Regulamento dos Mercados Digitais) e o Regulamento (UE) 2022/2065 do Parlamento Europeu e do Conselho, de 19 de outubro de 2022, relativo a um mercado único para os serviços digitais e que altera a Diretiva 2000/31/CE (Regulamento dos Serviços Digitais), *in JO* L 265, 12.10.2022, pp. 1-66 e *JO* L 277, 27.10.2022, pp. 1-102, respectivamente.

[35] Cfr. o "considerando" 68 do Regulamento dos Serviços Digitais, bem como o seu artigo 26º, nº 1, alínea *d)*. V. o que dizemos *infra*, nº 3 a) e c), sobre esta matéria.

[36] V. o "considerando" 69 e o artigo 26º, nº 3, do Regulamento dos Serviços Digitais. No que diz respeito aos utilizadores menores de idade, nos termos do artigo 28º, nº 2, do mesmo Regulamento, "os fornecedores de plataformas em linha não podem exibir anúncios publicitários na sua interface com base na definição de perfis (...) utilizando dados pessoais" (v., a este propósito, o "considerando" 71 do Regulamento dos Serviços Digitais).

Relativamente aos sistemas de recomendação – de viagens, hotéis, percursos, bens de consumo, entre outros – também os utilizadores deverão ser informados dos "critérios mais importantes para determinar as informações sugeridas", bem como "as razões da sua importância respetiva, nomeadamente se as informações forem consideradas prioritárias com base na definição de perfis e no seu comportamento em linha"[37].

Por sua vez, o Regulamento dos Mercados Digitais impõe às grandes plataformas "controladoras de acesso" (*gatekeepers*) obrigações de auditoria relativamente às técnicas de definição de perfis de consumidores utilizadas nos serviços que prestam[38].

3. Direitos de personalidade convocados na definição de perfis

A definição de perfis assenta na recolha, tratamento e cruzamento de grandes quantidades de dados dos visados, dados que contendem com diferentes bens da personalidade, da imagem e da palavra à vida privada, passando, naturalmente, pela protecção de dados e, até, pela liberdade e igualdade. Mas, em primeira linha, e com especial ênfase, estará em causa o carácter dos avaliados, descoberto e, eventualmente, exposto.

3.1. Direito ao carácter

O fluxo de dados que potencialmente podem ser hoje recolhidos através de meios electrónicos aliado a uma grande capacidade computacional para o seu tratamento constituem um especial desafio para o *direito ao carácter*, enquanto direito especial de personalidade, com o conteúdo de um direito a não ser submetido a avaliações de carácter, fora do contexto estrito das perícias médico-psiquiátricas previstas nas leis do processo[39]. Hoje, os problemas que se levantam relativamente ao direito ao carácter extravasam largamente os testes psicotécnicos ou exames grafológicos não consentidos,

[37] Cfr. o "considerando" 70 do Regulamento dos Serviços Digitais, bem como o seu artigo 27º. Sobre o tema, ainda que com base na proposta de 2020 que antecedeu o Regulamento dos Serviços Digitais, v. Esther Arroyo Amayuelas, "El derecho de las plataformas em la Unión europea", *in* Esther Arroyo Amayuelas, Yolanda Martínez Mata, Mariola Rodrìguez Font e Marc Tarrés Vives, *Servicios en plataforma. Estrategias regulatorias*, Madrid, Marcial Pons, pp. 47-48.

[38] Cfr. o artigo 15º do Regulamento dos Mercados Digitais, bem como o seu "considerando" 72.

[39] O direito ao carácter surge, de acordo com a classificação dos direitos especiais de personalidade adoptada por Orlando de Carvalho, como uma projecção do direito à inviolabilidade pessoal, incidindo sobre uma vertente "vital" deste direito, ao lado do direito à reserva da vida privada, do direito à história pessoal e do direito à verdade profunda: Orlando de Carvalho, *Teoria Geral do Direito Civil*, Coimbra, Gestlegal, 2021, pp. 267-268, nota 69.

como já tivemos a oportunidade de salientar noutro lugar[40]. Novas formas de avaliação da personalidade são perpetradas de uma forma subterrânea, sem que o avaliado se aperceba de que está a ser alvo de uma perquirição de carácter e, muito menos, sem que tenha prestado o seu consentimento para o efeito – pelo menos de uma forma consciente, ainda que possa ter aderido a uma cláusula contratual geral com esse objecto, sobretudo no contexto de um contrato electrónico[41].

A conduta que adoptamos *online* permite reunir informação significativa sobre os interesses que cultivamos, os nossos hábitos de consumo e de lazer, através das "pesquisas" que realizamos nos "motores de busca", dos vídeos que visualizamos, das fotografias que partilhamos, das compras que fazemos, das redes sociais que frequentamos e, até, através do conteúdo das mensagens de correio electrónico que enviamos e recebemos. Acresce a possibilidade de identificar os dispositivos electrónicos que utilizamos, móveis ou não, e as coordenadas geográficas onde nos encontramos – factores muitas vezes tidos em conta para a prática de *dynamic* e de *personalised pricing*[42].

Às informações assim recolhidas – e que contendem, muitas vezes, com aspectos da vida pessoal ou privada – somam-se os dados processados por objectos inteligentes, como relógios, televisões, automóveis, frigoríficos, aspi-

[40] Seguimos de perto, quanto a esta questão, o que escrevemos em Maria Raquel Guimarães, "A tutela da pessoa e da sua personalidade como fundamento e objecto da disciplina civilística. Questões actuais", *in XX Estudos comemorativos dos 20 anos da FDUP*, volume II, Coimbra, Almedina, 2017, p. 277 ss.

[41] Estas cláusulas, na medida em que sejam redigidas unilateralmente, de forma rígida e se destinem a um conjunto indeterminado de destinatários, estão sujeitas aos controlos de inclusão e de conteúdo previstos no Decreto-Lei nº 446/85, de 25 de Outubro, independentemente da fiscalização da validade do consentimento assim prestado para o tratamento de dados, nos termos dos artigos 7º a 9-º do RGPD.

[42] As técnicas de *dynamic* e de *personalised pricing* são abordadas pelo legislador europeu na Directiva (UE) 2019/2161 (Diretiva Omnibus) do Parlamento Europeu e do Conselho de 27 de novembro de 2019 que altera a Diretiva 93/13/CEE do Conselho e as Diretivas 98/6/CE, 2005/29/CE e 2011/83/UE do Parlamento Europeu e do Conselho a fim de assegurar uma melhor aplicação e a modernização das regras da União em matéria de defesa dos consumidores [in *JO* L 328 de 18/12/2019, pp. 7-28], concretamente no seu "considerando" 45, onde expressamente se diz que "[o]s profissionais podem personalizar o preço das suas ofertas para consumidores específicos ou categorias específicas de consumidores, com base em decisões automatizadas e na definição de perfis de comportamento dos consumidores, de molde a permitir-lhes avaliar o poder de compra do consumidor". Ao mesmo tempo, esta directiva veio impor um dever de informação do consumidor sempre que lhe seja apresentado um preço personalizado com base numa decisão automatizada, inserido no artigo 6º, nº 1, alínea *e-A)*, da Diretiva 2011/83/UE (v. o artigo 4º, nº 4, da Directiva Omnibus), e transposto, para o direito nacional, pelo artigo 4º, nº 1, alínea *l)*, do Decreto-Lei nº 24/2014, de 14 de Fevereiro.

radores, e até, de uma forma global, casas inteligentes. Estes dados já não dependem necessariamente de um comportamento activo do seu titular, de um "fazer algo", mas são coligidos não obstante a inacção do sujeito, na medida em que podem mesmo resultar dessa ausência de um comportamento: pense-se nos dados biométricos recolhidos por um *smart watch* durante o sono ou por uma casa inteligente, que assim reconhece padrões de actividade/inactividade do sujeito, podendo também inferir informações em função, por exemplo, da regulação da intensidade das luzes e do volume e tipo de música seleccionada.

Todos estes dados "em bruto" são depois objecto de tratamento, de "lapidação", através de algoritmos que procuram padrões de comportamentos, correlações, gerando informação com um valor económico importante, e que poderá ser utilizada, na melhor, das hipóteses, para efeitos de publicidade personalizada, mas, eventualmente também, para a conformação ou manipulação da conduta dos visados[43].

Percebe-se, por outro lado, o exponencial aumento da pressão sobre o direito ao carácter, se a estes dados recolhidos através de serviços da sociedade da informação[44] se puder somar a possibilidade de ainda agregar dados recolhidos na via pública, através de câmaras com reconhecimento facial, ou dados detidos pelo Estado, fiscais, médicos, laborais.

Os riscos associados à criação de perfis levaram o legislador nacional a consagrar expressamente o direito à "protecção do perfil" no domínio específico da utilização de plataformas digitais, no artigo 14º da Carta portuguesa de direitos humanos na era digital, aprovada pela Lei nº 27/2021, de 17 de Maio. Assim, na utilização de plataformas digitais, todos têm o direito de "proteção do seu perfil, incluindo a sua recuperação se necessário, bem como de obter cópia dos dados pessoais que lhes digam respeito nos termos previstos na lei"[45].

[43] Ver o "considerando" 69 do Regulamento dos Serviços Digitais. Sobre o tema, com desenvolvimento, v. Inês da Silva Costa, "A proteção da pessoa na era dos *big data*: a opacidade do algoritmo e as decisões automatizadas", *RED – Revista Electrónica de Direito*, vol. 24, nº 1, Fevereiro 2021, pp. 38-40, 42-44, disponível em https://cije.up.pt//client/files/0000000001/4-ines-costa_1677.pdf [consultado em 3/06/2022].

[44] Utilizamos aqui a expressão "serviços da sociedade da informação" no sentido consagrado no nº 1 do artigo 3º do Decreto-Lei nº 7/2004, de 7 de Janeiro, que transpôs para a ordem jurídica nacional a Directiva nº 2000/31/CE, do Parlamento Europeu e do Conselho, de 8 de Junho de 2000, relativa a certos aspectos legais dos serviços da sociedade de informação, em especial do comércio electrónico, no mercado interno [in *JO* L 178 de 17/7/2000, p. 1-16], abrangendo "qualquer serviço prestado a distância por via electrónica, mediante remuneração ou pelo menos no âmbito de uma actividade económica na sequência de pedido individual do destinatário".

[45] Cfr. a alínea *c)* do nº 1 do artigo 14º da Carta portuguesa de direitos humanos na era digital.

A positivação da protecção "do perfil" denota a preocupação do legislador relativamente à realidade brevemente exposta, não obstante o âmbito restrito da protecção consagrada. Ainda assim, o direito ao carácter, ameaçado ou violado em contextos distintos do da utilização de plataformas digitais, terá sempre a cobertura da cláusula geral de tutela de personalidade, contida no artigo 70º do Código Civil, que consagra um *direito geral de personalidade*, direito fundacional que protege a personalidade no seu todo, "direito que abrange todas as manifestações previsíveis e imprevisíveis da personalidade humana, pois é, a um tempo, direito à pessoa-ser e à pessoa-devir, ou melhor, à pessoa-ser em devir, entidade não estática mas dinâmica e com jus à sua «liberdade de desabrochar»", tal como configurado por Orlando de Carvalho[46].

3.2. Direitos à imagem, à palavra e à reserva da vida privada

Os dados recolhidos com vista à elaboração de perfis podem contender com os direitos à imagem, à palavra e à reserva da vida privada dos visados.

Os direitos à imagem e à palavra, enquanto projecções físicas de uma inviolabilidade pessoal[47], são direitos disponíveis pelo seu titular, que poderá consentir na captação e divulgação do "retrato" e dos sons que permitam a sua identificação. O direito à imagem está especialmente previsto no artigo 79º do Código Civil, com um regime extensível, por analogia, ao direito à palavra[48], e a Carta portuguesa de direitos humanos na era digital reconhece-os expressamente no artigo 12º, nº 1, "em ambiente digital".

O tratamento de fotografias e de vídeos publicados nas redes sociais ou a gravação e o reconhecimento de sons por assistentes virtuais têm, em princípio, na sua base a participação do próprio titular dos direitos visados, que "partilha" as imagens ou que instala os assistentes virtuais. No entanto, o tratamento dos dados assim recolhidos para a elaboração de perfis só será lícito mediante o consentimento do titular, prestado em concreto para esses fins,

[46] Orlando de Carvalho, *Teoria Geral do Direito Civil*, cit., p. 205; v., também, p. 33 (nota 16) e p. 267.

[47] Orlando de Carvalho, *Teoria Geral do Direito Civil*, cit., pp. 267-268 (v. nota 69).

[48] Posição adoptada por Orlando de Carvalho, no seu ensino oral. Integrando também a protecção da voz ("voz falada") no artigo 79º, no contexto específico das suas possíveis violações por sistemas de inteligência artificial, com o argumento de que a voz "está diretamente ligada à imagem da própria pessoa" e, portanto, parecendo não autonomizar este direito à voz como um direito especial de personalidade, v. Vítor Palmela Fidalgo, "§ 11. Inteligência artificial e direitos de imagem", cit., p. 141. Sobre a tutela do direito à palavra/à voz, ver, com desenvolvimento, a obra monográfica de Julia Ammerman Yebra, *El derecho a la propia voz como derecho de la personalidad*, A Coruña, Colex, 2021.

previamente especificados nos contratos de prestação de serviços celebrados com as respectivas plataformas, não podendo este consentimento ser presumido a partir da publicação ou divulgação das imagens, ou da interacção voluntária com um assistente virtual. E, é claro, será ilícita, também penalmente, a captação e o tratamento de imagens e sons sem a colaboração do titular dos respectivos direitos, à sua revelia, através de programas ou aplicações que accionam as câmaras ou microfones dos computadores e dispositivos móveis sem que o utilizador disso se aperceba.

Acresce que o direito à imagem poderá também ser violado sem que haja captação do "retrato" da pessoa, na medida em que este seja criado por programas de inteligência artificial. No caso das *deepfakes* ou de hologramas que possam interagir com terceiros, a violação do direito à imagem será cumulada com a violação do direito à verdade pessoal[49].

Já a captação de imagens de automóveis e, eventualmente, também dos seus condutores e acompanhantes, por sistemas de pagamento automático em autoestradas ou parques de estacionamento, é necessária para a execução do próprio contrato de prestação do serviço subjacente mas a licitude do tratamento dessas imagens para os fins próprios do contrato não se estende a outras finalidades e, certamente, não compreende a definição do perfil do titular do direito à imagem (e à reserva da vida privada).

A lei prevê, por outro lado, limitações ao direito à imagem, aplicáveis analogicamente ao direito à palavra, motivadas por razões de ordem subjectiva – como a notoriedade do titular do direito ou o cargo por ele desempenhado –, e razões objectivas, justificadas por exigências de polícia e de justiça, finalidades científicas, didácticas e culturais, mas também pelo facto de as imagens serem recolhidas em lugares públicos, dizerem respeito a factos de interesse público ou decorridos publicamente.

Estas limitações poderão justificar a captação de imagens por câmaras de vigilância instaladas na via pública, com fundamento em razões de prevenção e repressão criminais. No entanto, o potencial de violação não só do

[49] Vítor Palmela Fidalgo, "§ 11. Inteligência artificial e direitos de imagem", cit., pp. 141-142, refere também a hipótese de um robot reproduzir a imagem de uma pessoa, pronunciando-se sobre o caso americano *White vs. Samsung*, embora não distinguindo o direito à imagem e o direito à verdade pessoal. Para uma autonomização do direito à verdade pessoal, v. Orlando de Carvalho, *Teoria Geral do Direito Civil*, cit., pp. 267-268, nota 70. De acordo com a classificação adoptada pelo Autor, o direito à verdade pessoal surge compreendido num direito mais amplo à identidade pessoal, ao lado do direito ao nome, no sentido de "que da pessoa não se afirme o que não seja verdade". Na medida em que a imagem de uma pessoa seja reproduzida e utilizada em situações ficcionadas haverá uma violação cumulativa do direito à imagem e do direito à verdade pessoal.

direito à imagem mas também do direito à reserva da vida privada – uma vez que não há uma sobreposição exacta entre esfera privada e espaço privado, havendo também protecção da vida privada no espaço público, para além de nem sempre ser clara a delimitação jurídica e física destes espaços –, é particularmente preocupante na medida em que se verifique uma proliferação de câmaras e a ausência de uma regulação precisa do seu funcionamento[50].

A Proposta de Lei nº 111/XIV/2ª, que antecedeu a nova Lei nº 95/2021, de 29 de Dezembro, que regula a utilização e o acesso pelas forças e serviços de segurança a sistemas de videovigilância para captação, gravação e tratamento de imagem e som, representava um "cheque em branco", nas palavras da CNPD, à intrusão na vida privada dos cidadãos, para além de permitir, "com nula densificação normativa", a utilização de tecnologias de inteligência artificial e, em especial, de reconhecimento facial[51].

Recorde-se que esta proposta, da iniciativa do Governo, surge imediatamente depois da aprovação da Carta portuguesa de direitos humanos na era digital e da Proposta de um Regulamento Inteligência Artificial, tornando evidente que as garantias nestes domínios nunca são demasiadas e que, de facto, "o que acontece na China", pode, muito bem, "não ficar na China", e chegar a Portugal.

Entre as muitas críticas de que o regime proposto foi alvo – como o alargamento das finalidades que justificavam a utilização dos meios de videovigilância, nomeadamente em caso de "elevada circulação ou concentração de pessoas" e "ocorrência de facto suscetível de perturbação da ordem pública", e a não delimitação das finalidades previstas em relação aos diferentes dispositivos de videovigilância, com desigual potencial de violação da privacidade, como câmaras móveis instaladas em *drones*[52] –, salientamos a questão da utilização de meios de inteligência artificial para a visualização e tratamento de dados. De acordo com o nº 1 do artigo 18º da Proposta de Lei a "visualização e o tratamento dos dados pod[iam] ter subjacente um sistema de gestão ana-

[50] Neste sentido, v. o já citado Parecer da Comissão Nacional de Protecção de Dados (CNPD) 2021/143, de 4 de Novembro de 2021, sobre a Proposta de Lei nº 111/XIV/2.a., p. 1v, nº 8. Como assinalou a CNPD, a utilização de sistemas de videovigilância em espaço público "representa sempre uma ingerência sobre os direitos fundamentais, máxime dos direitos ao respeito pela vida privada e familiar e à proteção de dados pessoais".

[51] *Idem*, p. 2, nºs 9 e 10.

[52] Cfr. o artigo 3º da Proposta de Lei nº 111/XIV/2.a., bem como o Parecer da CNPD 2021/143, pp. 4-5, nºs 36-47. Note-se que as finalidades referidas no texto, que constavam das subalíneas *ii)* e *iii)* da alínea *d)* do artigo 3º da Proposta de Lei não foram incluídas no artigo 3º da Lei nº 95/2021, de 29 de Dezembro.

lítica dos dados captados, por aplicação de critérios técnicos de acordo com os fins a que os sistemas se destina[va]m", ao que acrescia a permissão para a captação de dados biométricos, no nº 2. Assim, previa-se em termos genéricos a utilização de tecnologia com um grande potencial de lesão dos direitos de personalidade dos visados sem delimitar de forma precisa os termos dessa utilização e sem a restringir a determinadas finalidades da videovigilância, previstas no artigo 3º, como assinalámos, com grande amplitude[53].

Como salientou a CNPD no seu Parecer 2021/143, esta previsão genérica do uso de sistemas de "gestão analítica dos dados captados", aliada à possibilidade de tratamento de dados biométricos, resultava numa "norma legislativa que, de forma subtil e encoberta, d[ava] abertura à incorporação de tecnologia de reconhecimento facial nos sistemas de vídeo vigilância em espaço público" e em espaços privados de acesso público[54].

E, acrescentou ainda a CNPD, de uma forma impressiva e particularmente pertinente face às reflexões que aqui vertemos sobre os sistemas de *profiling* estatais, pelo que transcrevemos:

> "Desconhece a CNPD se o legislador nacional está ciente das consequências reais da utilização deste tipo de tecnologia em sistemas de videovigilância no espaço público e no espaço privado de acesso ao público. Trata-se, na realidade, de dar luz verde à vigilância em massa pelas forças e serviços de segurança, negando qualquer dimensão de privacidade que ainda pudesse restar no espaço público (e no espaço privado aberto ao público). Ela permite o rastreamento dos cidadãos potenciado pela possibilidade de relacionamento das informações disponíveis nos sistemas de videovigilância dos estabelecimentos públicos e privados e demais espaços privados abertos ao público, a que se soma a utilização na atividade diária das forças e serviços de segurança das câmaras portáteis também por via do recurso a *drones*.
>
> Sendo evidente o impacto que tal controlo pode ter sobre qualquer sociedade democrática, pela facilidade com que esta ferramenta é utilizável como meio de repressão das liberdades de expressão, de manifestação e de reunião, como exemplos recentes noutros pontos do mundo têm demonstrado"[55].

A reserva da vida privada é ainda posta em causa, simultaneamente com o direito à imagem, sempre que as imagens processadas para efeitos de definição de perfis contendem com situações da vida pessoal, privada ou com

[53] Parecer da CNPD 2021/143, p. 12, nºs 121-122.
[54] Idem, p. 12v, nºs 123-124.
[55] *Idem*, p. 12v, nºs 126-127.

a reserva de segredo dos visados. E a privacidade é também ameaçada pela captação de informações em mensagens, na utilização de motores de busca, redes sociais, plataformas de comércio electrónico e, em geral, com o acesso a informação relativa à conduta dos visados *online*. Acresce que o perigo de violação do direito à reserva da vida privada aumenta proporcionalmente com o aumento das informações que são recolhidas e tratadas[56], a par com o risco de violação do direito à protecção de dados, pelo potencial de sobreposição que as informações relativas à vida privada têm relativamente aos dados pessoais.

3.3. Direito à protecção de dados

Às limitações introduzidas nos direitos à imagem, à palavra e à reserva da vida privada pela criação de perfis – para além da compressão, em primeira linha, do direito ao carácter – acresce a tensão imposta ao direito à protecção de dados pessoais.

As imagens, as palavras, e as informações recolhidas com vista à definição de perfis são simultaneamente tuteladas como dados pessoais, na medida em que permitam a identificação do seu titular[57]. O próprio RGPD refere-se à definição de perfis como tratamento de dados, ainda que incidentalmente, a propósito da aplicação territorial do Regulamento (artigo 3º), no seu "considerando" 24:

> "A fim de determinar se uma atividade de tratamento pode ser considerada «controlo do comportamento» de titulares de dados, deverá determinar-se se essas pessoas são seguidas na Internet e a potencial utilização subsequente de técnicas de tratamento de dados pessoais que consistem em definir o perfil de uma pessoa singular, especialmente para tomar decisões relativas a essa pessoa ou analisar ou prever as suas preferências, o seu comportamento e as suas atitudes".

Assim, à protecção conferida aos direitos de personalidade já identificados, soma-se o disposto no RGPD em matéria de licitude, lealdade e transparência do tratamento dos dados, limitação das finalidades e da conservação, e minimização dos dados[58]. Haverá também que ter em conta os limites intro-

[56] Assim, Mafalda Miranda Barbosa, *Inteligência artificial...*, cit., p. 133.

[57] Cfr. o artigo 4º, nº 1, do RGPD. Para uma distinção entre informação e dados, e entre informação pessoal e dados pessoais, v. Marcin Betkier, *Privacy online, law and the effective regulation of online services*, Cambridge/Antwerp/Chicago, Intersentia, 2019, pp. 9-13.

[58] Cfr. os artigos 5º e 6º do RGPD, bem como o seu "considerando" 39. Também a Carta portuguesa de direitos humanos na era digital, prevê no nº 2 do seu artigo 8º, com a epígrafe "Direito à privacidade em ambiente digital", o direito à proteção de dados pessoais, "incluindo o

duzidos pela lei nacional da proteção de dados pessoais, Lei nº 58/2019, de 8 de Agosto, em matéria de videovigilância, "cuja finalidade seja a proteção de pessoas e bens"[59], e que acresce ao disposto em matéria de videovigilância por razões de segurança pública, bem como as obrigações de informação já aludidas impostas às plataformas em linha pelo novo Regulamento dos Serviços Digitais.

Se os dados processados pelos sistemas de IA não permitirem a identificação directa ou indirecta do seu titular, na medida em que sejam dados anonimizados, desligados de um identificador – entendido num sentido amplo, como um nome, morada, *email*, mas também informações sócio-económicas, psicológicas, filosóficas –, então o RGPD não será convocado[60].

Mais questionável será a aplicação do Regulamento quando está em causa o tratamento de dados que não foram fornecidos pelo seu titular nem recolhidos directamente a partir do seu comportamento, mas sim inferidos a partir destes, dados derivados, resultantes do processamento de dados pessoais através de programas ou aplicações de IA[61]. No entanto, o potencial de lesão destes dados no que respeita à intimidade da vida privada parece aconselhar a sua compreensão na noção de dados pessoais.

controlo sobre a sua recolha, o registo, a organização, a estruturação, a conservação, a adaptação ou alteração, a recuperação, a consulta, a utilização, a divulgação por transmissão, difusão ou qualquer outra forma de disponibilização, a comparação ou interconexão, a limitação, o apagamento ou a destruição", numa espécie de súmula do disposto no RGPD.

[59] Cfr. o artigo 19º, nºs 2-4, da Lei nº 58/2019, de 8 de Agosto, que prevê, nomeadamente, que as câmaras não podem incidir sobre "vias públicas, propriedades limítrofes ou outros locais que não sejam do domínio exclusivo do responsável", o "interior de áreas reservadas a clientes ou utentes onde deva ser respeitada a privacidade" e o "interior de áreas reservadas aos trabalhadores, designadamente zonas de refeição, vestiários, ginásios, instalações sanitárias e zonas exclusivamente afetas ao seu descanso" e que, nos casos em que é admitida a videovigilância, é sempre proibida a captação de som, excepto autorização da CNPD ou durante o período em que as instalações vigiadas estejam encerradas.

[60] V. Gabriele Mazzini, "Q. A system of governance for artificial intelligence through the lens of emerging intersections between AI and EU law", *in* De Franceschi / Schulze, *Digital Revolution – New challenges for Law*, München/lveBaden-Baden, Beck/Nomos, 2019, pp. 281-282. Um sentido amplo de identificação de um titular de dados foi adoptado pelo GT29 (Grupo de Trabalho do Artigo 29º), Opinião 4/2007, de 20/06 (*Opinion 4/2007 on the concept of personal data*), disponível em https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf [consultado em 11/06/2022], p. 14, ratificada pelo órgão que substituiu o GT29, o *European Data Protection Board* (Comité Europeu para a proteção de Dados), na sua primeira reunião plenária: https://edpb.europa.eu/news/news/2018/endorsement-gdpr-wp29-guidelines-edpb_en [consultado em 11/06/2022].

[61] Sobre estes dados, as fronteiras da noção de "dados pessoais" e a posição do GT29 sobre o tema, v. Gabriele Mazzini, "Q. A system of governance for artificial intelligence through the lens of emerging intersections between AI and EU law", cit., pp. 283-285.

A actividade de *profiling* alimenta-se, por outro lado, de *big data* o que coloca dificuldades sérias no que respeita ao cumprimento dos princípios estabelecidos no RGPD[62]. O princípio da limitação das finalidades para as quais os dados são fornecidos impede a sua utilização para outros fins encobertos e, desde logo, para a avaliação do perfil do seu titular, não especificada aquando da recolha. E, sobretudo, o princípio da minimização dos dados, com o conteúdo da necessária adequação e limitação dos dados às finalidades definidas para o tratamento, "vive mal" com a ideia de *big data*[63]. *Os algoritmos de profiling* que integram IA funcionam como uma caixa negra que se alimenta de dados e infere dados novos dos dados fornecidos, sem que se conheça à partida quais os dados que irão ser recolhidos e que serão suficientes para os resultados a alcançar e mesmo que resultados serão alcançados e se corresponderão com os resultados previamente considerados pelo programador.

Acresce que, nos termos do artigo 22º do RGPD, "o titular dos dados tem o direito de não ficar sujeito a nenhuma decisão tomada exclusivamente com base no tratamento automatizado, incluindo a definição de perfis, que produza efeitos na sua esfera jurídica ou que o afete significativamente de forma similar"[64]. O legislador europeu associou, desta forma, a criação de perfis às decisões automatizadas, pelo risco acrescido para o titular dos dados da conjugação destes dois fenómenos, ainda que eles não andem necessariamente associados (embora a criação de perfis tenda a resultar em decisões automatizadas e, por outro lado, essas decisões automatizadas, com frequência, sujam em consequência de uma definição de perfis)[65].

[62] Sobre o conceito de *big data*, v. Inês da Silva Costa, "A proteção da pessoa na era dos *big data*: a opacidade do algoritmo e as decisões automatizadas", cit., p. 38. Na doutrina estrangeira, v., entre outros, Marcin Betkier, *Privacy online...*, cit., pp. 13-14, referindo-se ao "modelo dos 3Vs" para caracterizar os *big data*: volume, variedade e velocidade (estendido para "5Vs", incluindo valor e veracidade, de acordo com alguns autores).

[63] A expressão é de Lourenço Noronha dos Santos, "§ 12. Inteligência artificial e privacidade", *in* Manuel Lopes Rocha e Rui Soares Pereira (coords.), *Inteligência artificial & Direito*, Coimbra, Almedina, 2020, p. 152.

[64] O "considerando" 71 do RGPD exemplifica estas consequências relevantes das decisões automatizadas com "a recusa automática de um pedido de crédito por via eletrónica ou práticas de recrutamento eletrónico sem qualquer intervenção humana". Para mais desenvolvimentos sobre este ponto, v. Alba Soriano Arnanz, *Data protection for the prevention of algorithmic discrimination*, cit., pp. 142-144, Gabriele Mazzini, "Q. A system of governance for artificial intelligence through the lens of emerging intersections between AI and EU law", cit., pp. 286-287.

[65] Assim, Alba Soriano Arnanz, *Data protection for the prevention of algorithmic discrimination*, cit., p. 137, e Mafalda Miranda Barbosa, *Inteligência artificial...*, cit., p. 141. Também de acordo com Inês da Silva Costa, "A proteção da pessoa na era dos *big data*: a opacidade do algoritmo e as decisões automatizadas", cit., p. 56, "[m]uito embora a definição de perfis não se confunda com

O Regulamento prevê, ainda assim, excepções a esta proibição geral de ficar sujeito a decisões automatizadas, nomeadamente nos casos em que o titular dos dados tenha dado o seu consentimento "explícito". Isto significa que a pessoa não pode ficar sujeita a uma decisão tomada sem qualquer controlo ou intervenção humana – que não seja meramente aparente, de ratificação de uma tomada de posição processada de forma automatizada[66] – em resultado de uma avaliação do seu perfil, não legitimada pelo seu consentimento prévio[67]. E, mesmo nos casos em que essa decisão automatizada é lícita, o titular dos dados tem o direito de ser informado relativamente à sua existência bem como "à lógica subjacente" ao algoritmo, e ainda quanto às consequências de tal tratamento na sua esfera jurídica[68].

Também o legislador nacional foi sensível a esta necessidade de informação nestes casos, prevendo no nº 2 do artigo 9º da Carta portuguesa de direitos humanos na era digital, que "[a]s decisões com impacto significativo na esfera dos destinatários que sejam tomadas mediante o uso de algoritmos devem ser comunicadas aos interessados, sendo suscetíveis de recurso e auditáveis, nos termos previstos na lei".

4. O direito à igualdade e o direito à liberdade: reflexões conclusivas

Realizado este breve excurso pelos sistemas de *profiling*, em especial com recurso a IA, e apontados os riscos que estes sistemas geram para diferentes

as decisões automatizadas, a verdade é que não é fácil imaginar a existência de uma atividade de definição de perfis que não culmine numa decisão automatizada e, por outro lado, a maioria das decisões automatizadas surge como consequência de uma atividade de definição de perfis. Por este motivo, o legislador optou pela sua regulamentação conjunta".

[66] As hipóteses em que a intervenção humana se limita a "carimbar" a decisão automatizada (a expressão é de Gabriele Mazzini, "Q. A system of governance for artificial intelligence through the lens of emerging intersections between AI and EU law", cit., p. 285) são, assim, tomadas como decisões sem intervenção humana. Assim, também, Inês da Silva Costa, "A proteção da pessoa na era dos *big data*: a opacidade do algoritmo e as decisões automatizadas", cit., p. 58, com indicações (nota 172), Lourenço Noronha dos Santos, "§ 12. Inteligência artificial e privacidade", cit., p. 154, e Mafalda Miranda Barbosa, *Inteligência artificial...*, cit., pp. 143-144.

[67] Excepcionam-se, também, as hipóteses de a decisão automatizada ser "necessária para a celebração ou a execução de um contrato entre o titular dos dados e um responsável pelo tratamento", ou ser "autorizada pelo direito da União ou do Estado-Membro a que o responsável pelo tratamento estiver sujeito, e na qual estejam igualmente previstas medidas adequadas para salvaguardar os direitos e liberdades e os legítimos interesses do titular dos dados", desde que não tenha por base dados sensíveis, conforme o disposto nos nºs 2 e 4 do artigo 22º do RGPD. Sobre estas excepções, v. Alba Soriano Arnanz, *Data protection for the prevention of algorithmic discrimination*, cit., pp. 137-139, e Lourenço Noronha dos Santos, "§ 12. Inteligência artificial e privacidade", cit., pp. 154-156. V., também, o "considerando" 71 do RGPD.

[68] Cfr. o artigo 14º, nº 2, alínea *g*), do RGPD.

bens da personalidade, como a imagem, a palavra, a vida privada, os dados pessoais e o carácter, chama-se, por fim, a atenção para o direito à igualdade, em causa sempre que sejam utilizados algoritmos discriminatórios, e o direito à liberdade, limitado com o condicionamento ou mesmo manipulação da conduta da pessoa[69].

O direito à igualdade, enquanto direito de personalidade, impõe-se nas relações jurídico-privadas, no sentido de proteger a pessoa contra formas de discriminação em função da raça, religião, etnia, sexo, idade, convicções políticas ou ideológicas, filiação sindical, ou outras, tendo que ser compatibilizado com o princípio da autonomia privada e da liberdade contratual. No que respeita à utilização de algoritmos, tem sido demonstrado que os algoritmos empregues para processar dados pessoais conduzem a resultados discriminatórios, na medida em que reproduzem posições ancestrais que desfavorecem os membros de grupos vulneráveis[70]. Os algoritmos são criados por pessoas e alimentam-se de informações do mundo real pelo que tendencialmente irão replicar o enviesamento de que a sociedade padece, nomeadamente em processos de selecção de trabalhadores ou de concessão de crédito.

Por outro lado, a definição de um dado perfil irá condicionar as ofertas de produtos e serviços que o visado irá receber *online*, mas também as informações e notícias que lhe serão disponibilizadas, acabando por criar uma "bolha" de "realidade artificial" dentro da qual a pessoa passa a viver, mais ou menos afastada da realidade dos demais. O seu comportamento poderá ser assim manipulado, nomeadamente para fins políticos, e, em última instância, a sua liberdade é limitada[71].

Estas preocupações estiveram presentes na mente do legislador europeu do Regulamento dos Serviços Digitais, a propósito da utilização de técnicas de publicidade e de recomendações personalizadas, optimizadas para corresponder aos interesses dos utilizadores e "apelar potencialmente às suas

[69] Sobre o potencial de violação dos direitos à igualdade e à liberdade pelo tratamento de dados, v. Mafalda Miranda Barbosa, *Inteligência artificial*..., cit., pp. 133-135.

[70] Assim, v., por todos, Alba Soriano Arnanz, *Data protection for the prevention of algorithmic discrimination*, cit., pp. 71-72, e, no que concerne à discriminação na conformação da oferta contratual, v. Julio Álvarez Rubio, "Inteligencia artificial y protección jurídica de los consumidores", *in* José Ignacio Solar Cayón (ed.), *Dimensiones éticas y jurídicas de la inteligencia artificial en el marco del Estado de derecho*, Madrid, Editorial Universidad de Alcalá, 2020, p. 282 ss. Entre nós, v. Inês da Silva Costa, "A proteção da pessoa na era dos *big data*: a opacidade do algoritmo e as decisões automatizadas", cit., p. 46 ss.

[71] Sobre o tema, v. Inês da Silva Costa, "A proteção da pessoa na era dos *big data*: a opacidade do algoritmo e as decisões automatizadas", cit., pp. 45-46, e Mafalda Miranda Barbosa, *Inteligência artificial*..., cit., pp. 134-135.

vulnerabilidades", com "efeitos negativos particularmente graves"[72]. Como se pode ler nos "considerandos" que antecedem o articulado do diploma, "(e)m certos casos, as técnicas manipuladoras podem ter um impacto negativo em grupos inteiros e amplificar os danos sociais, por exemplo contribuindo para campanhas de desinformação ou discriminando determinados grupos"[73].

Estes riscos associados aos sistemas de *profiling*, potenciados pela IA, que se somam aos riscos já identificados que se impõem a outros direitos de personalidade, deverão deixar o legislador alerta no momento de intervir nestas matérias. Na Carta portuguesa de direitos humanos na era digital afirma-se que "[t]odos têm direito a que os conteúdos transmitidos e recebidos em ambiente digital não sejam sujeitos a discriminação, restrição ou interferência em relação ao remetente, ao destinatário, ao tipo ou conteúdo da informação, ao dispositivo ou aplicações utilizados, ou, em geral, a escolhas legítimas das pessoas"[74].

É, no entanto, necessário que estas proclamações não sejam letra morta ou, pelo menos, que o legislador não ofereça com uma mão aquilo que retira com a outra. Até porque, como dissemos em cima, estamos longe de poder assegurar que "*what happens in China, stays in China*"...

Bibliografia

Álvarez Rubio, Julio, "Inteligencia artificial y protección jurídica de los consumidores", *in* José Ignacio Solar Cayón (ed.), *Dimensiones éticas y jurídicas de la inteligencia artificial en el marco del Estado de derecho*, Madrid, Editorial Universidad de Alcalá, 2020, pp. 275-333

Ammerman Yebra, Julia, *El derecho a la propia voz como derecho de la personalidad*, A Coruña, Colex, 2021

Arroyo Amayuelas, Esther, "El derecho de las plataformas em la Unión europea", *in* Esther Arroyo Amayuelas, Yolanda Martínez Mata, Mariola Rodrìguez Font e Marc Tarrés Vives, *Servicios en plataforma. Estrategias regulatórias*, Madrid, Marcial Pons, pp. 21-69

Barbosa, Mafalda Miranda, *Inteligência artificial. Entre a utopia e a distopia, alguns problemas jurídicos*, Coimbra, Gestlegal, 2021

Betkier, Marcin, *Privacy online, law and the effective regulation of online services*, Cambridge/Antwerp/Chicago, Intersentia, 2019

Bradford, Laura / Aboy, Mateo / Liddell, Kathleen, "COVID-19 Contact Tracing Apps: A Stress Test for Privacy, the GDPR and Data Protection Regimes", *in*

[72] Cfr. os "considerandos" 69 e 70 do Regulamento dos Serviços Digitais.

[73] *Idem*, *ibidem*.

[74] Cfr. o artigo 10º da Carta portuguesa de direitos humanos na era digital, com a epígrafe "Direito à neutralidade da internet".

Journal of Law and the Biosciences, volume 7, 1, January-June 2020, pp. 1-21, disponível em https://doi.org/10.1093/jlb/lsaa034 [consultado em 12/06/2022]

Carvalho, Orlando de, *Teoria Geral do Direito Civil*, Coimbra, Gestlegal, 2021

Cho, Hyunghoon / Ippolito, Daphne / Yu, Yun William, "Contact Tracing Mobile Apps for COVID-19: Privacy Considerations and Related Trade-offs", *in Cornell University* arXiv:2003.11511v2 [cs.CR], 2020, pp. 1-12, disponível em https://arxiv.org/abs/2003.11511v2 [consultado em 12/06/2022]

CNPD, *Parecer da Comissão Nacional de Protecção de Dados (CNPD) 2021/143*, de 4 de Novembro de 2021, sobre a Proposta de Lei nº 111/XIV/2.a.

Costa, Inês da Silva, "A proteção da pessoa na era dos *big data*: a opacidade do algoritmo e as decisões automatizadas", *in RED – Revista Electrónica de Direito*, vol. 24, nº 1, Fevereiro 2021, pp. 34-82, disponível em https://cije.up.pt//client/files/0000000001/4--ines-costa_1677.pdf [consultado em 3/06/2022].

Fidalgo, Vítor Palmela, "§ 11. Inteligência artificial e direitos de imagem", *in* Manuel Lopes Rocha e Rui Soares Pereira (coords.), *Inteligência artificial & Direito*, Coimbra, Almedina, 2020, pp. 137-146

GT29 (Grupo de Trabalho do Artigo 29º), *Opinião 4/2007, de 20/06* (*Opinion 4/2007 on the concept of personal data*), disponível em https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf [consultado em 11/06/2022]

Guimarães, Maria Raquel, "A tutela da pessoa e da sua personalidade como fundamento e objecto da disciplina civilística. Questões actuais", *in XX Estudos comemorativos dos 20 anos da FDUP*, volume II, Coimbra, Almedina, 2017, pp. 291-312

Mazzini, Gabriele, "Q. A system of governance for artificial intelligence through the lens of emerging intersections between AI and EU law", *in* De Franceschi / Schulze, *Digital Revolution – New challenges for Law*, München/lveBaden-Baden, Beck/Nomos, 2019, pp. 1-55, disponível em https://ssrn.com/abstract=3369266 [consultado em 16/03/2023]

Orgad, Liav / Reijers, Wessel, "How to Make the Perfect Citizen? Lessons from China's Social Credit System", *in Vanderbilt Journal of Transnational Law*, 54, no. 5, November 2021, pp. 1087-1122

Resta, Giorgio, "La protezione dei dati personali nel diritto dell'emergenza Covid-19", *Giustizia civile.com*, 5, 2020, disponível em http://giustiziacivile.com/pdfpage/2262 [consultado em 12/06/2022]

Rettinger, Lizzy, "The Human Rights Implications of China's Social Credit System", *in Journal of High Technology Law*, 21, no. 1, 2021, pp. 1-33, disponível em https://sites.suffolk.edu/jhtl/home/publications/volume-xxi-number-1/ [consultado em 16/03/2023]

Santos, Lourenço Noronha dos, "§ 12. Inteligência artificial e privacidade", *in* Manuel Lopes Rocha e Rui Soares Pereira (coords.), *Inteligência artificial & Direito*, Coimbra, Almedina, 2020, pp. 147-159

Silva, Diogo Rodrigues da, "Consequences of ratings/reviews on sharing economy platforms", *in* M. Regina Redinha / M. Raquel Guimarães / F. Liberal Fernandes (coords.), *Sharing Economy: Legal Problems of a Permutations and Combinations Society*, Newcastle upon Tyne, Cambridge Scholars, 2019, pp. 382-387

Síthigh, Daithí Mac / Siems, Mathias, "The chinese social credit system: a model for other countries?", European University Institute, Department of Law, Working Paper LAW 2019/01, disponível em https://cadmus.eui.eu/bitstream/handle/1814/60424/LAW_2019_01.pdf [consultado em 12/06/2022]

Solar Cayón, José Ignacio, *La inteligencia artificial jurídica, El impacto de la innovación tecnológica en la práctica del derecho y el mercado de servicios jurídicos*, Pamplona, Aranzadi, 2019

Soriano Arnanz, Alba, *Data protection for the prevention of algorithmic discrimination*, Cizur Menor, Aranzadi, 2021

Weaver, John Frank, "Everything Is Not *Terminator*. Is China's Social Credit System the Future?", *in RAIL: The Journal of Robotics, Artificial Intelligence & Law*, 2, no. 6 (November-December), 2019, pp. 445-451, disponível em https://mclane.com/wp-content/uploads/Everything_Is_Not_Terminator_-_Is_Chinas_Social_Credit_System_the_Future.pdf [consultado em 16/03/2023]

Robots, dignidad y Derechos Humanos

Robots, dignity and Human Rights

Martín González López*

RESUMO: Os direitos humanos estão a enfrentar os desafios da era digital. Assim, surgiram correntes que advogam a superação do homo sapiens através da ciência e tecnologia ou a iminente superação do nosso potencial humano pela Inteligência Artificial.

Neste cenário, a teoria dos direitos humanos deve abordar os debates do nosso tempo: a validade dos direitos humanos de 1948, a abertura a novos sujeitos não humanos como os robôs ou o papel da Inteligência Artificial na garantia dos direitos dos cidadãos.

Este artigo irá abordar algumas das posições acima mencionadas e as suas consequências para os direitos humanos, incluindo a banalização dos direitos humanos.

PALAVRAS-CHAVE: Direitos Humanos, Inteligência Artificial, robôs, era digital, trivialização

ABSTRACT: Human rights are facing the challenges of the digital age. Thus, currents have emerged that advocate the overcoming of homo sapiens through science and technology or the imminent overcoming of our human potential by Artificial Intelligence.

* Doctorando de la Universidad de Burgos (UBU). m.gonzalez.lopez1996@gmail.com

In this scenario, human rights theory must address the debates of our time: the validity of the 1948 human rights, the opening up to new non-human subjects such as robots or the role of Artificial Intelligence in guaranteeing citizens' rights.

This article will address some of the aforementioned positions and their consequences for human rights, including the trivialisation of human rights.

KEYWORDS: Human rights, Artificial Intelligence, robots, digital age, trivialisation

SUMÁRIO: 1. Introducción 2. Poshumanismo y Transhumanismo 3. El camino hacia la creación 4. ¿Robots humanizados? 5. Sociedad IA 6. El premio de la dignidad 7. Derechos de los robots 8. Banalización de los derechos humanos

1. Introducción

La lucha por el reconocimiento de derechos a los animales lleva desarrollándose durante muchos años. Su argumento insignia hace referencia a la moralidad y el sinsentido de propiciar sufrimiento innecesario a animales no humanos. Los experimentos en animales, la industria alimenticia, la manufactura textil o su mala utilización en espectáculos de ocio constituyen ejemplos de un trato injustificado y cruel hacia estos seres.

La sintiencia, esto es, la capacidad de sufrir y sentir de los animales fue el primer paso para cuestionar el sufrimiento anteriormente mencionado. A la par, la respuesta de los no defensores de los derechos de los animales fue su no capacidad de comunicarse o, si se quiere, la falta de lenguaje. Esto se desmontó simplemente poniendo de relieve la capacidad de comunicación y lenguaje entre los animales de una misma especie. Cosa diferente es que este lenguaje no coincida o no sea entendido por el género humano[1]. Asimismo, la posibilidad de comunicarse y transmitir sentimientos o dolor de la que constan los seres humanos se puso en duda con dos ejemplos: ¿Acaso los recién nacidos, enfermos vegetales o acianos con demencia senil poseen esa capacidad de transmitir? ¿Qué diferencia a cualquiera de ellos con un animal adulto?

La respuesta no es más que la pertenencia a otra especie. Surgió entonces el término especismo, haciendo referencia a un trato favorable a los miembros de una especie en contra de otra, únicamente por el hecho de pertenencia[2].

[1] P. Singer, *Liberación animal: el clásico definitivo del movimiento animalista*, Taurus, 2018.

[2] Oscar Horta, "Términos básicos para el análisis del especismo "Los libros de la catarata, Madrid, 107-118, 2008, Pg. 2. Nuria Belloso Martín, "Un intento de fundamentar derechos de los

En este debate entran en juego argumentos de índole moral y antropocéntrica que, en parte, tienen su justificación filosófica. Es cierto que todo aquello a lo que podemos hacer referencia como humanos es fruto de nuestra evolución y concepción del mundo. Cuando argumentos en contra de un reconocimiento de estatus a los animales hablan de autonomía, conciencia, fraternidad, justifica, amor, etc. están haciendo acopio de un lenguaje y contenido que ha estado vinculado históricamente solo al ser humano. En ese sentido, proyectar en los animales no humanos esos términos resulta extraño[3].

No obstante, este movimiento no pretende una igualdad absoluta entre humanos y animales, más bien busca el fin del sufrimiento injustificado y un trato acorde a la naturaleza de cada ser[4]. Reconocer el derecho a voto a un elefante no tiene justificación dado que ese mamífero no posee una capacidad de razonamiento, conciencia, pertenencia o participación suficiente como para tomar parte en las decisiones políticas que favorecen el desarrollo democrático de una comunidad. Lo mismo ocurre con la libertad de expresión o el derecho a la educación. Se tratan de derechos con una clara referencia al ser humano.

El movimiento animalista reivindica aquello más cercano al derecho a la vida y a la no agresión, derechos que sintetizan los males que a lo largo de la historia han sufrido los animales[5] [6].

A su vez, también es posible alegar que un mal trato hacia los animales no humanos desemboca en un empeoramiento de las condiciones de vida de los humanos y, por lo tanto, el ser humano debe dedicar esfuerzos a su protección.

Podemos sacar a colación otro debate que intenta atribuir derechos o, por lo menos, un estatus jurídico a otro sujeto no humano: la naturaleza.[7] En este sentido, los argumentos no se centran en el sufrimiento o la vida sino en la preservación y conservación del medio donde habitamos, con el fin de conseguir unas condiciones materiales de vida adecuadas.

no-humanos (derechos de la Naturaleza) a partir del desarrollo sostenible", *Revista Catalana de Dret Ambiental* (RCDA), 2022, vol. 13, nº 1 junio. Tarragona, Universitat Rovira i Virgili, pp. 1-46. https://doi.org/10.17345/rcda3198 https://revistes.urv.cat/index.php/rcda/issue/current

[3] Martha C. Nussbaum, *Las fronteras de la justicia. Consideraciones sobre la exclusión*, Paidós, 2021, Pg. 104.

[4] Op Cit Pg 334. (2021). También, vid. Adela Cortina, *Las fronteras de la persona: el valor de los animales, la dignidad de los humanos*, Taurus, 2009.

[5] Manuel Atienza, Sobre la dignidad, Trotta, 2022, Pg. 150.

[6] DerechoUBA. Sobre la dignidad. Youtube (2021) Consultado 23/09/2022 1h 43min

[7] Nuria Belloso Martín, *El debate sobre la tutela institucional: generaciones futuras y derechos de la naturaleza*, Cuadernos de la Cátedra de Democracia y Derechos Humanos de la Universidad de Alcalá y el Defensor del Pueblo, nº14, 2018.

Sin profundizar demasiado en los argumentos a favor y en contra de nuevos derechos o sujetos jurídicos es interesante preguntarse acerca de catalogar este proceso como parte de la especificación de los derechos humanos.

Los derechos humanos tienen como fin la consecución y garantía de la dignidad humana, esto es, entender al ser humano como fin en sí mismo y, por lo tanto, un ser con derecho a unos elementos mínimos que posibiliten una vida digna. Es por ello por lo que se le reconoce al ser humano el derecho a la vida, a una vivienda, a una educación, vivienda, vestimenta, se prohíben los malos tratos, el sufrimiento y la tortura y se le blindan una serie de derechos sobre su propiedad y libertad a la hora de pensar, expresarse y perseguir ideas.

Siguiendo la clásica diferenciación de las etapas de surgimiento de los derechos humanos), el proceso de especificación viene a poner de manifiesto la concreción o aportación de nuevos elementos que matizan lo anterior. Este proceso puede hacer referencia tanto al reconocimiento de nuevos sujetos como de nuevos derechos[8]. Actualmente, la concreción y aportación a la que hacemos referencia ha puesto el foco de actuación únicamente en el género humano. Así se han reconocido derechos a las mujeres, los niños o grupos étnicos, entre otros. Hablar, por lo tanto, de incluir en este proceso tanto a sujetos no humanos como derechos cuyo contenido no incluye al ser humano resulta cuanto menos innovador y polémico.

El camino más básico de aceptar esta tesis conlleva que el reconocimiento de derechos que enriquezcan tanto la vida de los animales como el cuidado del medio ambiente confluyen en optimizar la vida del ser humano. En otras palabras, la ausencia de protección jurídica al medio ambiente o a los animales no humanos influye directamente en la degradación del medio donde habita el ser humano y, consiguientemente, nuestra vida se ve perjudicada. Esta protección forma parte del proceso de especificación de los derechos humanos[9]. Detrás de esta simplificación del debate encontramos muchos frentes abiertos a la hora de la considerar a la naturaleza y los animales como sujetos: si son un medio para el ser humano o un fin en sí mismo, si se puede hablar de dignidad en alguno de los dos sujetos, si tenemos que hablar propiamente de derechos o únicamente de deberes para el ser humano, si algo que no siente puede tener dignidad y, por lo tanto, derechos o si la dignidad puede graduarse.

[8] Vanesa Morente Parra, *Nuevos retos biotecnológicos para viejos derechos fundamentales: la intimidad y la integridad personal*, Universidad Carlos III de Madrid, Instituto de Derechos Humanos Bartolomé de las Casas, 2011, Pg. 198.

[9] Gregorio Peces Barba, *Lecciones de derechos fundamentales*, 2004, Pg. 124.

A pesar de rescatar más adelante algunos de los mismos argumentos que suelen esgrimirse para la defensa de derechos o deberes con los animales y la naturaleza, en este artículo pretendo llevar el debate a un escenario que está a caballo entre la realidad y la ficción o, si se quiere, entre el presente y el futuro: los robots, la Inteligencia Artificial (IA en adelante) y la superación del *homo sapiens*.

2. Poshumanismo y Transhumanismo

Las posturas transhumanistas y poshumanistas han abierto el camino hacia una sociedad futura en la que todos los pilares que conforman nuestra sociedad (jurídicos, sociales, políticos, económicos, laborales, de ocio o personales) se ponen en duda. Simplificadamente, desde estas perspectivas se trabaja con un cambio sustancial en la naturaleza del *homo sapiens*. Los avances en biotecnología, nanotecnología e IA llevarán consigo una mejora en las capacidades del ser humano y este superará su estado evolutivo dejando atrás al *homo sapiens*[10].

Todo ello no se centra únicamente en el ser humano, sino que pone de relieve el aumento cualitativo de la capacidad de las máquinas para formar parte de la sociedad. Los avances en IA permitirán a las máquinas no solo igualar sino superar la inteligencia del ser humano, logrando así un desarrollo de la sociedad nunca visto[11].

A día de hoy el trabajo es el ámbito de la sociedad donde mayor determinación han tenido las máquinas. A fin de cuentas, han sido creadas comúnmente para facilitar o mejorar el trabajo. Ligado al desarrollo de los robots y sus habilidades para desempeñar ciertos trabajos ya podemos encontrar artículos que se preguntan acerca de la cotización de estos en la seguridad social[12].

Entre los precedentes que marcan estas preguntas encontramos las resoluciones de la Unión Europea acerca de las repercusiones en el mercado laboral o la apuesta por una personalidad electrónica[13] o los estudios acerca de la desaparición de puestos de trabajos gracias a la automatización[14].

[10] Fernando H. Llano Alonso, *Homo excelsior. Los limites ético-jurídicos del transhumanismo*, 2018, Pg. 25.

[11] Nick Bostrom, *Superinteligencia. Caminos, peligros, estrategias*, 2016, Pg. 93.

[12] Sergio Saiz, *¿Deben cotizar los robots como si fueran trabajadores?*, Expansión, 2016; Carlos Javier Galán, *¿Deben cotizar los robots a la seguridad social?*, El País, 2019; El confidencial, Los robots deben cotizar a la Seguridad Social, UGT, 2019.

[13] Parlamento Europeo Resolución del Parlamento Europeo, de 16 de febrero de 2017, con recomendaciones destinadas a la Comisión sobre normas de Derecho civil sobre robótica, 2017.

[14] Frey Carl Benedikt & Michael Osborne, *The Future of Employment*, Oxford, 2013.

Si bien aceptamos esa realidad en la que cerca de un 50% de los trabajos actuales serán realizados por robots y a su vez, estas máquinas pasan a cotizar en la seguridad social ¿Tendrán derecho a todo aquello que la seguridad social de un estado concede?

El artículo 22 de la Declaración Universal de los Derechos Humanos de 1948 (DUDDHH en adelante) dice lo siguiente:

"Toda persona, como miembro de la sociedad, tiene derecho a la seguridad social, y a obtener, mediante el esfuerzo nacional y la cooperación internacional, habida cuenta de la organización y los recursos de cada Estado, la satisfacción de los derechos económicos, sociales y culturales, indispensables a su dignidad y al libre desarrollo de su personalidad."

Entendemos la Seguridad Social no únicamente como un deber de los ciudadanos a cotizar en un estado por su trabajo sino como un derecho a las prestaciones y beneficios que la seguridad social provee a las personas, todo lo referente a los derechos económicos sociales y culturales, véase la educación, sanidad, seguridad, infraestructuras, justicia, etc.

En ese sentido, si reconocemos la personalidad electrónica de la que habla la UE y sentamos las bases para una cotización de las máquinas en un puesto de trabajo ¿Deberíamos también reconocerles ciertos derechos si están contribuyendo al enriquecimiento y desarrollo del Estado?

La respuesta parece evidente y va en línea a qué derechos se le deben reconocer a los animales. Al igual que no tiene sentido reconocer ciertos derechos a los animales tampoco tiene hacerlo con las máquinas. Hablar de un derecho a la sanidad para los robots, a un proceso justo, a una vivienda digna o a una educación que desarrolle la personalidad carece de sentido.

Además, tampoco acaba de convencer un derecho a la vida a las máquinas dado que carecen tanto de la capacidad de sentir como de tener conciencia de sí mismos.

No obstante, ¿Qué ocurre si los avances a los que hemos hecho referencia anteriormente consiguen dotar a las máquinas de lenguaje, comprensión, conciencia, sentimientos y, en definitiva, adaptabilidad a la vida del ser humano? ¿Qué ocurre si la IA consigue desarrollar robots con características similares al ser humano, tanto en apariencia como en la forma de comportarse? En definitiva, ¿la creación de robots humanizados posibilitaría el reconocimiento de derechos a los robots? ¿Podríamos hablar de dignidad en los robots?

3. El camino hacia la creación

En 2011, Yuval Noah Harari público un *best seller* internacional con más de 21 millones de copias vendidas. Este libro, tal y como sitúa el subtítulo, pretende

ser una "breve historia de la humanidad". Para ello, Harari lo estructuró a través de tres grandes revoluciones: cognitiva, agrícola y científica. Es curioso situar la influencia religiosa que reside en este esquema histórico. La revolución cognitiva guarda relación con el mito de Adán y Eva junto al árbol del conocimiento y la sabiduría, la revolución agrícola no es más que una relectura del mito de Caín y Abel y la revolución científica se equipara al mito de la torre de Babel, donde los seres humanos intentaron construir una torre que llegase al cielo con la intención de convertirse en dioses[15].

En esto último sintetiza su obra Harari: los seres humanos, gracias a los avances tecnológicos están convirtiéndose en dioses y llegando al final del *homo sapiens*. Esta idea fue desarrollada en 2015 con la publicación de Homo Deus. En este libro, el mismo autor presenta una imagen del futuro que deja atrás muchas de las limitaciones biológicas del homo sapiens. Esto, en consonancia con la IA favorecerá a un nuevo Homo desarrollado en base a los últimos avances científicos.

Otros autores como Nick Bostrom (2016), Max Tegmark (2017) o Ray Kurzweil (2020) han trabajado acerca de un futuro similar. El común denominador que encontramos en sus obras es la creación de una IA superior a la del ser humano, concretamente, una explosión de Inteligencia por parte de las máquinas que consiga eclipsar el conocimiento humano y, por lo tanto, se modifiquen los pilares de nuestras sociedades actuales. Es lo que se conoce como Singularidad.

En principio, esta explosión de inteligencia no afecta a nuestro planteamiento inicial. Bien se puede pensar que estos avances únicamente tendrán consecuencias en el plano puramente científico o, a lo sumo, en el ocio o el trabajo, actualizando estos entornos. No obstante, los autores ya mencionados describen un cambio radical en la vida humana tal y como la conocemos. Un cambio en las formas de hacer política, en la estructuración y funcionamiento de los Estados, de las relaciones internacionales, de la proyección de la economía, de atajar los problemas medioambientales o bélicos, de restructuración de las relaciones sociales, de ámbito laboral o incluso en nuestra toma de decisiones.

Todo ello, en consonancia al desarrollo de los robots dotados de IA, seres con la capacidad de comprender el comportamiento humano y mimetizarse en la sociedad.

[15] Ernesto Castro, *Yuval Noah Harari, tú antes molabas*, 2020, 59 min 35 s https://www.youtube.com/watch?v=OiykBRAxDyE&t=3650s

4. ¿Robots humanizados?

Gracias a la ciencia ficción podemos hacernos una idea del mundo con determinados avances que, a día de hoy, parecen imposibles: un futuro de vida en otros planetas, despertarse de un sueño y darse cuenta de que vivimos en una simulación informática, un apocalipsis en el que las maquinas aniquilen a los humanos, un futuro en el que la ciencia pueda predecir el futuro, viajes en el tiempo u otras historias que han dado lugar a libros, películas y series de ciencia ficción.

En 2017 se publicó una entrevista al Dr. Ishiguro, jefe del departamento de robótica de la universidad de Osaka. El profesor es famoso por haber "dado vida" a un doble de sí mismo y a un androide con apariencia de una mujer joven. En algunos casos esos androides han sido parte de películas y en otros se les ha utilizado para impartir conferencias como si de su creador se tratasen. Todo ello, reproduciendo aquello que previamente se les ha programado.

Al margen de estos usos, el doctor hace referencia a un futuro de convivencia con los androides, esto es, un escenario en el que robots humanizados hayan conseguido mimetizarse en la sociedad a través del desempeño de trabajos como el cuidado de las personas con ciertas necesidades o trabajos que no necesitan trabajo humano, sino que se caracterizan por su mecanización y repetición.

A su vez, los androides a los que se hace referencia no solo ocuparán estos puestos, sino que aprenderán a reproducir sentimientos humanos como la felicidad o la tristeza y se comportan como nosotros. Al escuchar la entrevista la primera pregunta que me vino a la cabeza fue ¿Para qué? Para que se necesitan crear robots idénticos al ser humano y capaces de reproducir nuestras formas de sentir o comunicarnos. Ishiguro lo justificó argumentando que la creación de estos robots era un paso más para conocer mejor al ser humano. Los límites de la ética y la investigación con humanos o animales impiden experimentar con animales o seres humanos: la posibilidad de tener algo idéntico al ser humano puede ayudarnos conocernos mejor.

Precisamente, uno de los argumentos de crear un doble idéntico de sí mismo fue la intención de observarse a sí mismo desde fuera, con total objetividad.

Estos argumentos están sujetos a algunas críticas. La principal radica en resaltar que, independientemente de crear una máquina idéntica al ser humano, con pelo humano y piel sintética, a pesar de que reproduzca idénticamente nuestros sentimientos y formas de comportamiento, las máquinas son una creación humana, creadas con materiales artificiales y programadas para actuar tal y como los han programado. Por lo tanto, ese robot será una proyección más del punto de vista humano.

Por otro lado, el robot afrontará situaciones con los conocimientos que le han sido dados o, a lo sumo, a través del aprendizaje en base a unas reglas mínimas de funcionamiento que también le han sido insertadas. En definitiva, se comportará tal y como sus creadores han permitido que sea, dentro de unos cálculos determinado. No existe creatividad o innovación fuera del marco para el que se crean.

No obstante, Ishiguro desarrolló esta idea años atrás en sus artículos académicos. La comprensión entre máquinas y humanos se da gracias al esfuerzo que los creadores de las primeras ponen en traducir el lenguaje informático a un lenguaje comprensible. Aquello que leemos en una pantalla o lo que escuchamos cuando nuestro GPS nos guía en la carretera no es más que un tipo de traductor que hace comprensible lo que la máquina procesa.

Para el profesor nipón el siguiente paso en esta relación androide-humano es acercar las traducciones máquina-humano dando una apariencia humana a determinados robots, sobre todo aquellos que desempeñen trabajos históricamente realizados por humanos, bien la hostelería, el transporte o bien el cuidado de personas. En base a la experimentación se ha constatado que el ser humano comprende y acepta mejor una relación con una máquina si esta comparte ciertas características. En primer lugar, fue el lenguaje y, en un futuro también será la apariencia y el comportamiento[16].

En definitiva, podemos decir que el objetivo de la creación de robots humanizados hace referencia a la adaptación que el ser humano vivirá en los próximos años a un medio plagado de IA y robots. Con el fin de pacificar esta adaptación o reducir las posibilidades de un rechazo generalizado, se cree que la empatía con estos sujetos, que comparten nuestras formas de actuación, será positiva.

Ahora bien, llegados a ese punto, en sociedad con androides humanizados desempeñando trabajos humanos y conviviendo con los seres humanos ¿Se deberá hablar a su vez de derechos para los robots?

5. Sociedad IA

La comprensión del mundo del hoy y del mañana depende del enfoque. Vida 3.0, era digital, régimen de la información, era del big data o tercera revolución industrial son algunos de los enunciados que alumbran la nueva sociedad en la que vivimos. Es común encontrar autores que defienden los beneficios que traerá la incorporación de la IA en la vida, tanto en las facetas cotidianas

[16] H. Ishiguro, *Building artificial humans to understand humans*, Journal of Artificial Organs.

de la sociedad como en los extremos más cercanos a la ciencia ficción[17]; otros investigadores, a la par de sorprenderse por los avances en robótica, nanotecnología e IA pretenden advertir y prepararse para las repercusiones que esto causará[18]; en otro punto, también podemos leer a aquellos que muestran preocupación por la degeneración de la democracia, las relaciones humanas, la igualdad o la libertad debido a estas nuevas revoluciones tecnológicas[19]; por último, es común también entender estos cambios a modo de un simple proceso histórico, natural, que se desarrollará a consecuencia del proceso económico y político de los últimos años[20].

Una perspectiva a tener en cuenta es aquella defendida por Marta Peirano (2019) o Shoshana Zuboff (2020). La información es hoy en día lo que maneja el poder, el control de los datos y la exposición continua de la sociedad a través de la electrónica (smartphone, ordenadores, tablets). Las formas de dominación tradicionales, la vigilancia o la coacción y represión de los organismos estatales se está viendo desplazada por la psicopolítica y la autoexpotación.

Nuestra exposición continua en la red, el tráfico de nuestros datos y la constante vigilancia a los demás se traduce en una sensación de libertad que abraza cada vez más un escenario aparentemente libre y neutral que nos permita consumir y elegir libremente.

Con todo, entre las consecuencias de este nuevo régimen, está la desinteriorización de las personas[21] o el efecto de conformidad, como si cada uno vigilara al otro. En otras palabras, la deshumanización del individuo o de la jurisdicción[22], la pertenencia a una sociedad global por medio de lo digital, una pérdida de lo corpóreo que se sustituye por una vinculación con lo ficticio. Esto tiene consecuencias en la democracia, en el trabajo, en el ocio y en el desarrollo personal. Se pasa de la biopolítica a la psicopolítica.

Pues bien, si a este régimen de la información, que incluye los avances en IA, le sumamos la posibilidad de convivir con robots humanizados que, a su vez reproduzcan nuestras pautas de comportamiento y ocupen espacios en la

[17] Rafael Yuste, Cuando ya no esté: Rafael Yuste (Parte 2/2), 2016 | #0 Entrevistado por Iñaki Gabilondo. Movistar. (2007) 2 min 55s

[18] A. Ortega, *La imparable marcha de los robots,* Alianza Editorial, 2016.

[19] Byung-Chul Han, *Infocracia. La digitalización y la crisis de la democracia*, Taurus, 2022.

[20] Yuval Noah Harari & Slavoj Žižek, *Should We Trust Nature More than Ourselves?,* Youtube, 2022.

[21] Byung-Chul Han, *Psicopolítica*, Herder, 2021, Pg. 21.

[22] Rafael De Asís, *Derechos humanos e Inteligencia Artificial*, Seminario Permanente Gregorio Peces-Barba, Universidad Carlos III de Madrid, Grupo de Investigación Derechos Humanos, Estado de Derecho y Democracia, 2020, Pg. 16.

sociedad como si de humanos se tratase, la deshumanización a la que hacemos referencia no hará más que acrecentarse.

La proyección que se hace de este tema en los derechos humanos no puede tener más consecuencias que la propia deshumanización de los mismos ligado a su banalización.

Como hemos situado al principio del artículo, el debate de reconocimiento a sujetos no humanos de ciertos derechos universales no solo puede hacer referencia a elementos intrínsecos en los propios sujetos, sino que tiene repercusiones en la vida del ser humano. En último término, ese reconocimiento es necesario para una vida digna en la tierra.

Lo que se está poniendo de relieve no es más que la creación de cosas con apariencia humana que escenifiquen la vida del ser humano. Nunca se podrá crear una figura que iguale al ser humano partiendo de una naturaleza no orgánica. Todo lo que se desarrolle en esa línea deberá tener un reproche moral, bien por superar los límites éticos que protegen la clonación o bien por ser un síntoma de la decadencia del propio ser humano.

6. El premio de la dignidad

La lucha por los derechos humanos se articula en un proceso histórico de cientos de años. Desde la Grecia clásica, pasando por el imperio Romano, la lectura cristiana de los derechos naturales, la revolución francesa y sus declaraciones hasta la positivación de la DUDDHH es un recorrido que lleva tras de sí un amplio debate teórico ligado a la práctica política.

Los derechos humanos son fruto de los procesos políticos, económicos y sociales que se han ido concatenando a lo largo de la historia y que culminaron con la imperante necesidad de acordar una lista mínima de derechos a nivel mundial, evitando o previniendo otro conflicto similar a la II Guerra Mundial[23]. Desde entonces, los organismos mundiales que nacieron a raíz de la DUDDHH, acuerdos, protocolos e instituciones han buscado su aplicación en todos los países de la tierra.

Esto último se debe, entre otras razones, a que todos los Estados reconocidos en el mundo han aceptado y firmado esta Declaración. Con ello, podemos afirmar que todos los Estados del mundo coinciden en que todo ser humano tiene dignidad y, en consecuencia, se deben cumplir todos los derechos humanos que posibilitan esa vida digna.

[23] Mauricio Iván Del Toro, *La Declaración Universal de Derechos Humanos: un texto multidimensional*, Comisión Nacional de los Derechos Humanos México, México D.F, 2012, Pg. 32.

No obstante, sobre esta visión optimista de los derechos humanos se han vertido críticas que desconfían de su teoría. En primer lugar, es posible observar una clara diferencia entre el discurso de los derechos humanos y el contenido. En ese sentido, la mayoría de la población mundial es sujeto del discurso de los derechos humanos, pero no de su contenido[24]. Es decir, el disfrute de estos derechos es una tarea pendiente en un amplio sector de la población mundial y, a pesar de formar parte de las constituciones de todos los países y de los acuerdos internacionales, aun no se han articulado propuestas prácticas para hacerlos cumplir.

En segundo lugar, desde la teoría decolonial se ha puesto de relieve que el elenco de derechos humanos aprobado en el 48 únicamente hace referencia al contexto occidental y que peca de un sesgo eurocéntrico. De este modo, a la hora de aplicar el contenido de esos derechos a realidades distintas brotan dificultades que imponen formas de actuar ajenas. Se extrapolan formas de gobierno, formulas económicas o modos de comportamiento de un contexto a otro, borrando así cualquier elemento que no case con los cánones europeos[25].

Una tercera visión contrahegemónica surge de los países de tradición islámica. En 1990 se firmó en El Cairo la Declaración Universal de Derechos Humanos en el islam. Este acuerdo fue dirigido por la Organización para la cooperación Islámica y parte de una clara influencia religiosa en su entendimiento de los derechos humanos. Entre los argumentos que la apoyan se encuentra la contraposición a la DUDDHH por su influencia judeocristiana que, si se intenta aplicar en sus países, se topará con graves problemas de índole cultural[26].

Por otro lado, en la misma línea de la crítica eurocéntrica, se suele calificar a la DUDDHH como un instrumento que beneficia a occidente para imponer sus políticas en todo el mundo. De esta forma, occidente utiliza los derechos humanos para derrocar gobiernos, invadir países y dominar la política internacional[27].

Incluso las posturas que apoyan y aceptan los derechos humanos lanzan una crítica a su lenta implementación en el mundo. Estas posturas se han multiplicado a raíz de la crisis del Covid-19 y sus consecuencias en todo lo relacionado con los derechos humanos[28].

[24] Boaventura De Sousa Santos, *Derechos humanos, democracia y desarrollo*, Colección Dejusticia, 2014, Pg. 23.

[25] Enrique Dussel, *Filosofías del Sur. Descolonización y Transmodernidad*, Akal, 2015, Pg. 92.

[26] Erma Ivan Carrazco Nuñez, *Derechos humanos en el Islam. Una perspectiva comparada*, Revista de Relaciones Internacionales de la UNAM, 2018, Pg. 9.

[27] S. Žižek, "Contra los derechos humanos", New Left Review, 34, 2005.

[28] Amnistía Internacional, *La situación de los derechos humanos en el mundo*, 2021.

Con todo ello, en este contexto de lucha argumental que hace tambalear los pilares de los derechos humanos, que desconfía de sus ideas principales e incluso las ataca, han ido brotando debates paralelos que buscan engordar la lista de los mismos. Ya contamos con una tercera generación de derechos humanos que incide en sujetos cada vez más abstractos (naturaleza, minorías o la paz). Algunos estudios ya hablan de una cuarta generación y al calor de estas discusiones han ganado peso las posturas en favor de los derechos de los animales.

Cada uno de estos debates, al margen de beber de una voluntad optimista y humanitaria, desplazan el foco de los derechos humanos a debates teóricos en vez de focalizar la práctica de lo ya reconocido. En este orden, es posible que de cumplirse los derechos de la DUDDHH se tornase necesario el cuidado de otros ámbitos de la vida y, por lo tanto, no fuese necesario un debate que eclipsase la faceta activa de los derechos humanos.

Ahora bien, podemos aceptar que un reconocimiento exhaustivo de los derechos humanos de la naturaleza, de los animales, derechos a la paz, al desarrollo, a internet o a una democracia, no hacen más que sumar para conseguir una humanidad digna donde los derechos humanos sean una realidad generalizada. No obstante, mientras este camino se desarrolla con sus infinitas trabas teóricas y prácticas parece que nace un debate que sobrevuela la teoría de los derechos humanos, un debate que pretende ser fruto de la era digital y que este ligado a la era de las innovaciones tecnológicas, esto es, si debemos brindar derechos a las máquinas o si son necesarios nuevos derechos humanos que atajen las consecuencias de la neurotecnología[29].

En esta misma línea, las corrientes posthumanistas y transhumanistas persiguen el horizonte superador del ser humano y con él, el fin de todos los problemas mundanos. La solución ya no pasa por garantizar aquello que todo el mundo reconoce, por erradicar los males económicos del mundo y por garantizar vidas dignas. Su hipótesis no es más que evolucionar y dejar atrás el homo sapiens, el ser humano y, por lo tanto, los problemas ya no serán los de este ser sino otros aun por conocer. En definitiva, si el ser humano no existe tampoco existen los problemas de este.

En el camino hacia la posthumanidad nacerán debates interconectados con esas propuestas, véase los derechos de aquellos seres humanos que quieran

[29] Rafael De Asís, *Sobre la propuesta de los Neuroderechos,* DERECHOS Y LIBERTADES, Número 47, 2022, Pg. 5. Nuria Belloso Martín, *A vueltas con la igualdad y la no-discriminación en los neuroderechos. El reto del acceso equitativo a la mejora cognitiva*, en: Clovis Gorzevski (Organizador), *Direitos Humanos e participacâo política*, Volumen XIII, Porto Alegre (Brasil), Editora Imprensa Livre, 2022.

modificar su cuerpo con innovaciones tecnológicas, conocidos como ciborgs o también los derechos de las máquinas dotadas de cierta autonomía y capacidad de actuación: los robots.

7. Derechos de los robots

El Derecho como rama del conocimiento se suele subdividir en áreas que analizan situaciones concretas ligadas a ciertos ámbitos de la vida. El derecho laboral centra su estudio en las relaciones que se producen en el mercado del trabajo. El derecho penal articula los principios, normas y penas que se desligan de las conductas criminales dentro de un estado. Lo mismo ocurre con el Derecho Civil, Tributario o Internacional.

Respecto a los robots y su regulación ya se han pronunciado algunos estudiosos de la materia vislumbrando la inminente construcción de un área del derecho relativa a los robots[30]. El derecho de los robots será un área más del Derecho con una tarea unificadora e innovadora en tanto en cuanto deberá centrar unos principios sólidos que casen con los pilares de las sociedades en las que se aplique. Además, deberá mantener en todo momento una vista hacia el frente dado que día a día se evoluciona en lo relativo a la tecnología.

En este sentido, el derecho de los robots se entiende no tanto en atribuir derechos equiparables a los garantizados a los seres humanos sino al procedimiento y normas que vinculan el fabricante, productor o poseedor de un robot a la hora de juzgar un hecho en el que este último ha sido participe. Se habla entonces de responsabilidad de daños, de materiales defectuosos, responsabilidad en accidentes o todo aquello relativo al mal uso de un robot para el desempeño de una actividad. Abarcaría también todo lo relacionado con la propiedad intelectual o de fabricación de aquello que ha sido creado por un ente robótico; lo relacionado con el ámbito económico y financiero; o aquello que se desprende de la practica sanitaria, en particular todas aquellas situaciones en las que participe un robot y este en juego la integridad de un sujeto humano.

Ahora bien, el área de estudio y práctica de esta nueva área jurídica debe tener claras sus líneas de actuación y las fronteras que lo delimitan. Un peligro que puede derivar de este nuevo estudio es la pretensión de reconocer a estas máquinas derechos humanos o, en última instancia, dignidad[31].

Desde el posthumanismo y transhumanimo ya se han dado las primeras pinceladas a un reconocimiento de esta naturaleza. Se entiende que de ser creado un robot con apariencia humana que además se mimetice con los seres

[30] Andrés M. Barrio, *Derecho de los Robots*, Madrid, La Ley, 2019.
[31] Manuel Atienza, *Sobre la dignidad*, Trotta, 2022.

humanos de una sociedad, no existirían razones para no reconocerle un estatus equiparable al ser humano.

En este sentido, el catálogo de derechos humanos reconocidos hasta ahora deberían reformularse o, por lo menos, adaptarse a estas nuevas vicisitudes tecnológicas. Con todo ello, junto con el debate aun abierto sobre el reconocimiento de derechos a sujetos no humanos como la Naturaleza o los animales, encontraríamos los derechos a los robots, reconociéndoles tanto como sujetos autónomos y como sujetos de derechos.

Para entrar en lo que a mi modo de ver sería el reto al que se enfrentarían los derechos humanos de cumplirse esta situación es necesario matizar la veracidad de lo que se plantea.

En lo relativo a los debates que estudian los posibles derechos de los robots se suele advertir de que esta realidad aún no se ha consumado y, por lo tanto, aún se está al calor de lo que las innovaciones tecnológicas vayan a desarrollar. Hoy en día no contamos con robots que coticen a la seguridad social, no contamos con robots culpables en accidentes con humanos ni mucho menos contamos con robots que tengan reconocidos derechos humanos. No obstante, si es posible encontrar ejemplos de robots que han comenzado a recorrer este camino, aunque solo sea con fines mediáticos[32].

Con todo, el estudio que presentamos a continuación no tiene que ver con ningún hecho objetivo que este tomando relevancia en la actualidad, sino que se basa en las obras de autores que vaticinan un futuro de convivencia con seres no vivos.

8. Banalización de los derechos humanos

Actualmente, según OXFAM, 155 millones de personas viven una situación de crisis alimentaria; 783 millones de personas viven bajo el umbral de la pobreza (1.90 dólares diarios)[33] y los niños que nacen pobres tienen el doble de posibilidades de morir antes de los 5 años. No obstante, se está buscando regular un nuevo derecho humano con relación en la equidad a la hora de acceder a la mejora de las capacidades cerebrales[34].

La ONU ha puesto de relieve que "la brecha salarial de género en todo el mundo se sitúa en el 23%[35]. Aun así, el debate acerca del aumento de dere-

[32] Redacción Sophia, la robot que tiene más derechos que las mujeres en Arabia Saudita, BBC, 2017.
[33] OXFAM, El virus del hambre se multiplica. Conflictos, Covid-19 y cambio climático: una combinación mortal que agrava el hambre en el mundo, 2021.
[34] R. Yuste, J. Genser y S. Herrmann, "It's Time for Neuro-Rights", HORIZONS, 2021, Pg. 5.
[35] ONU, ONU Mujeres afirma que la brecha salarial del 23% entre mujeres y hombres es un robo, 2017.

chos humanos en relación con los avances en la era digital se está centrando en la no discriminación a partir de los datos obtenidos por medio de la neurotecnología.

En 2007 se recogió el derecho a la democracia igualitaria, plural, participativa y demás atribuciones en la Declaración Universal de Derechos Emergentes. Según un estudio del *V-DEM institute*, de la universidad de Gotemburgo los Estados que cuentan con democracias liberales han disminuido en los últimos años y más de la mitad de la población mundial no vive en democracias[36]. Asimismo, se busca positivar un derecho humano que garantice a las personas tomar decisiones libremente, es decir, sin manipulación neurotecnológica.

Estos son tres de los cinco Neuroderechos propuestos para su reconocimiento como derechos humanos en 2021. Entre sus defensores encontramos a Rafael Yuste, Marcello Ienca y Wrye Sententia (científicos) o J. Genser y S. Herrmann (juristas).

Entre los detractores de este proyecto encontramos doctores de la Universidad de Valparaiso, Chile, que ven en estos intentos de nuevos derechos humanos "cautelar nuevas amenazas a viejos derechos humanos". Es decir, no podemos hablar todavía de una nueva realidad que reclame para si nuevos derechos, sino que estamos ante una actualización de los peligros que pueden tambalear los derechos ya reconocidos. A modo de ejemplo "que surjan nuevas formas de matar no altera el contenido del derecho a la vida ni es fundamento para la creación jurídica de nuevos derechos"[37].

Otras criticas hacen referencia a la inflación en materia de derechos humanos, a la "incompetencia jurídica" de quienes promueven la propuesta, "la falta de debate académico", la infravaloración del potencial de los derechos humanos, "la falta de comprobación de la hipótesis de partida" o no haber demostrado aun los riesgos que acarrea el desarrollo de lo "neuro"[38].

A estas contundentes críticas podemos sumar la no garantía de los derechos ya reconocidos. En ese sentido, cuáles son los motivos de buscar una nueva regulación de derechos cuando ni siquiera podemos hablar de un cumplimiento de los ya existentes. No solo eso, sino que ese cumplimiento dista entre países, acrecentándose la diferencia norte-sur.

[36] Vanessa A. Boese, Nazifa Alizada, Martin Lundstedt, Kelly Morrison, Natalia Natsika, Yuko Sato, Hugo Tai, and Staffan I. Lindberg, *Autocratization Changing Nature? Democracy Report*, Varieties of Democracy Institute (V-Dem), 2022, Pg. 12.

[37] Alejandra Zúñiga Fajauri, Luis Villavicencio Miranda, Ricardo Salas Venegas, *¿Neuroderechos? Razones para no legislar*, CIPER/Academia, 2020.

[38] De Asís Sobre la propuesta de los Neuroderechos. DERECHOS Y LIBERTADES. Número 47. (2022) Pg. 14.

A la hora de abordar el debate es necesario fijar el ámbito y la proyección en la que nos adentramos. El problema no es, en este caso, la búsqueda e implementación de nuevos derechos o nuevas herramientas que permitan encajar en las sociedades avanzadas las repercusiones de las innovaciones tecnológicas. El problema radica en articular todo ello bajo la lógica de los derechos humanos, cuyo contenido construye sus cimientos en unos principios universales que puedan extenderse a todos los seres humanos.

Es interesante dar respuesta al encaje de la era digital en la cotidianeidad de la vida. Con ello, tarde o temprano será necesaria una regulación que mida los avances positivos y negativos de las tecnologías, en relación a los territorios donde se produzca, incluido todo lo relacionado con la neurociencia.

Alpha Condé, expresidente de Guinea, expuso en Davos qué "**algunas de las cosas que los países occidentales dan por sentado, como el alumbrado público o la alimentación para cargar el portátil o un teléfono inteligente en el que estes leyendo un artículo, demasiado a menudo son lujos en los países en desarrollo". Asimismo, en los países africanos se está buscando superar "**problemas más fundamentales, como la pobreza, la falta de energía o infraestructura"[39].

Los problemas que puso de relieve el expresidente guineano tienen que ver con derechos recogidos en la Declaración Universal de Derechos Humanos. Los problemas a los que atienden los Neuroderechos guardan relación con desafíos de las sociedades occidentales y, dentro de estas, con aquellas clases sociales que pueden estar al alcance de las mejoras tecnológicas.

En el mismo sentido, anualmente se "pasa revisión" a la garantía de los derechos humanos en el mundo y a los esfuerzos que ponen los países en conseguir implantar unas mejoras condiciones de vida en sus fronteras. Esto se hace a través de informes e indicadores con la supervisión de la ONU. Resulta complejo buscar que países o porcentajes de la población serán sujetos de disfrute de las mejoras tecnológicas, quienes accederán a las modificaciones cerebrales o quienes sufrirán la manipulación neurotecnológica.

Si ponemos la lupa en las obras posthumanistas resulta aún más difícil determinar quienes podrán superar la muerte o conseguir evolucionar del *Homo sapiens*.

Si desde los sectores más avanzados de las sociedades desarrolladas se avanza en poner de relieve la imperante necesidad de regular, a nivel universal, las mejoras tecnológicas cobrarán sentido el "yo europeo", donde Europa

[39] Alpha Condé, *How the technology revolution will transform Africa*, World Economic Forum, DAVOS, 2016.

"pretende descubrirse a sí mismo como universal, ultimo, que se sabe a sí mismo, y que puede reconstruir desde el mismo todo el mundo"[40].

Sin duda, la articulación de normas que adapten las innovaciones tecnológicas a la sociedad debe estar precedidas por el enfoque de los derechos humanos[41]. Esta defensa argumental demuestra la generalizada insuficiencia de los derechos humanos en el mundo. Tal y como se ha apuntado anteriormente todos los Estados reconocidos han suscrito la DUDDHH y los acuerdos internacionales que buscan su cumplimiento. Resulta lógico por lo tanto que todo aquello que se desarrolle en el marco legal de un Estado busque el compromiso y garantía de los derechos humanos y que, en última instancia, son estos derechos universales los que priman.

Resaltar la necesidad del enfoque de los derechos humanos significa que a priori no se están teniendo en cuenta estos para el desarrollo de las tecnologías. Poner de relieve que existen derechos humanos que pueden verse afectados con la aparición de nuevas tecnologías significa que no se han tenido en cuenta en el proceso de desarrollo. Recalcar que existen derechos humanos que ya protegen aspectos de la vida que pueden verse dañados por los usos de la IA o la neurotecnología significa que la DUDDHH no es tenida en cuenta cuando se habla de proteger al ser humano.

Todo esto sin duda es preocupante para la vigencia de los derechos humanos básicos que tan cuidadosamente han servido para atender a todos los seres humanos. A través de un contenido mínimo, universal y, en parte, abstracto han conseguido que cada uno de sus derechos sea aplicable a cada uno de los seres humanos. Es ahí donde reside su riqueza jurídica y donde se asientan los Estados contemporáneos[42].

Las innovaciones tecnológicas, la aplicación de la IA a cada vez más facetas de la vida o las predicciones de sociedades humano-robóticas encierran un reto para los derechos humanos. Lejos de pretender un análisis apocalíptico, no estamos hablando de una sociedad superada por las maquinas sino una sociedad en la que los pilares teóricos y los principios básicos de las normas queden obsoletos por el gran avance de la tecnociencia.

La consecuencia de desarrollar y avanzar en estas materias primero y, preguntarse después que consecuencias tendrán en nuestras vidas no hace más que quitar importancia a los grandes valores que edifican los derechos huma-

[40] Enrique Dussel, *Filosofías del Sur. Descolonización y Transmodernidad*, Akal, 2015, Pg. 90.
[41] Rafael De Asís, *Una mirada a la robótica desde los derechos humanos*, Madrid, Dykinson, 2014.
[42] Mauricio Iván Del Toro, *La Declaración Universal de Derechos Humanos: un texto multidimensional*, Comisión Nacional de los Derechos Humanos México, México D.F., 2012, Pg. 99.

nos, en definitiva, permanecer a la zafa de la IA, la robótica y su inclusión en la sociedad no hace más que banalizar tanto los derechos humanos como el potencial jurídico de los estados.

La banalización de un mal, como en este caso puede ser la banalización de no tener en cuenta los derechos humanos, no tiene lugar de la noche a la mañana, no tiene por qué percibirse en un hecho concreto o a través de un sujeto determinado, tiene que ver con un proceso paulatino, generalizado y aceptado por cada sociedad. Un proceso que actúa en cada capa de la misma y termina por calar en nuestro día a día dando lugar a un desprecio o, aún peor, a un olvido de los pilares fundamentales de la sociedad.

Bibliografía

Amnistía Internacional, *La situación de los derechos humanos en el mundo*, https://www.amnesty.org/es/documents/pol10/3202/2021/es/, 2021, Consultado 23/09/2022.

Asamblea General de la ONU, *Declaración Universal de los Derechos Humanos*, Paris, 1948.

Atienza, Manuel, *Sobre la dignidad*, Trotta, 2022.

Barrio, Andrés M., *Derecho de los Robots*, Madrid, La Ley, 2018.

Benedikt, Frey Carl & Osborne, Michael, *The Future of Employment*, Oxford Martin Programme on Technology and Employment, 2013.

Belloso Martín, Nuria, *El debate sobre la tutela institucional: generaciones futuras y derechos de la naturaleza*, Cuadernos de la Cátedra de Democracia y Derechos Humanos de la Universidad de Alcalá y el Defensor del Pueblo, nº14, 2018.

—, *Un intento de fundamentar derechos de los no-humanos (derechos de la Naturaleza) a partir del desarrollo sostenible*. Revista Catalana de Dret Ambiental (RCDA), 2022, vol. 13, nº1 junio, Tarragona, Universitat Rovira i Virgili, pp. 1-46.

—, *A vueltas con la igualdad y la no-discriminación en los neuroderechos. El reto del acceso equitativo a la mejora cognitiva,* en: Clovis Gorzevski (Organizador), *Direitos Humanos e participacâo política*, Volumen XIII, Porto Alegre (Brasil), Editora Imprensa Livre, 2022.

Bostrom, Nick, *Superinteligencia. Caminos, peligros, estrategias*, TEEl, 2016.

Carrazco, Nuñez, Erma, Ivan, *Derechos humanos en el Islam. Una perspectiva comparad,* Revista de Relaciones Internacionales de la UNAM, núm. 132, septiembre-diciembre de 2018, pp. 93-121.

Castro, Ernesto, *Yuval Noah Harari, tú antes molabas*, 2020, https://www.youtube.com/watch?v=OiykBRAxDyE&t=3650s, Consultado 22/09/2022.

Condé, Alpha, *How the technology revolution will transform Africa*, World Economic Forum, DAVOS, 2016, https://www.weforum.org/agenda/2016/01/what-does-the-fourth--industrial-revolution-mean-for-africa/, Consultado 23/09/2022.

Cortina, Adela, *Las fronteras de la persona: el valor de los animales, la dignidad de los humanos*, Taurus, 2009.

De Asís, Rafael, *Una mirada a la robótica desde los derechos humanos*, Madrid, Dykinson, 2014.

—, *Derechos humanos e Inteligencia Artificial*, Seminario Permanente Gregorio Peces-Barba, Universidad Carlos III de Madrid, Grupo de Investigación Derechos Humanos, Estado de Derecho y Democracia, 2020.

—, *Sobre la propuesta de los Neuroderechos*, DERECHOS Y LIBERTADES, Número 47, Época II, junio 2022, pp. 51-70.

De Sousa Santos, Boaventura, *Derechos humanos, democracia y desarrollo*, Colección Dejusticia, 2014.

Del Toro, Mauricio Iván, *La Declaración Universal de Derechos Humanos: un texto multidimensional*, Comisión Nacional de los Derechos Humanos México, México D.F., 2012.

Declaración de los Derechos Humanos en el Islam, Conferencia Islámica de El Cairo, Disponible en https://www.refworld.org/cgi-bin/texis/vtx/rwmain/opendocpdf.pdf?reldoc=y&docid=50acbf1c2, 1990, Consultado el 23/09/2022.

DerechoUBA, *Sobre la dignidad*, https://youtu.be/CH09AaKTCXY, 2021, Consultado 23/09/2022.

Dussel, Enrique, *Filosofías del Sur. Descolonización y Transmodernidad*, Akal, 2015.

El confidencial Los robots deben cotizar a la Seguridad Social. UGT. https://www.ugt.es/los-robots-deben-cotizar-la-seguridad-social (25/02/2019). Consultado 22/09/2022.

Galan, Carlos Javier, *¿Deben cotizar los robots a la seguridad social?*, El País, https://elpais.com/retina/2019/04/12/tendencias/1555063168_443364.html (15/04/2019), Consultado 22/09/2022.

Han, Byung-Chul, *Infocracia. La digitalización y la crisis de la democracia*, Taurus, 2022.

—, *Psicopolítica*, Herder, 2021.

Horta, Oscar, "*Términos básicos para el análisis del especismo*" en González, Marta I., Riechmann, Jorge, Rodríguez Carreño, Jimena y Tafalla, Marta (coords.), *Razonar y actuar en defensa de los animales*, Los libros de la catarata, Madrid, 2008, 107-118.

Ishiguro H., *Building artificial humans to understand humans,* Journal of Artificial Organs, 2007.

Ishiguro H., *Cuando ya no esté: Dr. Ishiguro, el Quijote de la robótica*. Entrevistado por Iñaki Gabilondo, Movistar, 2017, https://www.youtube.com/watch?v=e4D99prZlYg, Consultado a 22/09/2022.

Kurzweil, Ray, *La singularidad está cerca*, Lola Books, 2020.

Llano Alonso, Fernando H., *Homo excelsior. Los limites ético-jurídicos del transhumanismo,* Tirant lo blanch, 2018.

Noah Harari, Yuval, *Sapiens. Una breve historia de la humanidad*, Debate, 2016.

—, *Homo Deus. Breve historia del mañana*, Debate, 2020.

Morente Parra, Vanesa, *Nuevos retos biotecnológicos para viejos derechos fundamentales: la intimidad y la integridad personal,* Universidad Carlos III de Madrid, Instituto de Derechos Humanos Bartolomé de las Casas, 2011.

Noah Harari, Yuval & Žižek, Slavoj, *Slavoj Žižek & Yuval Noah Harari | Should We Trust Nature More than Ourselves?*, Youtube, https://youtu.be/3jjRq-CW1dc, 2022, Consultado a 23/09/2022.

Nussbaum, Martha C, *Las fronteras de la justicia. Consideraciones sobre la exclusión*, Paidós, 2021.

ONU, *ONU Mujeres afirma que la brecha salarial del 23% entre mujeres y hombres es un robo*, https://www.un.org/sustainabledevelopment/es/2017/03/onu-mujeres-afirma-que--la-brecha-salarial-del-23-entre-mujeres-y-hombres-es-un-robo/, 2017, Consultado 23/09/2022.

Ortega, A., *La imparable marcha de los robots*, Alianza Editorial, 2016.

OXFAM, El virus del hambre se multiplica. Conflictos, Covid-19 y cambio climático: una combinación mortal que agrava el hambre en el mundo https://lac.oxfam.org/latest/press-release/el-virus-del-hambre-se-multiplica-conflictos-covid-19-y-cambio--clim%C3%A1tico-una, 2021, Consultado 23/09/2022.

Parlamento Europeo, Resolución del Parlamento Europeo, de 16 de febrero de 2017, con recomendaciones destinadas a la Comisión sobre normas de Derecho civil sobre robótica, Estrasburgo, https://www.europarl.europa.eu/doceo/document/TA-8-2017-0051_ES.html?redirect#title1, 2017, Consultado 22/09/2022.

Peces Barba, Gregorio, *Lecciones de derechos fundamentales*, 1º ed, Dykinson, 2004.

Peirano, M., *El enemigo conoce el sistema: Manipulación de ideas, personas e influencias después de la economía de la atención*, Debate, 2019.

Redacción Sophia, la robot que tiene más derechos que las mujeres en Arabia Saudita. BBC https://www.bbc.com/mundo/noticias-41803576, 2017, Consultado a 23/09/2022.

Saiz, Sergio, *¿Deben cotizar los robots como si fueran trabajadores?*, Expansión, https://www.expansion.com/juridico/actualidad-tendencias/2016/12/26/585d681aca4741ec378b45e4.html (26/12/2016), Consultado 22/09/2022.

Singer, P., *Liberación animal: el clásico definitivo del movimiento animalista*, Taurus, 2018.

Tegmark, Max, *Vida 3.0. Qué significa ser humano en la era de la Inteligencia Artificial*, Taurus 2017.

Vanessa A. Boese, Nazifa Alizada, Martin Lundstedt, Kelly Morrison, Natalia Natsika, Yuko Sato, Hugo Tai, and Staffan I. Lindberg, *Autocratization Changing Nature? Democracy Report*, Varieties of Democracy Institute (V-Dem), 2022.

Yuste, Rafael, Cuando ya no esté: Rafael Yuste (Parte 2/2) | #0 Entrevistado por Iñaki Gabilondo, Movistar, https://youtu.be/XnzUkYfeU4g Consultado 23/09/2022, 2016.

Yuste, R., Genser J. y Herrmann S., "*It's Time for Neuro-Rights*", HORIZONS, Winter 2021 / Issue Nº 18, 2021.

Žižek, S., "*Contra los derechos humanos*", *New Left Review*, 34, 85-100. https://newleftreview.es/issues/34/articles/slavoj-zizek-contra-los-derechos-humanos.pdf, 2005, Consultado el 23/09/2022.

Zuboff, S., *Capitalismo de la vigilancia*, DEBATE, 2020.

Zúñiga, Fajauri Alejandra, Villavicencio, Miranda Luis, Salas Venegas, Ricardo, *¿Neuroderechos? Razones para no legislar,* CIPER/Academia, 2020.

IA e Robótica: a caminho da personalidade jurídica?[1]

IA & Robotics: towards legal personality?

Sónia Moreira[*]

RESUMO: A Inteligência Artificial pode definir-se como um ramo das ciências da computação que visa dotar um agente de *software* de capacidade para receber estímulos externos do seu meio ambiente (dados) para resolver determinado problema de forma autónoma, ou seja, sem intervenção humana. Para tanto, é criado um código, através da elaboração de algoritmos, que vão determinar a forma de actuar do agente de *software*.

Este agente de *software* pode ser dotado de capacidade de autoaprendizagem (*machine learning* ou mesmo *deep learning*), que lhe permitirá ir além da sua programação inicial, podendo até, eventualmente, tomar decisões para as quais não foi programado.

Esta autonomia pode fazer-nos levantar as mais variadas questões: pode um agente autónomo ser considerado imputável? Pode ser responsabilizado caso tome uma decisão ou actue no mundo físico e venha a provocar danos a alguém? Podem os agentes autónomos possuir estados intencionais? E, mais polémico, pode defender-se a atribuição de personalidade jurídica a agentes autónomos?

PALAVRAS-CHAVE: Inteligência Artificial; *Robots*; Personalidade Jurídica

[*] Profª Auxiliar da Escola de Direito da Universidade de Minho. Investigadora do JusGov.

[1] A versão inglesa do presente texto foi entregue para publicação no Yearbook do E-TEC de 2022, disponível em https://www.jusgov.uminho.pt/pt-pt/publicacoes/anuario-etec-2020-2-2/ [consultado a 17/07/2023].

ABSTRACT: Artificial Intelligence can be defined as a branch of computer science which aims to provide a software agent with the ability to receive external *stimuli* from its environment (data) to solve a given problem autonomously, i.e., without human intervention. In order to do so, a code is created, through the elaboration of algorithms, which will determine how the software agent will act.

This software agent can be endowed with self-learning capacity (machine learning or even deep learning), which will allow it to go beyond its initial programming, being able, eventually, to take decisions for which it was not programmed.

This autonomy may make us raise the most varied questions: can an autonomous agent be considered imputable? Can it be held responsible if it decides or acts in the physical world and causes harm to someone? Can autonomous agents possess intentional states? And, more controversially, can the attribution of legal personality to autonomous agents be defended?

KEYWORDS: Artificial Intelligence; Robots; Legal Personality

1. Conceptualização

1.1. IA e Robótica: conceitos introdutórios

Para uma entidade ser considerada inteligente, precisa de possuir pelo menos cinco características: ser capaz de comunicar ("the easier it is to communicate with an entity, the more intelligent the entity seems[2]"); possuir conhecimento interno (ter algum conhecimento sobre si própria); possuir conhecimento externo (conhecer o mundo exterior, aprender sobre ele e utilizar essa informação); ser capaz de agir para atingir determinados objectivos; possuir criatividade (ou seja, ser capaz de encontrar soluções alternativas quando a sua acção inicial não lhe permite atingir os objectivos em questão)[3].

É habitual classificar a IA de acordo com três tipos de inteligência: Fraca, Média e Forte. O primeiro – ANI (*Artificial Narrow Intelligence*) – é aquele que

[2] Gabriel Hallevy, "The criminal liability of artificial intelligence entities – from Science fiction to legal social control", *Akron Intellectual Property Journal*, vol. 4, 2, 2010, p. 175 (pp. 171-201), disponível em *https://ideaexchange.uakron.edu/akronintellectualproperty/vol4/iss2/1* [consultado a 09/06/2022].

[3] *Idem*, pp. 175-176.

é especializado numa única área, tendo um certo objectivo (por exemplo, máquinas treinadas para jogar xadrez); o segundo – AGI (*Artificial General Intelligence*) – imita a mente humana, sendo capaz de compreender conceitos complexos e resolver problemas aprendendo com a sua própria experiência; o terceiro – ASI (*Artificial Super Intelligence*) – já possui competências sociais, igualando ou mesmo superando o cérebro humano[4]. No contexto tecnológico actual, esta última, a também denominada IA Forte, ainda não existe.

É importante, a este respeito, distinguir um agente autónomo de um *robot*. Muitas vezes utiliza-se a expressão "robot" como sinónimo de "máquina inteligente". Já ouvimos falar de *Robot-advisors* (que são aplicações de *software* que vários bancos utilizam e até disponibilizam aos seus clientes, para os ajudar a investir nos mercados financeiros[5]), do Juiz-*robot* (programas de computador que auxiliam a tomada de decisão do juiz, criando, até propostas de sentença, após análise dos dados carreados para o processo e da jurisprudência anterior semelhante ao caso concreto[6]), de Veículos Autónomos (automóveis que se espera que venham a circular sem condutor, ou sem que alguém seja chamado à tarefa da condução[7]). De todos estes agentes autónomos – alguns ainda em fase de teste, outros já implementados em maior ou menor grau – só o último exemplo é que poderá classificar-se como um *robot*.

[4] Marcos Ehrhardt Júnior/ Gabriela Buarque Pereira Silva, "Pessoa e Sujeito de Direito: Reflexões sobre a Proposta Europeia de Personalidade Jurídica Eletrônica", *RJLB*, Ano 7, 2021, nº 2, pp. 1100-1101 (pp. 1089-1117), disponível em https://www.cidp.pt/revistas/rjlb/2021/2/2021_02_1089_1117.pdf [consultado a 13/06/2022].

[5] O *Robot-Advisor* é apenas um dos instrumentos usados na chamada *FinTech* ("financial technology"), i.e. o uso das novas tecnologias (*Tech*) no sector financeiro (*Fin*). A. Barreto Menezes Cordeiro, "Inteligência Artificial e Consultoria Robótica", in António Menezes Cordeiro/Ana Perestrelo de Oliveira/Diogo Pereira Duarte (coords.), *FinTech: Desafios da tecnologia financeira*, 2ª ed., Coimbra, Almedina, 2019, p. 221.

[6] Sobre o Juiz-*Robot*, v. Sónia Moreira, "Artificial Intelligence: Brief considerations regarding the Robot-Judge", in Maria Miguel Carvalho/Sónia Moreira (eds.), *Governance & Technology – E-Tec Yearbook*, JusGov – Research Centre for Justice and Governance/University of Minho – School of Law, 2021, pp. 297-313, disponível em https://www.jusgov.uminho.pt/publicacoes/etec-yearbook-2021-2/ [consultado a 13/06/2022].

[7] Sobre algumas das questões levantadas por estes veículos na área da responsabilidade civil, v. Eva Sónia Moreira da Silva, "Considerations regarding Artificial Intelligence and Civil Liability: the case of autonomous vehicles", *SSRN – JusGov Research Paper Series* nº 2022–02 (April 14, 2022), pp. 1-12, disponível em SSRN: https://ssrn.com/abstract=4083771 [consultado a 13/06/2022] e Sofia Patrícia Travassos de Freitas Alcaide, *A Responsabilidade Civil por Danos Causados por Veículos Autónomos*, Coimbra, Almedina, 2021.

Nas palavras de Patrick Hubbard, um *robot* é "embodied software"[8], ou seja, é um programa de computador que possui um corpo físico por via do qual interage com o mundo, sem o controlo constante e/ou direto de um ser humano[9]. Os *Robot-advisors* e os juízes-*robots*, para já, ao menos, são apenas aquilo que se chama de "bots"[10], ou seja, agentes autónomos (agentes de *software*, programas de computador) que não possuem corpo físico, mas foram criados com vista à realização de uma determinada tarefa, seja ela propor uma solução de investimento financeiro ou uma sentença, seja outro tipo de interacções, como aquelas que temos com *chatbots* como a *Siri* ou a *Alexa*, ou com a nossa *box* da televisão, que nos apresenta sugestões ou recomendações de filmes ou séries, atendendo à análise que faz das nossas visualizações anteriores e das nossas preferências.

Assim, basicamente, os *bots* estão preparados para analisar dados, detectar padrões e resolver o problema para o qual foram criados. Os *robots* fazem o mesmo, mas, como possuem uma componente de *hardware*, intervêm no mundo físico.

Não se pense que, pelo facto de o agente de *software* não estar corporizado é menos perigoso. É certo que um *robot* de uma linha de montagem pode matar alguém se os seus sensores não detetarem a sua presença; um Veículo Autónomo pode atropelar uma pessoa se não a identificar como tal ou não identificar que o semáforo ficou vermelho. Mas se um programa de computador começar a fazer investimentos financeiros autonomamente, com base no perfil de risco do cliente e do seu histórico de investimentos, pode acarretar danos patrimoniais gravíssimos; e se a *Alexa* começar a decidir fazer as compras no Continente *online* por nós, usando o nosso cartão de crédito, não estará em causa só a nossa conta bancária, mas a nossa autodeterminação.

De todo o modo, não há como negar que a figura cinematográfica do *robot* – uma máquina antropomórfica, como a *Robot Sofia* – não corresponde à realidade, ou melhor, não é a única realidade a ter em conta no que toca aos agentes autónomos.

[8] AA.VV., *Robot Law*, Ryan Calo/ A. Michael Froomkin/ Laurie Silvers/ Mitchell Rubenstein (eds.), Edward Elgar, 2016, p. 59.

[9] Michael Froomkin define *robot* como "a man-made object capable of responding to external stimuli and acting on the world without requiring direct – some might say constant – human control". AA.VV., *Robot Law*, Ryan Calo/ A. Michael Froomkin/ Laurie Silvers/ Mitchell Rubenstein (eds.), Edward Elgar, 2016, p. XI.

[10] Paulo Novais/Pedro Miguel Freitas, *Inteligência Artificial e Regulação de Algoritmos*, Diálogos, União Europeia-Brasil, 2018, p. 17 (pp. 1-91), disponível em http://www.sectordialogues.org/documentos/noticias/adjuntos/ef9c1b_Intelig%C3%AAncia%20Artificial%20e%20Regula%C3%A7%C3%A3o%20de%20Algoritmos.pdf [consultado a 14/08/2021].

1.2. Personalidade Jurídica

Desde o primeiro ano da licenciatura em Direito que nos deparamos com este conceito. O conceito de personalidade jurídica é um conceito criado pelo Homem e ao serviço do Homem. Pessoa em sentido jurídico é todo o ente que pode ser sujeito de relações jurídicas, ou seja, que pode ser titular de direitos e de obrigações[11].

Durante largos anos este conceito não correspondeu ao conceito de pessoa em sentido ético – referimo-nos à figura da escravidão, que durante milénios foi perfeitamente aceite[12]. No entanto, apesar de a escravatura ter sido abolida e de se ter reconhecido que todos os seres humanos, pelo simples facto de serem pessoas em sentido ético, são também pessoas em sentido jurídico – uma conquista que mais não é do que reconhecimento do estado natural das coisas ou, se quisermos, um reconhecimento daquilo que decorre do Direito Natural[13] – ainda há bem pouco tempo, vimos ordens jurídicas afirmar – de forma sustentada em conceitos técnico-jurídicos – que nem todas as pessoas possuíam (os mesmos) direitos, justificando tratamentos desumanos e genocídio[14/15]. Recordemos a segunda guerra mundial, pois se esquecermos a história, corremos o risco de a repetirmos. Mas não precisamos de ir tão longe: ainda hoje, há ordens jurídicas que não reconhecem a todas as pessoas em sentido ético o mesmo estatuto jurídico. Pensemos no fundamentalismo

[11] Sobre o conceito de personalidade jurídica, v., e.g., Heinrich Ewald Hörster /Eva Sónia Moreira da Silva, *A Parte Geral do Código Civil Português*, 2ª ed., Coimbra, Almedina, 2019, pp. 315-316; Carlos Alberto da Mota Pinto, *Teoria Geral do Direito Civil*, 5ª ed. (por António Pinto Monteiro/Paulo Mota Pinto), Coimbra, Coimbra Editora, 2020, pp. 193 e 201; Rabindranath Capelo de Sousa, *Teoria Geral do Direito Civil*, Vol. I, Coimbra, Coimbra Editora, 2003, pp. 249-250.

[12] Em Portugal, por exemplo, a escravidão foi abolida em 1836. Cf. Heinrich Ewald Hörster /Eva Sónia Moreira da Silva, *A Parte Geral do Código Civil Português*, *cit.*, p. 182, n. 246.

[13] Afirmando que "o Direito não pode deixar de reconhecer às pessoas humanas a personalidade, assim como não lhes pode recusar a dignidade humana" porque "[e]stá fora do seu alcance por Direito Natural", Pedro Pais de Vasconcelos, *Teoria Geral do Direito Civil*, 9th ed., Coimbra, Almedina, 2019, p. 39.

[14] Dando nota de que "[n]as ordens jurídicas colectivistas (...) a personalidade não é uma qualidade inata da pessoa, mas é atribuída aos homens (...) de acordo, aliás, com o carácter positivista daquelas ordens", Heinrich Ewald Hörster /Eva Sónia Moreira da Silva, *op. cit.*, p. 316.

[15] Pedro Pais de Vasconcelos considera que o entendimento de que a personalidade deriva da "qualidade de [se] ser pessoa" atribui ao conceito de personalidade jurídica uma "dimensão ética", defendendo "as pessoas contra os perigos, historicamente já experimentados, de condicionamento e manipulação ou mesmo de recusa de personalidade a pessoas individualmente consideradas ou grupos de pessoas com base em critérios rácicos, ou religiosos". Cf. Pedro Pais de Vasconcelos, Teoria Geral do Direito Civil, cit., pp. 38-39.

islâmico e no tratamento diferenciado que concede às mulheres, por exemplo. Reparem que não me refiro ao tratamento misógino ou desigualitário que ainda hoje se reconhece existir *de facto* nos países ocidentais e com o qual lutamos todos os dias, mas, da assunção de um estatuto *jurídico* diferente para diferentes grupos de pessoas ou etnias.

De todo o modo, a cultura ocidental em geral e a nossa ordem jurídica em particular podem orgulhar-se desta conquista: a do reconhecimento de que *não é a ordem jurídica* que atribui ao ser humano a prerrogativa da personalidade jurídica[16]. A personalidade jurídica *é inerente* a todo o ser humano, só pelo facto de ter nascido[17] (ou para alguns autores, de ter sido concebido[18]). O artigo 66º, nº 1, do nosso CC, é uma manifestação claríssima do princípio da igualdade e, neste sentido e acima de tudo, uma expressão do princípio fundamental da dignidade da pessoa humana.

Todos os seres humanos, pelo simples facto de o serem, são sujeitos de direitos e de obrigações, alguns dos quais inalienáveis, irrenunciáveis e relativamente indisponíveis. Referimo-nos, evidentemente, aos direitos de personalidade, de que todos somos detentores desde o nascimento[19]. Historicamente, os direitos de personalidade emergem como "direitos inatos e originários da pessoa, alicerçados na natureza humana", encontrando "o seu fundamento

[16] Neste sentido, considerando que, pelo contrário, "a personalidade jurídica é a projecção na lei (...) da personalidade humana", optando, assim, pela nomenclatura "pessoas humanas – pessoas jurídicas", em vez de "pessoas singulares – pessoas colectivas", Orlando de Carvalho, *Teoria Geral do Direito Civil*, 4ª ed. por Francisco Liberal Fernandes/ Maria Raquel Guimarães/ Maria Regina Redinha, Coimbra, Gestlegal, 2021, pp. 191, ss.

[17] Neste sentido, Heinrich Ewald Hörster /Eva Sónia Moreira da Silva, *op. cit.*, p. 315.

[18] A este respeito, e.g., Rabindranath V. A. Capelo de Sousa, *O Direito Geral de Personalidade,* Coimbra, Coimbra Editora, 1995, p. 364; José de Oliveira Ascensão, *Direito Civil – Teoria Geral*, Vol. I, *Introdução, as Pessoas, os Bens*, 2ª ed. Coimbra, Coimbra Editora, 2000, p. 55; Manuel António Carneiro da Frada, "A protecção juscivil da vida pré-natal – Sobre o estatuto jurídico do embrião", in Joana Liberal Arnaut (org.), *Direito e Justiça – Verdade, Pessoa Humana e Ordem Político-Jurídica, Colóquio Internacional em Homenagem a Mário Emílio Forte Bigotte Chorão*, Faculdade Católica, 2008, pp. 153-154; Menezes Cordeiro, *Tratado de Direito Civil*, IV, *Parte geral – Pessoas,* Coimbra Almedina, 2011, pp. 363-365. Afirmando que a criança ainda não nascida tem personalidade jurídica por ser uma pessoa, embora os seus direitos dependam do seu nascimento, "com excepção daqueles que efetivamente [sejam] imprescindíveis para assegurar a incolumidade do nascituro", Mafalda Miranda Barbosa, *Lições de Teoria Geral do Direito Civil*, Coimbra, Gestlegal, 2021, pp. 264 ss., em especial, p. 297.

[19] Sobre esta matéria, v., e.g., Pedro Pais de Vasconcelos, *Direito de Personalidade*, Coimbra, Almedina, 2019 (reimpr.); Rabindranath Capelo de Sousa, *O Direito Geral de Personalidade*, Coimbra, Coimbra Editora, 1995.

último" na dignidade humana[20], encontrando-se a sua constitucionalização na figura de diversos direitos fundamentais[21/22].

2. A caminho da personalidade jurídica dos agentes dotados de IA?

2.1. Argumentos a favor

Se a personalidade jurídica é algo que é inerente ao ser humano, como é possível que se possa ponderar a sua atribuição a entes não humanos?

Na verdade, o Direito já o faz. Referimo-nos, naturalmente, às pessoas colectivas. As pessoas colectivas são substratos (pessoais ou patrimoniais) criados por seres humanos, dotados de uma orgânica própria, que lhes permite atingir o fim para o qual foram criados, um fim demasiado grande para ser prosseguido por pessoas singulares individualmente[23]. As vantagens da atribuição de personalidade jurídica a estes substratos são inegáveis, desde logo, no que toca à separação patrimonial entre os patrimónios dos seus membros e o património da própria pessoa colectiva.

Estas razões de ordem económica e social justificam a criação de entidades separadas das pessoas singulares, mas que podem actuar no mundo jurídico. Como se justifica, no entanto, a sua criação à luz do Direito? Aqui podemos socorrer-nos do princípio da autonomia privada: se as pessoas singulares podem prosseguir os seus interesses de forma individual, porque não o poderão fazer em conjunto, organizando-se de acordo com certos estatutos, com uma orgânica própria e independente[24]? *Et voilá*: eis que vimos nascer as pessoas colectivas, uma realidade incontornável no mundo de hoje: associações, fundações, sociedades comerciais, sociedades civis sob forma comercial, etc.

Ora, se a lei reconhece a possibilidade de atribuição de personalidade jurídica a estes substratos, porque não fazê-lo aos agentes autónomos?

Como vimos, actualmente já há agentes que conseguem evoluir para além da sua programação e actuar de forma completamente autónoma. Há rela-

[20] Ana Filipa Morais Antunes, *Comentário aos artigos 70º a 81º do Código Civil (Direitos de Personalidade)*, Lisboa, Universidade Católica Editora, 2012, p. 13.

[21] Sobre esta questão, v., e.g., Paulo Mota Pinto, *Direitos de Personalidade e Direitos Fundamentais. Estudos*, Coimbra, Gestlegal, 2018.

[22] O crescente reconhecimento da sua importância pode ver-se também na jurisprudência portuguesa, que tem vindo a concretizar os seus conceitos indeterminados. Cf. Guilherme Machado Dray, *Direitos de Personalidade. Anotações ao Código Civil e ao Código o Trabalho*, Coimbra, Almedina, 2006, p. 7.

[23] Cfr. Heinrich Ewald Hörster /Eva Sónia Moreira da Silva, *op. cit.*, pp. 401 ss.; Carlos Alberto da Mota Pinto, *Teoria Geral do Direito Civil*, *cit.*, pp. 269 ss.

[24] Neste sentido, Heinrich Ewald Hörster /Eva Sónia Moreira da Silva, *op. cit.*, pp. 403-404.

tos de programas de computador que começaram a tomar decisões inexplicáveis em face da sua programação original (tendo sido, preventivamente, desligados)[25]. Através dos mecanismos de *Machine learning*, o agente autónomo recolhe informação do meio ambiente, de outros agentes (através de interacção com outros agentes autónomos, como aparelhos domésticos *smart*, câmaras de videovigilância, etc.) e da própria *internet* e de bases de dados a que tenha acesso, aumentando os dados originais de que dispunha. Por outras palavras, a máquina aprende por si própria, autonomamente[26]. Se o agente for dotado de *Deep learning*, esta capacidade de autoaprendizagem aproxima-se da do ser humano, pois replica a nossa rede neuronal.

Esta autonomia torna extremamente difícil imputar os danos provocados por um agente autónomo a pessoas singulares[27], surgindo o problema do *liability gap*[28]. Quem responde? A pessoa singular não tem culpa se não podia prever a actuação lesiva do agente autónomo, pelo que não pode ser responsabilizada; o agente autónomo não pode responder porque, mesmo que seja considerado inteligente o suficiente para se lhe reconhecer um estado intencional, não possui personalidade jurídica, não tem direitos (e, portanto, não tem património) nem obrigações (como a de indemnizar).

Em face destas questões, a UE chegou a admitir a possibilidade de se atribuir ou de se criar uma "personalidade jurídica" aos agentes artificialmente inteligentes. A Resolução do Parlamento Europeu, de 16 de fevereiro de 2017, que contém recomendações à Comissão sobre disposições de Direito Civil sobre Robótica (2015/2103(INL))[29], no Ponto 1. dos *Princípios gerais relativos ao desenvolvimento da robótica e da inteligência artificial para utilização civil*,

[25] Um caso exemplar é o do *robot Gaak*, um *robot* que ficou sem vigilância por quinze minutos, escapou da arena onde devia interpretar o papel de "caçador" ou "presa" de forma a testar-se o princípio da sobrevivência do mais apto para os *robots* dotados de IA, e encontrou uma forma de sair atravessando a parede das instalações sem ter sido programado para o fazer. Cf. Marcos Ehrhardt Júnior/ Gabriela Buarque Pereira Silva, "Pessoa e Sujeito de Direito: Reflexões sobre a Proposta Europeia de Personalidade Jurídica Eletrônica", *cit.*, p. 1105.

[26] *Idem*, p. 1103.

[27] Mafalda Miranda Barbosa, "O futuro da responsabilidade civil desafiada pela inteligência artificial: as dificuldades dos modelos tradicionais e caminhos de solução", *Revista de Direito Civil*, V, nº 2, 2020, p. 265.

[28] Steven S. Gouveia, "O problema da lacuna da responsabilidade na Inteligência Artificial", in Manuel Curado/ Ana Elisabete Ferreira/ André Dias Pereira (eds.), *Vanguardas da Responsabilidade – Direito, Neurociências e Inteligência Artificial*, Faculdade de Direito da Universidade de Coimbra, Petrony, 2019, pp. 172-173.

[29] Disponível em https://www.europarl.europa.eu/doceo/document/TA-8-2017-0051_PT.html [consultado a 31/10/2022].

"[i]nsta a Comissão a propor definições comuns à escala da União de sistemas ciberfísicos, de sistemas autónomos, de robôs autónomos inteligentes e das suas subcategorias, tendo em consideração as seguintes características de um robô inteligente:

- aquisição de autonomia através de sensores e/ou da troca de dados com o seu ambiente (interconetividade) e da troca e análise desses dados;
- autoaprendizagem com a experiência e com a interação (critério opcional);
- um suporte físico mínimo;
- adaptação do seu comportamento e das suas ações ao ambiente;
- inexistência de vida no sentido biológico do termo".

Na al. f) do Ponto 59., a Comissão também é chamada a "[c]riar um estatuto jurídico específico para os robôs a longo prazo, de modo a que, pelo menos, os robôs autónomos mais sofisticados possam ser determinados como detentores do *estatuto de pessoas eletrónicas* responsáveis por sanar quaisquer danos que possam causar e, eventualmente, aplicar a *personalidade eletrónica* a casos em que os robôs tomam decisões autónomas ou em que interagem por qualquer outro modo com terceiros de forma independente" (itálico nosso).

No entanto, esta tomada de posição foi recebida criticamente, pelo que, a UE, nos mais recentes documentos sobre IA, voltou atrás: nada a este respeito é defendido no Livro Branco sobre Inteligência Artificial da Comissão Europeia[30], nem na Proposta de Regulamento sobre Inteligência Artificial do Parlamento Europeu e do Conselho (Artificial Intelligence Act)[31]. O mesmo se diga da recentíssima Proposta de Diretiva Sobre Responsabilidade Civil da Inteligência Artificial do Parlamento Europeu e do Conselho[32].

[30] European Comission, *White Paper on Artificial Intelligence – A European approach to excellence and trust*, p. 2 (disponível em https://ec.europa.eu/info/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en [consultado a 1/10/2020].

[31] Disponível em https://eur-lex.europa.eu/legal-content/PT/TXT/?uri=CELEX:52021PC0206 [consultado a 22/12/2021]. Na verdade, a Resolução do PE com Recomendações sobre uma Proposta de Regulamento, de 2020, até diz expressamente o contrário no Considerando (6) da Proposta de Regulamento sobre Responsabilidade Civil: "os sistemas de IA não têm personalidade jurídica nem consciência humana e (...) a sua única missão é servir a humanidade".

[32] Proposta de Directiva do PE e do Conselho (Directiva Responsabilidade da IA) de 28/09/2022, Disponível em https://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=CELEX:52022PC0496&qid=1665770433787&from=PT [consultada a 14/10/2022].

2.2. Argumentos contra

Apesar de haver autores que defendem a possibilidade da existência de estados intencionais de certos *robots* e agentes autónomos[33], a verdade é que a doutrina discute o conceito de autonomia destes agentes.

Concordamos com Mafalda Miranda Barbosa, que afirma que estamos perante uma mera "autonomia tecnológica", porque a "inteligência artificial [se] baseia (...) na acumulação de conhecimento, sendo incapaz de interpretações criativas ou de julgamentos sobre o que é certo ou errado (...), [sendo] sempre condicionada pelos *inputs* do programador"[34]. Portanto, estamos perante uma "autonomia algorítmica", já que as "decisões" do agente autónomo são sempre pré-determinadas pelas diretrizes dadas pelo programador. Assim, a autonomia destes agentes não se confunde com a autonomia humana, com a autonomia privada e, seguramente, com a autodeterminação do ser humano. Um agente autónomo não possui livre-arbítrio, não determina o que quer fazer da sua "vida", não tem sonhos, aspirações, propósitos, não determina o seu próprio fim.

Na verdade, uma capacidade de aprendizagem e entendimento semelhantes à humana, incluindo capacidade decisória verdadeiramente autónoma (ou seja, que não dependa dos *inputs* originais do programador), possuindo criatividade e, até, sentimentos, é algo que ainda não existe no actual estado da tecnologia: a chamada IA Forte. Os cientistas divergem quanto à possibilidade de este estado de desenvolvimento se poder atingir a não ser através de uma *interface* Homem-Máquina, ou seja, no campo da cibernética (e do tão falado transumanismo)[35], mas, a ser assim, não haverá que discutir sobre a persona-

[33] Giovanni Sartor, "Cognitive Automata and the Law: electronic contracting and the intentionality of software agents", *Artificial Intelligence and Law*, nº 17, 2009, Springer, pp. 253-290; Pedro Miguel Freitas/ Francisco Andrade/ Paulo Novais, "Criminal Liability of Autonomous Agents: from the unthinkable to the plausible", in Pompeu Casanovas et al. (eds.), *AICOL IV/V 2013, LNAI 8929*, Springer, 2014, pp. 145-156.

[34] Mafalda Miranda Barbosa, "O futuro da responsabilidade civil desafiada pela inteligência artificial: as dificuldades dos modelos tradicionais e caminhos de solução", *cit.*, pp. 291.

[35] Há cientistas que prevêem que se atinja um nível de desenvolvimento tecnológico entre 2030 e 2045 que permita a criação de um sistema intelectualmente semelhante ao ser humano. APDSI (Associação Para a Promoção e Desenvolvimento da Sociedade da Informação), *No Limiar na Autodeterminação da Inteligência Artificial?*, Printinglovers, s.d., p. 24. Apesar destas previsões, outros autores consideram que este nível de desenvolvimento da IA (a chamada HLAI – *Human Level Artificial Intelligence* – uma máquina capaz de pensar e agir como um ser humano, incluindo nos domínios social, cultural e emocional, possuindo, nomeadamente, criatividade e a capacidade de pensar "fora da caixa") apenas será possível se se basear numa *interface* ente homem e máquina, na figura do ciborgue. *Idem*, pp. 32-35.

lidade jurídica, porque sempre estaremos perante uma pessoa singular, ainda que esta tenha componentes electrónicos incorporados em si. Contudo, ainda que seja possível uma máquina atingir um nível de desenvolvimento que lhe permita tomar consciência de si própria, possuir um raciocínio semelhante ao humano, incluindo criatividade e, até, sentimentos, ou seja, ainda que uma máquina possua IA Forte, é legítimo ponderar-se a atribuição de personalidade jurídica nestes casos?

O Livro Branco sobre IA da Comissão Europeia e a Proposta de Regulamento sobre IA do PE e do Conselho determinam que o funcionamento da IA deve submeter-se sempre ao respeito pelos direitos fundamentais dos cidadãos, nomeadamente, ao respeito pela dignidade da pessoa humana e à proteção da sua privacidade. Por isso, a pergunta que temos de nos fazer é a seguinte: atribuir personalidade jurídica a um agente dotado de IA Forte não violará o princípio da dignidade da pessoa humana? Não deveria ser apenas o ser humano a possuir personalidade jurídica?

Podíamos dizer que isso já acontece actualmente, uma vez que o Direito atribui personalidade jurídica às pessoas colectivas, apesar de estas não serem pessoas em sentido ético. Contudo, como vimos, a criação das pessoas colectivas visa a prossecução de interesses humanos, demasiado grandes para que as pessoas singulares os possam prosseguir por si, isoladamente; as pessoas singulares, ao abrigo da sua autonomia privada, podem organizar-se entre si e criar substratos que as ajudem a prosseguir os seus interesses. Portanto, a personalidade jurídica das pessoas colectivas mais não é do que um expediente técnico-jurídico ao serviço das pessoas singulares[36].

Nas palavras de Orlando de Carvalho, "[s]ó há personalidade jurídica porque existe personalidade humana. (...) Há personalidade jurídica quando existe (logo que existe e enquanto existe) personalidade humana. (...) Há personalidade jurídica até onde e *só até onde o exija a personalidade humana*" (itálico nosso). "As outras "personalidades jurídicas" são meramente analógicas e instrumentais" [37], como é o caso das pessoas colectivas.

[36] Mafalda Miranda Barbosa discorda que se possa fazer uma analogia entre agentes autónomos e pessoas colectivas, uma vez que estas últimas foram criadas para que os interesses humanos colectivos ou comuns pudessem ser perseguidos (ou para que isto pudesse ser feito de uma forma mais eficiente), o que não acontece no caso dos *robots*, o que apenas permitiria que o seu proprietário fosse exonerado de responsabilidade; em qualquer caso, a autora conclui que, mesmo que esta exoneração de responsabilidade pudesse ser considerada um interesse humano, não resolveria qualquer problema, uma vez que os *robots* não têm bens. V. Mafalda Miranda Barbosa, *op. cit.*, pp. 294-295.

[37] Orlando de Carvalho, *Teoria Geral do Direito Civil*, *op. cit.*, pp. 192 e 193.

Assim é necessário ponderar o seguinte: a atribuição de personalidade jurídica aos agentes dotados de IA (Forte) serve quem? É necessária? É imprescindível? É útil ao ser humano?

Não nos parece.

É certo que poderia evitar-se o *liability gap*, visto que estes agentes, sendo dotados de personalidade jurídica, poderiam ser obrigados a indemnizar. Contudo, para isso, teriam de possuir património. Ora, como iriam estes agentes adquirir bens? Possivelmente, teria de ser criado um fundo patrimonial por parte do seu produtor ou do seu dono. Ora, não se atinge o mesmo resultado através da criação de seguros de responsabilidade civil a cargo do seu dono ou do produtor (ou eventualmente do programador) e através da criação de um regime de responsabilidade objectiva destas pessoas a favor do lesado, já que, sendo elas quem retira vantagens da sua criação e comercialização (responsabilidade do produtor) ou da sua utilização (responsabilidade do dono/utilizador), deveriam ser também elas a arcar com a desvantagem de ter de assumir as indemnizações pelos danos causados por estes agentes ao abrigo do princípio *ubi commoda, ibi incommoda*[38]?

Por outro lado, poder-se-ia falar de "dono"? Sendo o agente autónomo um sujeito de relações jurídicas, poderia ser, ao mesmo tempo, objeto do direito de propriedade?

Por último, faz sentido que um agente autónomo seja titular do direito à vida, à integridade física, à imagem, à honra... à semelhança de um ser humano? Como defender a existência, por exemplo, do direito à autodeterminação ou do direito ao livre desenvolvimento da personalidade de uma máquina[39]? Atribuir a uma máquina um estatuto legal similar ao do ser humano é objectificar o ser humano, é diminuir o ser humano, ofendendo a sua dignidade.

É de referir que o nosso sistema legal não criou um estatuto legal semelhante ao do ser humano no que toca aos animais. Apesar de reconhecer que se trata de seres vivos que possuem sensibilidade e, por isso, são considerados

[38] Igualmente neste sentido, Marcos Ehrhardt Júnior/ Gabriela Buarque Pereira Silva, "Pessoa e Sujeito de Direito: Reflexões sobre a Proposta Europeia de Personalidade Jurídica Eletrônica", *cit.*, pp. 1111 ss. Os autores questionam a proporcionalidade e necessidade desta personificação, uma vez que há outros mecanismos capazes de garantir as devidas indemnizações, tal como seguros obrigatórios. Os autores concluem que, pelo menos para já, não há bases antropológica e axiológica suficientes para justificar a criação de uma personalidade electrónica, pelo que os *robots* devem ser tratados como uma coisa. *Idem*, pp. 1116 e 1117.

[39] Sobre o direito ao livre desenvolvimento da personalidade, v. Paulo Mota Pinto, *Direitos de Personalidade e Direitos Fundamentais. Estudos*, *cit.*, pp. 7 e ss., nomeadamente o facto de que deriva da dignidade humana e de que é a base de um "direito geral de liberdade". *Idem*, p. 11.

objecto de protecção, a nossa lei não os considera sujeitos de direitos (ou de obrigações, naturalmente). Na verdade, apesar de o legislador ter criado um subtítulo relativo aos animais dentro do Título que trata dos elementos da relação jurídica – ou seja, o subtítulo I-A (sendo o subtítulo I "Das Pessoas" e o subtítulo II "Das Coisas") – em nenhum dos artigos deste novo subtítulo se afirma que estes sejam detentores de direitos. Pelo contrário, o art. 201º-D determina que se lhes aplique, como regime subsidiário, o regime das coisas, desde que tal não seja incompatível com a sua natureza; além disso, o art. 1302º, nº 2, afirma peremptoriamente que os animais são *objecto* do direito de propriedade, embora os distinga das restantes coisas corpóreas (às quais se aplica o art. 1301º, nº 1). Por outras palavras, embora a lei tenha excluído os animais das coisas em termos formais (uma vez que excluiu a sua regulamentação da regulamentação das coisas), a verdade é que continua a qualificá-los como objecto de direitos e não como sujeito de direitos. Assim, considerando a definição de coisa do art. 202º, n.os 1 e 2, do Código Civil Português, concordamos com os autores que qualificam os animais como coisas (= objecto de direitos privados), embora coisas *sui generis*, uma vez que possuem um regime diferente das restantes coisas corpóreas[40]. O elemento sistemático de interpretação da lei, nomeadamente no que se refere ao lugar onde o legislador colocou o regime dos animais no Código Civil, não nos parece suficiente para lhes reconhecer outra qualificação, uma vez que o que é mais relevante é o seu regime jurídico, ou seja, um argumento de natureza substancial e não meramente formal[41].

Em suma, a atribuição de personalidade jurídica aos agentes autónomos traz mais problemas do que aqueles que resolve, sendo certo que aqueles que resolve podem encontrar soluções muito mais consentâneas com os princípios fundamentais da nossa ordem jurídica, nomeadamente, com o princípio

[40] Heinrich Ewald Hörster "A propósito da Lei nº 8/2017, de 3 de Março: os animais ainda serão coisas (objectos da relação jurídica)?", *Revista Jurídica Portucalense*, vol. nº 22, 2017, pp. 66-76, onde o autor ainda explica que o legislador partiu da premissa errada de que o regime alemão das coisas era semelhante ao português, o que não sucede. No mesmo sentido, Cristina Dias, "O divórcio e o novo Estatuto Jurídico dos Animais, introduzido pela Lei nº 8/2017, de 3 de Março – quem fica com o animal de companhia?", in Regina Beatriz Tavares Da Silva/Úrsula Cristina Basset (coords.), *Família e Pessoa: uma Questão de Princípios*, Academia Iberoamericana de Derecho de Familia e de las Personas/ADFAS, p. 289, n. 1.

[41] Pelo contrário, considerando que os animais constituem um *tertium genus* (entre as pessoas e as coisas), Filipe Albuquerque Matos/Ana Mafalda Miranda Barbosa, *O novo estatuto jurídico dos animais*, Coimbra, Gestlegal, 2017, p. 7, e Luís Manuel Teles de Menezes Leitão, *Direitos Reais*, 9ª ed. Coimbra, Almedina, 2020, pp. 78-79.

da dignidade da pessoa humana e com os princípios que regem o instituto da responsabilidade civil.

3. Conclusões

Cremos que, neste momento, não se justifica a atribuição de personalidade jurídica aos agentes autónomos, por vários motivos:

a) Em primeiro lugar, porque ainda não atingiram um grau de autonomia ou de autoconsciência semelhante às humanas;
b) Em segundo lugar, porque tal facto não serve os interesses humanos, ou seja, ao contrário do que acontece com as pessoas colectivas, a personalização destes agentes não traz nenhum benefício que não possa ser atingido por uma via menos controversa, pelo que não estamos perante uma situação análoga à das pessoas colectivas;
c) Em terceiro lugar, tratar uma "máquina" da mesma forma que um ser humano é violar o princípio da dignidade da pessoa humana;
d) Finalmente, mesmo que esta "máquina" venha a ter sensibilidade, não se justificará que se lhe atribua personalidade jurídica. Veja-se o regime jurídico dos animais, que são objecto de protecção em virtude de serem seres com sensibilidade, mas não são dotados de personalidade jurídica, sendo apenas considerados um objecto *sui generis* de relações jurídicas, com regime próprio.

No entanto, reservo-me a possibilidade de vir a mudar de opinião quando estivermos perante uma IA Forte. Veremos quais serão, efectivamente, as suas características, as suas capacidades, a sua autopercepção de si mesma, a sua autonomia, a sua sensibilidade. Mas, como já discutia Azimov na obra "O Homem Bicentenário"[42], continuo a pensar se podemos, de facto, considerar humano um ser que potencialmente seja infinito e não esteja sujeito, como todos nós, à mais certa das regras da natureza: a da mortalidade.

Bibliografia

Aa.Vv., *Robot Law*, Ryan Calo/ A. Michael Froomkin/ Laurie Silvers/ Mitchell Rubenstein (eds.), Edward Elgar, 2016

Alcaide, Sofia Patrícia Travassos de Freitas, *A Responsabilidade Civil por Danos Causados por Veículos Autónomos*, Coimbra, Almedina, 2021

Antunes, Ana Filipa Morais, *Comentário aos artigos 70º a 81º do Código Civil (Direitos de Personalidade)*, Lisboa, Universidade Católica Editora, 2012

[42] Isaac Azimov, *The Bicentennial Man*, Gollancz, 2020.

Ascensão, José de Oliveira, *Direito Civil – Teoria Geral*, Vol. I, *Introdução, as Pessoas, os Bens*, 2ª ed. Coimbra, Coimbra Editora, 2000

Barbosa, Mafalda Miranda, "O futuro da responsabilidade civil desafiada pela inteligência artificial: as dificuldades dos modelos tradicionais e caminhos de solução", *Revista de Direito Civil*, V, nº 2, 2020

CORDEIRO, A. Barreto Menezes, "Inteligência Artificial e Consultoria Robótica", in António Menezes Cordeiro/Ana Perestrelo De Oliveira/Diogo Pereira Duarte (coords.), *FinTech: Desafios da tecnologia financeira*, 2ª ed., Coimbra, Almedina, 2019

Da frada, Manuel António Carneiro, "A protecção juscivil da vida pré-natal – Sobre o estatuto jurídico do embrião", in Joana Liberal Arnaut (org.), *Direito e Justiça – Verdade, Pessoa Humana e Ordem Político-Jurídica, Colóquio Internacional em Homenagem a Mário Emílio Forte Bigotte Chorão*, Faculdade Católica, 2008

Da silva, Eva Sónia Moreira "Considerations regarding Artificial Intelligence and Civil Liability: the case of autonomous vehicles", *SSRN – JusGov Research Paper Series* nº 2022–02 (April 14, 2022), pp. 1-12, disponível em SSRN: https://ssrn.com/abstract=4083771 [consultado a 13/06/2022]

De carvalho, Orlando, *Teoria Geral do Direito Civil*, 4ª ed. por Francisco Liberal Fernandes/ Maria Raquel Guimarães/ Maria Regina Redinha, Coimbra, Gestlegal, 2021

De sousa, Rabindranath Capelo, *O Direito Geral de Personalidade*, Coimbra, Coimbra Editora, 1995

De sousa, Rabindranath Capelo, *Teoria Geral do Direito Civil*, Vol. I, Coimbra, Coimbra Editora, 2003

De sousa, Rabindranath V. A. Capelo, *O Direito Geral de Personalidade*, Coimbra, Coimbra Editora, 1995

De vasconcelos, Pedro Pais, *Direito de Personalidade*, Coimbra, Almedina, 2019 (reimpr.)

Dias, Cristina, "O divórcio e o novo Estatuto Jurídico dos Animais, introduzido pela Lei nº 8/2017, de 3 de Março – quem fica com o animal de companhia?", in regina beatriz tavares da silva/úrsula cristina basset (coords.), *Família e Pessoa: uma Questão de Princípios*, Academia Iberoamericana de Derecho de Familia e de las Personas/ADFAS, p. 289, n. 1

Dray, Guilherme Machado, *Direitos de Personalidade. Anotações ao Código Civil e ao Código o Trabalho*, Coimbra, Almedina, 2006

Ehrhardt júnior, Marcos / silva, Gabriela Buarque Pereira, "Pessoa e Sujeito de Direito: Reflexões sobre a Proposta Europeia de Personalidade Jurídica Eletrônica", *RJLB*, Ano 7, 2021, nº 2, pp. 1089-1117, disponível em https://www.cidp.pt/revistas/rjlb/2021/2/2021_02_1089_1117.pdf [consultado a 13/06/2022]

European Comission, *White Paper on Artificial Intelligence – A European approach to excellence and trust*, p. 2 (disponível em https://ec.europa.eu/info/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en [consultado a 1/10/2020]

Freitas, Pedro Miguel / andrade, Francisco / novais, Paulo, "Criminal Liability of Autonomous Agents: from the unthinkable to the plausible", in Pompeu Casanovas et al. (eds.), *AICOL IV/V 2013, LNAI 8929*, Springer, 2014, pp. 145-156

GOUVEIA, STEVEN S., "O problema da lacuna da responsabilidade na Inteligência Artificial", in MANUEL CURADO/ ANA ELISABETE FERREIRA/ ANDRÉ DIAS PEREIRA (eds.), *Vanguardas da Responsabilidade – Direito, Neurociências e Inteligência Artificial*, Faculdade de Direito da Universidade de Coimbra, Petrony, 2019

HALLEVY, GABRIEL, "The criminal liability of artificial intelligence entities – from Science fiction to legal social control", *Akron Intellectual Property Journal*, vol. 4, 2, 2010, pp. 171-201, disponível em *https://ideaexchange.uakron.edu/akronintellectualproperty/vol4/iss2/1* [consultado a 09/06/2022]

HÖRSTER, HEINRICH EWALD / DA SILVA, EVA SÓNIA MOREIRA, *A Parte Geral do Código Civil Português*, 2ª ed., Coimbra, Almedina, 2019

HÖRSTER, HEINRICH EWALD, "A propósito da Lei nº 8/2017, de 3 de Março: os animais ainda serão coisas (objectos da relação jurídica)?", *Revista Jurídica Portucalense*, vol. nº 22, 2017

LEITÃO, LUÍS MANUEL TELES DE MENEZES, *Direitos Reais,* 9ª ed. Coimbra, Almedina, 2020

MATOS, FILIPE ALBUQUERQUE / BARBOSA, ANA MAFALDA MIRANDA, *O novo estatuto jurídico dos animais*, Coimbra, Gestlegal, 2017

MENEZES CORDEIRO, ANTÓNIO, *Tratado de Direito Civil,* IV, *Parte geral – Pessoas,* Coimbra Almedina, 2011

MOREIRA, SÓNIA, "Artificial Intelligence: Brief considerations regarding the Robot-Judge", in Maria Miguel Carvalho/Sónia Moreira (eds.), *Governance & Technology – E-Tec Yearbook*, JUSGOV – Research Centre for Justice and Governance/University of Minho – School of Law, 2021, pp. 297-313, disponível em https://www.jusgov.uminho.pt/publicacoes/etec-yearbook-2021-2/ [consultado a 13/06/2022]

NOVAIS, PAULO / FREITAS, PEDRO MIGUEL, *Inteligência Artificial e Regulação de Algoritmos*, Diálogos, União Europeia-Brasil, 2018, pp. 1-91, disponível em http://www.sectordialogues.org/documentos/noticias/adjuntos/ef9c1b_Intelig%C3%AAncia%20Artificial%20e%20Regula%C3%A7%C3%A3o%20de%20Algoritmos.pdf [consultado a 14/08/2021]

PINTO, CARLOS ALBERTO DA MOTA, *Teoria Geral do Direito Civil*, 5ª ed. (por ANTÓNIO PINTO MONTEIRO/PAULO MOTA PINTO), Coimbra, Coimbra Editora, 2020

PINTO, PAULO MOTA, *Direitos de Personalidade e Direitos Fundamentais. Estudos*, Coimbra, Gestlegal, 2018

SARTOR, GIOVANNI, "Cognitive Automata and the Law: electronic contracting and the intentionality of software agents", *Artificial Intelligence and Law*, nº 17, 2009, Springer, pp. 253-290

Um Direito Civil sem pessoa? Apontamento sobre a sua (im)possibilidade na era da automação

A Civil Law without human agents? A brief note about its impossibility in an age of automation

Tiago Azevedo Ramalho*

RESUMO: O presente estudo procura reflectir sobre se o recurso a decisões automatizadas é compatível com a metódica de realização do Direito Civil. Conclui estarmos diante de mais uma manifestação do problema da relação entre "ser humano" e "tecnologia" por si criada.

PALAVRAS-CHAVE: Direito Civil, Inteligência Artificial, Função Judicial, Tecnologia, Metodologia Jurídica

ABSTRACT: The present study seeks to expose how the use of automated decisions is compatible with Civil Law and his methodology. It concludes that we are facing yet another manifestation of the problem of the relationship between "human being" and "technology".

KEYWORDS: Civil Law, Artificial Intelligence, Judiciary, Technology, Legal Methodology

* Faculdade de Direito da Universidade do Porto (FDUP). Centro de Investigação Jurídica (CIJ/FDUP).

SUMÁRIO: 1. Introdução ao tema e *iter*. 2. Campos de incidência da inteligência artificial no direito. 3. Um Direito Civil sem pessoa? 4. A alternativa à decisão automatizada: a automação sem decisão. 5. A relação do ser humano com a tecnologia.

1. Introdução ao tema e *iter*

Tem a presente comunicação por título "Um Direito Civil sem pessoa? Apontamento sobre a sua (im)possibilidade na era da automação"[1]. Trata-se, por conseguinte, de uma reflexão que tem por objecto principal o *Direito Civil* e a sua *possibilidade* de afirmação diante do desenvolvimento – e amplo emprego – de formas automatizadas de processamento de dados, que, por operarem de modo que ao ser humano não é possível reconstituir por inteiro, são considerados de "inteligência artificial". Atendendo à índole específica de semelhantes modos de tratamento de dados, são eles, para este efeito, designados sumariamente por "automação", uma vez que, iniciado o respectivo processamento, nele não intervém a mediação humana.

Ocasiona este desafio de enquadramento ao Direito Civil, portanto, esse conjunto de desenvolvimentos que dão pelo nome de "inteligência artificial". Estamos então diante de um domínio que, reconhece-se, ultrapassa de forma manifesta o campo de competência profissional de um jurista, ao menos quando aja nessa precisa qualidade. Aliás, a entrada no mundo das técnicas de inteligência artificial, cultuada nos domínios da engenharia informática, dos sistemas de comunicação, dos grandes números, etc., solicita como *senha* de acesso o cultivo de um conjunto de faculdades humanas não raro nos antípodas daquelas que são especialmente mobilizadas no estudo do Direito; e decerto como *contrasenha* que os seus cultores façam próprias intenções significativamente diferentes (por ex., a expansão do espaço de intervenção manipulativa sobre o mundo[2]) daquelas que devem mobilizar o orbe jurídico (por ex., a ordenação equitativa das relações humanas em contextos plurais).

[1] Constitui o presente texto a versão escrita da comunicação apresentada no Congresso Internacional sobre "Inteligência Artificial e Direito" decorrido na Faculdade de Direito da Universidade do Porto a 12 e 13 de Maio de 2022. Sem prejuízo do desenvolvimento de alguns dos pontos então considerados, o texto conserva as marcas de concisão e de coloquialidade próprias do registo comunicativo que determinou a respectiva feitura. Conforme o teor do estudo patenteará, apresenta-se somente um princípio de reflexão sobre o seu objecto.

[2] Cf. o auxílio interpretativo da "Modernidade" oferecido por HARTMUT ROSA, *Unverfügbarkeit,* Suhrkamp: Berlin, 2020, pp. 21 e ss., que nela identifica como móbil a tentativa de tornar o mundo "disponível".

Sem prejuízo, a irrupção de formas de inteligência artificial no circuito das relações humanas, uma vez tornadas parte do *habitat* cultural do homem contemporâneo, conduz a inevitáveis interpelações dirigidas à forma jurídica de ordenação das relações humanas. A partir de então, não deve o Direito desinteressar-se ao menos da decisão a respeito de se, e, em caso afirmativo, em que medida, lhes oferece um regime de enquadramento.

Mesmo se o domínio das técnicas da inteligência artificial ultrapassa o campo próprio de competência do jurídico, já não lhe escapa, portanto, a definição do lugar e do estatuto fundamental que se lhes deve reservar. É desde uma tal perspectiva que se justifica – mais: que se impõe – uma especial atenção do Direito, olhando o fenómeno das técnicas automáticas da inteligência artificial, não desde o seu exacto modo de funcionamento (questão propriamente técnica), mas dos pressupostos em que se estribam e dos efeitos que delas decorrem (questão eminentemente jurídica). E, atentando no primeiro e ponderando os segundos, delimitar um regime jurídico que lhes sirva de enquadramento[3].

Dividirei a minha apresentação nas seguintes partes. Começarei por considerar quais os campos de incidência da inteligência artificial no Direito (**2.**); depois verei em que termos se pode equacionar, isto se o puder ser, um Direito Civil sem uma referência pessoal (**3.**); farei menção a alternativas a processos decisórios criadas por meios automatizados (**4.**); finalmente, procurar-se-á reconduzir o tema da inteligência artificial à questão, mais ampla e vasta, da relação entre o ser humano e a tecnologia (**5.**).

2. Campos de incidência da *inteligência artificial* no Direito

Que as técnicas de inteligência artificial interferirão necessariamente no Direito é evidente[4]. Mas não é sempre igual o nó de interferência. São diferentes, com efeito, as possibilidades de cruzamento entre as primeiras (as

[3] Como guia orientador no mundo da inteligência artificial, servi-me da recente obra de Gaspard Koenig, *La fin de l'individu. Voyage d'un philosophe au pays de l'intelligence artificelle*, Paris, Le Point/Éditions de L'Observatoire, 2019. Foi obra que se afigurou de especial valia: o Autor, professor de Filosofia na *Sciences Po* de Paris, propõe-se justamente a servir de *guia* ao "novo mundo" da inteligência artificial, mediando o acesso às ciências da computação desde o chão comum das ciências do espírito. A obra, com certo sabor jornalístico, é resultado de mais de uma centena de entrevistas a diferentes tipos de intervenientes no mundo da inteligência artificial, entre professores, empresários (start-ups), think-tanks, intelectuais públicos, agentes políticos, etc.

[4] Especificamente para o processo de realização do Direito, pode confrontar-se, com brevidade, cf. a introdução de Wolfgang Hoffmann-Riem, "Verhaltenssteuerung durch Algorithmen – Eine Herausforderung für das Recht", *Archiv des öffentlichen Rechts* 142 (2017), pp. 1-42, Mario Martini, "Algorithmen als Herausforderung für die Rechtsordnung", *Juristenzeitung* 72 (2017), pp. 1017-1025;

técnicas de inteligência artificial) e o segundo (o Direito), levando ao surgimento de problemas de índole diferenciada.

Distinguiria os cinco seguintes planos em que se podem dar tais cruzamentos.

(a) As técnicas de *inteligência artificial* são um meio de criação de novos objectos, que, logo que considerados pelo Direito, se tornam objectos jurídicos, isto é, sujeitos a uma regulação que lhes modela o respectivo estatuto.

Dentro do estatuto relativo ao bem pode considerar-se o regime jurídico de definição da sua natureza – criam-se bens coisificáveis? –, do seu conteúdo juridicamente relevante – com que limites? –, da sua titularidade ou dos frutos da sua utilização – em proveito de quem? –; mas também da repercussão do objecto sobre aqueles que o dominem ou com eles interajam, conduzindo a possíveis questões de responsabilidade civil – quem responde por eles? –. As diferentes *engenharias* criam constantemente novos *produtos* que pedem ao regime jurídico um estatuto que os receba.

Veja-se, a título de exemplo, o conjunto de questões possivelmente levantadas pelo recurso a engenhos robóticos.

(b) As técnicas de *inteligência artificial* são *factor* de *condicionamento* da pessoa humana[5].

Isto é, não só criam novos objectos para a pessoa humana (a), como fazem da pessoa humana um seu objecto. A inteligência artificial aproveita agora as intenções da *psicologia* e da *economia comportamental* – dotando-as aliás de novos dados e de técnicas de processamento que permitem atingir novas "conclusões" – com vista a obter com mais eficácia os propósitos que visa. Tais finalidades podem oscilar entre o auxílio (por ex., para aprimorar o resultado de uma pesquisa por intermédio de um motor de busca), a manipulação (por ex., publicidade orientada), ou a viciação (por ex., videojogos com níveis de dificuldade variável, do mais fácil ao mais difícil, como meio de criação de dependência; promoção do consumo impulsivo).

Tal condicionamento pode operar desde um outro ponto de vista: permitindo o tratamento de dados relativos à pessoa que condicionem as decisões que possam ser tomadas a respectivo respeito (pense-se, por ex., na prática de *credit scoring*)[6].

Gabriele Buchholtz, "Legal Tech. Chancen und Risiken der digitalen Rechtsanwendung", *Juristische Schulung* 10/2017, pp. 955 e ss.

[5] Cf., Hoffmann-Riem, "Verhaltenssteuerung durch Algorithmen...", cit., pp. 11-14.

[6] Cf. Martini, "Algorithmen als Herausforderung...", cit., p. 1018, quando nota que não poucos cidadãos temem mais as práticas de *scoring* – que podem condicionar por ex. o acesso a crédito bancário, a concessão de seguros, etc. – do que as medidas de coacção estadual.

Uma regulação jurídica da inteligência artificial intervirá, neste âmbito, desde a perspectiva de protecção dos próprios destinatários diante do contacto com os métodos manipulativos; ou, então, colocando barreiras ao respectivo modo de funcionamento.

(c) As técnicas de *inteligência artificial* são *factor* de *condicionamento* da *comunidade política*[7].

O mesmo tipo de condicionamento que pode ter lugar à escala pessoal pode ocorrer ao nível social. Aliás, o emprego da inteligência artificial está mesmo vocacionado a conseguir fazê-lo, uma vez que, pelos meios a que recorre, consegue dirigir-se a grandes números de pessoas. Oferece, agora, a possibilidade de condicionamento da opinião pública, o que pode ocorrer para uma multiplicidade de efeitos: afirmação de certas linhas de opinião, propósitos eleitorais, etc.

Desta feita, a *inteligência artificial* cruza-se especialmente com a economia, a demografia, a política.

Em qualquer um dos últimos dois pontos, o impacto da *inteligência artificial* está estreitamente ligado a um conjunto de invenções técnicas que lhe permitem colocar em *rede* grandes números de pessoas: os meios informáticos de computação, que são enormes instrumentos de cálculo; a *internet;* os *smartphones,* com quanto implicaram de exposição ininterrupta a uma rede na qual mecanismos de inteligência artificial sistematicamente intervêm.

A regulação jurídica há-de ser feita nos mesmos termos da vertente anterior.

(d) As técnicas de *inteligência artificial* – e não só os seus produtos, como em (a) –, são, elas próprias, *objecto* de regulação jurídica, sobretudo pelas consequências antes referidas em (b) e (c).

A forte assimetria entre o poder de condicionamento obtido por meios de inteligência artificial e a posição da pessoa impõe que se balizem os termos mediante os quais opera. Se, nas duas vertentes anteriores, a regulação olhava a inteligência artificial desde os respectivos os *efeitos,* nesta última olha desde o seu próprio "alimento", por ex., mediante o estabelecimento de regras relativas à recolha e processamento de dados.

(e) Finalmente, as técnicas de *inteligência artificial* pode constituir um *meio* ou uma *técnica* de regulação jurídica[8] – a chamada (na designação anglo-saxónica) *Legal Tech* ou *Legal Technology*.

[7] Cf. Hoffmann-Riem, "Verhaltenssteuerung durch Algorithmen...", cit., pp. 14-15.

[8] A inteligência artificial é somente a ponta da lança da intervenção da informática no Direito. De facto, são muitas as formas de recurso à electrónica ou informática ao serviço do Processo Civil, e que oferecem seguramente fortes ganhos processuais. Pense-se, a título de exemplo, no

Presentemente, equacionam-se as seguintes possibilidades: por um lado, programas informáticos de análise de documentos jurídicos, de auxílio à função de julgar, ou mesmo de sua substituição (em âmbitos mais restritos: indemnizações em âmbito aéreo; acidentes de viação; previsão de reincidência[9]); por outro, em verdadeiras técnicas autónomas de resolução de litígios (por ex., no âmbito de plataformas electrónicas) e que, ao menos no plano fáctico, constituem espaços privados de afirmação do Direito[10].

É nestes cinco cruzamentos entre as técnicas de inteligência artificial e o Direito que terá lugar a *juridificação* deste fenómeno, que se encontra ainda largamente por realizar[11].

Perante estes cinco campos de relação, vê-se que a relação entre Direito e inteligência artificial (ou o que por ela se pretende realizar) não é unidireccional: não se encontra o Direito somente fora do mundo, regulando autocrática e independentemente um novo fenómeno que lhe é apresentado. Se tal teoricamente poderia ocorrer em dois planos, o (a) e o (d), já nos restantes poderiam ser as técnicas de a inteligência artificial (na verdade, *quem* delas dispõe) que indirecta, como em (b) e (c), ou mesmo directamente, como em (e), se substituem e se configuram como uma alternativa ao Direito.

Destes cinco planos, é perante o último que se situa a presente reflexão. Em especial, incide a reflexão sobre o emprego de decisões com recurso a inteligência artificial, que têm lugar quando "surja autonomamente uma decisão automática e a estrutura do processo decisório já não possa ser reconstituída"[12].

regime apresentação de actos processuais a juízo por via electrónica (art. 144º, 1, do Código de Processo Civil) ou de inquirição por meio tecnológico (art. 502º do Código de Processo Civil). Simplesmente, não interferem – ou não o fazem com tanta intensidade – no processo metódico de realização do Direito. Tal é o *nervo* em que toca a inteligência artificial, e que lhe oferece clara singularidade: não a eficiência do processo, mas o próprio método jurídico. Sobre a *Legal Tech,* cf. as notas de Hoffmann-Riem, "Verhaltenssteuerung durch Algorithmen...", cit., pp. 15-17. Sobre a relação entre a introdução destes meios tecnológicos e a afectação do núcleo do método jurídico, cf. Susan Hähnchen/ Paul T. Schrader/ Frank Weiler/ Thomas Wischmeyer, "Legal Tech. Rechtsanwendung durch Menschen als Auslaufmodell?", *Juristische Schulung* 7/2020, p. 627.

[9] Cf. o software designado *Compas: Correctional Offender Management Profiling for Alternative Sanctions.*

[10] Cf. Martin Fries, "Paypal Law und Legal Tech – Was macht die Digitalisierung mit dem Privatrecht?", *Neue Juristische Wochenschrift* 39/2016, pp. 2860 e ss.

[11] É disso sinal a obra de Peter Mankowski, *Rechtskultur,* Mohr Siebeck: Tübigen, 2016, que, no retrato que propõe da cultura jurídica contemporânea, não oferece especial destaque a esta matéria.

[12] Christian Ernst, "Algorithmische Entscheidungsfindung und personbezogene Daten", *Juristenzeitung* (21/2017), p. 1027.

3. Um Direito Civil sem pessoa?

Tal delimitação leva a que se coloque a seguinte pergunta: será possível um Direito Civil sem pessoa?

Posto que o sentido de uma tal interrogação não é unívoco, há-de ser precisado. Com efeito, a referência à pessoa pode tomar por referência o *objecto* da regulação ou o *agente* comparticipante na realização do Direito Civil.

Se a questão for colocada na primeira das acepções, não parece adquirir especial sentido. Do ponto de vista do objecto da regulação jurídica, seja ela a regulação jurídico-civil ou outra, haverá – enquanto houver humanidade plural – sempre um *alguém*. Pouco importa de que modo semelhante regulação adquire a luz do dia ou é levada a cabo: ela há-de incidir sobre pessoas e as suas relações.

A pergunta relativa a um Direito Civil sem pessoa adquire significado desde a segunda perspectiva, ou seja, dos agentes comparticipantes na realização do Direito Civil. Onde quer que os meios de realização do Direito Civil sejam substituídos por técnicas automatizadas de inteligência artificial, substituindo-se – ou secundarizando-se de modo significativo – o juízo humano em favor de outros meios de resolução da controvérsias[13], aquilo que teremos será ainda um *Direito Civil* ou já uma outra nova forma de ordenação jurídica?

Ponderamos, para efeitos da presente reflexão, o recurso a programas decisórios que operem integral ou praticamente como substitutivos de uma decisão judicial humana: integralmente, na hipótese de não haver nenhuma participação humana (no momento decisório propriamente dito); praticamente, quando a haja, mas sem capacidade efectiva de formular as mais das vezes um juízo autónomo em relação ao proposto pelo programa auxiliar de decisão[14].

Dir-se-á que a inteligência artificial está destituída de muitas das capacidades básicas do ser humano, o que, por si só, denunciaria o sem sentido daquela

[13] O Regulamento Geral da Protecção de Dados (Regulamento (UE) nº 679/2016, de 27 de Abril) não oferece, nos termos do art. 22º, senão uma limitada protecção contra decisões automatizadas tomadas com base em "perfis"; e não exclui a possibilidade de decisão tomada com o auxílio de dados automaticamente trabalhados. Sobre este regime, cf. Ernst, "Algorithmische Entscheidungsfindung...", cit., pp. 1029-1032 e Martini, "Algorithmen als Herausforderung...", cit., p. 1120.

[14] Realmente, as diferenças entre uma decisão puramente automatizada ou ainda com mediação humana (mas recurso a meios auxiliares) podem ser muito estreitas – tão mais estreitas (mesmo inadvertíveis) quão mais intenso for o grau de confiabilidade em meios automatizados e menor a disponibilidade temporal para um autónomo juízo judicativo. Cf. o apontamento convergente de Hoffmann-Riem, "Verhaltenssteuerung durch Algorithmen...", cit., p. 36; e veja-se também, do mesmo autor, "Der Umgang mit Wissen bei der digitalisierten Rechtsanwendung", *Archiv des öffentlichen Rechts* 145 (2020), pp. 1 e ss., 24-25.

interrogação[15]. De entre as capacidades de que não dispõe, está uma que é, aliás, essencial ao ser humano e ao seu pensar, a capacidade de *conceptualização*: de, a partir de objectos diferenciados, chegar à formação de *conceitos* que sirvam de grelhas de interpretação e intelecção de uma realidade infinitamente complexa. Semelhante capacidade de conceptualização, de abstracção da diferença, é meio mediante o qual o ser humano expressa a sua racionalidade pela palavra, o seu *lógos*. Bem diferente é, portanto, a inteligência artificial, que, se tem uma *lógica* – a do *processamento* –, contudo não tem nenhum *lógos*: opera por associação, agregação, cálculo, a partir de dados que lhe são fornecidos.

Mas a comparação não deve ser feita mediante o confronto dos termos em que a razão humana e a automação da inteligência artificial operam – a conceptualização humana *vs*. o processamento de cálculo automatizado –; mas em termos simplesmente funcionais[16]. Um *automóvel* é certamente capaz de menos diversidade de movimento do que a pessoa humana (não sabe saltar, nem fazer a roda, nem dar uma cambalhota; etc.); mas os que faz, fá-los com bem mais eficácia.

A questão, portanto, não é o que a inteligência artificial *não é capaz* de fazer[17]; mas se *o que é capaz de fazer* está apto a substituir a capacidade propriamente humana, tornada inútil ou irrelevante perante o novo modo de resolução de controvérsias. Que o ser humano é apto a conviver num mundo radicalmente desconhecido, porém funcional, é ponto manifestado pela experiência quotidiana de realização de uma viagem com o auxílio do *Google maps*[18].

[15] Para uma ponderação entre as (incomparáveis) técnicas de decisão humana e com recurso à inteligência artificial, cf. Ernst, "Algorithmische Entscheidungsfindung...", cit., pp. 1026-1036, e Wolfgang Hoffmann-Riem, "Der Umgang mit Wissen...", cit., pp. 1-39, Hähnchen/ Schrader/ Weiler/ Wischmeyer, "Legal Tech...", cit., pp. 625 e ss. Kyriakos N. Kotsoglou, "Subsumptionsautomat 2.0. Über die (Un-)Möglichkeit einer Algorithmisierung der Rechtserzeugung", *Juristenzeitung* 69 (9/2014), pp. 451-457 – com *Erwiderung* de Martin Engel, "Algorithmisierte Rechtsfindung als juristische Arbeitshilfe", *Juristenzeitung* 69 (22/2014), pp. 1096-1100 –, e Daniel Timmermann/ Katharina Gelbrich, "Können Algorithmen subsumieren? Möglichkeiten und Grenzen von Legal Tech", *Neue Juristische Wochenschrift* 1-2/2022, "Können Algorithmen subsumieren?...", cit., 25 e ss.

[16] As sociedades modernas, conforme aponta Rosa, *Unverfügbarkeit...*, cit., pp. 14-15, procuram o seu equilíbrio no *dinamismo* da evolução técnica. Tal pede funcionalidade, não verdade.

[17] Na verdade, a inteligência artificial só é "inteligência" no confronto com uma perspectiva muito redutiva do que é a cognição humana, pressupondo um radical e insustentável dualismo entre corpo e capacidade cognitiva/ intelectual. Cf. os esclarecimentos de Oliver Schlaudt, *Das Technozän*, Frankfurt a.M., Klostermann, 2022, pp. 130-131.

[18] Cf. já as célebres páginas de Max Weber, *Ciência como Vocação*, trad. de Artur Mourão, s.l., disponível em lusosofia.net, pp. 13-14. Tem por base conferência proferida em 1917 e publicada em 1919.

Dito de outro modo: parece certo que a inteligência artificial não se consegue substituir ao modo humano de pensar e julgar. Mas pode porventura *funcionar* de um modo tal que se aceite como substitutivo do próprio das valências humanas. E, por isso, torna-se possível que o ser humano simplesmente abdique desse seu modo humano de proceder e de julgar, e opte por se submeter aos mecanismos de processamento alternativos da inteligência artificial – como em tantos outros momentos históricos optou já por substituir o emprego de algumas das suas faculdades pela utilização de meios tecnológicos.

A tentação é muito forte.

Uma das formas privilegiadas de pensar o Direito na modernidade, quando não a mais frequente em discurso teorético, foi a de o tomar como *técnica de transformação* da sociedade de acordo com uma *vontade* que se expressa em *legislação*. Mediante essa legislação, poderia o soberano recriar a ordem de um mundo cujos referentes objectivos se viam a erodir[19]. Neste tipo ideal, o momento de afirmação do Direito seria o instante legislativo[20]: a partir daí caberia apenas aplicá-lo e impô-lo com eficácia. Como seus valores fundamentais, estariam a "vontade" legislativa, a montante, e a certeza e a segurança na sua aplicação, a jusante. O Direito, portanto, é pensado como *código* aplicável a uma pluralidade humana, de uma forma tal que as regras que balizam a sua interacção se tornam claras, previsíveis, unívocas[21]; e é calculada a sua maior ou menor valia em função da eficácia do seu *enforcement*.

Max Weber descreve uma tal perspectivação da ordem jurídica nos seguintes termos:

> "Pretendemos falar de ordem jurídica quando se prevê a aplicação de quaisquer meios de coacção físicos ou psíquicos que são exercidos por um aparelho de coacção, isto é, por uma ou mais pessoas que estão prontas para esta função para a eventualidade de se verificar a previsão [das normas que determinam a

[19] Cf. o texto fundamental de Jean Bodin, *Les six livres de la Republique,* Lyon, Imprimerie de Jean de Tournes, 1579 (especialmente capítulos 8 e 10 do livro I). Para enquadramento, ver Thomas Leinkauf, *Geschichte der Philosophie,* 6, *Die Philosophie des Humanismus und der Renaissance* (Hrsg. Wolfgang Röd), München, C.H. Beck, 2021, pp. 329 e ss.

[20] Só numa semelhante tradição se poderá dizer, como Heck, "Gesetzesauslegung und Interessenjurisprudenz", *Archiv für die civilistische Praxis* 112 (1914), 17 (*§2.6*), que a "Política não é outra coisa do que tomar parte na actividade legislativa" ("Denn Politik ist nichts anderes als Teilnahme an der Gesetzgebung").

[21] Cf. Zygmunt Bauman, *Modernidade e Ambivalência,* trad. de Marcus Penchel, Lisboa, Relógio d'Água, 2007.

aplicação de meios de coacção], quando, portanto, um certo tipo de socialização existe com vista ao exercício da coacção jurídica"[22].

Assim se pense o Direito, fica visível até que ponto a inteligência artificial pode constituir uma alternativa real – se o nao for já – a outras formas de pensar a regulação jurídica. Bastará pedir sucessivas operações de cálculo que optimizem os resultados alcançados. O que o uso a técnicas de inteligência artificial promete é apenas um exercício optimizado de um modo afinal não tão inédito de pensar o Direito.

Duvidoso é que aquele modo de pensar o Direito, ao jeito de algumas das feições próprias da modernidade, reflicta realmente o processo metódico que se tem por mais desejável[23]. A aspiração de conceber o Direito enquanto código demiurgicamente pré-posto – *prescrito* – que reforma e transforma a vida política no exacto sentido intencionado pelo agente legiferante (que, por essa razão, tenderá a pretender a identificação entre a ordem jurídica e o acto legislativo) está bem sedimentada. Mas se ela chegou a ser realmente recebida com semelhante radicalidade, ou se é sequer desejável que o seja, é diferente questão[24]. De há muito, com efeito, que se adquiriu a consciência da especial importância do momento jurisdicional, como um instante autónomo em relação ao legislativo de realização do Direito, porque visando dar respostas a *casos* de complexidade irredutível à simplicidade dos critérios normativos de regras de decisão "pré-postas".

Um Direito politicamente enraizado subsiste, na verdade, não pela incessante renovação de momentos autocráticos de criação de novas regras decisórias, mas através da *mediação* desses novos critérios pela dogmática jurídica e da sua *recepção* pela comunidade jurídica, em particular por intermédio de uma rede descentralizada e plural de instâncias jurisdicionais, se e em quanto reconhecem validade ao critério pré-posto[25]. Subsiste, portanto, na medida em

[22] Max Weber, *Wirtschaft und Gesellschaft. Recht* (Hrsg. Werner Gephart/ Siegfried Hermes), Tübingen, Mohr Siebeck, 2014, pp. 369-370 (tradução minha).

[23] Cf. Timmermann/ Gelbrich, "Können Algorithmen subsumieren?...", cit., pp. 25, 29, 30.

[24] Cf. as referências realizadas, sp., na n. 15.

[25] Do *Decretum* de Graciano, c. 1140, parte integrante do *Corpus Iuris Canonici,* consta uma formulação muito clara do modo como se articulam o momento da criação e o momento da recepção: "Leges instituuntur, cum promulgantur, firmatur, cum moribus utentium approbantur" (IV, III). Uma tradução possível seria: "As leis são instituídas mediante a respectiva promulgação; adquirem firmeza quando são recebidas pelo comportamento dos seus destinatários". O *Codex Iuris Canonici,* ora na versão de 1917, ora de 1983, apenas explicitou a primeira parte do princípio. Cf. o cân. 8, §1 do Código de Direito Canónico de 1917 ("Leges instituuntur, cum promulgantur"); e o cân.

que o momento da "prática" (do Direito) recusa ser absorvido pelo momento da "criação"[26]; em que o *caso* se impõe na sua autónoma fisionomia à consideração julgador[27].

Semelhante modo de realização do Direito tem por corolário compossibilitar que os destinatários da decisão tenham a real possibilidade de efectivamente comodelarem os termos em que a ordem jurídica se lhes aplica, mediante uma rede – uma rede de relações humanas, entenda-se – que permite servir de crivo no concreto processo de decisão. Em lugar da "uniformidade" que resultaria de reconduzir todo o fenómeno do jurídico ao momento da criação de regras legais, ou do "casuísmo" resultante inversamente de tudo apostar somente no momento resolutivo, eis o equilíbrio resultante da adequada correlação entre momento de criação normativa e momento de recepção judicativa.

Desta circunstância resultará, a meu ver, parte da força do Direito Civil, enquanto Direito cujos critérios normativos, para conseguirem uma efectiva aplicação prática, devem passar pelo filtro de um amplíssimo conjunto de *mediadores,* que se conseguem representar na posição de possíveis destinatários da aplicação dessas mesmas normas, e, desde esse exercício empático, ajuizar da respectiva justeza concreta (*epieíkeia*). O Direito Civil é o Direito comum da pessoa comum também por que lhe pertence: não no sentido de cada um se elevar à condição de juiz (ou legislador) em causa própria, que poria em causa uma objectivação da ordem jurídica dependente da existência de uma instância decisória distinta das partes, mas pela garantia de que o destinatário de aplicação de cada norma – as pessoas constituídas enquanto sujeitas aos efeitos de uma decisão com valor obrigatório – conta com que a norma que será aplicada ao seu caso verá o seu sentido determinado por um outro-eu, uma outra pessoa comum, que controla o respectivo ajuste como fonte do critério de resolução do caso.

Quer dizer: a garantia de acesso a um *processo* não tem por objecto apenas uma "solução" da causa, mas uma definição dos termos de composição do litígio mediante um *julgador* que possa ser interpelado, e que no seu processo de decisão *pondere* a adequação do quadro normativo ao caso concreto.

do Código de Direito Canónico de 1983, com a única diferença de dele constar uma redacção no singular ("lex instituitur cum promulgatur").

[26] Para estas distinções, cf. Hannah Arendt, *A Condição Humana,* trad. de Roberto Raposo, Lisboa, Relógio D'Água, 2001, pp. 271 e ss. (nº 31).

[27] Notando como o modelo algorítmico é, por via das suas regras de funcionamento, incapaz de olhar o caso na sua singular irredutibilidade, cf. Ernst, "Algorithmische Entscheidungsfindung...", cit., p. 1028.

Dizia há pouco que uma das formas privilegiadas de pensar o Direito na modernidade, quando não a relevante, foi a de o olhar como *técnica de transformação* da sociedade de acordo com uma *vontade* que se expressa em *legislação*. É certo. Mas também certo é que, sob a semelhante discurso legitimador, pelo menos o Direito Civil viveu, afirmou-se e foi tranquilamente aceite de um outro modo, realizando-se ao abrigo de diferentes esquemas metódicos e com base em pressuposições distintas da redução do fenómeno do Direito ao instante legislativo, mas não por isso menos aceites (ainda que não devidamente tematizadas).

Ora, é este equilíbrio que é ou pode ser fortemente colocado em causa pelas técnicas de *inteligência artificial*, enquanto elas contribuam para eliminar o momento propriamente decisório, conduzindo a que, em contrapartida, prepondere o momento de definição das regras jurídicas aplicáveis. Onde tal transição ocorra, prevalece uma configuração do Direito enquanto forma *imposta* à sociedade, eliminando-se o momento da recepção, e reduzindo-se o Direito somente a *projecção* do poder político. Com a subsequente elisão da função do jurista[28] e a mutação da índole própria do momento de realização do Direito: de acto *social* passa a automatismo tecnológico[29].

Claro que também o julgador se pode demitir da sua função mediadora, recebendo acriticamente as regras pré-dadas, sem grande esforço de precisão do seu sentido. Mas aí, muito correctamente, dizemos que já está a agir roboticamente. Ao menos perante um *robot* humano sobra sempre a esperança de o conseguir fazer acordar da letargia em que se encontra.

Num quadro robótico e automatizado de configuração do Direito, portanto, o Direito Civil tal como historicamente o reconhecemos – enquanto Direito assente em ponderações, em equilíbrios, em juízos de sobreposição de entendimentos – não chega a ter lugar, sendo substituído por um diferente esquema de resolução de controvérsias que sacrifica a concertação de acções humanas plurais à afirmação uniformizadora de um unitário poder criador.

Numa perspectiva, temos o Direito como operando apenas com "dados": aplicação de um "dado normativo", a norma, a certos "dados factuais", gerando a estatuição prevista pela norma. Na outra, temos a clara consciência de que a realização do Direito é uma *acção humana,* um julgamento *histórico* aconte-

[28] Cf. igualmente as reflexões de Matthias Kilian, "Die Zukunft des Juristen. Weniger, anders, weiblicher, spezialierter, alternativer – und entbehrlicher?", *Neue Juristische Wochenschrift* 42/2017, pp. 3043 e ss.

[29] Nestes últimos termos, contrapondo a índole social do acto de realização do Direito ao recurso a meios tecnológicos, cf. Hoffmann-Riem, "Verhaltenssteuerung durch Algorithmen...", cit., pp. 26-31 e Buchholtz, "Legal Tech...", cit., pp. 956-957.

cido no tempo e no espaço, que solicita a responsabilidade do decisor pelo dizer da sua actuação.

4. A alternativa à decisão automatizada: a automação sem decisão

A interferência dos meios automáticos no Direito Civil pode, porém, dar-se de um diferente modo. Não já, agora, mediante o recurso a meios alternativos *de* decisão, mas de alternativas *à* decisão.

Servem de exemplo os instrumentos procedimentais destinados a excluir do âmbito de juízos judicatórios um amplo conjunto de relações jurídicas, procedendo ao respectivo "acertamento" mediante puros processos automatizados. É o que acontece com a injunção.

Vejamos sumariamente o respectivo regime, tal como presentemente regulado, entre nós, no Decreto-Lei 269/98, de 1 de Setembro[30]. Requerendo alguém, dentro do âmbito de aplicação do respectivo regime, a interpelação de outrem para realizar o cumprimento de uma obrigação, é este interpelado com a cominação de, nada dizendo, o requerimento de injunção adquirir valor executório (cf. os arts. 7º, 9º e 12º do Anexo ao diploma). O silêncio do requerido é, nos termos da lei, suficiente para se considerar automaticamente acertada, para efeitos executivos, a relação jurídica que determinou o recurso ao procedimento injuntivo (art. 14º, 1). É um regime que contrasta de modo acentuado com o previsto para a revelia do Réu em processo de declaração: aí não se prevê a consequência imediata de condenação no pedido, mas ainda a apreciação do litígio pelo julgador (art. 567º do Código de Processo Civil).

Estamos, pois, perante uma técnica de "acertamento" de uma controvérsia que prescinde explicitamente de um qualquer juízo – mas cujos efeitos, nomeadamente para efeitos executivos, serão equiparáveis (com algumas excepções) aos de uma sentença judicial (cf. o art. 857º e 731º, *a contrario,* do Código de Processo Civil). Pode ser, aliás, modelada de uma forma tal que a aposição da fórmula executória venha a ter lugar automaticamente, por simples funcionamento de meios informáticos. Por essa razão, é expediente altamente promovido no âmbito de relações contratuais de massas, como meio prático para tramitar grandes volumes de pedidos de tutela judicial.

[30] Modificado, por último, pela Lei nº 117/2019, de 13/09. Cf. tb. o Regulamento (CE) nº 1896/2006 do Parlamento Europeu e do Conselho de 12 de Dezembro de 2006 que cria um procedimento europeu de injunção de pagamento Notam igualmente a relação entre estes regimes e o problema da "automação" Martin Engel, "Algorithmisierte Rechtsfindung...", cit., p. 1100 e Buchholtz, "Legal Tech...", cit., p. 959.

Serve o exemplo da injunção – e o mesmo se poderá dizer, *mutatis mutandis,* para os chamados *títulos judiciais impróprios* – para ilustrar de que modo, como antes dissemos, o recurso a meios automatizados, mesmo não provendo aquilo que pode ser facultado por um julgador, pode, porém, constituir uma sua alternativa. Mesmo que essa alternativa passe pela negação da intervenção de um julgador, a negação de um processo, a negação de um juízo, a negação de uma metódica, a negação de uma certa ideia de Direito.

5. A relação do ser humano com a tecnologia

Estamos, pois, em condições de avançar para o momento conclusivo desta reflexão. Que tem em vista responder à pergunta: qual a *índole específica* do problema que a *inteligência artificial* solicita ao Direito?

O quadro de fundo desta questão é a relação entre o ser humano e a tecnologia, relação que, aliás, codeterminou a própria evolução humana[31]. E que suscita (ao menos) três tipos de problemas:

(a) O da *relação entre a vida humana* e o *uso de artifícios* por si criados.

Tal relação pode ser fecunda, quando os artifícios desenvolvidos pelo ser humano promovem aquilo que o ser humano reconhece por valioso; ou prejudicial, no caso inverso[32].

(b) O da repartição do *domínio* sobre os novos artifícios gerados pelo ser humano.

Sobretudo numa fase em que os novos artifícios – neste caso, a inteligência artificial – adquirem uma saliente capacidade de dominação da vida pública (e privada), o domínio sobre os novos artifícios torna-se uma questão política de relevância central – como era a do domínio sobre *latifúndios* numa sociedade agrícola ou sobre os *meios de produção* numa sociedade industrial.

(c) O da assimetria de conhecimento tecnológico entre quem domina as novas técnicas (...anteontem os escribas das cidades mesopotâmicas, que dominavam uma selecta forma de registo, a escrita...) e quem as desconhece, mas é por elas afectado.

Tudo ponderado, estamos, porém, diante de problemas que assumem, não uma índole técnica que possa ser resolvida mediante mais técnica, mas

[31] Cf. Schlaudt, *Das Technozän...*, cit., pp. 45 e ss.

[32] Sobre este tema da relação entre a pessoa humana e a tecnologia, é particularmente interessante a reflexão que a personalidade de Ivan Illich conduziu ao longo da respectiva vida. Cf., para introdução à respectiva obra, David Cayley, *Ivan Illich in Conversation,* Anansi: Toronto, 1992, David Cayley, *The Rivers North of the Future. The Testament of Ivan Illich as told to David Cayley,* pref. de Charles Taylor, Anansi: Toronto, 2005 e David Cayley, *Ivan Illich – An Intellectual Journey,* The Pennsylvania State University Press: Pennsylvania, 2021.

política: problemas que respeitam, não propriamente ao modo o ser humano cria, desenvolve e optimiza os *artifícios* e as *técnicas* por si gerados, mas, depois de os gerar, ao modo como deseja com eles conviver. Compreende-se o fascínio gerado pelo desenvolvimento da técnica[33]. Mas transcendental a ela está determinar quais os âmbitos do seu emprego, e qual a confiança que nela se aceita poder depositar. Tal é a questão que pode ser propriamente política, e que para existir basta ser colocada.

Não tenho grandes dúvidas de que uma pura *era da automação* dos juízos decisórios será de impossibilidade de um Direito Civil, que é um Direito *de pessoas* no duplo sentido que um genitivo pode revestir: *objectivo* (Direito que tem por objecto pessoas) e subjectivo (Direito que, na sua determinação concreta, é da *autoria* de pessoas). Mas, mesmo se essa hora chegar, estará aí enquanto realidade meramente política (mesmo que a partir de inovações tecnológicas), sob a ameaça que se abate sobre qualquer realidade humana – a temporalidade que a sujeita à possibilidade da história e da mudança.

Dizia: não tenho grandes dúvidas de que uma *era da automação* será a da impossibilidade de um Direito Civil. Se, e em que medida, chegámos, chegaremos ou desejaremos sequer chegar a uma tal era da automação, que, por nova que seja, ameaça realizar anquilosados projectos transformação social, é, claro, uma diferente questão. Mas é questão que, a meu ver, merece três vezes uma resposta negativa.

Bibliografia

Arendt, Hannah, *A Condição Humana,* trad. de Roberto Raposo, Lisboa, Relógio D'Água, 2001

Bauman, Zygmunt *Modernidade e Ambivalência,* trad. de Marcus Penchel, Lisboa, Relógio d'Água, 2007

Bodin, Jean, *Les six livres de la Republique,* Lyon, Imprimerie de Jean de Tournes, 1579

Buchholtz, Gabriele, "Legal Tech. Chancen und Risiken der digitalen Rechtsanwendung", *Juristische Schulung* 10/2017

Cayley, David, *Ivan Illich – An Intellectual Journey,* The Pennsylvania State University Press: Pennsylvania, 2021

Cayley, David, *Ivan Illich in Conversation,* Anansi: Toronto, 1992

Cayley, David, *The Rivers North of the Future. The Testament of Ivan Illich as told to David Cayley,* pref. de Charles Taylor, Anansi: Toronto, 2005

Ernst, Christian, "Algorithmische Entscheidungsfindung und personbezogene Daten", *Juristenzeitung* 21, 2017

[33] Cf. as reflexões de H. Arendt, *A Condição Humana*..., cit., pp. 175 e ss. (nº 18).

Fries, Martin, "Paypal Law und Legal Tech – Was macht die Digitalisierung mit dem Privatrecht?", *Neue Juristische Wochenschrift* 39/2016

Hähnchen, Susan / Schrader, Paul T. / Weiler, Frank / Wischmeyer, Thomas, "Legal Tech. Rechtsanwendung durch Menschen als Auslaufmodell?", *Juristische Schulung* 7, 2020

Heck, Philipp, "Gesetzesauslegung und Interessenjurisprudenz", *Archiv für die civilistische Praxis* 112, 1914

Hoffmann-Riem, Wolfgang, "Der Umgang mit Wissen bei der digitalisierten Rechtsanwendung", *Archiv des öffentlichen Rechts* 145, 2020

Hoffmann-Riem, Wolfgang, "Verhaltenssteuerung durch Algorithmen – Eine Herausforderung für das Recht", *Archiv des öffentlichen Rechts* 142, 2017

Kilian, Matthias, "Die Zukunft des Juristen. Weniger, anders, weiblicher, spezialisierter, alternativer – und entbehrlicher?", *Neue Juristische Wochenschrift* 42/2017

Koenig, Gaspard, *La fin de l'individu. Voyage d'un philosophe au pays de l'intelligence artificelle,* Paris, Le Point/ Éditions de L'Observatoire, 2019

Kotsoglou, Kyriakos N., "Subsumptionsautomat 2.0. Über die (Un-)Möglichkeit einer Algorithmisierung der Rechtserzeugung", *Juristenzeitung* 69, 9, 2014

Leinkauf, Thomas, *Geschichte der Philosophie,* 6, *Die Philosophie des Humanismus und der Renaissance* (Hrsg. Wolfgang Röd), München, C.H. Beck, 2021

Mankowski, Peter, *Rechtskultur,* Mohr Siebeck: Tübigen, 2016

Martini, Mario, "Algorithmen als Herausforderung für die Rechtsordnung", *Juristenzeitung* 72, 2017

Rosa, Hartmut, *Unverfügbarkeit,* Suhrkamp: Berlin, 2020

Schlaudt, Oliver, *Das Technozän*, Frankfurt a.M., Klostermann, 2022

Timmermann, Daniel / Gelbrich, Katharina, "Können Algorithmen subsumieren? Möglichkeiten und Grenzen von Legal Tech", *Neue Juristische Wochenschrift* 1-2, 2022

IV
QUESTÕES DE RESPONSABILIDADE CIVIL

Aplicação da IA às DD: da responsabilidade civil do advogado?

Use of AI and DD: potential effects on lawyer's civil liability?

Cláudia Isabel Costa*

RESUMO: A advocacia está a ser *invadida* por vários produtos e serviços baseados em IA que almejam auxiliar o advogado na análise e tomada de decisões jurídicas. Em certas matérias estão já disponíveis no mercado aplicações que interagem diretamente com potenciais clientes, substituindo o advogado humano na resolução de questões jurídicas mais elementares. Na área do M&A (operações de fusão e aquisição) têm sido comercializadas várias soluções que prometem ajudar os advogados na análise de documentos e elaboração de relatórios. Contudo, parece-nos que esta comercialização deverá ser acompanhada de uma profunda reflexão ética e jurídica sobre a introdução destas tecnologias na prática de atos tipicamente jurídicos. É hora de analisarmos o novo tipo de relação que está a ser construído entre a IA, o advogado e os seus clientes. É hora de enfrentarmos a disrupção que a IA irá provocar nas normas deontológicas. Propomo-nos assim,

* Licenciada em Criminologia e Direito pela FDUP – Faculdade de Direito da Universidade do Porto. Mestranda na FDUL – Faculdade de Direito da Universidade de Lisboa e frequência de pós-graduação em AI&LAW. Advogada desde 2019 e atual Associada na Abreu Advogados. Intervém sobretudo nas áreas de M&A, Societário, Comercial e Contencioso/Arbitragem. http://linkedin.com/in/cláudia-isabel-costa-49b416150. https://abreuadvogados.com/pessoas/claudia-isabel-costa/

neste artigo, a uma brevíssima e introdutória reflexão sobre uma potencial e inovadora repartição de tarefas jurídicas entre a IA, o Advogado e o Cliente, bem como sobre o impacto que terá na responsabilidade civil profissional do advogado por danos causados ao cliente.

PALAVRAS-CHAVE: IA; Advocacia; M&A; DD; Normas Deontológicas; Responsabilidade Civil.

ABSTRACT: The legal profession is being invaded by several AI-based products and services that aim to assist the lawyer in legal analysis and decision-making. In certain matters, AI-applications are already available in the market that interact directly with potential clients, replacing the human lawyer in the resolution of more basic legal issues. In the area of M&A (mergers and acquisitions) several solutions have been marketed that promise to help lawyers in the analysis of documents and preparation of reports. However, it seems to us that this commercialization should be accompanied by a deep ethical and legal reflection on the introduction of these technologies in the practice of typically legal acts. It is time for us to analyze the new type of relationship that is being built between AI, the lawyer and their clients. It is time for us to face the disruption that AI will cause in deontological norms. Thus, in this article, we propose a very brief and introductory reflection on a potential and innovative division of legal tasks between AI, the lawyer and the client, as well as on the impact it will have on the lawyer's professional civil liability for damages caused to the client.

KEYWORDS: AI; Legal Practice; M&A; DD; Deontological Norms; Professional Civil Liability.

1. Introdução

Hoje em dia, o advogado tem ao seu dispor mecanismos de IA que podem alegadamente ajudá-lo a aumentar a eficiência do seu trabalho, a reduzir os seus custos e a libertá-lo das tarefas mais automáticas e rotineiras, para dedicar-se a tarefas mais complexas e especializadas.

Este tipo de soluções têm particular relevância na realização de *Due Diligence* ("DD") nas operações de M&A, conhecida por ser uma fase caracterizada pela revisão de centenas de documentos, através de um conjunto de atos que podem ser simples, rotineiros e automáticos.

Contudo, é necessário compreender até que ponto estes mecanismos podem ser utilizados. Além disso, também é necessário perceber quem será responsabilizado caso sejam causados danos por erros e omissões cometidas durante a DD.

Para responder a estas questões, começamos por tecer considerações breves sobre o a relação entre o Direito e a IA, passaremos à aplicação da IA às profissões jurídicas, e afunilaremos, para abordar em específico, o processamento de linguagem natural ("PNL") e a sua relação com a linguagem jurídica. Feita esta análise, passaremos a explicar o que é uma DD, as suas finalidades e procedimentos. Após o que, analisaremos eventuais situações de responsabilidade civil dos advogados e, por fim, as dificuldades que se colocam à imputação de danos ao Advogado, para daí retirarmos as nossas principais conclusões.

2. IA e Direito

2.1. Considerações gerais

Antes de mais, devemos esclarecer qual é o nosso posicionamento relativamente ao que se entende por IA e ao modo como esta se relaciona com o Direito.

A nosso ver, o termo IA para além de ser falacioso, tudo diz e nada diz sobre o que realmente é ou existe de inteligência e de artificial.

De um modo geral, numa primeira abordagem, qualquer um de nós é levado a interpretar o termo IA como uma inteligência semelhante ou igual à humana, em alguns casos até superior, mas existente em máquinas.

Do nosso ponto de vista, não poderíamos estar mais errados, entendemos que a IA é tudo sobre humanos e muito pouco sobre máquinas.

A IA é apenas um termo guarda-chuva que abriga múltiplas técnicas que se desenvolvem um campo amplo e interdisciplinar[1] que, em comum, têm como objetivo encontrar soluções práticas mais eficientes para determinados problemas[2]. Isto permite libertar o Ser Humano de determinadas tarefas,

[1] Neste sentido, Johannes J. Frühbauer, *Künstliche In telligenz, Autonomie und Verantwortung Erkundungen im maschinen- und roboterethischen Reflexionskontext*, p. 4, *in* <https://books.ub.uni-heidelberg.de/heibooks/reader/download/945/945-4-95750-2-10-20211109.pdf> (30.10.2022).

[2] Assim, Julian Nida-Rümelin e Nathalie Weidenfeld, *Digitaler Humanismus*, Munique, Max Planck Forschung, 2018, pp. 3 e 4.

dispondo de mais tempo para especializar-se em áreas de maior interesse. A esmagadora maioria das técnicas, senão todas, abrigadas neste guarda-chuva procura apenas replicar algumas características do Ser Humano para aplicá-las à resolução destes problemas[3].

Contudo, devemos ter em mente que, esta replicação ou imitação é feita por meio dos dados que lhe fornecemos e de sofisticados cálculos matemáticos[4].

Considerando o que se disse, estaremos perante uma inteligência que em comparação com a humana será sempre limitada, pois a máquina vê o mundo através de dados e cálculos. Ela é alimentada pelos nossos dados, não tem mundo interior, não tem emoções ou sentimentos (apenas é treinada para identificá-los e reagir).

Dito isto, importa esclarecer que vemos as várias técnicas IA apenas como um conjunto de instrumentos sofisticados[5] que o Homem tem ao seu dispor.

A nossa visão mantém-se ainda que estejamos a falar da sua forma mais evoluída e por conta da aclamada revolução da computação quântica que trará mais velocidade e eficiência a estas máquinas. De modo nenhum nos parece que temos motivos para humanizá-las, nem para abandonar uma abordagem antropocêntrica no Direito. Contudo, isso não significa que que não se deva analisar o impacto que esta tem nas normas jurídicas e procurar soluções que acomodem de forma mais justa os interesses das partes envolvidas.

No centro desta discussão está o termo algoritmo e esse conceito é aquele que vale apena analisar e compreender, pois apenas assim é que conseguiremos concluir com segurança como deverá o Direito reagir à disrupção que tem acontecido.

O algoritmo começou por ser um conjunto preciso de instruções ou regras, ou uma série metódica de passos que um programador dava a um computador para fazer cálculos, resolver problemas e tomar decisões[6].

Sucede que, nas últimas décadas os algoritmos têm sido desenvolvidos de forma a conseguirem analisar os grandes volumes de dados e informação que estão disponíveis online ("*Big Data*").

[3] *Idem,* pp. 3 e 4.

[4] Cf. Mark Lycett, "*Datafication": making sense of (big) data in a complex world' in* European Journal of Information Systems, 2013, p. 381 *in <https://link.springer.com/article/10.1057/ejis.2013.10>* (30.10.2022).

[5] Ver Joachim von Braun, Margaret S. Archer, Gregory M. Reichberg, Marcelo Sanchez Sorondo, Robotics, *AI and Humanity*, Science, Ethics, and Policy, Springer Cham, 2021, pp. 3 e 18.

[6] Cf. Juan Gustavo Corvalán, *Inteligencia artificial: retos, desafíos y oportunidades – Prometea: la primera inteligencia artificial de Latinoamérica al servicio de la Justicia,* p. 2 *in* < https://revistas.ufpr.br/rinc/article/view/55334> (30.10.2022).

A forma de desenvolvimento conhecida baseou-se na transição entre "*dar instruções precisas ao algoritmo*", do género, se acontecer a e b, faz c, para "*lhe darmos dados de treino*", ou seja, alimentarmos o algoritmo com uma grande quantidade de dados e através de cálculos matemáticos, ele analisa rapidamente toda essa informação, encontra padrões e dá-nos um resultado em forma de probabilidade. O papel do programador "(...) *não é dizer ao algoritmo o que deve fazer. É dizer ao algoritmo como treinar-se a si próprio para o que deve fazer, recorrendo aos dados e às leis das probabilidades*"[7]. É isto que nos dá a sensação artificial de estarmos perante uma máquina inteligente, autónoma, capaz de aprender e se adaptar. Elas não estão, contudo, isentas de erros e podem provocar danos. O problema é que a IA é uma *black box*, ela poderá devolver-nos determinados resultados, sem que saibamos que cálculos, padrões, informações a possam ter influenciado. Apesar de estamos perante instrumentos que na sua sofisticação podem escapar parcialmente à vontade e controlo do Humano que o usa ou beneficia dela, não negamos o problema central subjacente, a capacidade para provocar danos em grande escala e de forma totalmente imprevisível[8].

O uso destes mecanismos relaciona-se com o Direito a dois níveis. Por um lado, criam inúmeros obstáculos ao intérprete-aplicador na interpretação e aplicação de princípios jurídicos consolidados na ordem jurídica, bem como de institutos e normas jurídicas existentes, de aplicação à comunidade jurídica em geral. Exemplo disso, é a discussão em torno do reconhecimento e atribuição da personalidade eletrónica, da imputação de danos aos robôs, atribuição de direitos de autor a robôs ou exercício da intermediação financeira por sistemas de aconselhamento automático e plataformas de negociação algorítmica. Por outro lado, estes mecanismos podem ser utilizados no âmbito do exercício das profissões jurídicas, designadamente do juiz e do advogado, criando dificuldades de adaptação e aplicação das normas processuais e deontológicas na administração da justiça.

Roger Brownsword[9] desenvolve a este propósito a teoria da *Law 3.0* para explicar de que forma o Direito aborda este fenómeno disruptivo dos princípios e normas jurídicas.

[7] Cf. Nick Polson e James Scott, *Inteligência Artificial, Como funciona e como podemos usá-la para criar um mundo melhor*, Vogais, Lisboa, 2020, p. 11.

[8] Cf. Yavar Bathaee, *The artificial intelligence black box and the failure of intent and causation*, Harvard Journal of Law & Technology, vol. 31, nº 2, Spring, 2018, p. 913 *in* < https://jolt.law.harvard.edu/assets/articlePDFs/v31/The-Artificial-Intelligence-Black-Box-and-the-Failure-of-Intent-and-Causation-Yavar-Bathaee.pdf > (30.10.2022).

[9] Ver *Law 3.0*, Nova Iorque, Glasshousebook, 2021, p. 12.

De acordo com este autor, nós teríamos diferentes abordagens a este fenómeno, a *Law 1.0*, *Law 2.0* e *Law 3.0*.

Na primeira abordagem, procura-se aplicar as normas jurídicas existentes, tentando manter a coerência interna de aplicação dos princípios gerais e das normas jurídicas.

Contudo, as normas existentes poderão não ser adequadas ou podem ser omissas, seria então necessário revê-las e alterá-las para que pudessem cumprir de forma adequada o seu propósito, o que nos leva à segunda abordagem. Simultaneamente, órgãos de regulação e de aplicação da lei teriam também que ser atualizados e adaptados para garantir o cumprimento destas normas alteradas.

Neste tipo de abordagem, vive-se entre extremos, considera-se que as normas não serão adequadas se implicam uma excessiva regulação impedindo o desenvolvimento e aplicação destas novas tecnologias, mas também não serão adequadas se não regularem de forma suficiente e expuseram a comunidade jurídica a riscos inaceitáveis, procurando-se a melhor forma de gerir risco.

Passa-se assim à terceira bordagem, procura de soluções tecnológicas que possam ajudar a combater os riscos que surgem associados ao uso da IA.

Muitas outras abordagens poderiam ser referidas, mas parece-nos que esta capta e agrupa de forma particularmente certeira os vários discursos e discussões que existem a propósito do uso de mecanismos de IA.

Feito este enquadramento, centrar-nos-emos agora na forma como a IA tem influenciado a advocacia.

2.2. Uso de IA na advocacia

a) Fases evolutivas

O uso da IA aplicada ao Direito é um campo de estudo que existe desde 1980[10], tendo sido cunhado com o nome de *legaltech*.

A aplicação da IA à advocacia conhece três níveis de desenvolvimento.

O primeiro nível centra-se no uso da IA no *case flow management*, na revisão contratual e pesquisa jurídica, por serem tarefas que consumem muitas horas e comportam tarefas simples e repetitivas[11]. O objetivo será desenvolver estas técnicas de tal forma que possa libertar-se advogados de tarefas simples e rotineiras, para que estes possam focar-se em casos judiciais mais complexos e que exijam uma maior análise e atenção[12].

[10] Cf. Kevin Ashley, *Artificial Intelligence and Legal Analytics, New Tools for Law Practice in the Digital Age*, Cambridge, University Press, Reino Unido, 2017, p. 30.

[11] Christy Ng, *AI in the legal profession, in The Cambridge Handbook of Artificial Intelligence Global Perspectives on Law and Ethics*, Cambridge, Cambridge University Press, 2022, p. 38.

[12] *Idem*, p. 39

Neste nível incluem-se técnicas como a *cloud-based tools*, usada para armazenamento inteligente de documentos para que seja mais fácil procurá-los, programas computacionais que sejam capazes de fazer pesquisas mais compreensivas em entidades públicas, uso de mecanismos de processamento de linguagem natural ("PNL") para revisão de contratos, mecanismos que ajudem na gestão de assuntos como o *workflow automation*, *e-discovery tools* e extração de dados[13].

O segundo nível centra-se na substituição dos advogados por mecanismos de IA em questões mais substanciais que vão além de meras tarefas simples e rotineiras. Este nível de aplicação não está tão desenvolvido, não é tão publicitado e encontra fortes entraves ao seu desenvolvimento, tais como, inexistência de informação jurídica suficientemente estruturada para treinar os algoritmos. Neste caso, inclui-se o advogado-robô que seja capaz de interpretar contratos e avaliar o risco comercial, de pensar como um advogado e ler documentos de forma a conseguir elaborar relatórios com os pontos críticos encontrados na realização de uma DD[14].

Existem já alguns advogados-robô no mercado, o DoNotPay[15] e o LISA[16], o primeiro é uma aplicação que utiliza um algoritmo que decifra a linguagem quotidiana e ajuda a resolver questões jurídicas como a impugnação de coimas por excesso de velocidade. Este software torna-se mais sofisticado com o tempo, quanto mais informação o software analisa, melhor pode corrigir erros passados. O segundo cria um acordo de confidencialidade ("NDA"). O utilizador pode aceder à aplicação e a LISA faz uma série de perguntas, o NDA é criado e enviado para o utilizador. O utilizador faz no LISA as alterações que entender. Quando os interesses das partes estiverem todos acomodados, o documento está pronto para ser assinado. A LISA também pode aplicar-se à redação de contratos de arrendamento.

O terceiro nível de desenvolvimento inclui o uso de instrumentos de análise preditiva para ajudar na procura inteligente e previsão do resultado de um processo judicial usando dados agregados a partir de decisões judiciais[17]. Estes programas tentam imitar e executar determinadas caraterísticas do raciocínio humano na análise de uma situação, na resposta a uma questão

[13] *Ibidem,* p. 40.

[14] Christy Ng, *AI in the legal profession ...*, p. 40.

[15] Consultar https://donotpay.com/.

[16] Consultar https://robotlawyerlisa.com/nda/.

[17] Cf. Christy Ng, *AI in the legal profession, in The Cambridge Handbook of Artificial Intelligence Global Perspectives on Law and Ethics*, Cambridge, Cambridge University Press, 2022, p. 40.

legal ou na previsão de um determinado resultado. O elemento-chave nestes programas está em conseguirem traduzir o raciocínio jurídico em linguagem computacional[18].

b) Técnicas

No ponto anterior, verificamos que a *legaltech* tem conhecido várias fases de desenvolvimento e avançamos já com a enunciação de algumas técnicas. Iremos agora aproveitar para analisar com mais detalhe algumas destas técnicas para entendermos como é que funcionam na prática e de que forma estas auxiliam os advogados no seu dia-a-dia.

Comecemos com o *eDiscovery (eletronic discovery)* é um sistema de IA que pode ser usado para encontrar, entender e apresentar de forma eletrónica informação relevante para uma ação judicial. O advogado que queira usar este sistema deve, em primeiro lugar, seleciona uma amostra de e-mails e dá indicação ao sistema de palavras que sejam relevantes. Depois disso, a IA "analisa" o conteúdo de e-mails "percebendo" quem é que o envia, quem recebe, o que diz, para aprender os dados que o advogado indicou como sendo relevantes naquelas mensagens. Depois da IA aprender com esta amostra, o advogado "dá-lhe acesso" aos restantes e-mails para que a IA possa analisá-los e tente encontrar aqueles que preenchem os critérios que aprendeu. Desta forma, a IA faz logo uma pré-seleção dos e-mails que serão relevantes e que devem ser revistos pelo advogado, poupando tempo e recursos. De forma a assegurar que não ficou nada de relevante de fora, o advogado poder fazer uma revisão rápida aos e-mails deixados de fora, vendo apenas o assunto, por exemplo[19]. Este sistema inclui outras valências além da pesquisa jurídica, exemplo disso é a extração de texto, a categorização (agrupa de forma automática o que for semelhante) e a ferramenta de visualização, todas estas valências pode ser usadas simultânea ou separadamente[20].

A pesquisa jurídica é uma atividade de grande relevância na prática jurídica. Usar a IA para ajudar nesta tarefa significa que é preciso "ensinar" à IA o que é o direito e ensiná-la a encontrar informação que seja relevante para usar no aconselhamento jurídico, usar nos processos judiciais ou na elaboração dos contratos[21].

[18] *Idem,* p. 40.

[19] Explicação dada por Noah Waisberg e Alexander Hudek, *AI for Lawyers, How Artificial Intelligence is adding value, amplifying expertise, and transforming careers,* Canada, Wiley, 2021, p. 101.

[20] *Idem*, p. 105.

[21] *Ibidem*, p. 107.

Nesta técnica a linguagem assume um papel nuclear, quer na tentativa de compreendê-la, quer na tentativa de criá-la. Para que esta técnica seja bem sucedida, é necessário dar contexto e "ensinar" os conceitos à IA. A IA deve "entender" quais são as particularidades do caso, por exemplo, partes envolvidas, jurisdição competente, lei aplicável, qual o pedido, no caso de uma ação judicial, e outros dados que sejam relevantes. Adicionalmente, o sistema tem que "entender" o que procura para encontrar o que é relevante, conseguindo identificar matéria relevante mesmo quando a linguagem usada para iniciar a pesquisa é diferente daquele que é usada nos textos analisados. Neste último caso, será necessário "explicar" à IA quais as várias palavras ou frases que podem ser utilizadas com o mesmo significado[22].

Encontra-se ainda disponível no mercado, a *litigation analytics*, através da qual se introduz no sistema os factos de um caso judicial concreto e o sistema compara esses factos com todos os casos que tenham sido decididos pelo tribunal para identificar aqueles que são semelhantes, depois avalia o peso que cada um dos fatores teve na decisão tomada. Isto significa que a informação que dou ao sistema tem que ser da melhor qualidade possível para que os resultados também tenham a maior qualidade possível[23].

Adicionalmente, temos os *softwares* de revisão de contratos, os advogados introduzem os contratos no sistema, depois explica-se ao sistema o que deve procurar, que cláusulas deve encontrar, o sistema analisa encontra todas essas cláusulas, depois faz um sumário com os seus *findings*[24].

Por último, temos os expert systems através dos quais se recorre a técnicas de inferência para automatizar conhecimento através da lei substantiva, documentos e processos. Nestes sistemas, os padrões, os factos e as inferências que podem ser feitas são realizadas por humanos. Dito de outro modo, os especialistas mapeiam o tema em questão construindo um conjunto de regras que depois são introduzidas no *software*. Estes sistemas pode ser usados internamente para garantir a qualidade dos serviços que é prestada[25].

Antes de prosseguirmos com esta análise, focar-nos-emos agora numa técnica que será especialmente relevante nas DD e nas especiais dificuldades que encontra na sua implementação.

[22] Noah Waisberg e Alexander Hudek, *AI for Lawyers, How Artificial Intelligence is adding value, amplifying expertise, and transforming careers*, Canada, Wiley, 2021, pp. 111 e 112.

[23] *Idem*, p. 126.

[24] *Ibidem*, p. 136.

[25] Noah Waisberg E Alexander Hudek, *AI for Lawyers, How Artificial Intelligence is adding value, amplifying expertise, and transforming careers*, Canada, Wiley, 2021, p.147.

2.3. Linguagem Jurídica e PNL

A PNL é um campo multidisciplinar de estudo que associa a IA à linguística computacional[26], procurando que as técnicas consigam analisar a linguagem humana e identifiquem padrões nos textos escritos[27].

O grande desafio está em ensinar aos algoritmos como analisar a sintaxe e a semântica dos textos escritos pelos Seres Humanos[28], de modo a que o processem, entendam e produzam representações da linguagem da mesma forma que os Seres Humanos[29].

As teorias modernas da PLN foram impulsionadas por Noan Chomsky através da sua obra *Estruturas Sintáticas* e das suas observações a esse propósito.

Noan Chomsky explicava que a gramática seria um dispositivo que gerava todas as sequências gramaticais da linguagem e deixaria de fora todas as sequências não gramaticais. Assim, o primeiro passo, estaria sempre em distinguir entre as frases gramaticais das frases não gramaticais. O autor notou ainda que gramática era independente do significado da frase, sendo possível termos frases gramaticalmente corretas sem que estas tivessem qualquer significado. Ele acaba por analisar a descrição linguística como um sistema com vários níveis de representação[30].

Nos anos 80, a PLN começou a focar-se nos modelos probabilísticos e na estatística, após o que se começaram a usar os modelos de aprendizagem automática nos anos 90, conhecida por *machine learning*[31].

Do ponto de vista da PNL, o Direito pode ser visto como um conjunto de regras que podem expressar-se de forma lógica e ser transpostas para linguagem computacional.

O problema é que as normas jurídicas caracterizam-se por sere m gerais e abstratas o que dificulta a aplicação deste tipo de técnicas[32]. A PNL encontra também inúmeras dificuldades perante o facto de termos diferentes interpretações das mesmas normas jurídicas[33]. Sucede ainda que, muitas normas

[26] Cf. Yoav Goldberg, *A Primer on Neural Network Models for Natural Language Processing*, Journal of Artificial Intelligence Research, 57, 2016 p. 346.
[27] *Idem*, p. 347.
[28] *Ibidem*, p. 347.
[29] Cf. Yoav Goldberg, *A Primer on Neural Network...*, p. 348.
[30] Cf. Yoav Goldberg, *A Primer on Neural Network ...*, p. 349.
[31] Cf. Yoav Goldberg, *A Primer on Neural Network ...*, p. 349.
[32] Cf. Kevin Ashley, *Artificial Intelligence and Legal Analytics, New Tools for Law Practice in the Digital Age*, Cambridge, University Press, Reino Unido, 2017, p. 48.
[33] *Idem*, p. 49.

jurídicas têm remissões ou preveem exceções que estas técnicas têm dificuldade em identificar e interpretar[34].

Para contornar algumas destas dificuldades são utilizadas algumas técnicas como os processos de normalização sistemática. Esta técnica é aplicada da seguinte forma: procede-se à seleção de uma frase legal, à identificação das suas diferentes componentes, à verificação da sintaxe e reorganização da frase de forma a torná-la mais "*apta*" a ser transposta para linguagem informática[35].

Também têm sido utilizadas as redes neuronais[36], pois estas, conseguem lidar com a criatividade associada à linguagem e conseguem associar o significado da linguagem ao contexto devido[37].

Agora que ficamos a conhecer melhor as técnicas existentes, as suas dificuldades e o caminho que seguem, passaremos a abordar em concreto a DD.

3. Responsabilidade Civil do Advogado, IA e DD

3.1. DD no M&A

As operações de fusão e aquisição, comumente conhecidas por operações de M&A ("Fusões e Aquisições") ou "*Mergers and Acquisitions*") referem-se aos processos que conduzem à fusão, cisão e aquisição de empresas.

Em traços gerais, a fusão é uma operação jurídica através da qual, duas ou mais empresas se juntam para formar uma só, traduzindo-se num fenómeno de concentração económica de empresas[38]. Por seu lado, a cisão parte de uma sociedade para dar origem a outras duas sociedades, através da reestruturação descentralizada da sociedade inicial[39].

Estas duas operações jurídicas subdividem-se em várias categorias. Assim, dentro da fusão, teremos a fusão por incorporação e a fusão por concentração. Na primeira, a sociedade pré-existente mantém-se e incorpora outra, na segunda, duas ou mais sociedades pré-existentes, transferem as suas posições jurídicas para uma nova entidade criada para esse efeito[40].

[34] *Ibidem*, p. 50.

[35] Cf. Kevin Ashley, *Artificial Intelligence ...*, p. 51.

[36] As redes neuronais são é uma técnica de IA que podem ser definidas como uma rede composta por elementos de processamento simples, denominados de unidades, que estão interligados entre si, de constituição semelhante à do cérebro humano *in* Kevin Gurney, *An introduction to neural networks*, Londres, Taylor & Francis e-Library, 2004, p. 14 *in* <https://www.inf.ed.ac.uk/teaching/courses/nlu/assets/reading/Gurney_et_al.pdf> (30.10.2022).

[37] Cf. Yoav Goldberg, *A Primer on Neural Network...*, p. 347.

[38] Cf. António Menezes Cordeiro, *Direito das Sociedades, I, Parte Geral*, 3ª edição comentada e atualizada, Coimbra, Almedina, 2016, p. 1125.

[39] *Idem*, p. 1126.

[40] *Ibidem*, p. 1127.

A cisão poderá ser simples, cisão-dissolução e cisão-fusão. Na primeira categoria, a sociedade-mãe destaca parte do seu património e com ele constitui-se uma nova sociedade, no segundo caso, uma sociedade dissolve-se e divide-se o seu património, sendo cada uma das partes destinadas a constituir uma nova sociedade, e a última modalidade combina a fusão com a cisão.

A aquisição de empresas poderá ser feito através de *asset deals* ou *share deals*, a primeira incide sobre a compra da empresa ou sobre ativos e passivos considerados individualmente, e a segunda sobre a compra das participações sociais numa sociedade comercial[41].

A concretização com sucesso de cada uma destas operações exige que se tome um conjunto de decisões que devem ser prévia e devidamente ponderadas e avaliadas.

A configuração das operações de fusão e aquisição depende muito da localização da empresa adquirente e da empresa-alvo, da dimensão do negócio, da natureza e complexidade das atividades desenvolvidas pelas empresas envolvidas, informação que se dispõe sobre as empresas, dos modos de financiamento da operação[42], da análise dos intangíveis e dos seus recursos, da antecipação e avaliação de riscos e a sua minimização[43].

Nestas operações desempenha um papel fundamental a DD, em português, auditoria ou diligência pré-contratual[44].

O conceito *Due Diligence* vem do termo romano *diligentia,* no qual se distinguia *diligentia quam suis rebus*, referindo-se ao cuidado que uma pessoa normal deve utilizar na condução e gestão dos seus negócios, e *diligencia exactíssima* ou *diligentia boni patrisfamilia,* cuidado equivalente ao do chefe de família[45].

O termo começou depois a ser utilizado pelos intermediários financeiros norte-americanos depois da aprovação do *Securities Act* de 1933. Este processo de *Due Diligence* servia de fundamento de defesa aos intermediários financeiros na responsabilidade civil pelo prospeto. Os intermediários financeiros teriam que atuar com a devida diligência na análise que faziam à sociedade emitente e teriam de ter segurança nas conclusões que seriam divulgadas aos investidores. Depois esta prática passou para as operações de M&A[46].

[41] Cf. José Gomes Ferreira, *M&A, Aquisição de empresas e de participações sociais*, Lisboa, AEFDUL Editora, 2022, p. 33.

[42] Cf. Domingos Ferreira, *Fusões, aquisições, cisões e outras reestruturações de empresas – abordagem jurídica e "due diligence" integral*, volume 2, Rei dos Livros, p. 353.

[43] *Idem,* p. 355.

[44] *Ibidem,* p. 354.

[45] Cf. Domingos Ferreira, *Fusões, aquisições, ...,* p. 354.

[46] Cf. José Ferreira Gomes, *M&A – Aquisição de empresas e de participações sociais ...,* p. 34.

Hoje em dia, pode ser entendida como um procedimento que compreende diversas atividades de investigação, através do qual se procede à análise detalhada de todos os elementos considerados essenciais para avaliar e fixar o preço do negócio[47] e se garante que depois da concretização da operação não surgem contingências das quais não se tinha conhecimento[48].

A DD tem como finalidade imediata que o comprador conheça a situação da empresa, da sociedade que detém e das participações sociais que serão transmitidas e tem como finalidade mediata avaliar a relação risco-benefício associado à transação[49].

Podem ser realizados vários tipos de auditoria: legal, financeira, estratégica e de mercado, cultural, sobre as tecnologias e sistemas de informação, ambiental ou reputacional.

Na DD legal pretende-se proceder à análise de contratos, direitos corpóreos ou tangíveis, marcas, patentes, existência de litígios, fiscal. Desta forma, analisa-se a validade dos contratos, responsabilidades da empresa-alvo, indemnizações que esta deva pagar e perdas de valor, potenciais problemas legais que possam surgir no futuro e possam ser resolvidos antes da assinatura do contrato, avaliação de custos e possibilidade de contencioso, identificação de áreas de risco, com maior instabilidade e que possa ter impacto na transação[50].

As DD podem ser promovidas pelo vendedor, comprador, pode ser uma DD inicial ou confirmatória.

Normalmente, uma DD legal abrange áreas como societário, comercial, seguros, laboral, fiscal, imobiliário, regulatório, concorrência, licenciamento, ambiente, contencioso, proteção de dados e propriedade intelectual.

Devido aos mais recentes desenvolvimentos do ESG, o âmbito da DD poderá ser alargado para se verificar se os parâmetros de *environment, social* e *governance*, onde se inclui, por exemplo, o respeito pelos direitos humanos e a prossecução de objetivos sustentáveis, são cumpridos.

Além disso, com a intensificação do mercado tecnológico devido ao desenvolvimento da IA, *blockchain*, metaverso, também se poderá ampliar o seu grau de abrangência ao nível da propriedade intelectual/industrial e direitos de autor.

Este procedimento começa pela partilha de uma *checklist* que inclui pedidos de documentação específicos relativos às várias áreas acima mencionadas,

[47] Cf. José Ferreira Gomes, *M&A – Aquisição de empresas e de participações sociais*..., p. 35.
[48] Cf. Domingos Ferreira, *Fusões, aquisições* ..., p. 354.
[49] Cf. José Ferreira Gomes, *M&A – Aquisição de empresas e de participações sociais*..., p. 35.
[50] Cf. Domingos Ferreira, *Fusões, aquisições* ..., p. 441.

após o que a empresa-alvo recolhe os documentos solicitados e faz o *upload* dos mesmos para a chamada *Virtual Data Room* ("VDR"), estando depois disponíveis para a empresa adquirente e os seus advogados acederem e analisarem.

Depois de acederem aos documentos, o adquirente e os seus advogados analisam os documentos para identificar potenciais riscos que possam afetar a transação e o valor do negócio. Estes riscos são catalogados desde aqueles que podem ser resolvidos previamente à transação até aqueles que a impedem de acontecer.

Estas tarefas podem ser muito difíceis porque os documentos podem estar espalhados por muitos locais, levando a que informações importantes sejam negligenciadas ou difíceis de encontrar e considerando que a empresa-alvo, normalmente, precisa de digitalizar muitos documentos físicos[51].

Ao longo do processo de DD, as partes envolvidas negoceiam os termos da transação, sendo que, o documento final terá normalmente um conjunto de representações e garantias ou indemnizações específicas, muitas delas relativas aos riscos já encontrados na DD[52], até que conclui a DD e a transação.

Compreendida a DD e as suas finalidades, verifiquemos agora o papel fundamental do advogado.

3.2. Responsabilidade Civil do Advogado e DD

Os advogados estão sujeitos a vasto conjunto de regras deontológicas no exercício da sua profissão plasmadas no Estatuto da Ordem dos Advogados ("EOA"), aprovada pela Lei nº 145/2015, de 09 de Setembro

O EOA começa por dizer que a Ordem dos Advogados é uma associação pública representativa dos profissionais que exercem a advocacia (cf. artigo 1º do EAO). Apenas os advogados com inscrição em vigor na ordem podem praticar em Portugal atos próprios da advocacia (cf. artigo 66º, nº 1 do EAO).

Os atos próprios do advogados surgem elencados na Lei nº 49/2004, de 24 de agosto ("Lei dos atos próprios dos advogados"), e incluem, o exercício do mandato forense, a consulta jurídica, a elaboração de contratos e a prática dos atos preparatórios tendentes à constituição, alteração ou extinção de negócios jurídicos, designadamente os praticados junto de conservatórias e cartórios notariais, a negociação tendente à cobrança de créditos e o exercício do mandato no âmbito de reclamação ou impugnação de atos administrativos ou tributários.

[51] Cf. Ben Klaber, *Artificial Intelligence and Transactional Law: Automated M&A Due Diligence,* p.3 in <http://users.umiacs.umd.edu/~oard/desi5/additional/Klaber.pdf> (30.10.2022).

[52] Cf. Ben Klaber, *Artificial Intelligence* ..., p. 4.

O título profissional de advogado está exclusivamente reservado aos licenciados em Direito com inscrição em vigor na Ordem dos Advogados, bem como a quem, nos termos do respetivo estatuto, reúne as condições necessárias para o adquirir, sendo que quem praticar atos próprios dos advogados sem qualificação para o fazer comete o crime de procuradoria ilícita.

O crime de procuradoria ilícita visa tutelar o especial interesse público aqui está em causa, o da manutenção da *integridade ou a intangibilidade do sistema oficial instituído para a prática de atos próprios das profissões dos Advogados e Solicitadores*[53], e o da boa administração da justiça, assegurando de que os clientes apenas são assistidos por profissionais qualificados para o efeito.

De acordo com o EOA, o advogado é indispensável à administração da justiça e, como tal, deve ter um comportamento público e profissional adequado à dignidade e responsabilidades da função que exerce.

Ele deve ainda manter a honestidade, probidade, retidão, lealdade, cortesia e sinceridade no exercício da sua profissão (cf. artigo 88º do EOA).

A relação do advogado com o seu cliente deve fundar-se na confiança recíproca, devendo o advogado agir de forma a defender os interesses legítimos do cliente (cf. artigo 97º do EOA).

O advogado não deve aceitar o patrocínio de uma questão se souber, ou dever saber, que não tem competência ou disponibilidade para dela se ocupar prontamente, a menos que atue conjuntamente com outro advogado com competência e disponibilidade para o efeito (cf. artigo 98º do EOA).

Está ainda obrigado a estudar com cuidado e tratar com zelo a questão de que seja incumbido, utilizando para o efeito todos os recursos da sua experiência, saber e atividade (cf. Artigo 100º do EOA).

O advogado pode ser civilmente responsabilizado por ação ou omissão, se causou danos ao seu cliente, violando de forma culposa os deveres acima elencados.

Para que haja lugar a responsabilidade civil, será necessário que se verifique a culpa de quem presta o serviço, o dano de quem o serviço é prestado e o nexo de causalidade e a existência de um facto ilícito[54].

Quando se verifica uma violação do contrato por incumprimento ou incumprimento defeituoso das obrigações a que está adstrito, no âmbito do contrato que tem com o cliente, haverá lugar a responsabilidade contratual.

[53] Consultar em http://www.dgsi.pt/jtrc.nsf/8fe0e606d8f56b22802576c0005637dc/b199421521d5b646802585180043b51c?OpenDocument.

[54] Cf. Moutinho de Almeida, *Responsabilidade Civil dos advogados,* Coimbra, Coimbra Editora, 1985, p. 20.

Pode também existir uma situação de responsabilidade aquiliana, quando temos situações em que o advogado se locupleta com o dinheiro entregue pelo seu cliente para o pagamento de custas, algo que está fora do seu mandato[55].

É de grande relevância abordarmos aqui, o chamado erro de ofício, este resulta da violação da obrigação do advogado estudar com cuidado e tratar com zelo a questão de que seja incumbido, utilizando todos os recursos, experiência, saber e atividade[56].

Poderá ainda haver negligência se houve descuido em adquirir ou atualizar a técnica e os conhecimentos necessários ao bom desempenho da profissão ou quando exerce a profissão distraidamente, sem cuidados, cautelas e zelos e ainda quando há imprudência de achar que se tem uma capacidade que não possui[57].

O erro de ofício apenas se verifica quando haja lugar a uma ação por ignorância, inépcia ou negligência[58].

Todas as obrigações acima elencadas devem ser observadas no âmbito da assessoria prestada pelos advogados na realização da DD.

A DD serve para identificar riscos e para estabelecer passos subsequentes na transação. Desta forma, se houver erros cometidos durante a realização da DD estes poderão conduzir à responsabilidade civil do advogado que a realizou, sempre que, em consequência de erros ou omissões do relatório, o Cliente não tivesse adquirido a empresa ou pago o preço que aceitou.

A sujeição a esta responsabilidade enfatiza a importância da inventariação exata dos documentos revistos e dos pressupostos e limitações tidos em conta na execução da DD.

Surge aqui a importância de uma declaração inicial, esclarecendo o que está incluido e não incluído no processo de DD, como por exemplo, não procura de informação adicional relativamente aos documentos disponibilizados, assume-se que os documentos disponibilizados são verídicos e que constituem cópias fiéis dos respetivos originais, entre outros.

Esclarecidos os deveres do advogado e a imputação de danos por si causados na realização de DD, verifiquemos agora onde pode entrar a IA.

3.3. Aplicação da IA à DD

Os advogados revêm durante os processos de DD centenas ou milhares de documentos, o que torna o trabalho moroso e repetitivo.

[55] *Idem*, p. 21.
[56] *Ibidem*, p.21.
[57] Cf. Moutinho de Almeida, *Responsabilidade Civil ...*, p.22.
[58] Cf. Moutinho de Almeida, *Responsabilidade Civil ...*, p. 22.

As técnicas de IA podem ajudar a identificar, classificar, organizar, hierarquizar e destacar documentos que devem ser disponibilizados ao comprador, com maior eficácia, rapidez e menos custos.

Além disso, os contratos celebrados pela sociedade-alvo da transação serão sempre revistos, sejam comerciais, financeiros, relacionados com direitos de propriedade intelectual ou industrial, celebrados com partes relacionadas ou com terceiros.

A IA pode ser treinada para encontrar cláusulas que normalmente são de grande relevância, tais como, as cláusulas de confidencialidade, não concorrência, indemnização, e resolução de litígios[59]. Temos ainda as cláusulas de *change of control* ou *ownership clauses* que têm sempre impacto nas transações, pois obrigam à obtenção de autorização de entidades bancárias para a concretização da transação ou à obtenção de *waivers* que permitam avançar com a transação.

As tecnologias de IA, designadamente, a PNL, podem ajudar na pesquisa e identificação da existência deste tipo de cláusulas nos vários contratos que estão a ser objeto de revisão[60]. Aliás, os algoritmos de categorização podem ser treinados e testados num conjunto de contratos para aprender a identificar documentos que contenham determinadas expressões[61].

Atualmente existe o Kira Systems[62] que extrai e analisa automaticamente pontos importantes de um contrato que podem ser úteis na realização da DD e elabora relatórios.

Se juntamos a IA à atividade do advogado, como ficará a sua responsabilidade?

3.4. Danos: Quem é o advogado, quem é o responsável?

No âmbito de uma operação de M&A, a IA poderá ter uma intervenção mais elementar ou uma intervenção mais complexa considerando-se a existência de advogados-robô.

De uma forma mais elementar, a IA pode ajudar com tarefas simples e rotineiras auxiliando na procura, categorização e sistematização dos documentos que devem ser objeto de análise durante a DD. Além disso, como já vimos acima, a IA pode depois ser treinada para identificar um conjunto de cláusulas nos contratos que estão a ser revistos e, a um nível mais avançado, elaborar relatórios com as contingências encontradas na realização de uma DD.

[59] Cf. Ben Klaber, *Artificial Intelligence and Transactional Law...*, p. 8.
[60] *Idem*, p. 9.
[61] *Ibidem*, p. 9.
[62] Disponível em https://kirasystems.com/.

Se como referimos acima, durante o processo de DD foram cometidos erros ou omissões que conduziram à má formação do preço ou a que a transação não se realizasse, o advogado poderá ser responsabilizado pelos danos que causou, se agiu ilícita e culposamente.

Imagine-se agora que o advogado serve-se de uma IA para categorizar os documentos, rever os documentos e elaborar um relatório, funções que até à data de hoje são desempenhadas por advogados. Imagine-se que não foram identificadas contingências de grande relevo para a transação? Imagine-se que é divulgada informação sensível provocada por algum vírus ou ataque informático? Se forem causados danos, haverá lugar a responsabilidade civil? Se sim, de quem?

Para que possamos responder a esta primeira questão, deveremos responder a uma outra, nestes casos estaremos perante atos próprios praticados dos advogados?

Vejamos, a título de exemplo, o caso Lola v. Skadden[63], David Lola propôs uma ação contra a sociedade de advogados Skadden, Arps, Meagher, Slate & Flom LLP (Skadden), pedindo uma compensação pela violação das normas relativas à prestação de trabalho suplementar. No que aqui nos interessa, Lola alegou que o seu trabalho estava limitado a três tarefas essenciais: analisar os documentos para ver que termos de pesquisa apareceriam, marcar os documentos em categorias pré-determinadas e identificar certos documentos com base em protocolos específicos. Discutiu-se em tribunal se este trabalho de revisão seria ou não trabalho jurídico. O tribunal considerou que não era porque seria um trabalho que não implicaria qualquer tipo de raciocínio jurídico, mas apenas a aplicação de critérios desenvolvidos por outros para agrupar os documentos em diferentes categorias, tarefa que poderia ter sido perfeitamente realizada por uma máquina.

Não estamos assim tão convencidos de que esta pudesse ser a solução no ordenamento jurídico português.

Com efeito, à luz do artigo 1º da Lei nº 49/2004, de 24 de agosto, são atos próprios dos advogados a prática dos atos preparatórios tendentes à constituição de negócios jurídicos. Parece-nos que todo o trabalho desenvolvido na DD são atos preparatórios tendentes à constituição de negócios jurídicos e, como tal, serão atos próprios dos advogados.

A categorização de documentos, identificação de cláusulas e elaboração de relatórios que identifiquem as contingências parecem à primeira vista tare-

[63] Ver detalhes em https://law.justia.com/cases/federal/appellate-courts/ca2/14-3845/14-3845-2015-07-23.html.

fas relativamente simples e repetitivas, que poderiam ser desenvolvidos por não advogados, sejam eles administrativos ou IA, mas isso não corresponde inteiramente à verdade.

O processo de negociação e aquisição de empresas é complexo e próprio a cada empresa em causa, pelo que, todo o processo desde a simples categorização de documentos até à elaboração do relatório deverá sempre ser executado com o máximo zelo, cuidado e diligência por parte do advogado, para que as reais contingências que possam impactar o negócio sejam identificadas e tratadas pelas partes envolvidas.

Dito isto, parece-nos que, é vedado pela lei que se use a IA para substituir o advogado nestas tarefas sem qualquer tipo de supervisão. Nessa medida, as empresas que produzem e oferecem este tipo de serviços devem sempre procurar entender o que lhes está ou não vedado e deve sempre trabalhar com cooperação com advogados. Se pensarmos no LISA e no DoNotPay, estes sistemas seriam ilegais em Portugal.

Não obstante isto, não se deve negar os potenciais benefícios que uma IA possa ter no aumento de eficiência e redução de custos da realização deste tipo de atos, admitindo-se, portanto, o seu uso pelos Advogados, desde que cumpra com as normas deontológicas a que está vinculado.

À luz da lei atual um advogado humano que utilize AI para automatizar parte da sua prática deverá ser responsabilizado pela violação da responsabilidade profissional daí resultante[64].

O Ser Humano é o último responsável pela solução legal, o sistema apenas ajuda no acesso à informação de que precisa para construir a solução. Este sistema seleciona, ordena, sublinha e sintetiza a informação, mas também explora a informação e interage com os dados[65].

Escusado será dizer que a linguagem jurídica é uma linguagem muito própria, mesmo que se apliquem mecanismos de normalização sistemática, não é garantido que a IA consiga de facto *entender* e dar uma *resposta* adequada. Será sempre preciso assegurar-se que os interesses dos clientes estão verdadeiramente salvaguardados com o seu uso.

O advogado que recorre à IA deverá ponderar de forma cuidada como pode ou não usar a IA, estar seguro das vantagens e dos riscos para os seus clientes e implementar todas as estratégias que sejam necessárias para salvaguardar os interesses dos clientes, assegurar que a IA é devidamente treinada

[64] Cf. *Brian S. Haney, Applied Natural Language Processing For Law Practice*, B.C. Intell. Prop. & Tech. F., 2020, p. 10 *in <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3476351>* (30.10.2022).
[65] Cf. Kevin Ashley, *Artificial Intelligence and Legal Analytics ...*, p. 30.

com informação de qualidade. De outro modo, poderá ser responsabilizado por não ter agido com o cuidado, zelo e diligência necessário e adequado.

Aliás, lembre-se que à luz da *Law 3.0* não deveremos apenas focar-nos na lei existente, mas também nas soluções tecnológicas que podem e devem ser implementadas que ajudam na prevenção da ocorrência de danos.

Apesar das afirmações que fazemos acima, admitimos que perante um caso concreto a imputação de danos ao advogado poderá não ser assim tão clara. Deveremos sempre ter em consideração o tipo de técnica utilizada, a tarefa que esta está efetivamente a executar e o tipo e nível de intervenção do advogado.

Também devemos contar um outro conjunto de questões, é preciso perceber se os clientes admitem o uso de IA ou não, o uso da IA pode ser exigido pelos clientes porque reduz o tempo e os custos. Nesta reflexão é preciso também contar com a nova dinâmica que surgirá entre o advogado e o cliente. Há quem já questione se o advogado do futuro não poderá ser obrigado a ser uma especialista em IA[66].

Por último, a questão mais delicada, imagine-se que o advogado foi zeloso, cuidadoso, supervisionou a IA adequadamente, mas mesmo assim se cometerem erros de análise.

Como tem vindo a ser referido na literatura sobre estas matérias, a IA é caracterizada por ser uma *black box*[67], sendo muito difícil que quem a produziu ou a utiliza, compreenda porque é que, em determinadas situações, ela tomou determinadas decisões. Pode dar-se o caso de o advogado conseguir demonstrar que cumpriu com todas as normas deontológicas e mesmo assim a IA agiu de forma imprevisível.

Os danos provocados ficam por ressarcir? Será necessária a atribuição de personalidade eletrónica à IA e responsabilizá-la diretamente? Atribuindo-se personalidade eletrónica, poderá dizer-se que, em alguns casos, a IA pode ser considerada um profissional qualificado e integrar a Ordem dos Advogados?

[66] Assim, Chris Chambers Goodman, *AI/Esq.: Impacts of Artificial Intelligence in Lawyer-Client Relationships*, 72, 1, 2019, p. 183 *in* <https://core.ac.uk/download/pdf/226775434.pdf> (30.10.2022); ver ainda, Drew Simshaw, *Ethical Issues in Robo-Lawyering: The Need for Guidance on Developing and Using Artificial Intelligence in the Practice of Law*, 70, 1, Hastings Law Journal, 2019, *in* < https://repository.uchastings.edu/cgi/viewcontent.cgi?article=3837&context=hastings_law_journal> (30.10.2022).

[67] Yavar Bathae, *The artificial inteligence black box and the failure of intent and causation*, Harvard Journal of Law & Technology, vol. 31, number 2, Spring 2018, p. 891; Cf. Henrique Sousa Antunes, *Direito e Inteligência Artificial*, Universidade Católica Editora, p. 39.

As questões são múltiplas e as respostas são escassas ainda, apenas perante casos concretos poder-se-á avaliar devidamente cada uma questões, mas deixamos abaixo algumas linhas para reflexão futura.

Alguns autores explicam que hoje em dia o Ser Humano faz parte de um todo interconectado entre si e que contribui para a tomada das suas decisões, por isso mesmo, se tem vindo a desenvolver os conceitos de cognição distribuída[68] e moralidade distribuída[69], chamando a atenção para a quebra na vontade e na intencionalidade individual e o seu desligamento das consequências que ocorram.

Relacionado com isto, está ainda o ressurgimento da responsabilidade coletiva[70], a revolução provocada pela IA estaria a revelar uma realidade mais profunda, a necessidade voltarmos a sistemas de responsabilidade coletiva. Estamos perante mecanismos que trazem benefícios para todos, portanto, todos nós deveríamos responder pelos riscos do uso. Neste caso, teríamos o produtor da IA, o utilizador, ou seja, o advogado, mas também o cliente, caso tenha concordado com o seu uso, etc.

Há autores que referem a necessidade de nos aproximarmos de um sistema mais securitário[71], no qual se gere em conjunto os riscos do uso da IA e se criam fundos próprios para ressarcir os danos. Nesse caso, poderia equacionar-se a hipótese de os advogados usarem IA em determinados casos, mas garantirem a existência de seguros adequados para que se pudessem ressarcir os clientes pelos danos causados.

Ficam estas linhas para reflexão e desenvolvimento futuro ...

4. Conclusão

O uso de IA na área legal traz inúmeras vantagens, mas o seu desenvolvimento deverá ser feito de forma sustentada e sempre no respeito pelo inte-

[68] Cf. Lígia Bernardino, Marinela Freitas e Ricardo Gil Soeiro, *Pós-Humano. Que Futuro?* Antologia De Textos Teóricos, Edições Humus Lda., Vila Nova De Famalicão, 2020, p.23.

[69] Luciano Floridi, *Faultless responsibility: on the nature and allocation of moral responsibility for distributed moral actions*, p. 5, *in* <https://royalsocietypublishing.org/doi/epdf/10.1098/rsta.2016.0112> (30.10.2022).

[70] Cf. Hannelore Bublitz/Roman Marek/Christina L. Steinmann/Hartmut Winkler, „Einleitung", indies. (Hg.), *Automatismen*, München, 2010, S. 9-16. Siehe auch: Tobias Conradi, Breaking News. *Automatismen in der Repräsentation von Krisen- und Katastrophenereignissen*, Paderborn, 2015, S. 29-36.

[71] Cf. Emiliano Marchisio, *Proposal toward "no-fault" civil liability regulation following Artificial Intelligence evolution in health-care*, Rivista di diritto dei media 2/2020, p.170, *in <https://www.medialaws.eu/rivista/proposal-toward-no-fault-civil-liability-regulation-following-artificial-intelligence-evolution-in-health-care/>* (30.10.2022).

resse público subjacente que se traduz na boa administração da justiça e salvaguarda dos melhores interesses dos clientes.

A IA deve ser usada apenas como um instrumento e não como um substituto do advogado, para além do mesmo ser estatuariamente vedado, a IA não possui todas as competências sociais e técnicas necessárias à boa resolução dos casos em concreto. Deverá sempre apostar-se numa abordagem colaborativa e manter-se sempre a supervisão do advogado.

Esta é uma discussão que deverá ser desenvolvida considerando as possíveis evoluções futuras da IA e da sua relação com advogados, clientes e empresas que fornecem estes serviços devem dialogar, promovendo uma análise mais compreensiva do fenómeno.

Deve ainda equacionar-se a hipótese de estarmos perante uma situação que se enquadra num cenário maior que nos obrigue a aproximar o nosso sistema de um sistema mais securitário e de gestão de risco ou estudar de forma mais detalhada hipóteses mais compreensivas de responsabilidade coletiva.

Bibliografia

Antunes, Henrique Sousa, *Inteligência Artificial e Responsabilidade Civil*, in Inteligência Artificial & Direito, coordenado por Manuel Lopes Rocha e Rui Soares Pereira, Almedina, Coimbra, 2020.

Almeida, Moutinho, de *Responsabilidade Civil dos advogados,* Coimbra, Coimbra Editora, 1985.

Ashley, Kevin, *Artificial Intelligence and Legal Analytics, New Tools for Law Practice in the Digital Age,* Cambridge, University Press, Reino Unido, 2017.

Bathaee, Yavar, *The artificial intelligence black box and the failure of intent and causation*, Harvard Journal of Law & Technology, vol. 31, nº 2, Spring, 2018, p. 913 *in* < https://jolt.law.harvard.edu/assets/articlePDFs/v31/The-Artificial-Intelligence-Black-Box--and-the-Failure-of-Intent-and-Causation-Yavar-Bathaee.pdf > (30.10.2022).

Bernardino, Lígia; Freitas, Marinela e Soeiro, Ricardo Gil, *Pós-Humano. Que Futuro?* Antologia De Textos Teóricos, Edições Humus Lda., Vila Nova De Famalicão, 2020.

Braun, Joachim von; Archer, Margaret S.; Reichberg, Gregory M.; Sorondo, Marcelo Sanchez, *Robotics, AI and Humanity*, Science, Ethics, and Policy, Springer Cham, 2021.

Brownsword, Roger, *Law 3.0*, Nova Iorque, Glasshousebook, 2021.

Bublitz, Hannelore; Marek, Roman; Steinmann Christina L.; Winkler, Hartmut, „*Einleitung*“, indies. (Hg.), *Automatismen*, München, 2010, S. 9-16. Siehe auch: Tobias Conradi, Breaking News. *Automatismen in der Repräsentation von Krisen- und Katastrophenereignissen*, Paderborn, 2015, S. pp. 29-36.

Cordeiro, António Menezes, *Direito das Sociedades, I, Parte Geral,* 3ª edição comentada e atualizada, Coimbra, Almedina, 2016.

Corvalán, Juan Gustavo, *Inteligencia artificial: retos, desafíos y oportunidades – Prometea: la primera inteligencia artificial de Latinoamérica al servicio de la Justicia,* p. 2 *in* < https://revistas.ufpr.br/rinc/article/view/55334> (30.10.2022).

Ferreira, Domingos, *Fusões, aquisições, cisões e outras reestruturações de empresas – abordagem jurídica e "due diligence" integral,* volume 2, Rei dos Livros, 2017.

Floridi, Luciano, *Faultless responsibility: on the nature and allocation of moral responsibility for distributed moral actions,* p. 5, *in* <https://royalsocietypublishing.org/doi/epdf/10.1098/rsta.2016.0112> (30.10.2022).

Fruhbauer, Johannes J., *Künstliche In telligenz, Autonomie und Verantwortung Erkundungen im maschinen- und roboterethischen Reflexionskontext*, *in* https://books.ub.uni-heidelberg.de/heibooks/reader/download/945/945-4-95750-2-10-20211109.pdf (30.10.2022).

Goldberg, Yoav, *A Primer on Neural Network Models for Natural Language Processing,* in Journal of Artificial Intelligence Research 57 (2016), pp. 345-420.

Gomes, José Ferreira, *M&A, Aquisição de empresas e de participações sociais,* Lisboa, AEFDUL Editora, 2022.

Goodman, Chris Chambers *AI/Esq.: Impacts of Artificial Intelligence in Lawyer-Client Relationships*, 72, 1, 2019, *in* <https://core.ac.uk/download/pdf/226775434.pdf> (30.10.2022).

Haney, Brian S., *Applied Natural Language Processing For Law Practice*, B.C. Intell. Prop. & Tech. F., 2020, *in <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3476351>* (30.10.2022).

Klaber, Ben, *Artificial Intelligence and Transactional Law: Automated M&A Due Diligence* in http://users.umiacs.umd.edu/~oard/desi5/additional/Klaber.pdf (30.10.2022).

Lycett, Mark, "*Datafication": making sense of (big) data in a complex world' in* European Journal of Information Systems, 2013 *in https://link.springer.com/article/10.1057/ejis.2013.10* (30.10.2022).

Marchisio, Emiliano, *Proposal toward "no-fault" civil liability regulation following Artificial Intelligence evolution in health-care*, pp.153-171, Rivista di diritto dei media 2/2020, *in <https://www.medialaws.eu/rivista/proposal-toward-no-fault-civil-liability-regulation-following-artificial-intelligence-evolution-in-health-care/>* (30.10.2022).

Ng, Christy, *AI in the legal profession,* in *The Cambridge Handbook of Artificial Intelligence Global Perspectives on Law and Ethics*, Cambridge, Cambridge University Press, 2022.

Nida-Rümelin, Julian e Weidenfeld, Nathalie, *Digitaler Humanismus,* Munique, Max Planck Forschung, 2020.

Pires, Catarina Monteiro, *Aquisição de empresas e de Participações Acionistas, Problemas e litígios,* Coimbra, Almedina, 2020.

Polson, Nick, e Scott, James, *Inteligência Artificial, Como funciona e como podemos usá-la para criar um mundo melhor*, Vogais, Lisboa, 2020.

Waisberg, Noah e Hudek, Alexander, *AI for Lawyers, How Artificial Intelligence is adding value, amplifying expertise, and transforming careers,* Canada, Wiley, 2021.

A culpa na responsabilidade civil contratual por acto de agentes de software autónomos no direito português – alguns problemas

The principle of fault in contractual civil liability for actions performed by autonomous software agents in portuguese law – a few issues

Miguel do Carmo Mota*

resumo: A emergência de tecnologias que recorrem a técnicas comummente reconduzidas à chamada "inteligência artificial" coloca diversos desafios no contexto do Direito Civil, dada a sua *autonomia*. Um desses problemas prende-se justamente com a apreciação da culpa no contexto da responsabilidade civil contratual quando os respectivos actos de cumprimento da obrigação estejam a cargo de um destes sistemas ditos *autónomos*. O presente texto pretende definir os contornos do problema no contexto da ordem jurídica portuguesa, avançando apenas de forma breve os possíveis caminhos de solução.

palavras-chave: agentes autónomos; princípio da culpa; responsabilidade civil contratual; incumprimento contratual; inteligência artificial; autonomia

abstract: The emergence of technology which resorts to techniques commonly grouped under the moniker "artificial intelligence" challenges long-standing tenets of Civil Law, given their *autonomy*. One of these challenges occurs in

* Doutorando. Assistente Convidado na Católica Global School of Law (Lisboa). Membro do Católica Research Centre for the future of Law. migueldocarmomota@gmail.com

the context of contractual civil liability, whenever one of these so-called *autonomous* systems is tasked with executing the necessary acts of obligational performance. The following text intends to frame the issue at hand in the specific context of the Portuguese legal order, offering possible paths of solution in a cursory way.

KEYWORDS: autonomous agents; the fault principle; contractual civil liability; contractual default; artificial intelligence; autonomy

1. Introdução

O tema do qual este artigo trata é, na verdade, um problema antigo revestido de uma roupagem nova. A responsabilidade civil contratual e os seus requisitos já foram amplamente discutidos na doutrina, pelo que pouco mais de inovador haveria a acrescentar, não fosse a circunstância do constante progresso tecnológico desafiar as fronteiras e a adequação das soluções jurídicas actualmente vigentes. Uma dessas novas questões prende-se justamente com a avaliação e atribuição da *culpa* na responsabilidade civil contratual quando, no (in)cumprimento da obrigação estejam em causa sistemas autónomos (vulgo, que impliquem a intervenção de tecnologias que se incluam dentro da chamada inteligência artificial[1]). Se existem já desenvolvimentos relevantes no campo da responsabilidade civil extracontratual[2], a questão permanece, por

[1] O termo "inteligência artificial" abrange um vasto leque de tecnologias e aplicações da mesma, pelo que o seu uso, por si só, é demasiado amplo para ser operacional. Veja-se, por exemplo, o Art. 3º/1 da Proposta de Regulamento do Parlamento Europeu e do Conselho em matéria de Inteligência Artificial, que define um sistema de inteligência artificial como sendo "um programa informático desenvolvido com uma ou várias das técnicas e abordagens enumeradas no anexo [à proposta], capaz de, tendo em vista um determinado conjunto de objetivos definidos por seres humanos, criar resultados, tais como conteúdos, previsões, recomendações ou decisões, que influenciam os ambientes com os quais interage".

[2] Veja-se, por exemplo, a Proposta de Regulamento do Parlamento Europeu e do Conselho em matéria de Inteligência Artificial (2021/0106(COD)), ou a Resolução do Parlamento Europeu de Outubro de 2020 em matéria de responsabilidade civil extracontratual aplicável à inteligência artificial ((2020/2014(INL)). Para um comentário a esta última, vd. Henrique Sousa Antunes, "Civil liability applicable to artificial intelligence: a preliminar critique of the European Parliament Resolution of 2020", disponível em https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3743242, 2021

ora, sem qualquer plano de regulamentação (Europeia) explícita no que diz respeito à responsabilidade civil contratual. Num mundo onde sistemas autónomos já são utilizados na preparação e execução do cumprimento da obrigação[3], cumpre saber como proceder à avaliação e atribuição da culpa nestes casos de incumprimento contratual. O presente texto pretende delinear a problemática geral da questão, apresentando os principais pontos de tensão entre os princípios gerais aplicáveis à culpa na responsabilidade civil contratual tal como existente no ordenamento jurídico português e um conjunto de realidades tecnológicas novas, sem paralelo até agora. Ainda assim, este estudo não pretende ser uma análise exaustiva da questão, nem avançar com uma proposta de solução definitiva dos referidos problemas, mas tão somente expor quais os desafios jurídicos desta nova factualidade para a ordem jurídica, e quais os pontos de reflexão que se impõem neste contexto.

2. O Estado da Arte

2.1. Do Princípio da Culpa na Responsabilidade Civil Contratual no Código Civil Português

Em bom rigor, o princípio da culpa – recordemos, base fundamental do direito da responsabilidade civil português[4] – e suas ramificações no contexto da responsabilidade civil contratual já foi objecto de estudo ao longo do vasto período de desenvolvimento científico do Direito Civil. Na tradição jurídica portuguesa (entre outras[5]), a culpa é um dos pressupostos de responsabilização do devedor que incumpriu a obrigação contratual; é o que resulta do

[3] Em termos gerais, pode-se falar em *industry bots*, que são utilizados na melhoria e aumento de eficiência em processos/cadeias de produção, no seu uso em determinadas prestações de serviços (médicos, por exemplo), ou em *bots* utilizados no contexto do e-commerce. A este propósito veja-se, por exemplo, Isabelle Wildhaber/Melinda Lohmann, "Roboterrecht – eine Einleitung", *in Aktuelle Juristische Praxis/Pratique juridique actuelle*, vol. 2, 2017, p. 136-138), Jörg Müller, *The Design of Intelligent Agents – A Layered Approach*, Springer-Verlag, Heidelberg, 1996, p. 124-126, Thomas Schulz, *Verantwortlichkeit bei autonomy agierenden Systemen*, Nomos Verlagsgesellschaft, Baden-Baden, 2014, p. 58/59, Julia Grizinger, "Der Einsatz Künstlicher Intelligenz in Vertragsverhältnissen", in Beyer, *et al*, *Privatrecht 2050 – Blick in die digitale Zukunft*, Baden-Baden, Nomos Verlag, 2019, p. 156.

[4] João de Matos Antunes Varela, *Das Obrigações em Geral*, vol. II, 8ª ed., 2000, Coimbra, Almedina, p. 96/97, Fernando Pessoa Jorge, *Ensaio sobre os Pressupostos da Responsabilidade Civil*, 1968, Lisboa, *Ciência e Técnica Fiscal*, p. 354, Mário Júlio Almeida Costa, *Direito das Obrigações*, 12ª ed., 2011, Coimbra, Almedina, pp. 1038, Catarina Monteiro Pires, *Contratos I – Perturbações na Execução*, 2019, Coimbra, Almedina, p. 103.

[5] A título de exemplo, o *Bundesgesetzbuch* alemão prevê, nos seus artigos §276 e §280, a culpa do devedor faltoso como pressuposto para a existência de responsabilidade civil, tal como o artigo 97 da Lei Suíça das Obrigações, e o artigo 6:74 do Código Civil dos Países Baixos.

disposto no artigo 798º do Código Civil[6], ao imputar a obrigação de indemnização pelos danos causados "àquele que *culposamente* incumpriu a obrigação" (itálico nosso), sendo a referida exigência concretizada pelo artigo 799º do mesmo diploma, ao submeter a apreciação da culpa, neste contexto, à conformidade da conduta em causa com o critério *bonus paterfamilias*, por remissão para o critério geral previsto no artigo 487º, em matéria de responsabilidade civil extracontratual, estabelecendo também uma presunção de culpa contra o devedor incumpridor, de modo a facilitar a responsabilização do mesmo[7].

Daqui resulta que, sem prejuízo de alguns casos contados[8], a responsabilidade civil do devedor assume um pendor predominantemente subjectivista. Esta é uma opção legislativa que, não sendo unânime em todos os sistemas jurídicos – as ordens jurídicas de *common law* atribuem responsabilidade civil por mero incumprimento contratual, independentemente da existência ou não de culpa[9] – se justifica pela intenção de moralizar o sistema de responsabilidade civil, promovendo um comportamento ético e conforme com as exigências e diligências que a vida coletiva impõem, humanizando a imputação de responsabilidade. Cria-se um padrão de comportamento que pretende incutir um cunho moralizante, mas não impositivo de um padrão de comportamento específico à responsabilidade civil: um guia de comportamento, mas não uma imposição inflexível de conformação do comportamento humano conforme o modelo de um homem ideal[10]. Esta intenção torna-se clara ao considerar o método de avaliação da culpa segundo o referido critério *bonus paterfamilias*: avalia-se a conduta do devedor faltoso segundo um componente *interno* e um componente *externo*: respetivamente, o que o deve-

[6] Considerar-se-ão futuras menções a normativos sem indicação expressa do diploma a que pertencem como integrando o Código Civil (Português).

[7] Antunes Varela (2000), p. 101. A ideia fundamental é a de que já estando o elenco dos deveres que vinculam as partes contratuais definido com exactidão, ao contrário do que sucede na responsabilidade civil extracontratual, cabe à parte que a cujo cumprimento dos mesmos ela faltou que tem de justificar o motivo pelo qual não o logrou fazer. Monteiro Pires (2019), p. 105/106 acrescenta que a prova do motivo referido é muito mais custosa, se não mesmo impossível, para o credor, que se veria obrigado a escrutinar a conduta da contraparte faltosa.

[8] A título de exemplo, Monteiro Pires (2019), p. 105, em que a Autora menciona a possibilidade de se estipular no contrato uma obrigação "de garantia", que afastaria a necessidade de culpa do devedor.

[9] Cf, por exemplo, Hein Kötz, Europäisches Vertragsrecht, 2ª ed., 2015, Tübingen, Mohr Siebeck, pp. 369

[10] Nas palavras de Henrique Sousa Antunes, "Anotação ao artigo 487º do Código Civil", in *Comentário ao Código Civil – Das Obrigações em Geral*, 2018, Lisboa, Universidade Católica Editora, p. 301

dor representou mentalmente relativamente ao caso concreto e às potenciais consequências das suas acções, e quais as correspondentes decisões que este tomou com base nessa mesma representação. É mediante estes critérios que se atribui a culpa ao devedor, sendo cada um destes elementos constitutivos da distinção entre *dolo* e *negligência,* conforme a representação (ou falta dela) do devedor, bem como a influência dessa representação nas suas decisões[11]. Mais concretamente, o devedor poderá ilidir a presunção de culpabilidade que sobre ele recai se conseguir provar com sucesso que, dadas as circunstâncias, não poderia possivelmente ter agido de forma a que a obrigação pudesse ser cumprida com sucesso, ou que dele não seria exigido um grau de diligência superior àquele que seria exigido da pessoa média que se encontrasse em semelhante papel[12]. Tipicamente, esta exclusão da culpa reconduz-se a casos em que existe uma qualquer circunstância que, sendo imprevisível e inevitável pelo devedor, acabe por frustrar o cabal cumprimento da obrigação.

2.2. Conceito de Sistema Autónomo

O que torna estas vetustas ideias novamente dignas de reexame é a circunstância de, nos tempos que correm, como referido acima, ser possível recorrer a chamados sistemas de software autónomos para que estes se encarreguem de executar os actos de cumprimento da obrigação em causa. Com "sistemas de software autónomo" referimo-nos, em primeiro lugar, a *software* sem existência corpórea, como seja o caso de *auction bots*, que são programas que licitam em leilões *online* em nome de um utilizador humano. Adicionalmente, incluímos também no conceito acima todo o *software* que tem uma corporização no mundo físico – em linguagem corrente, robôs[13]. O exemplo mais ilustrativo da aplicação destes sistemas autónomos no cumprimento de obrigações será em fábricas ou armazéns, como os armazéns da Amazon, que preparam e executam a entrega da encomenda feita pelo comprador – contribuindo-se desta forma para o cumprimento do contrato de compra e venda subjacente – ou na sua utilização ao longo de cadeias de produção complexas, em que seja exequível pensar-se na sua autonomização mediante o uso destes agentes[14]. Poderemos também pensar em casos em que um determinado contrato de

[11] Antunes Varela (2000), p. 97-100, Pessoa Jorge (1986), pp. 356.

[12] *Ibidem*.

[13] O Art. 3º, a) da Proposta de Regulamento em matéria de responsabilidade civil extracontratual define sistema de IA abrangendo tanto aqueles exclusivamente existentes sobre a forma de *software* como aqueles com suporte físico.

[14] Ver exemplos citados acima, referidos na nota de rodapé nr. 3.

transporte seja executado por um veículo que circule autonomamente, sem necessitar de um condutor humano que controle a viatura, levando assim a cabo a sua tarefa sem necessidade de intervenção humana directa.

Assim, são duas das características deste tipo de agentes mais frequentemente apontadas que revestem maior interesse para o tema em causa: a sua *autonomia* e, potencialmente, a sua *capacidade de aprendizagem*. Esta autonomia traduz-se, em termos sumários, na capacidade que estes sistemas têm de apreender o contexto em que se situam, e, com base nessas impressões, agir sem um constante e directo controlo da parte de um agente (humano)[15]. É certo que as ações destes sistemas não são arbitrárias nem aleatórias, obedecendo estas a diretrizes pré-estabelecidas, balizadas por instruções ou procedimentos também eles pré-estabelecidos por um determinado utilizador humano[16]. Bem assim, o conceito de autonomia não é uma medida absoluta ou sequer qualitativa, sendo antes uma escala ao longo da qual os diversos exemplos de software autónomo se posicionam, sendo dotados de maior ou menor autonomia, dependendo das características específicas de cada sistema, bem como do contexto específico em que operam[17]. Não obstante, o que acaba por caracterizar de forma relevante estes agentes é, entre outras características, justamente o facto de que, em maior ou menor grau, eles terem a capacidade de *decidir* e *agir* de forma independente de qualquer agente humano, na medida em que não é necessário que um humano esteja presente em todo o momento a nortear as acções do sistema, nem tão pouco a tomar as decisões necessárias para o desempenho da sua tarefa[18].

Os problemas que esta autonomia implica agudizam-se nos casos em que os sistemas autónomos em causa têm o potencial de poderem adaptar os seus processos "cognitivos" e a sua forma de agir em conformidade com os dados

[15] Veja-se, a título de exemplo, Walter Brenner/Rüdiger Zarnekow/Harmut Wittig, *Intelligent Software Agents – Foundations and Applications*, 1998, Heidelberg, Springer-Verlag, p. 25/26, Torsten Eymann, *Digitale Geschäftsagenten – Softwareagenten Im Einsatz*, 2003, Heidelberg, Springer-Verlag, p. 21, bem como , Jenay Beer/Arthur Fisk/Wendy Rogers, "Toward a Framework for Levels of Robot Autonomy in Human-Robot Interaction", *in* Journal of Human-Robot Interaction, 2014, p. 76, em que os Autores elencam várias definições de autonomia (técnica).

[16] Michael Luck/Mark D'Inverno/Steve Munroe, "Autonomy: Variable and Generative", *in* Henry Hexmoor/ Cristiano Castelfranchi/Rino Falcone (eds.), *Agent Autonomy*, 2003, Heidelberg, Springer Science+Business Media, p. 14/15.

[17] Margaret Boden, "Autonomy and Artificiality", in Margaret Boden (ed.), *The Philosophy of Artificial Life*, 1996, Oxford, Oxford University Press, p. 102, Phillip Hacker, "Verhaltens- und Wissenszurechnung beim Einsatz von Künstlicher Intelligenz", *in Neue Juristische Zeitschrift,* vol. 3, Munique, C.H. Beck, 2018, p. 252-254

[18] Brenner/Zarnekow/Wittig (1998), p. 25/26; Eymann (2003), p. 21

adquiridos pela sua experiência de funcionamento[19]. Ainda que esta capacidade de aprendizagem seja, regra geral, dirigida à optimização de recursos ou à comunicação com outros agentes[20], o sentido concreto desta evolução e adaptação é, por definição, desconhecido à partida, tendo em conta que esta capacidade permite ao sistema adaptar a sua forma de agir de acordo com os dados recolhidos durante o seu funcionamento de modo a alcançar determinado resultado da melhor forma – ainda que dentro dos parâmetros originalmente definidos pelo utilizador humano, como referido acima. Assim, deixa de ser necessário que o sistema autónomo em causa esteja na presença de todos os dados relevantes para o seu melhor funcionamento *ab initio.* Fala-se, assim, na *capacidade de aprendizagem* que estes sistemas apresentam, tendo estes a possibilidade de recolher dados ao longo do seu funcionamento, utilizando os mesmos para alterar o seu comportamento. Ainda que a função primordial desta característica seja, logicamente, promover uma maior eficiência no funcionamento do agente de software, a imprevisibilidade na concretização desse objetivo faz com que esse resultado possa implicar outras consequências – imprevisíveis – para o agente humano[21].

2.3. A Erosão do Princípio da Culpa?

Estas duas características (em particular, existindo outras que, sendo corolários ou desenvolvimentos das duas acima descritas, também poderão revestir interesse[22]) fazem com que estes sistemas não sejam necessariamente ferramentas inertes, à disposição da ação humana, à semelhança da grande maioria das coisas utilizadas pelo ser humano até ora no cumprimento das obrigações, situando-se estes ainda assim longe de uma putativa personificação jurídica deste tipo de sistemas[23]. Esta circunstância desafia os cânones do Direito Civil, que, no contexto em causa, não prevê qualquer categoria "intermédia" entre

[19] Martin Sommer; *Haftung für autonome Systeme*, Baden-Baden, Nomos-Verlag, 2020, p. 38/39

[20] Rüdiger Zarnekow, *Softwareagenten und elektronische Kaufprozesse: Referenzmodelle zur Integration*, Wiesbaden, Deutscher Universitäts-Verlag, 1999, p. 22/23

[21] Herbert Zech, "Liability for Autonomous Systems: Tackling Specific Risks of Modern IT", in Sebastian Lohsse/Reiner Schulze/Dirk Staudenmayer (eds.), *Liability for Artificial Intelligence and the Internet of Things*, Oxford, Hart Publishing, 2019, p. 189.

[22] Como sejam a *reatividade* destes agentes, que se traduz na capacidade de agir mediante a verificação de determinadas circunstâncias, a sua *proatividade,* que implica que estes agentes tomem a iniciativa de agir sem que para isso sejam determinados, ou a *capacidade de comunicação*, característica fundamental que possibilita a interacção do agente com outros agentes, humanos ou não. Cf., a título de exemplo, Brenner/Zarnekow/Wittig (1998), p. 24-27.

[23] Discutindo a questão, veja-se Mafalda Miranda Barbosa, "Inteligência Artificial, *E-Persons* e Direito: Desafios e Perspetivas", in *Estudos de Direito do Consumidor,* nº 16, edição especial, 2020, pp. 57

ferramenta e ser humano. Neste contexto, cumpre lançar a interrogação: como imputar uma conduta (ou omissão) culposa ao devedor que recorra a um sistema destes para executar os actos de cumprimento da obrigação? Ou, visto de outra forma, o que será exigível ao devedor faltoso demonstrar para ilidir a presunção de culpa que sobre ele impende? Contando que o conceito de culpa em causa depende, como referido acima, de uma avaliação da representação e correspondente conduta do devedor, e, como tal, assumindo (justificadamente) a presença apenas de seres humanos no contexto do cumprimento da prestação obrigacional, como poderá qualquer culpa ser atribuída num contexto em que os actos de cumprimento estão fora da disponibilidade directa do referido devedor? Com efeito, a partir do momento em que o cumprimento da obrigação fica a cargo de um sistema autónomo do género descrito acima, não existe qualquer acção ou omissão humanas que possam ser avaliadas para efeitos de atribuição de culpa, gerando-se, consequentemente, o problema de saber *como* fazer esta avaliação. Sendo verdade que a presunção de culpa que recai sobre o devedor faltoso tenderá a promover uma responsabilização do mesmo, cumpre, ainda assim, saber quais os actos que estarão na disponibilidade do devedor para que este possa tomar todas as diligências exigíveis de modo a ilidir essa mesma presunção admitindo-se a hipótese do devedor conseguir provar a sua falta de culpa em casos como estes. A questão coloca-se, assim, em torno do apuramento da culpa, que é um conceito demarcado da própria cognição e experiência humanas, quando os actos cuja culpabilidade (ou falta dela) está a ser avaliada provém de um sistema autónomo, separado e fora do âmbito de controlo do devedor (em todos os momentos do processo de cumprimento da obrigação), dono do sistema autónomo.

A questão assume contornos mais delicados quando ponderarmos a aplicação do regime aplicável aos auxiliares do cumprimento (art. 800º), e a sua possível aplicação ao caso em questão[24]. Com efeito, a norma pressupõe que o auxiliar em causa seja um agente humano; daí que se impute a conduta (e suas implicações) do auxiliar ao devedor que a ele recorreu. Neste contexto, surgem duas questões. Em primeiro lugar, cumpre saber se é possível considerar um sistema deste tipo como sendo um verdadeiro auxiliar ao cumprimento, no sentido dado pela norma. Será o grau de autonomia destes sistemas suficiente para que se justifique a aplicação deste preceito? Como visto, esta pressupõe que o auxiliar em causa seja um agente humano, dotado da inerente autonomia que assiste a todo o ser humano, daí que a sua conduta seja

[24] Discutindo a questão, vd. António Pinto Monteiro, ""*Qui facit per alium, facit per se*" – *Será Ainda Assim na Era da Robótica?*", *in Estudos de Direito do Consumidor*, nº 16, edição especial, 2020, pp. 11

avaliada em termos semelhantes àqueles em que o devedor em si seria avaliado, se, hipoteticamente, tivesse sido o próprio devedor a providenciar pelo cumprimento da obrigação, sem prejuízo da potencial responsabilidade *in elegendo* do devedor pela escolha/direcção do auxiliar[25].

Tendo em conta a factualidade ora em causa, cumpre saber se será possível (e adequado) aplicar o disposto no art. 800º CC para regular casos deste género. Em caso afirmativo, cumpre ainda saber como aplicar o regime correspondente, já que este implica que se impute a conduta do auxiliar humano *ta quale* ao devedor. Voltamos, assim, e mais uma vez, à questão de saber como aplicar à conduta de um sistema autónomo a bitola estabelecida para a avaliação da culpa, que, no contexto do nosso sistema jurídico, para lá de ser um conceito decalcado da experiência humana, assume uma base predominantemente subjectiva. Será possível aplicar a referida norma a casos destes, ainda que potencialmente por analogia?[26]

3. Análise e Comentário. Possíveis Caminhos de Solução

É esta, em termos amplos, a problemática que a ligação do princípio da culpa no contexto da responsabilidade civil contratual traz à aplicação de sistemas autónomos. Todo o problema gira em torno da dificuldade em moldar um conceito moldado à experiência humana a uma situação, a uma realidade, que não é, por ora, sequer equiparável a uma verdadeira experiência humana; pelo menos, não em termos rigorosamente equivalentes. Assim sendo, como poderá o sistema jurídico reagir de modo a alcançar uma solução justa, e que concilie todos os interesses em causa?

Fazendo frente, em primeiro lugar, à dualidade entre uma culpa puramente pessoal e subjetiva e a aplicação do regime dos auxiliares de cumprimento, poder-se-ia pensar na convivência de ambos em simultâneo, de forma análoga ao que acontece, por exemplo, na Resolução do Parlamento Europeu com Recomendações em matéria de responsabilidade civil extracontratual no contexto da inteligência artificial (2020/2014(INL)), em que se estabelecem dois regimes paralelos: um regime de responsabilidade objetiva para actividades especialmente perigosas, tipificadas no próprio texto do Regulamento[27],

[25] Antunes Varela (2000), p. 101. Contra, Pessoa Jorge (1968), p. 146.

[26] Optando-se por uma resposta afirmativa a esta pergunta, cumpre também saber em que termos se poderá aplicar o regime de exclusão da responsabilidade previsto no art. 800º/2. Cf. Pinto Monteiro (2020), pp. 21.

[27] Cf. o Art. 3º, c) da Proposta, que define o conceito de "alto risco" todo o sistema que "um potencial importante de um sistema de IA que funcione de forma autónoma causar prejuízos ou danos a uma ou várias pessoas de forma aleatória e que vai além do que se pode razoavelmente esperar; a

e um outro de presunção de culpa para as restantes[28]. *Mutatis mutandis*, seria também possível distinguir casos em que conforme o grau de autonomia do sistema autónomo (por exemplo), poder-se-ia distinguir entre um sistema de responsabilidade objetiva em toda a linha, bem como um outro em que a culpa é meramente presumida, à semelhança do que acontece agora. Os critérios para alcançar tal distinção são, contudo, pouco evidentes, podendo-se ponderar em relevar factores como sejam a natureza da prestação, o tipo e características do sistema utilizado, ou o grau de intervenção exigido ou necessário da parte do agente humano. Esta diferenciação assentaria, assim, na imprevisibilidade e controlo à disposição do mesmo.

Em alternativa, poder-se-ia optar por um regime unitário que acolha, em toda a linha, uma destas soluções. Por um lado, pode-se optar pela aplicação de um regime que acolha a responsabilidade objetiva *tout court*, responsabilizando-se necessariamente o devedor faltoso em caso de incumprimento, à semelhança do que acontece em sistemas de *common law*. Esta seria uma opção que se poderia justificar pela elevada imprevisibilidade inerente a estes sistemas autónomos, considerando-se que o correspondente (elevado) risco de prejuízo justifica a desnecessidade de verificação de culpa para que possa haver uma verdadeira responsabilização do devedor faltoso e correspondente imputação da obrigação de indemnização, à semelhança de soluções equivalentes existentes no regime da responsabilidade civil extracontratual. Sendo esta uma solução extremamente protetora do credor, não é líquido que se justifique que se estabeleça uma responsabilização objetiva em termos tão perentórios, sob pena de se esvaziar a importância do princípio da culpa no contexto do direito da responsabilidade, que, como referido acima, continua a ser um baluarte do direito da responsabilidade português. Bem assim, esta seria uma solução que poderia possivelmente desincentivar o recurso a estas tecnologias, já que a perspetiva de uma eventual obrigação de indemnização seria, evidente, muito maior, desvirtuando-se assim a bitola de responsabilidade vigente em geral. Assim sendo, qualquer pessoa iria refletir com um cuidado acrescido antes de recorrer a um sistema deste tipo para executar o cumprimento das suas obrigações. Sendo esta, à primeira vista, uma mera

importância deste potencial depende da interligação entre a gravidade dos eventuais prejuízos ou danos, o grau de autonomia de decisão, a probabilidade de o risco se concretizar e a forma e o contexto em que o sistema de IA é utilizado" bem como o Art. 4º/2 da mesma, que estabelece a tipicidade dos sistemas que apresentam, na acepção do diploma em causa, esse mesmo alto risco.

[28] Art. 8º/1 + 2 da Proposta, que estabelece essa presunção de culpa, bem como os actos relevantes em matéria de ilisão da mesma.

consideração de conveniência pessoal, há que considerar a circunstância de um menor apelo ao uso destas tecnologias poderá potencialmente implicar um menor investimento da investigação e desenvolvimento das mesmas, desacelerando assim o progresso tecnológico. Sendo certo que este não deverá ser um valor a relevar em si mesmo, deverá este ser um daqueles a ponderar no difícil equilíbrio de considerações que ora se impõe.

Por outro lado, optando-se por um regime de responsabilidade subjetiva que contemple uma presunção de culpa contra o devedor, na esteira do que é actualmente estipulado pelo nosso Código Civil, é impossível contornar a questão da correspondente presunção de culpa, e de saber qual a fasquia exigida para que o devedor consiga ilidir a mesma. Assim, cumpre saber quais os actos concretamente considerados que o devedor poderá praticar para que este observe o seu dever de diligência. Poder-se-ia ponderar que existem medidas concretas que este poderá tomar para o efeito, relacionadas com a escolha do *software* adequado, com a sua manutenção e atualização, ou com um grau adequado de supervisão do desempenho do sistema autónomo, por exemplo[29]. Uma especificação dos actos concretos necessários para cumprir o dever de diligência imposto implicitamente pelo princípio da culpa implicaria, de certa forma, um afastamento de um conceito aberto de culpa como conceito não tipificante de condutas culposas, confrontadas com um critério abstracto previsto na lei (no caso, o critério *bonus paterfamilias* aplicável por força do Art. 799º/2) prevalente em grande parte dos sistemas de Direito Civil, que consagram uma responsabilidade civil contratual assente ainda no pressuposto subjetivo da culpa[30]. Tal opção justifica-se, inclusive, na medida em que o nosso sistema jurídico-civil, bem como aqueles que também acolhem a culpa como pressuposto de responsabilização, parte(m) de um pressuposto de liberdade e autonomia fundamentais, reconhecido a todo o ser humano[31]. Assim, a responsabilização pelas acções individuais voluntariamente tomadas

[29] Veja-se, a título de comparação, o referido Art. 8º/2 da Proposta de Regulamento em matéria de responsabilidade civil extracontratual aplicável à inteligência artificial, que define, a título de exemplo, o tipo de actos aptos a afastar a presunção de culpa. Para um maior desenvolvimento relativamente ao dever de diligência aplicável nestes casos, cf. Martin Sommer, Haftung für Autonome Systeme, pp. 92

[30] Para lá do referido artigo 798º CC, cf., noutras jurisdições e a título de exemplo, o §276 do BGB alemão, o Art. 97º da Lei das Obrigações suíça, o Art. 1218 do Código Civil Italiano, ou o Art. 6:74 do Código Civil holandês.

[31] Cf., a título de exemplo, Hannes Unberath, Die Vertragsverletzung, Tübingen, Mohr Siebeck, 2007, pp. 182, bem como Mafalda Miranda Barbosa, *Lições de Teoria Geral do Direito Civil*, Coimbra, Gestlegal, 2021, pp. 27

por cada um é o reverso da medalha do referido amplo espaço de liberdade; responsabilização essa que assenta num conceito amplo como o da culpa: é ilícito (e culposo) o acto que provém de um uso abusivo ou impróprio do espaço de liberdade inerente ao ser humano[32]. Bem assim, o critério de diligência referido acima visa acautelar a confiança generalizada na estabilidade do comércio jurídico, sendo, por isso, contextual, e como tal apurado de acordo com os contornos do caso concreto[33]. Não obstante esta função "estabilizadora" do comércio jurídico, o desrespeito desta fasquia continua a depender, no essencial, da avaliação da conduta concreta do devedor. Com uma mudança de paradigma para uma fasquia de diligência "tipificada", mediante a previsão explícita dos actos/omissões constitutivos de um comportamento culposo, passaríamos, destarte, para uma aproximação a uma apreciação da culpa de índole *regulatória*, estabelecendo-se critérios e parâmetros concretos cuja verificação irá determinar uma responsabilização, ou a falta dela, ao contrário de se avaliar o comportamento em causa subsumindo-o a uma cláusula aberta. Naturalmente, ainda que o (não-)cumprimento destes deveres concretos de diligência continue a assentar na ideia de responsabilidade por actos voluntários, a avaliação da culpa e consequente responsabilização poderá potencialmente passar por uma alteração significativa do método de avaliação da culpa do devedor. Acresce a isto a ideia bastante pragmática de que uma fasquia regulatória demasiado alta irá invalidar todo o propósito de se recorrer a este tipo de sistemas, que será justamente libertar os agentes humanos de um controlo constante da parte dos mesmos, tal como aludido acima a propósito de uma responsabilização objectiva. Sendo certo que o uso destas tecnologias não deverá implicar uma atenuação substancial da exigência estabelecida pela fasquia de responsabilização atualmente vigente, uma exigência demasiada alta nesse sentido implicará consequências semelhantes àquelas avançadas a propósito da responsabilização objetiva, referidas acima.

Poder-se-ia também ponderar até que ponto seria possível limitar ou afastar a culpa ou a responsabilidade convencionalmente, e em que termos tal convenção seria admitida, na medida em que a vontade – neste caso, das partes – poderia, num momento anterior, estabelecer a fasquia pela qual o dever de diligência do devedor se deverá estabelecer. Ou, em contrapartida, poderá ainda circunscrever ou suprimir a responsabilidade do mesmo como um todo, sem prejuízo desta limitação violar o disposto no Art. 809º e seguintes do

[32] Mafalda Miranda Barbosa (2021), pp. 63

[33] Cf. a título de exemplo, Dirk Looschelders, *Schuldrecht – Allgemeiner Teil*, Munique, Franz Vahlen Verlag, 2018, p. 183.

CC – de resto, também aplicável aos casos em que se considere estes sistemas autónomos como sendo auxiliares, nos termos do Art. 800º/2 do mesmo diploma. De forma semelhante, podemos também ponderar até que ponto é que as vicissitudes imprevisíveis próprias deste tipo de sistema poderão potencialmente justificar a exclusão da imputação por fundamento análogo ao do caso fortuito ou da força maior; em bom rigor, está em causa um evento pelo menos em certa medida imprevisível para o devedor faltoso; em que termos poderá esta circunstância justificar uma exclusão de qualquer imputação da conduta do sistema autónomo ao seu operador?

4. Conclusão

Em suma, e em jeito de conclusão, o problema apresentado ainda tem, por ora, mais incógnitas do que respostas, estando em causa o sempre difícil exercício de adaptar uma realidade nova aos velhos edifícios erigidos em tempos idos, e o princípio da culpa não é uma excepção a essa ideia. Sendo certo que este problema tem o condão de testar os limites e fronteiras do Direito Civil e da Responsabilidade, é também da responsabilidade dos juristas repensar esses fundamentos e as soluções em vigor com vista a alcançar, neste caso, uma melhor e mais justa repartição da culpa com o fim de alcançar um cabal equilíbrio entre todos os interesses em jogo: aqueles do lesado, do devedor que lança mão deste tipo de sistemas para agilizar os seus processos, e da sociedade como um todo, que, evidentemente, terá um interesse particular em fomentar o desenvolvimento tecnológico sem que tal se repercuta numa produção desregrada de danos para os cidadãos.

Bibliografia

Antunes, Henrique Sousa, "Anotação ao artigo 487º do Código Civil", in *Comentário ao Código Civil – Das Obrigações em Geral*, 2018, Lisboa, Universidade Católica Editora, p. 299-303

Antunes, Henrique Sousa, "Civil liability applicable to artificial intelligence: a preliminar critique of the European Parliament Resolution of 2020", Lisboa, 2021, *in* https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3743242, 2021

Barbosa, Mafalda Miranda, "Inteligência Artificial, E-Persons e Direito: Desafios e Perspetivas", in *Estudos de Direito do Consumidor*, nº 16, edição especial, 2020, pp. 57

Barbosa, Mafalda Miranda, *Lições de Teoria Geral do Direito Civil*, Coimbra, Gestlegal, 2021

Beer, Jenay/Fisk, Arthur/Rogers, Wendy, "Toward a Framework for Levels of Robot Autonomy in Human-Robot Interaction", in *Journal of Human-Robot Interaction*, nr. 74, vol. 3, 2014, p. 74-99

BODEN, MARGARET, "Autonomy and Artificiality", in BODEN, MARGARET (ed.), *The Philosophy of Artificial Life*, 1996, Oxford, Oxford University Press, p. 95-107

BRENNER, WALTER/ZARNEKOW, RÜDIGER/WITTIG, HARMUT, *Intelligent Software Agents – Foundations and Applications*, 1998, Heidelberg, Springer-Verlag

COSTA, MÁRIO JÚLIO ALMEIDA, *Direito das Obrigações*, 12ª ed., 2011, Coimbra, Almedina

EYMANN, TORSTEN, *Digitale Geschäftsagenten – Softwareagenten Im Einsatz*, 2003, Heidelberg, Springer-Verlag

GRIZINGER, JULIA, "Der Einsatz Künstlicher Intelligenz in Vertragsverhältnissen", in BEYER, et al, *Privatrecht 2050 – Blick in die digitale Zukunft*, Baden-Baden, Nomos Verlag, 2019, pp. 151-179

HACKER, PHILLIP, "Verhaltens- und Wissenszurechnung beim Einsatz von Künstlicher Intelligenz", in *Rechtswissenschaft*, ano 9, vol. 3, Munique, C.H. Beck, 2018, p. 243-289

JORGE, FERNANDO PESSOA, *Ensaio sobre os Pressupostos da Responsabilidade Civil*, 1968, Lisboa, Ciência e Técnica Fiscal

KÖTZ, HEIN, *Europäisches Vertragsrecht*, 2ª ed., 2015, Tübingen, Mohr Siebeck

LOOSCHELDERS, DIRK, *Schuldrecht – Allgemeiner Teil*, Munique, Franz Vahlen Verlag, 2018

LUCK, MICHAEL/D'INVERNO, MARK/MUNROE, STEVE, "Autonomy: Variable and Generative", in HEXMOOR, HENRY/CASTELFRANCHI, CRISTIANO/FALCONE, RINO (eds.), *Agent Autonomy*, 2003, Heidelberg, Springer Science+Business Media, p. 11-29

MONTEIRO, ANTÓNIO PINTO, ""Qui facit per alium, facit per se" – Será Ainda Assim na Era da Robótica?", in *Estudos de Direito do Consumidor*, nº 16, edição especial, 2020, p. 11-31

MÜLLER, JÖRG, *The Design of Intelligent Agents – A Layered Approach*, Heidelberg, Springer-Verlag, 1996

PIRES, CATARINA MONTEIRO, *Contratos I – Perturbações na Execução*, 2019, Coimbra, Almedina

SCHULZ, THOMAS, *Verantwortlichkeit bei autonom agierenden Systemen*, Baden-Baden, Nomos Verlagsgesellschaft, 2014

SOMMER, MARTIN; *Haftung für autonome Systeme*, Baden-Baden, Nomos-Verlag, 2020

UNBERATH, HANNES, *Die Vertragsverletzung*, Tübingen, Mohr Siebeck, 2007

VARELA, JOÃO DE MATOS ANTUNES, *Das Obrigações em Geral*, vol. II, 8ª ed., 2000, Coimbra, Almedina

WILDHABER, ISABELLE/LOHMANN, MELINDA, "Roboterrecht – eine Einleitung", in Aktuelle Juristische Praxis/Pratique juridique actuelle, vol. 2, 2017, p. 136-138

ZARNEKOW, RÜDIGER, *Softwareagenten und elektronische Kaufprozesse: Referenzmodelle zur Integration*, Wiesbaden, Deutscher Universitäts-Verlag, 1999

ZECH, HERBERT, "Liability for Autonomous Systems: Tackling Specific Risks of Modern IT", IN LOHSSE, SEBASTIAN/SCHULZE, REINER/STAUDENMAYER, Dirk (eds.), *Liability for Artificial Intelligence and the Internet of Things*, Oxford, Hart Publishing, 2019, p. 187-200

Inteligencia artificial y vehículos autónomos en la propuesta de Directiva de responsabilidad por productos defectuosos

Artificial intelligence and autonomous vehicles in the Proposal for the product liability directive

Monica Navarro-Michel*

RESUMO: Las normas vigentes de responsabilidad por productos defectuosos no están concebidas para la era digital, y la Comisión europea está trabajando para adaptarlas a los nuevos desafíos de la inteligencia artificial. Este trabajo analiza algunos de los retos de la normativa actual y cómo los resuelve la reforma europea en la propuesta de Directiva de 28 septiembre de 2022 sobre responsabilidad por los daños causados por productos defectuosos. El presente trabajo ilustra la aplicación de la normativa europea actual y la proyectada tomando como ejemplo un producto concreto, el vehículo automatizado, conectado, con proyección de ser totalmente autónomo. Presento las novedades introducidas en esa propuesta de directiva, y me centro en algunos problemas específicos, entre los que cabe destacar las vulnerabilidades de ciberseguridad y los riesgos de desarrollo.

PALAVRAS-CHAVE: Responsabilidad por producto defectuoso, vehículo autónomo, conducción autónoma, riesgos de desarrollo, inteligencia artificial, ciberseguridad.

* Profesora de Derecho Civil. Universitat de Barcelona.

ABSTRACT: Current product liability rules are not designed for the digital age, and the European Commission is working to adapt them to the new challenges of artificial intelligence. This paper discusses some of the challenges of the current regulation and how they are addressed by the European reform in the proposed Directive of 28 September 2022 on product liability. This paper illustrates the application of the current and projected European regulations taking as an example a specific product, the automated, connected vehicle, aimed to become fully autonomous. I highlight the new features introduced in this proposed directive, and focus on some specific problems, including cybersecurity vulnerabilities and development risk defence.

KEYWORDS: Liability for defective products, autonomous vehicles, development risk defence, artificial intelligence, cybersecurity

1. Introducción

El punto de partida de este trabajo es la ponencia presentada en el Congreso *Direito e inteligência artificial* organizado por la Faculdade de Direito da Universidade do Porto y CIJE- Centro de Investigaçao Jurídico Económica, celebrado los días 12 y 13 Mayo 2022, en la que tuve la oportunidad de plantear algunas de las dificultades de la aplicación de las normas sobre responsabilidad por productos defectuosos a los daños causados por vehículos automatizados, como ejemplo paradigmático de un producto que incorpora dispositivos de inteligencia artificial. Desde entonces, la Comisión europea ha elaborado propuestas normativas que inciden sobre el tema de la ponencia; en concreto, la propuesta de Directiva del Parlamento europeo y del Consejo sobre responsabilidad por los daños causados por productos defectuosos, de 28 de septiembre de 2022, COM(2022) 495 final (PDPD). Con el fin de evitar que este *paper* quedase obsoleto antes de su publicación, introduzco referencias al régimen propuesto por la Comisión europea, destacando los cambios que introduce, sus aciertos y posibles fallos.

La modificación del régimen de responsabilidad por productos defectuosos se enmarca en las intensas reformas a nivel europeo tendentes a preparar Europa para la era digital. Junto a ella, se ha presentado una Propuesta de Directiva del Parlamento europeo y del Consejo relativa a la adaptación de las normas de responsabilidad civil extracontractual a la inteligencia artificial (Directiva sobre responsabilidad en materia de IA), 28 de septiembre de 2022, COM(2022) 496 final. Aunque pueden surgir dificultades de coordinación entre ambas propuestas, lo cierto es que no debería haber solapamiento, en la medida en que esta vendría a constituir el régimen general de responsabilidad extracontractual, y la de productos defectuosos, el régimen especial.

Para ilustrar las cuestiones jurídicas que plantea la normativa actual de productos defectuosos, y cómo quedan resueltas en la propuesta de reforma, me centro en el vehículo autónomo (*rectius*, vehículo con sistemas automatizados, conectados, aún no totalmente autónomos). Es necesario hacer una precisión terminológica, acerca del uso del término vehículo autónomo, vehículo automatizado, vehículo conectado, pues no son enteramente coincidentes. La automatización, en sus niveles más altos, llevará a la autonomía de la conducción, pero aún no hay vehículos totalmente autónomos a la venta[1].

La Sociedad de ingenieros de automoción internacional (SAE, por su acrónimo en inglés) ha elaborado una clasificación de los niveles de automatización, para homogeneizar los estándares de la industria. El documento, publicado por primera vez en enero de 2014, ha sido actualizado en varias ocasiones, siendo la última versión la de abril de 2021[2]. La clasificación tiene en cuenta no solo lo que el vehículo es capaz de hacer o no hacer, sino, sobre todo, el papel que tiene el ocupante del vehículo, pues de conductor pasaría a ser pasajero, en el nivel de conducción totalmente autónomo.

En los niveles más bajos de automatización (niveles 0 a 2), el conductor debe realizar todas las tareas de conducción y monitorizar el entorno, y es compatible con algún mecanismo de seguridad (frenado automático) o de asistencia a la conducción (*adaptive cruise control, lane assist*) o al estacionamiento, pero quien tomas las decisiones de la conducción de manera continuada es el ser humano[3].

[1] La SAE es contundente al afirmar que, por el momento, no hay vehículos en el mercado que sean completamente autónomos. https://www.sae.org/news/2022/03/navigating-the-language-of-vehicle-autonomy

[2] *SAE Recommended Practice J3016. Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles*, 4ª edición, 2021.

[3] En España, la Instrucción 16 TV/89, de 20 de enero de 2016, de la Dirección General de Tráfico, sobre estacionamiento asistido de vehículos a motor aclara que, en estos casos, el conductor humano

A partir del nivel 3 (automatización condicional), el sistema monitoriza el entorno y toma las decisiones, de manera continuada, y el conductor pasa a ser solo conductor "de reserva" (*fallback ready user*). El usuario de vehículo únicamente deberá asumir la conducción cuando el vehículo solicite su intervención, si detecta un fallo en el sistema o una situación de peligro en el entorno. Eso sí, este conductor "de reserva" debe estar atento a los avisos del vehículo (*request to intervene*) y debe estar preparado para asumir la conducción en todo momento. Una vez hecha la solicitud de intervención, el vehículo generalmente reduce la velocidad durante unos segundos para que el conductor pueda asumir la conducción a partir de ese momento o, si lo considera oportuno, llevar al vehículo a una situación de riesgo mínimo (*minimal risk condition*), que típicamente consiste en conducir hasta el arcén y parar.

En el nivel 4 (automatización alta), el vehículo puede conducir de forma autónoma cuando concurren determinadas circunstancias geográficas (barrios residenciales, bases militares, campus universitarios, autopistas) o ambientales (meteorológicas y diurnas/nocturnas). En caso de emergencia, el vehículo mismo puede colocarse en una situación de riesgo mínimo. El nivel 5 es el de automatización completa, es el vehículo totalmente autónomo, en cualquier entorno. En el nivel 5 el ocupante es, siempre, un pasajero; en el nivel 4 únicamente cuando está activado el modo automático[4].

2. Ampliación del ámbito objetivo y subjetivo en la Propuesta de Directiva

2.1. El concepto de producto incluye el software

Un vehículo autónomo es un bien con elementos digitales, en la terminología de la Directiva (UE) 2019/771, del Parlamento Europeo y del Consejo, de 20 de mayo de 2019, relativa a determinados aspectos de los contratos de compraventa de bienes (DCCB en lo sucesivo), que lo define como "todo objeto mueble tangible que incorpore contenidos o servicios digitales o esté interconectado con ellos de tal modo que la ausencia de dichos contenidos o servicios digitales impediría que los bienes realizasen sus funciones" (art. 2.5.b) DCCB). Y es un producto a efectos de la Directiva 85/374/CEE del Consejo, de 25 de julio de 1985, relativa a la aproximación de las disposiciones legales,

sigue siendo el responsable del vehículo. Y ello aunque el control de los mandos del vehículo se realice desde fuera del vehículo. Porque la conducción remota no convierte la conducción en autónoma.

[4] Me he ocupado de la responsabilidad en caso de accidente de circulación en cada uno de los niveles de automatización en Mónica Navarro-Michel, "La aplicación de la normativa sobre accidentes de tráfico a los causados por vehículos automatizados y autónomos", *Cuadernos de Derecho Transnacional*, 2020, vol. 12, nº 1, pp. 941-961.

reglamentarias y administrativas de los Estados miembros en materia de responsabilidad por los daños causados por productos defectuosos. Según el art. 2, se considera producto cualquier bien mueble, aun cuando esté incorporado a otro bien mueble o inmueble, e incluye expresamente la electricidad.

El debate jurídico acerca del ámbito objetivo de la responsabilidad por producto defectuoso se ha centrado en la calificación de los programas informáticos (software), ya que pueden ser considerados productos o servicios. La Directiva 85/374 no exige la corporalidad del bien, por lo que la doctrina se ha decantado, en general, por incluir los programas informáticos en su ámbito de aplicación[5]. Cuando el programa informático está preinstalado en el producto final (*embedded software*), puede ser considerado un componente, aunque surgen dificultades de calificación cuando el programa informático es instalado con posterioridad, así como las actualizaciones posteriores, cuando son suministradas sin soporte material, de manera online, y acaso por empresa distinta a la del fabricante del vehículo (*stand alone software*). La propuesta de Directiva pone fin a esta discusión al añadir a la definición de producto, en su art. 4.1, dos nuevos productos intangibles: los denominados "archivos de fabricación digital"[6] y los programas informáticos[7].

La propuesta de Directiva distingue entre componente y servicio conexo. Un componente es "cualquier artículo, tangible o intangible, o cualquier servicio conexo, que está integrado en un producto o interconectado con él por el fabricante de ese producto o que esté bajo su control" (art. 4.3) y un servicio

[5] He abordado esta cuestión y otras que aquí no se incluyen en Mónica Navarro Michel, "Vehículos automatizados y responsabilidad por producto defectuoso", *Revista de Derecho Civil*, 2020, vol. VII, nº 5, pp. 173-223.

[6] El art. 4.2 los define como una versión digital o plantilla digital de un bien mueble. El cdo 14 resulta aclarador: "Los archivos de fabricación digital, que contienen la información funcional necesaria para producir un elemento tangible permitiendo el control automatizado de máquinas o herramientas, como taladros, tornos, molinos e impresoras 3D, deben considerarse productos, a fin de garantizar la protección de los consumidores en los casos en que esos archivos sean defectuosos."

[7] Sin embargo, no todos los programas informáticos están incluidos, ya que, por un lado, "el código fuente de los programas informáticos no debe considerarse un producto a efectos de la presente Directiva, ya que se trata de pura información" (cdo 12) y por otro, "A fin de no obstaculizar la innovación o la investigación, la presente Directiva no debe aplicarse a los programas informáticos libres y de código abierto desarrollados o suministrados fuera del transcurso de una actividad comercial. Este es el caso, en particular, de los programas informáticos, incluidos su código fuente y sus versiones modificadas, que se comparten abiertamente y son de libre acceso, utilizables, modificables y redistribuibles. Sin embargo, cuando los programas informáticos se suministren a cambio de un precio o los datos personales se utilicen de forma distinta a la de mejorar la seguridad, la compatibilidad o la interoperabilidad del programa informático y, por tanto, se suministren en el transcurso de una actividad comercial, debe aplicarse la Directiva" (cdo 13).

conexo es "un servicio digital que está integrado en un producto o interconectado con él, de tal manera que su ausencia impediría al producto realizar una o varias de sus funciones" (art. 4.4). Esta distinción pone de manifiesto que la distinción entre producto y servicio resulta poco nítida, y acaso deba ser superada, ya que un servicio conexo puede ser un producto.

La conclusión que cabe extraer ahora es que un servicio conexo puede ser un componente cuando esté *integrado en* el producto final o *interconectado con* él. La integración en el producto final es el concepto ya conocido de componente; la era digital obliga a tener en cuenta ahora la interconexión. Lo relevante es que esa integración o interconexión se haga o bien por el fabricante del producto final o bien bajo su control, aunque proceda de fuentes ajenas. Por tanto, la propuesta de Directiva permite afirmar, sin lugar a duda, que los programas informáticos son productos, a efectos de la responsabilidad por producto defectuoso.

2.2. El concepto de defecto

a) Las expectativas de seguridad

El concepto de defecto, vinculado a las expectativas legítimas de seguridad ("un producto es defectuoso cuando no ofrece la seguridad a la que una persona tiene legítimamente derecho", art. 6 Directiva 85/374) se mantiene en la propuesta de Directiva, aunque se añade la mención de que las expectativas a tener en cuenta son las del público en general[8] ("un producto se considerará defectuoso cuando no ofrece la seguridad que el público en general tiene derecho a esperar" (art. 6.1.h) PDPD), aunque también se tendrá en cuenta las del destinatario final.

Las expectativas de seguridad serán mayores o menores en función de la información suministrada por el fabricante sobre el producto. La información insuficiente o inadecuada acerca del uso correcto del vehículo, así como de sus riesgos, conlleva responsabilidad del fabricante (defecto de información). El uso seguro de los vehículos automatizados exige conocimiento de los diferentes niveles de seguridad. Sin embargo, existe una gran disparidad entre lo que los vehículos saben hacer y lo que el gran público cree que saben hacer, dada la percepción generalizada de que el vehículo es capaz de hacer de forma autónoma más tareas de las que realmente es capaz de hacer. El fabricante contribuye a generar las expectativas de seguridad, así como las expectativas de uso, mediante la publicidad y los términos con los que presentan los vehí-

[8] En la Directiva vigente esta referencia se encuentra en el considerando 6º ("gran público").

culos, que incluso puede llegar a constituir publicidad ilícita, en la medida en que induzca al conductor a una falsa o no justificada sensación de seguridad[9]. Algunos términos, como *autopilot* o *self-driving*, indican una autosuficiencia en la conducción, una independencia del control humano, y dado que la terminología imprecisa puede generar falsas expectativas y, por tanto, accidentes, los fabricantes están abandonando ciertas palabras en su publicidad[10].

El art. 6 de la propuesta de DPD amplía las circunstancias a tener en cuenta, acorde con la era digital. Son éstas: "a) la presentación del producto, incluidas las instrucciones de instalación, uso y mantenimiento; b) el uso razonablemente previsible y el uso indebido del producto; c) el efecto en el producto de la posibilidad de seguir aprendiendo después del despliegue; d) el efecto sobre el producto de otros productos que quepa esperar razonablemente que se utilicen junto con el producto; e) el momento en que el producto fue introducido en el mercado o puesto en servicio o, si el fabricante conserva el control sobre el producto después de ese momento, el momento en que el producto dejó el control del fabricante; f) los requisitos de seguridad del producto, incluidos los requisitos de ciberseguridad pertinentes para la seguridad; g) cualquier intervención de una autoridad reguladora o de un operador económico contemplado en el artículo 7 en relación con la seguridad de los productos; h) las expectativas específicas de los usuarios finales a los que se destina el producto."

Entre las circunstancias a tener en cuenta, cabe destacar ahora dos de ellas, que son novedades respecto a la regulación vigente. Una de ellas es la capacidad de aprendizaje posterior a la puesta en circulación del producto (*machine learning*). La posibilidad de que el producto siga aprendiendo después de su puesta en circulación es una de las características de los bienes con dispositivos de inteligencia artificial. Sin embargo, no aclara la Directiva en qué sentido debe tenerse en cuenta: por un lado, podríamos entender que esa capacidad de autoaprendizaje permite exonerar de responsabilidad al fabricante del producto capaz de aprendizaje autónomo, dada la imprevisibilidad de la conducta por el fabricante, que interrumpe la relación de causalidad; por otro lado, podríamos entender que la capacidad de autoaprendizaje no exonera de responsabilidad, ya que las expectativas de seguridad incluyen

[9] Como dispone el Real Decreto Legislativo 6/2015, de 30 de octubre, que aprueba el Texto refundido de la Ley de tráfico y circulación de los vehículos a motor y seguridad vial, en su art. 52.
[10] The Waymo Team, *Why You'll Hear Us Saying Fully Autonomous Driving Tech From Now On* (6 Enero 2021), https://blog.waymo.com/2021/01/why-youll-hear-us-say-autonomous-driving.html. [último acceso 5/12/2022].

la imposición de algunos límites a la capacidad de aprendizaje del producto, es decir, que el diseño incorpore salvaguardas para evitar comportamientos peligrosos. Entiendo que el sentido de la norma no puede ser exonerar de responsabilidad al fabricante, aunque habrá que ver qué juego tiene aquí la excepción de riesgos de desarrollo.

En cualquier caso, los vehículos automatizados utilizan sistemas de inteligencia artificial, pero su autonomía está muy limitada, ya que las reglas son estrictas (código de circulación), y los objetivos muy concretos (llegar al destino por la ruta más rápida posible). Pueden tener capacidad de aprendizaje, sobre todo de reconocimiento de entornos y de búsqueda de la respuesta más adecuada, pero encaja en lo que ha venido a denominarse inteligencia artificial débil. La inteligencia artificial fuerte es la que más fácilmente puede adoptar decisiones totalmente autónomas e imprevisibles.

La otra circunstancia gira en torno al "momento de puesta en circulación" de la regulación actual, que en la propuesta de Directiva pasa a ser introducción en el mercado, puesto en servicio o comercializado[11]. Estos momentos suelen coincidir y pueden fijarse en el tiempo; corresponden al momento en que el fabricante deja de tener el control del producto. La nueva propuesta de regulación añade, además, que "si el fabricante conserva el control sobre el producto después de ese momento, el momento en que el producto dejó el control del fabricante". La era digital resta claridad al momento exacto, ya que el producto no sale enteramente de la esfera de control del fabricante una vez entra en el mercado. Dada la posibilidad técnica, y la obligación, de los fabricantes de controlar la seguridad de los productos más allá de su puesta en circulación, los fabricantes deben seguir siendo responsables de las deficiencias de seguridad producidos por los programas informáticos o servicios conexos bajo su control. El fabricante de productos digitales, por tanto, seguirá siendo responsable de las vicisitudes del producto, más allá de la puesta en circulación. La fijación del memento es relevante, no solo en la definición de producto defectuoso, sino al describir las causas de exoneración del productor, como luego veremos.

[11] Para la definición de cada uno de ellos, véase los arts. 4.8, 4.9 y 4.10 PDPD.

b) Los defectos de ciberseguridad

La ciberseguridad implica la protección frente a ciberataques. El Reglamento sobre la ciberseguridad en la UE[12] la define no como un estado o condición de seguridad, sino como una actividad: "todas las actividades necesarias para la protección de las redes y sistemas de información, de los usuarios de tales sistemas y de otras personas afectadas por las ciberamenazas" (art. 2.1), y entiende por tales "cualquier situación potencial, hecho o acción que pueda dañar, perturbar o afectar desfavorablemente de otra manera las redes y los sistemas de información, a los usuarios de tales sistemas y a otras personas" (art. 2.8). El Reglamento europeo crea el certificado de ciberseguridad que deberán tener los productos, servicios o procesos, conforme han sido evaluados y cumplen los requisitos específicos de seguridad. Este certificado está llamado a generar confianza en los productos (cdo. 65)[13].

A medida que aumenta la automatización y la conectividad de los vehículos, aumenta también el riesgo de ciberataques. Un ciberataque puede tener varios efectos, que van desde el control remoto de la música o el aire acondicionado (que es molesto pero no afecta la seguridad de la conducción), al control remoto de la velocidad o el freno, o de los mandos de dirección (que sí genera riesgos de seguridad), y pueden ser accedidos con intencionalidad criminal (comisión de delitos de homicidio, robo mediante alunizaje o, en los casos más graves, ataque terrorista)[14]. Los expertos en seguridad advierten de la facilidad con la que es posible hackear un vehículo, y todos los fabricantes han tenido incidencias en este sentido[15].

[12] Reglamento (UE) 2019/881 del Parlamento europeo y del Consejo de 17 de abril de 2019 relativo a ENISA (Agencia de la Unión Europea para la Ciberseguridad) y a la certificación de la ciberseguridad de las tecnologías de la información y la comunicación y por el que se deroga el Reglamento (UE) nº 526/2013 ("Reglamento sobre la Ciberseguridad en la UE").

[13] Sobre las dificultades de certificación, Scott Mclachlan; Burkhard Schafer, Kudakwashe Dube, Evangelia Kyrimi & Norman Fenton, "Tempting the Fate of the furious: cyber security and autonomous cars", *International Review of Law, Computers & Technology*, 2022, vol. 36, nº 2, pp. 1-21, sobre todo p. 13 y ss.

[14] El acceso inconsentido también puede tener como finalidad la captación de datos personales del usuario. Esa vulneración de la privacidad tendrá consecuencias a efectos del Reglamento general de protección de datos, Reglamento (UE) 2016/679 del Parlamento europeo y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos y por el que se deroga la Directiva 95/46/CE.

[15] Para conocer algunos ejemplos, ver Steve Tengler, "Top 25 Auto Cybersecurity Hacks: Too Many Glass Houses To Be Throwing Stones", *Forbes*, 30 junio 2020. https://www.forbes.com/sites/stevetengler/2020/06/30/top-25-auto-cybersecurity-hacks-toomany-glass-houses-to-be-throwing-stones/

El 6 de julio de 2022 entró en vigor el Reglamento (UE) 2019/2144, de seguridad general de vehículos[16] que, para las cuestiones relativas a la ciberseguridad, se remite a los Reglamentos de las Naciones Unidas[17]. Este marco regulatorio sobre la ciberseguridad centra su atención en la necesidad de adoptar medidas tendentes a evitar el acceso no autorizado al sistema informático del vehículo, pero hay otras vulnerabilidades del sistema, como la interceptación o interferencia en las líneas de comunicación (V2V o V2I[18]), o su bloqueo, así como la instalación de *malware* tras la actualización de software *online*. El esfuerzo regulatorio se centra en el acceso no autorizado, a través de las redes y los sistemas de seguridad, como primera línea de defensa frente a los ciberataques; también sería útil que se adoptasen medidas para reducir el daño una vez producido el acceso inconsentido[19].

A efectos de responsabilidad, como es un riesgo conocido, previsible, los fabricantes deben adoptar medidas para evitar las vulnerabilidades del sistema. En todo caso, estas vulnerabilidades de ciberseguridad son defectos porque no ofrecen la seguridad que las personas pueden legítimamente esperar. Y esto ocurrirá también con las actualizaciones de software, ya que el perjudicado tiene la expectativa legítima de que no introducirán virus o facilitarán el acceso inconsentido.

[16] Reglamento (UE) 2019/2144, del Parlamento Europeo y del Consejo de 27 de noviembre de 2019 relativo a los requisitos de homologación de tipo de los vehículos de motor y de sus remolques, así como de los sistemas, componentes y unidades técnicas independientes destinados a esos vehículos, en lo que respecta a su seguridad general y a la protección de los ocupantes de los vehículos y de los usuarios vulnerables de la vía pública.

[17] Para los vehículos automatizados y conectados, son los Reglamentos de las Naciones Unidas nº 155 y 156. Reglamento nº 155 de la Comisión Económica para Europa (CEPE) de las Naciones Unidas – Disposiciones uniformes relativas a la homologación de los vehículos de motor en lo que respecta a la ciberseguridad y al sistema de gestión de esta; y Reglamento nº 156 de las Naciones Unidas – Disposiciones uniformes relativas a la homologación de vehículos en lo que respecta a las actualizaciones de software y al sistema de gestión de actualizaciones de software [2021/388].

[18] Los vehículos autónomos están dotados de un sistema de comunicación para transferirse información entre ellos (V2V, *vehicle to vehicle*) o con las infraestructuras (V2I, *vehicle to infrastructure*). Colectivamente, y a medida que aumenta estas posibilidades de comuncación, V2X (*vehicle to everything*).

[19] Como sugiere Nynke E. Vellinga, "Connected and vulnerable: cybersecurity in vehicles", *International Review of Law, Computers & Technology*, 2022, vol. 36, nº 2, pp. 1-20, en p. 3-4. La autora propone algunas medidas, como la obligación de aislar los sistemas críticos de seguridad (velocidad, frenado, dirección) de los sistemas que no afectan a la seguridad de la conducción (música y entretenimiento) para evitar que estos sirvan de puerta de entrada a aquel, y la obligación de restaurar el estado de ciberseguridad, mediante el desarrollo de actualizaciones de software para eliminar las vulnerabilidades o la retirada de los vehículos con esa vulnerabilidad.

Pensemos en un supuesto concreto: un vehículo tiene un error de software porque el intermitente se activa cada vez que el vehículo se detiene. El fabricante, que tiene la obligación de monitorizar el producto, detecta el error, corrige el software, y avisa a los propietarios de la necesidad de actualizar el programa informático, que pueden hacer de manera online, sin necesidad de llevar el vehículo al concesionario. Imaginemos ahora que la actualización del software corrige efectivamente dicho error, pero introduce una vulnerabilidad en la ciberseguridad del vehículo, que un hacker explota, generando daños. ¿Debe responder el fabricante del vehículo de esos daños?

Es muy posible que el fabricante pueda alegar, como excepción, que el defecto no existía en el momento de puesta en circulación del producto (art. 7.b) Directiva 85/374, art. 10.1.c) PDPD). Esta causa de exoneración de responsabilidad está vinculada a la interrupción del nexo causal y a la falta de control del fabricante a partir de un determinado momento. Por un lado, si el defecto no es originario, sino que fue introducido con posterioridad, normalmente por un tercero (típicamente, falta de mantenimiento del producto en condiciones de conservación necesarias, error en la manipulación del producto), ello rompe la relación de causalidad. Por otro lado, exige la fijación del momento de puesta en circulación y ya hemos visto las dificultades para fijar con claridad ese momento.

La propuesta de Directiva se decanta por mantener la responsabilidad del fabricante en este caso. El art. 10.2 PDPD, que completa el art. 10.1.c), impide la liberación de responsabilidad cuando el defecto del producto se deba al programa informático, incluidas las actualizaciones o mejoras, siempre que esté bajo el control del fabricante. En su exposición de motivos señala que los programas informáticos o servicios conexos, ya sea en forma de mejoras, de actualizaciones o de algoritmos de aprendizaje automático, "deben considerarse bajo el control del fabricante cuando sean suministrados por él o cuando este los autorice o influya de otro modo en su suministro por un tercero" (cdo 37). Y sigue diciendo que "debe restringirse la posibilidad de que los operadores económicos eludan su responsabilidad demostrando que un defecto se produjo después de que introdujeran en el mercado el producto o lo pusieran en servicio cuando la defectuosidad de un producto consista en la falta de actualizaciones o mejoras de los programas informáticos que sean necesarias para abordar las vulnerabilidades de ciberseguridad y mantener la seguridad del producto" (cdo 38).

Por tanto, los fabricantes del producto responderán de los daños causados por no suministrar actualizaciones o mejoras de seguridad del software cuando sea indispensable para abordar las vulnerabilidades del producto,

teniendo en cuenta la evolución de los riesgos de ciberseguridad. Ahora bien, si el propietario del vehículo conectado ignora las advertencias del fabricante acerca de la necesidad de instalación de actualizaciones de software, y se producen daños, no podrán ser imputados al fabricante.

El fabricante podría intentar alegar que la intervención de un hacker interrumpe el nexo causal, y que el daño resultante no le puede ser atribuido. Sin embargo, es el defecto de ciberseguridad lo que el hacker explota, de manera que si el fabricante hubiese adoptado mayores medidas de ciberseguridad, el producto no hubiese sido vulnerable y el acceso inconsentido no se hubiera producido. La responsabilidad de un operador económico no se reduce cuando los daños sean causados tanto por el defecto como por un tercero (art. 8.1 Directiva 85/374 y art. 12.1 PDPD)[20]. Por si hubiera dudas, la Propuesta de Directiva se refiere expresamente al tercero que explota una vulnerabilidad de ciberseguridad del producto: "en aras de la protección de los consumidores, cuando un producto es defectuoso, por ejemplo, debido a una vulnerabilidad que hace que el producto sea menos seguro de lo que el público en general tiene derecho a esperar, la responsabilidad del operador económico no debe reducirse como consecuencia de tales actos u omisiones" (cdo 41). Solo podrá exonerarse cuando los propios perjudicados hayan contribuido a la causa de los daños, por ejemplo, si hacen caso omiso a las advertencias del fabricante de actualizar el software de seguridad, porque la culpa de la víctima sí puede reducir o eliminar la responsabilidad del operador económico (art. 8.2 Directiva actual, art. 12.2 PDPD).

2.3. Sujetos responsables

Los sujetos responsables por producto defectuoso, según la Directiva 85/374 vigente, son tres: el fabricante (tanto del producto final como del componente o de materia prima, tanto del real como del aparente), el importador y el proveedor o suministrador. En principio, la responsabilidad recae sobre el fabricante del producto final y, en algunos casos, sobre otros sujetos. Así, cuando no se pueda identificar al fabricante, la responsabilidad recae sobre el proveedor o suministrador, salvo que éste identifique, en un plazo prudencial, al fabricante; cuando el producto se haya importado desde fuera de la Unión Europea, la responsabilidad recae sobre el importador.

La idea que subyace a esta canalización de responsabilidad en el fabricante del producto final es que éste tiene el control final sobre el producto. En caso

[20] La directiva vigente deja a salvo expresamente la acción de repetición en el mismo precepto, y para la propuesta de Directiva, habrá que acudir al art. 2.3.b) que deja a salvo el derecho de repetición.

de daños, el perjudicado puede demandar al fabricante del producto final o al fabricante del componente, o a ambos, y es una garantía para el perjudicado poder canalizar la responsabilidad en el fabricante del producto final.

La propuesta de Directiva añade dos actores más en el escenario de posibles sujetos responsables: el representante autorizado[21] y el prestador de servicios de tramitación de pedidos. Para referirse a todos los sujetos, conjuntamente, emplea el término "operador económico" (4.16 PDPD).

Los procesos de fabricación de los años ochenta, cuando se elaboró la Directiva 85/374, eran propios de la fabricación en masa de la era industrial. En numerosas ocasiones, los fabricantes del producto final ensamblaban partes o componentes fabricados por otras empresas. Siguiendo este paradigma, el software puede ser concebido como un componente más. Sin embargo, en la era digital, esta idea debe ser revisada, en la medida en que el software puede ser elemento indispensable y, a diferencia de otros, obliga a adaptar el producto final a las necesidades del programa informático, y no al revés.

La normativa de responsabilidad por productos defectuosos permite canalizar la responsabilidad hacia el fabricante del producto final, lo que elimina la necesidad de averiguar el origen o la causa del defecto y la identificación del fabricante del componente. La distinción entre fabricante del producto final y fabricante del componente o parte integrante es relevante, ya que ambos reciben un trato jurídico distinto. Aunque el perjudicado puede demandar a ambos, existe una causa de exoneración específica para el productor de un componente, pues no será responsable si demuestra que el defecto es imputable a la concepción del producto al que ha sido incorporado o a las instrucciones dadas por el fabricante del producto final (art. 7.f Directiva 85/374, art. 10.1.f PDPD). Pero no ocurre al revés. El fabricante del producto acabado no puede exonerarse alegando que el defecto reside en una de las partes, componentes o materia prima suministrada. Eso sí, una vez pagada la indemnización, tendrá derecho a repetir frente a los demás.

Esta diferencia de trato jurídico está justificada por el proceso de fabricación en serie, que puede exigir alteraciones de los componentes en algunos casos. Por un lado, el componente se puede fabricar siguiendo las especificaciones o instrucciones del fabricante del producto final. Por otro lado, el componente normalmente no llega al consumidor y usuario de una manera inalterada, sino que su integración en el producto final implica alguna adapta-

[21] El representante autorizado es "toda persona física o jurídica establecida en la Unión que haya recibido un mandato por escrito de un fabricante para actuar en nombre de este en tareas específicas (art. 4.12 PDPD).

ción por parte del fabricante del producto final. Finalmente, el fabricante del componente a menudo no tiene medios prácticos o eficientes para supervisar el producto final. El fabricante del producto final es el que debe adoptar las decisiones de reducción de riesgos, es el que debe detectar y remediar los riesgos evitables del producto. Lo que justifica que el fabricante del componente no responda de los defectos del producto final es que no tiene control sobre éste y, por tanto, su responsabilidad queda limitada a aquellos casos en que el componente ya era defectuoso en el momento de salir de su esfera de control.

En relación con los vehículos automatizados, las cosas funcionan de manera algo distinta, pues el programa informático (y, por tanto, su fabricante o programador) tiene una importancia superior a la de otros componentes. Las necesidades del programador informático han llevado a cambios en la manera de diseñar los vehículos. El software no es lo accesorio que se une al producto final, sino que el producto final se adapta a las necesidades del software. El hecho de que las empresas digitales (Google, por ejemplo) hayan irrumpido en el mercado de la automoción revela que, en ocasiones, el producto final (el *hardware*) puede ser lo accesorio.

Por otro lado, los fabricantes de software siguen supervisando el funcionamiento del vehículo con posterioridad a su puesta en circulación, y si algo falla, lo más probable es que, en última instancia, sea atribuible al fabricante de software, no al fabricante del producto final, pues la tecnología controla el vehículo, no al revés. El fabricante de la tecnología digital está en una mejor posición para ajustar el software, que el fabricante del producto final no controla.

Visto lo anterior, seguir manteniendo que el software es un componente del vehículo, y equipararlo a un accesorio, sobre todo en los más altos niveles de automatización, puede no ser acertado. Las empresas tradicionales en el sector de la automoción se han aliado con empresas tecnológicas para fabricar vehículos automatizados, y habitualmente celebran acuerdos de distribución de riesgos, que no son similares a los que acuerdan con otros proveedores. El fabricante del software es el que supervisa el funcionamiento postventa, el que corrige los errores, y como el software es más fácil de arreglar si algo falla, los fabricantes del vehículo final no están preocupados por los gastos de retirada del producto. Eso permite concluir que en relación con los vehículos altamente automatizados, la distinción entre fabricante de producto final y fabricante de componente puede quedar superada, y la empresa que proporciona el software puede responder como fabricante de producto final[22].

[22] En EEUU, el fabricante del componente puede ser considerado fabricante del producto final si "participa sustancialmente" en la integración del componente en el diseño del producto final,

En todo caso, con independencia del reparto interno de riesgos, focalizar la responsabilidad en el fabricante del producto final tiene una ventaja práctica para el perjudicado, que no debe preocuparse por identificar o detallar el origen del defecto, ni la identidad de los fabricantes, más allá del que sí es inmediatamente conocido, que es el fabricante del producto final.

3. La carga de la prueba

Uno de los problemas a los que se enfrenta el perjudicado cuando interpone una acción judicial de reclamación de una indemnización por daños causados por productos defectuosos gira en torno a la prueba que debe aportar. El perjudicado deberá probar el daño, el defecto y la relación causal entre este y aquel (art. 4 Directiva 85/374, art. 9.1 PPDP).

A pesar de que la propuesta de Directiva parte de la misma idea, seguidamente reduce las dificultades probatorias del perjudicado mediante la introducción de unas presunciones *iuris tantum*[23] de defectuosidad (art. 9.2) y de relación de causalidad (art. 9.3). Cabe valorar positivamente que la propuesta de Directiva incluya expresamente estas disposiciones para aligerar la carga de la prueba que recae sobre el demandante, que puede resultar injusta por excesiva, en la línea con lo sugerido por la Comisión europea[24] y del Parlamento europeo[25]. Finalmente termina el precepto con unas indicaciones sobre la facilidad y disponibilidad probatoria de las partes, que conviene destacar. Veamos.

Existen tres hechos de los que cabe presumir que el producto es defectuoso. En primer lugar, por incumplimiento de la obligación del demandado de exhibir las pruebas pertinentes que obren en su poder, cuando el órgano jurisdiccional ha ordenado su aportación al procedimiento judicial (art. 9.2.a) PDPD). Según el art. 8 de la propuesta de Directiva, los órganos jurisdiccionales pueden ordenar al demandado que aporte las pruebas de que disponga,

y esa integración causa el defecto del producto final. No se trata únicamente de que el fabricante del componente tenga mayores conocimientos sobre su producto (en este caso, el software), sino que participa en el diseño del producto final o tiene un control sobre el producto acabado. Cfr. *Restatement (Third) of Torts: Products Liability* (§5). El punto de partida es, igualmente, la exoneración de responsabilidad del fabricante del componente, salvo estas situaciones.

[23] El art. 9.5 PDPD permite al demandado refutar estas presunciones.

[24] Comisión europea, *Libro Blanco de la inteligencia artificial – un enfoque europeo orientado a la excelencia y la confianza*. Bruselas, 19 de febrero de 2020, COM(2020) 65 final.

[25] Parlamento europeo, Resolución de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un régimen de responsabilidad civil en materia de inteligencia artificial 2020/2014(INL).

a petición de la parte demandante, siempre que ésta haya aportado pruebas suficientes para acreditar la verosimilitud de la reclamación de indemnización (art. 8.1). Eso sí, los órganos jurisdiccionales deben velar por que la exhibición de pruebas se limite a las que sean necesarias y proporcionadas para respaldar la demanda (art. 8.2). Los jueces deberán tener en cuenta no solo los intereses de la parte demandante, sino los de la parte demandada y acaso de terceros, y la necesidad de proteger la información confidencial y los secretos comerciales (art. 8.3); a tal fin se adoptarán medidas para preservar la confidencialidad de la información en el procedimiento judicial (art. 8.4). Con esta presunción, se incentiva la aportación de prueba que obra en poder del demandante.

En segundo lugar, por incumplimiento de los requisitos de seguridad de la normativa, tanto nacional como europea, que protegen contra el riesgo de daño que se ha producido (art. 9.2.b) PDPD). Es sabido que el cumplimiento de la normativa de seguridad no exime al demandado de responsabilidad, ya que ésta puede ser inadecuada o insuficiente, pero ahora se puede afirmar claramente que su incumplimiento genera una presunción de defectuosidad. Ahora bien, para que opere la presunción, es necesario que el ámbito de protección de la normativa incumplida sea específicamente el riesgo de daño producido; si se trata de una normativa que nada tiene que ver con el daño producido, su cumplimiento o incumplimiento resulta irrelevante a estos efectos. Es el demandante quien deberá demostrar que el producto no cumple con los requisitos obligatorios de seguridad.

En tercer lugar, por mal funcionamiento evidente del producto que, durante su uso normal, ha causado daños (art. 9.2.c) PDPD). El concepto de defecto no está vinculado con la idea de funcionamiento adecuado, cuya falta puede generar acciones de incumplimiento contractual, pero no de daños. Si el defecto no se puede probar, pero el mal funcionamiento sí, cabe presumir aquel de este. Acaso la prueba del mal funcionamiento ya es prueba directa del defecto, por lo que no siempre será necesario acudir a una presunción.

La presunción de causalidad entre el defecto y el daño surge cuando se haya acreditado el carácter defectuoso del producto y que el tipo de daño sufrido es "compatible normalmente con el defecto en cuestión" (art. 9.3 PDPD). Esta presunción es muy ventajosa para el perjudicado, pues la relación de causalidad es, posiblemente, el elemento más complicado de demostrar en un proceso judicial.

Finalmente, el art. 9 PDPD admite la posibilidad de presumir el defecto, la causalidad, o ambas cosas, cuando el demandante se enfrenta a dificultades excesivas, dada la complejidad técnica o científica del caso (art. 9.4 PDPD).

Con ello se intenta hacer frente a la asimetría de información entre las partes en el proceso judicial. Para que pueda operar esta presunción, como indica el precepto, el demandante debe haber demostrado suficientemente que el producto "contribuyó" a los daños, y que es probable que el producto sea defectuoso o que su carácter defectuoso sea una causa probable del daño. Por tanto, si el demandante logra demostrar que el producto ha contribuido a la causación del daño, aunque no sabe en qué medida ni si es suficiente para ser considerado la causa adecuada del daño, aunque es probable que lo sea, pero no lo puede demostrar por la complejidad técnica del caso, estas dificultades probatorias del demandante, llevan a la presunción de relación de causalidad. En todo caso, el demandado está en mejores condiciones para refutar la presunción, por sus conocimientos expertos y la disponibilidad de la prueba.

4. Causas de exoneración. En particular, la excepción de riesgos de desarrollo

La propuesta de Directiva, en su art. 10, establece las causas de exoneración de responsabilidad civil de los operadores económicos. De todas ellas, que en su mayoría coinciden, adaptadas para el mundo digital, con las vigentes hoy, me centraré ahora en la excepción de riesgos de desarrollo, que mayor debate ha generado, centrado en la conveniencia o no de su mantenimiento en la era digital.

La conocida excepción por riesgos de desarrollo obliga a tener en consideración los avances científicos y tecnológicos y cómo afecta su evolución una vez el producto ha sido puesto en circulación[26]. Si el estado de los conocimientos de la ciencia y de la técnica en el momento de la puesta en circulación del producto no permite descubrir la existencia del defecto, el fabricante quedará exento de responsabilidad (art. 7.e) Directiva 85/374, art. 10.1.e) PDPD). Lo relevante es el estado de la cuestión, los conocimientos científicos y técnicos, en el momento de puesta en circulación del producto (no el existente en el momento de producción del daño) y la imposibilidad de detectar el defecto por parte de la ciencia y la técnica (no lo que puede apreciar el fabricante en cuestión).

[26] Bernhard A Koch, Jean-Sébastien Borghetti, Piotr Machnikowski, Pascal Pichonnaz, Teresa Rodríguez De Las Heras Ballell, Christian Twigg-Flesner, Christiane Wendehorst, "Response of the European Law Institute to the Public Consultation on Civil Liability – Adapting Liability Rules to the Digital Age and Artificial Intelligence", *Journal of European Tort Law*, 2022, vol.13 , nº 1, pp. 25–63, en p. 47, plantean que, de mantenerse la excepción de riesgos de desarrollo, debería ajustarse la noción de puesta en circulación y reclaman una mayor concreción del estado de conocimientos, dada la amplia expansión de la información disponible, particularmente *online*.

Para el fabricante de un vehículo automatizado, como ocurre con los productos que incorporan nuevas tecnologías, va a ser muy fácil alegar la excepción de los riesgos de desarrollo. Sin embargo, con el fin de proteger a las víctimas, y para potenciar la confianza en el nuevo sector de automoción, resultaría aconsejable excluir la posibilidad de alegación de esta excepción a los fabricantes de vehículos automatizados[27]. En este sentido se pronunció el Grupo de Expertos de Responsabilidad y Nuevas Tecnologías de la Comisión europea, en su informe de 21 de noviembre de 2019[28], lo cual parece razonable. Si los fabricantes buscan promover la confianza en el mercado de los vehículos automatizados, la excepción de riesgos de desarrollo puede minar esa confianza, pues desincentiva al consumidor o usuario conocedor de que no va a recibir una compensación en caso de daño causado por productos digitales novedosos. En contra, cabe alegar que la eliminación de la excepción de riesgos de desarrollo desincentiva la investigación o la innovación, en la medida en que el fabricante que investiga y descubre defectos del producto responde. Sin embargo, no parece que haya un incentivo económico en diseñar productos inseguros, ya que la seguridad va a ser elemento clave en el éxito de estos productos. Para que los vehículos autónomos tengan una buena acogida en el mercado, deben ser productos seguros.

El debate acerca de la conveniencia o no de mantener la excepción de riesgos de desarrollo en caso de productos con nuevas tecnologías inteligentes queda zanjado en la propuesta de Directiva, que mantiene la excepción por riesgos de desarrollo con carácter general. En realidad, la opción legislativa va más lejos. Y es que, a diferencia de la Directiva vigente[29], la nueva propuesta de Directiva no permite a los Estados miembros decidir si incorporan o no la exoneración por riesgos de desarrollo. Ello es acorde con el régimen de armonización máxima que establece la propuesta de Directiva (art. 3 PDPD), por lo que las legislaciones nacionales no pueden mantener o introducir disposiciones distintas, aunque sean más favorables para los perjudicados. La imposibilidad de mantener divergencias nacionales, unido a que las disposiciones

[27] Igualmente, Pilar Álvarez Olalla, "Responsabilidad civil en la circulación de vehículos autónomos", en *Inteligencia artificial y riesgos cibernéticos. Responsabilidades y aseguramiento*, Dir. Monterroso Casado, E., Coord. Muñoz Villarreal, A., Tirant lo Blanch, Valencia, 2019, pp. 145-170, en p. 164. En contra, Susana Navas Navarro, "Responsabilidad civil del fabricante y tecnología inteligente. Una mirada al futuro", *Diario La Ley*, nº 35, 27 diciembre 2019, pp. 1-11, en p. 4 (página impresa de la versión online).

[28] Grupo de Expertos de Responsabilidad y Nuevas Tecnologías, *Liability for Artificial Intelligence and other Emerging Digital Technologies*, pp. 42-43.

[29] Art. 15.1.b) Directiva.

en la propuesta de directiva son bastante detalladas, conduce a afirmar que posiblemente el instrumento más adecuado para regular esta materia hubiese sido no una directiva, sino más bien un Reglamento[30].

Si se mantiene la excepción de riesgos de desarrollo para todos los productos, con o sin componente digital, y los Estados miembros no tienen un margen para introducir modificaciones, las legislaciones nacionales que han introducido una distinción a efectos de la alegación de la excepción en función de la categoría del producto (España, Francia, Hungría) deberán eliminar esa distinción[31]. Se mantienen, eso sí, los derechos que puedan tener los perjudicados con arreglo a algún régimen especial de responsabilidad existente a fecha de 30 de julio de 1985 (art. 2.3.d) PDPD); aunque no lo dice expresamente, se refiere al sistema alemán de responsabilidad por medicamentos único existente con anterioridad a la entrada en vigor de la Directiva 85/374.

El estado de la ciencia y de la técnica a tener en cuenta es el existente en el momento de puesta en circulación el producto o, como dice ahora la propuesta de Directiva, en la introducción en el mercado. Según el considerando 37, "el momento de introducción en el mercado o de puesta en servicio es normalmente el momento en que un producto sale del control del fabricante, mientras que para los distribuidores es el momento en que lo comercializan. Por lo tanto, los fabricantes deben quedar exentos de responsabilidad cuando demuestren que es probable que la defectuosidad que causó los daños no existiera en el momento de la introducción en el mercado o puesta en servicio o que se produjo después de ese momento. Sin embargo, dado que las tecnologías digitales permiten a los fabricantes ejercer control más allá del momento de la introducción del producto en el mercado o de la puesta en servicio, los fabricantes deben seguir siendo responsables de las deficiencias que se produzcan después de ese momento como resultado de programas informáticos o servicios conexos que estén bajo su control, ya sea en forma de mejoras o actualizaciones o de algoritmos de aprendizaje automático. Estos programas informáticos o servicios conexos deben considerarse bajo el control del fabricante cuando sean suministrados por él o cuando este los autorice o influya de otro modo en su suministro por un tercero".

[30] Siguiendo los criterios de la *Guía práctica común del Parlamento Europeo, del Consejo y de la Comisión para la redacción de textos legislativos de la Unión Europea*, 2015, p. 12.

[31] En España, los productores de medicamentos, alimentos o productos alimentarios destinados al consumo humano no pueden alegar esta causa de exoneración (art. 140.3 TRLGDCU); en Francia, los elementos del cuerpo humano y los productos derivados del mismo (art. 1245-11 *Code civil*) y Hungría excluye los medicamentos (Código civil de Hungría, sección 6: 555 (3)).

Esto incide en la discusión en torno a si cada una de las actualizaciones postventa del programa informático constituye un producto nuevo (con un nuevo momento de introducción en el mercado) o si, por el contrario, se trata de un producto mejorado. La respuesta acaso dependa de si las actualizaciones son necesarias para el funcionamiento y/o seguridad de la conducción o si son mejora de un producto que funciona igualmente sin incorporar las nuevas actualizaciones. Un producto no podrá ser considerado defectuoso por el solo hecho de que posteriormente se ponga en circulación un producto de forma más perfeccionada" (art. 6.2 Directiva 85/374, art. 6.2 PDPD). Entiendo que si se trata de una modificación del software de seguridad, cada nueva actualización constituye una nueva puesta en circulación del producto. Las actualizaciones que afecten a simples mejoras sin incidencia en la seguridad, seguirán rigiéndose por el momento de puesta en circulación del vehículo anterior, al igual que el resto de elementos (ruedas, airbag, cinturones de seguridad). Esta es una cuestión que deberá ser objeto de revisión ya que, como manifiesta el Libro Blanco, la incorporación de programas informáticos en los productos puede modificar el funcionamiento de tales productos a lo largo de su ciclo de vida, y dar lugar a riesgos que "no se abordan adecuadamente en la legislación en vigor, que se centra sobre todo en los riesgos de seguridad en el momento de la comercialización."

Las actualizaciones de seguridad pueden hacerse online, de manera imperceptible para el propietario del vehículo, que sólo es informado de que la actualización ha tenido lugar, o puede ser necesario que el consumidor realice alguna acción (acudir al concesionario, que instalará la actualización, o conectarse a una página web). A efectos de responsabilidad, la falta de instalación de actualizaciones de software críticas para la seguridad puede constituir un supuesto de culpa exclusiva de la víctima.

5. Observaciones finales

Son muchos los aspectos positivos de la Propuesta de Directiva sobre la que trabaja la Comisión europea, en particular los que tienen que ver con la aligeración de la carga de la prueba del demandante, y las que refuerzan la idea de que el fabricante no puede desentenderse del producto una vez comercializado. El balance es, en general, positivo. Otros aspectos no han sido abordados aquí, que también merecen una valoración positiva, como la eliminación de la franquicia, la introducción de la categoría de daño para la salud psicológica (art. 4.6.a) PDPD), aunque no coincide exactamente con el daño moral, el

aumento del plazo de 10 años de duración de este régimen de responsabilidad a 15 años para atender a la latencia de los daños corporales (art. 14.3 PDPD)[32].

Sin embargo, espero que la Comisión reconsidere su posición respecto a la excepción de riesgos de desarrollo, que creo debería eliminar, al menos respecto de aquellos productos con inteligencia artificial fuerte. Solo así puede generarse la confianza necesaria para su aceptación. Cabe insistir en las ideas clave de la responsabilidad civil, que fundamentan la responsabilidad objetiva en el control (o posibilidad de control), sin olvidar aquí la idea de beneficio. La excepción de riesgo de desarrollo no puede ser un mecanismo tan fácil para quedar exonerado de responsabilidad. Y no se diga que ello desincentiva la investigación, ya que el incentivo proviene de la búsqueda del mejor producto, del más seguro, que permite ganar posiciones en el mercado, y no deriva de la liberación de las consecuencias dañosas que, no hay que olvidar, los vehículos automatizados están diseñados con la intención de reducir.

Bibliografía

Álvarez Olalla, Pilar, "Responsabilidad civil en la circulación de vehículos autónomos", en *Inteligencia artificial y riesgos cibernéticos. Responsabilidades y aseguramiento*, Dir. Monterroso Casado, E., Coord. Muñoz Villarreal, A., Tirant lo Blanch, Valencia, 2019, pp. 145-170.

Bellet, Thierry; Cunneen, Martin; Mullins, Martin; Murphy, Finbarr; Pütz, Fabian; Spickermann, Florian; Braendle, Claudia; Baumann, Martina Felicitas, "From semi to fully autonomous vehicles: New emerging risks and ethico-legal challenges for human-machine interactions", *Transportation Research. Part F*, 2019, nº 63, pp. 153–164.

Gómez Ligüerre, Carlos, "La Propuesta de Directiva sobre responsabilidad por daños causados por productos defectuosos", *InDret,* 2022, nº 4, pp. 1-7.

Kasap, Atilla, *Autonomous Vehicles. Tracing the Locus of Regulation and Liability*, Edward Elgar Publishing, 2022.

Koch, Bernhard A.; Borghetti, Jean-Sébastien; Machnikowski, Piotr; Pichonnaz, Pascal; Rodríguez De Las Heras Ballell, Teresa; Twigg-Flesner, Christian; Wendehorst, Christiane, "Response of the European Law Institute to the Public Consultation on Civil Liability – Adapting Liability Rules to the Digital Age and Artificial Intelligence", *Journal of European Tort Law*, 2022, vol.13, nº 1, pp. 25–63.

[32] Cabe entender que esa duración máxima de 10 años no es un plazo de prescripción y, como advierte Carlos Gómez Ligüerre, "La Propuesta de Directiva sobre responsabilidad por daños causados por productos defectuosos", *InDret*, 2022, nº 4, pp. 1-7, p. 7, debería cambiar el redactado porque si lo denomina plazo de prescripción, quedará sujeto a las interrupciones y suspensiones propias de las legislaciones nacionales.

Mclachlan, Scott; Schafer, Burkhard; Dube, Kudakwashe; Kyrimi, Evangelia & Fenton, Norman, "Tempting the Fate of the furious: cyber security and autonomous cars", *International Review of Law, Computers & Technology*, 2022, vol. 36, nº 2, pp. 1-21.

Navarro-Michel, Mónica, "La aplicación de la normativa sobre accidentes de tráfico a los causados por vehículos automatizados y autónomos", *Cuadernos de Derecho Transnacional*, 2020, vol. 12, nº 1, pp. 941-961.

Navarro Michel, Mónica, "Vehículos automatizados y responsabilidad por producto defectuoso", *Revista de Derecho Civil,* 2020, vol. VII, nº 5, pp. 173-223.

Navas Navarro, Susana, "Responsabilidad civil del fabricante y tecnología inteligente. Una mirada al futuro", *Diario La Ley*, nº 35, 27 diciembre 2019, pp. 1-11.

Tengler, Steve, "Top 25 Auto Cybersecurity Hacks: Too Many Glass Houses To Be Throwing Stones", *Forbes*, 2020. https://www.forbes.com/sites/stevetengler/2020/06/30/top-25-auto-cybersecurity-hacks-toomany-glass-houses-to-be-throwing-stones/

Vellinga, Nynke E., "Connected and vulnerable: cybersecurity in vehicles", *International Review of Law, Computers & Technology*, 2022, vol. 36, nº 2, pp. 1-20.

Breves reflexões sobre a reparação de danos causados na prestação de cuidados de saúde com utilização de robots*

Brief reflections on civil liability for damage caused in the context of provision of healthcare services using robots

Rute Teixeira Pedro**

RESUMO: As inovações relativas à inteligência artificial (IA) e à robótica são utilizadas, crescentemente na área da saúde, afetando, de forma significativa, a prestação de serviços nesse domínio. A par das múltiplas vantagens daí advenientes, existem inequívocos riscos que podem traduzir-se na concretização de danos. No presente trabalho, apresentaremos uma breve reflexão sobre alguns desafios jurídicos que emergem no que respeita à reparação de tais danos. A nossa atenção concentrar-se-á na responsabilidade dos prestadores dos cuidados de saúde, enquanto utilizadores de dispositivos dotados de IA que merecem a qualificação como robots. Destacaremos as características que a IA apresenta para identifi-

* Este texto encontra-se publicado, com o mesmo título, em *Inteligência artificial e robótica. Desafios para o direito do século XXI,* (coord. Eva Sónia Moreira, Pedro Freitas), Coimbra, Gestlegal, 2022, pp. 151- 185. Introduziram-se apenas referências pontuais, na medida em que nos pareceram pertinentes, à proposta apresentada pela Comissão Europeia, a 28 de setembro de 2022, de uma Diretiva relativa à adaptação das regras de responsabilidade civil extracontratual à inteligência artificial (também denominada Diretiva Responsabilidade da IA) e que contém propostas de soluções jurídicas que contendem com as questões objeto de reflexão neste trabalho.

** Prof. Auxiliar da Faculdade de Direito da Universidade do Porto.
Investigadora do CIJ – Centro de Investigação Jurídica.

carmos algumas questões que essas características colocam ao funcionamento da responsabilidade civil, considerando em especial quatro núcleos problemáticos: o do configuração do âmbito dos deveres de informação e de obtenção do consentimento esclarecido do doente; o da identificação dos títulos de imputação que podem ser convocados para fundar a responsabilidade do prestador de cuidados de saúde; o da identificação da(s) pessoa(s) a que pode ser imputada responsabilidade à luz da multiplicidade de intervenientes no iter que permite a utilização dos referidos dispositivos e, finalmente, o da operação de apuramento do nexo causal, atendendo ao contexto de acentuada multicausalidade. Debruçar-nos-emos sobre a possíveis respostas no plano de direito constituído, mas também sobre o possível sentido que as soluções podem vir a revestir no plano do direito a constituir, desde logo dos instrumentos jurídicos que podem ser adotados no contexto da União Europeia.

PALAVRAS-CHAVE: IA; Robot; Saúde; Responsabilidade Civil; Consentimento esclarecido do doente.

ABSTRACT: Innovations regarding Artificial intelligence (AI) and robotics are increasingly used in the healthcare field with significant effects on the provision of related services. Despite the multiple advantages arising therefrom, there are certain risks, and damage may often be produced. This work briefly analyses some legal challenges that emerge regarding its compensation. Our attention will be focused on the civil liability of healthcare providers when they use AI-enabled devices that can be qualified as robots. The AI characteristics AI will be highlighted in order to identify some difficulties that derive from those characteristics to the functioning of civil liability. Four problematic areas will be regarded: the scope of the duties to inform and to obtain informed consent from the patient; the identification of grounds upon which the civil liability of the health care provider may be based; the identification of the person(s) to whom liability may be attributed in light of the multiplicity of persons that intervene in the *iter* that allows the use of the above-mentioned devices and, finally, the assessment of the causal link, considering the context of accentuated multicausality. This work will focus on the possible answers in terms of the applicable law but also on the possible solutions that may be adopted in the future, namely as a consequence of the adoption of legal instruments in the context of the European Union.

KEYWORDS: AI; Robot; Health; Civil Liability; Patient`s informed consent.

1. Observações introdutórias

As inovações relativas à inteligência artificial (IA) e à robótica – que se destacam na assinalável evolução científica e tecnológica que se vive nas últimas décadas[1] – afetam transversalmente a sociedade com impactos múltiplos e de grande significado na produção e distribuição de bens e na prestação de serviços. A área dos cuidados de saúde não constitui exceção à transformação em curso[2].

No presente trabalho, apresentaremos uma breve reflexão sobre alguns desafios jurídicos que emergem no que respeita à reparação de danos causados no âmbito da prestação de cuidados de saúde quando existe o recurso à utilização de *robots*. Vamos centrar a nossa atenção na responsabilidade dos pres-

[1] Transformação apelidada de 4ª revolução industrial. Klaus Schwab, *The Fourth Industrial Revolution*, New York, Crown Business, 2017.

[2] Ainda que, nesta área, não se vislumbre como desejável ou verosímil a possibilidade de total substituição dos profissionais humanos por *robots* na prestação de cuidados de saúde, já que a estes dispositivos faltará – mesmo no nível mais avançado de autonomia e no estádio mais sofisticado de IA – o desenvolvimento (suficiente) da componente emocional (de empatia com o outro ser humano, ainda para mais num contexto de maior vulnerabilidade), essencial ao bom desempenho dos cuidados de saúde. Nesse sentido, Rembrandt Devillé, Nico Sergeyssels e Catherine Middag, "Basic Concepts of AI for legal scholars", in Jan De Bruyne e Cedric Vanleenhove (eds.), *Artificial Intelligence and the Law*, Cambridge, Intersentia, 2021, p. 18. Sendo muito difícil haver uma atuação empática por parte dos dispositivos dotados de IA e constituindo essa componente humana um elemento "crucial" na relação médico-doente, não haverá a substituição dos homens pelos *robots*, ainda que estes sejam cada vez mais importantes na prestação de cuidados de saúde. Por isso, os autores recorrem a uma citação para evidenciar o sentido da evolução nesta área – "to cite Langlotz: «AI will not replace radiologists, but radiologists who use AI will replace radiologists who do not»". *Idem*, *ibidem*.

tadores dos referidos cuidados – nomeadamente dos profissionais médicos, mas também de outras pessoas que, servindo-se desses profissionais, desenvolvem a sua atividade nesse setor – perante os doentes que sofrem danos, no âmbito do serviço que lhes é prestado com recurso a dispositivos dotados de IA que merecem a qualificação como *robots*. Concentrar-nos-emos, pois, na problemática atinente à responsabilidade civil do utilizador desses instrumentos, não nos detendo, precipuamente, na eventual responsabilidade de outras pessoas, nomeadamente aquelas que sejam qualificadas como produtores dos mesmos[3].

Tomaremos, para objeto do nosso estudo, as constelações fácticas em que o serviço de cuidados de saúde é prestado por particulares, no setor privado, e em que para a fundamentação de um pedido ressarcitório deve, portanto, ser convocada aplicação do regime jurídico de responsabilidade civil previsto no Código Civil (CC), seja aquele que é aplicável à responsabilidade obrigacional[4], seja o que se aplica no âmbito delitual[5].

Importa ter presente que a utilização de *robots* aparece numa linha evolutiva de transformação profunda na atividade de prestação de cuidados de saúde, nomeadamente no que respeita ao exercício da medicina, como destacaremos na secção 2 deste trabalho. Os desafios ao funcionamento da responsabilidade civil na área que demarcámos derivam, em larga medida, dos atributos que caracterizam os *robots* e as tecnologias emergentes (nomeadamente de IA) que neles se apresentam incorporadas e são, em muitos pontos, comuns a outros setores de atividade em que se verifica um idêntico movimento de robotização. Na secção 3 do presente trabalho, tendo em atenção as especificidades da prestação de cuidados de saúde, vamos destacar, de entre

[3] Não trataremos, portanto, do regime jurídico da responsabilidade do produtor. Sobre este ponto, veja-se Bernhard Koch, "Product liability 2.0 – mere update or new version?", in Sebastian Lohsee, Reiner Schulze e Dirk Satudenmayer (eds.), *Liability for artificial intelligence and the Internet of things*, Baden-Baden, Nomos, 2019, pp. 99 e ss, e, entre nós, Filipe Albuquerque Matos, "Responsabilidade por danos causados a terceiros por robôs", *Direito e Robótica, Número especial de Estudos de Direito do Consumidor,* agosto de 2020, pp. 179 e ss, e Mafalda Miranda Barbosa, *Inteligência Artificial. Entre a Utopia e a distopia, alguns problemas jurídicos*, Gestlegal, 2021, pp. 81 a 87.

[4] Como em princípio ocorrerá. Na verdade, nas situações que consideramos, há, em regra, a celebração de um contrato entre o paciente e o prestador do serviço (que pode ser o médico ou uma outra entidade, nomeadamente de natureza hospitalar). Veja-se o que se dirá já na próxima secção.

[5] Dado que, estando em causa a prestação de cuidados de saúde, há bens (vida, saúde, integridade física, por exemplo) protegidos delitualmente, nomeadamente através do reconhecimento de direitos absolutos ou de normas de proteção, que podem ser atingidos e assim suscitar a aplicação do art. 483º do CC, verificados que se encontrem os correspondentes requisitos.

a miríade de questões que podem ser assinaladas[6], quatro núcleos problemáticos candentes no âmbito da aplicação das regras de responsabilidade civil, buscando as respostas que o direito positivo português lhes oferece no presente. Partindo de uma análise que atenderá ao direito constituído, assinalaremos algumas debilidades que se detetam nesse plano[7], trazendo para a reflexão as perspetivas de evolução jurídica que se anteveem, nomeadamente no âmbito do direito da União Europeia.

2. Digitalização e robotização em curso da prestação de cuidados de saúde

A robotização da prestação de cuidados de saúde integra-se numa linha evolutiva de grande transformação que se deu no século passado neste setor de atividade com repercussões de grande relevo no plano do seu tratamento jurídico[8]. Na verdade, depois do abandono de uma perspetiva (quase) sacralizada que perdurou durante séculos e que blindava, em larga medida, esta área à sindicação pelo direito e, portanto, ao funcionamento dos regimes jurídicos de responsabilização, deu-se uma significativa assimilação da relação de prestação de cuidados de saúde a outras relações (contratuais) de consumo de serviços e à consequente aceitação da responsabilização (civil) dos seus prestadores[9]. Essa transformação foi acompanhada pela massificação

[6] Veja-se o relatório elaborado pelo Expert Group on Liability and New Technologies – New Technologies Formation, *Liability for artificial intelligence and other emerging digital technologies*, 2019, disponível em https://www.europarl.europa.eu/meetdocs/2014_2019/plmrep/COMMITTEES/JURI/DV/2020/01-09/AI-report_EN.pdf (pp. 19 a 31). [acedido, pela última vez, em 30-03-2023].

[7] Esta nova "obrigará à revisão dos pressupostos clássicos da responsabilidade civil", nas palavras de Henrique Sousa Antunes, "Inteligência Artificial e Responsabilidade Civil: Enquadramento", *Revista de Direito da Responsabilidade Civil*, Ano I, 2019, pp. 139 e 140, [acedido, pela última vez, em 30-03-2023], acessível em https://revistadireitoresponsabilidade.pt/2019/inteligencia-artificial-e-responsabilidade-civil-enquadramento/.

[8] Sobre essa evolução, veja-se Rute Teixeira Pedro, *A Responsabilidade Civil do Médico. Reflexões sobre a noção da perda de chance e a tutela do doente lesado*, Colecção do Centro de Direito Biomédico da Faculdade de Direito da Universidade de Coimbra nº 15, Coimbra, Coimbra Editora, 2008, pp. 30 e ss.

[9] Hoje esta relação é perspetivada como uma relação contratual que se estabelece entre um profissional especializado numa determinada área e o doente, ou entre uma entidade complexa dotada de uma estrutura empresarial que oferece ao público cuidados integrados de saúde a que o paciente recorre, formando-se entre eles um contrato. A afirmação de que a prestação de cuidados de saúde, nomeadamente médicos, se pode desenvolver através do mecanismo contratual e de que a responsabilidade que nesse âmbito se suscite apresenta (também) natureza contratual ou obrigacional, seguindo o regime previsto nos arts. 798º e ss. do CC, é, hoje, aceite, entre nós, na doutrina e jurisprudência. Para a evolução doutrinal, veja-se o nosso *A Responsabilidade Civil do Médico...*, *op. cit.*, pp. 56 e ss. No que respeita à jurisprudência, veja-se, a título ilustrativo, o Acórdão

e anonimização dos cuidados que representaram outros tantos desafios ao funcionamento dos figurinos legais, nomeadamente do instituto jurídico da responsabilidade civil, quando se verifica a ocorrência de um evento gerador de danos, cuja reparação o doente busca.

É neste contexto de transformação – que potenciou o crescimento acentuado de litígios nesta área de atividade – que se dá um movimento de crescente digitalização. Os recursos eletrónicos apareceram, desde logo, como instrumentos muito úteis para a elaboração, consulta, atualização e arquivo dos ficheiros clínicos contendo os dados relativos aos pacientes e à descrição dos cuidados que lhes são ministrados, assim como para a prescrição de fármacos[10]. Acresce que a evolução tecnológica ofereceu múltiplos instrumentos para que a prestação de cuidados de saúde à distância se pudesse concretizar. Os meios telemáticos passaram a estar presentes em muitos momentos do *iter* de prestação de cuidados de saúde, da teleconsulta à teleintervenção cirúrgica. Dá-se, então, a ampliação gradual do recurso à telemedicina[11] que, tendo acontecido primeiro gradualmente, recebeu, muito recentemente, o impulso ditado pelas medidas de contenção de transmissão do vírus SARS--CoV-2 e de combate à pandemia pela doença infeciosa causada por esse novo coronavírus (COVID-19). Por outro lado, a acentuação da digitalização da vida moderna potenciada pelo aumento exponencial da interconectividade refletiu-se também na área da saúde com a proliferação de *smart things* resultante do fenómeno denominado *Internet of Things*. Aparece, pois, o denominado s*mart patient*, ao alcance do qual se encontram múltiplos instrumentos de monitorização do estado de saúde e deteção de potenciais problemas carecidos de tratamento[12].

do Supremo Tribunal de Justiça (STJ) de 22 de março de 2018 (Processo nº 7053/12.7TBVNG.P1.S1), acessível em www.dgsi.pt.

[10] Pense-se nas ferramentas de prescrição eletrónica médica (PEM) que incluem algoritmos com regras de prescrição que derivam, nomeadamente, de normas emitidas pela Direção-Geral de Saúde e pelo INFARMED (Despacho nº 7979-P/2015, de 20 de julho), na sequência do Decreto-Lei nº 106 -A/2010, de 1 de outubro, que estabeleceu o princípio da obrigatoriedade da prescrição eletrónica, com vista à racionalização do acesso ao medicamento, no âmbito do Serviço Nacional de Saúde (SNS). A denominada "receita sem papel", que permitiu a "desmaterialização Eletrónica da Receita", é um modelo eletrónico que abrange várias etapas, nomeadamente a prescrição pelo médico e a dispensa na farmácia.

[11] Para a qual já se preveem regras particulares nos arts. 46º e ss. do Código Deontológico da Ordem dos Médicos (doravante, CDOM) constante do Regulamento nº 707/2016, publicado em *Diário da República*, na 2ª série, nº 139, de 21 de julho de 2016.

[12] Sobre a multiplicidade de componentes deste fenómeno que compreende plataformas móveis (onde se incluem os telemóveis, *tablets* e outros *gadgets* eletrónicos), sensores (uns externos, outros

Mais recentemente, a revolução digital atingiu um novo patamar com o recurso crescente a instrumentos que incorporam tecnologia sofisticada dotada de aptidão para agir com autonomia (ou, pelo menos, com graus consideráveis e variáveis de autonomia), replicando, até certo ponto, a inteligência humana. Pode ilustrar-se o fenómeno de que falamos com a referência à utilização de *robots* cirúrgicos[13]. Podem também referir-se os *robots* assistentes que cuidam do paciente internado, nomeadamente no período no pós-operatório (o levantam da cama, o ajudam a mover-se) ou ainda os *robots* associados à prescrição e administração de fármacos[14].

Da multiplicação do recurso à IA e da crescente robotização – também promovidas pelo contexto pandémico em que vimos vivendo desde 2020 – resultam grandes benefícios que têm de ser destacados e não podem ser desconsiderados quando se pondera a resposta que se entende dever ser oferecida pelo direito no que respeita à reparação dos danos causados por aqueles instrumentos. Na verdade, devem destacar-se muitas vantagens económicas e sociais de que a comunidade e cada um dos seus membros beneficiam. Na verdade, o recurso a estes dispositivos tecnológicos aumenta as possibilidades de se fazer chegar os cuidados de saúde a pessoas deles carecidas em zonas mais remotas e desfavorecidas[15]. Por outro lado, assinala-se a poupança

mais "invasivos"), *apps* e correspondentes ligações e os problemas jurídicos que eles levantam, veja-se Carolina Cunha, "O doente sem horário: breve anatomia dos problemas jurídicos suscitados pelas aplicações móveis na área da saúde", *Direito e Robótica*, *Número especial de Estudos de Direito do Consumidor*, agosto de 2020, pp. 43 e ss.

[13] O setor da saúde é uma área de propagação da inovação robótica, como afirma Erica Palmerini, que dá como exemplos, para além dos *robots* cirúrgicos e os *robots*-assistentes usados em entidades hospitalares, as cápsulas médicas inteligentes introduzidas no corpo humano, os sistemas robóticos de monotorização de parâmetros fisiológicos, as próteses biónicas avançadas introduzidas no sistema nervoso central e periférico e os *robots* usados na reabilitação de lesões de medula. Erica Palmerini, «Robotica e diritto: suggestioni, intersezioni, sviluppi a margine di una ricerca europea», *Responsabilità civile e previdenza*, Volume 81, Fascículo 6, 2016, pp. 1818, 1820 e 1832. Para uma descrição da utilização de IA no âmbito da saúde (desde a prestação de cuidados de saúde, à investigação e desenvolvimento de fármacos, à gestão e planificação de sistemas de saúde, à saúde pública e à vigilância de saúde pública), veja-se o documento produzido pela Organização Mundial de Saúde (OMS), *Ethics and governance of artificial intelligence for health: WHO guidance*, World Health Organization, 2021, [acedido, pela última vez, em 30-03-2023], disponível em https://www.who.int/publications/i/item/9789240029200 (pp. 6-16).

[14] Pode até haver uma conjugação de vários *robots* que interagem uns com os outros. Tome-se, como exemplo, aquele que foi dado pelo Doutor Francisco Andrade, na Conferência organizada na Escola de Direito do Minho, em 5 de novembro de 2021, em que o *robot* prescritor de fármacos de uma entidade hospitalar interage com o *robot* da entidade farmacêutica que dispensa os fármacos.

[15] Desde o início da introdução de elementos tecnológicos na medicina que se afirmou, aliás, a vantagem dos mesmos para as camadas mais desprotegidas da população e nomeadamente

de custos e de comportamentos repetitivos, técnicos e burocráticos em que os profissionais e técnicos por vezes se veem enredados[16], permitindo que concentrem tempo e energia nas tarefas em que as suas competências, saberes e habilidades mais fazem a diferença[17]. Acresce que, evitando-se o contacto entre humanos, se pode, assim, facilitar a prevenção de disseminação de doenças contagiosas[18]. Finalmente e muito particularmente, deve evidenciar-se como uma das vantagens mais significativas o aumento da fiabilidade de alguns resultados face àqueles que se alcançariam sem o uso desses dispositivos, desde a previsão de risco de ocorrência de certas doenças, da sua recorrência num dado paciente[19], à previsão das possibilidades de cura ou sobrevivência e à conclusão sobre a determinação da solução terapêutica que mais potencia essas possibilidades. Pode, pois, alcançar-se um maior grau de pessoalização do juízo que se faz sobre aquele concreto paciente e a situação em que ele se encontra[20], num dado momento, por a IA poder trabalhar com uma quantidade muito maior de dados do que a inteligência humana poderia fazer agregadamente e em tempo útil.

É, no entanto, incontornável a referência à circunstância de, do mesmo passo, se fomentar o aparecimento de novas espécies de eventos lesivos de onde podem promanar danos cuja ressarcibilidade deve ser equacionada. Na verdade, a fiabilidade dos resultados que se deixou assinalada como uma vantagem da utilização das novas tecnologias depende da quantidade e da qua-

para as que se encontram em zonas mais distantes dos principais centros hospitalares e, muito particularmente, nos países mais pobres. Gaëlle Marti, Lucie Cluzel-Métayer e Samir Merabet, "Droit et intelligence artificielle", *La semaine Juridique. Édition Générale*, n.os 51 e 52, 20 de dezembro de 2021, p. 1373.

[16] Essa é uma das "dez mudanças essenciais em paradigmas da sociedade atual" que a "IA e as demais tecnologias digitais emergentes vêm aprofundando ou ocasionando", segundo Henrique Sousa Antunes, *Direito e Inteligência Artificial*, Lisboa, Universidade Católica Portuguesa, 2020, pp. 13 e ss., em especial, p. 15.

[17] Há ganhos na organização e gestão de serviços clínicos, cirúrgicos, anestesiológicos e de administração de fármacos, por exemplo. Antonio Oliva *et alii*, "Management of Medico-Legal Risks in Digital Health Era: A Scoping Review», *Frontiers in medicine*, publicado em janeiro de 2022, https://doi.org/10.3389/fmed.2021.821756 [acedido em 25/05/2022], p. 1 [acedido, pela última vez, em 30-03-2023].

[18] O que se destacou no contexto pandémico vivido nos últimos anos.

[19] Por exemplo, quanto a doenças cardiovasculares ou ao glaucoma. Antonio Oliva *et alii*, "Management of Medico-Legal Risks in Digital Health Era: A Scoping Review", *op. cit.*, p. 1.

[20] A evolução da "massificação à personalização", com a promoção de "respostas tecnológicas dimensionadas às necessidades concretas do beneficiário", é outra das dez mudanças essenciais produzidas pela IA na enunciação de Henrique Sousa Antunes, *Direito e Inteligência Artificial*, *op. cit.*, pp. 20 e ss., em especial, p. 21.

lidade dos dados em que o juízo conclusivo do dispositivo de IA se baseou[21]. Vêm-se antecipando também os perigos associados à concretização de dois fenómenos que se vêm verificando e terão um impacto acentuado na prestação de cuidados de saúde, promovendo situações de ocorrência de *erros* de que podem resultar danos. Falamos, por um lado, de uma crescente perda de perícia e competência ("*deskilling*") dos profissionais[22] que deixarão de exercitar a técnica e habilidade necessárias ao desempenho de certos atos que os *robots* passam a praticar e, por outro lado, falamos do risco de descuidos decorrentes da confiança excessiva ("*overfaith*") depositada pelos mesmos na IA e nos *robots*[23]. Estas tendências não podem deixar de ser consideradas quando se equaciona, desde logo à luz de um regime de responsabilidade civil assente na prática de ilícitos culposos, o padrão de exigibilidade comportamental que deve considerar-se aplicável *in casu* e, consequentemente, a afirmação da censurabilidade (objetiva e subjetiva) de tais condutas para merecerem a qualificação como factos fundadores de responsabilidade civil.

3. Dos desafios ao funcionamento da responsabilidade civil por danos causados no âmbito da prestação de cuidados de saúde com recurso a robots

Os atributos específicos e caracterizadores da IA e dos *robots* conduzem, pois, à emergência de desafios jurídicos ao funcionamento da responsabilidade civil na área da prestação de cuidados de saúde, acentuando, aliás, núcleos proble-

[21] Trata-se de uma manifestação da "*data hungriness*". Antonio Oliva *et alii*, «Management of Medico-Legal Risks in Digital Health Era: A Scoping Review», *op. cit.*, p. 5. A sub-representação de certos grupos é uma das causas dos perigos do enviesamento das conclusões e da menor fiabilidade dos resultados a que chegam os dispositivos tecnológicos de saúde. Assim, Hussein Ibrahim, Xiaoxuan Liu, Nevine Zariffa, Andrew D. Morris e Alastair K. Denniston, "Health data poverty: an assailable barrier to equitable digital health care", *The Lancet Digital Health*, Vol. 3, nº 4, março de 2021, p. e260, disponível em www.thelancet.com/digital-health [acedido, pela última vez, em 30-03-2023]. Pense-se no caso da utilização de um dispositivo a propósito de diagnóstico de um cancro de pele desenvolvido com dados de pessoas com uma determinada cor de pele, que será inadequado para pessoas com outra cor de pele. Este exemplo de "*risque de biais*" é apresentado por Gaëlle Marti, Lucie Cluzel-Métayer e Samir Merabet, "Droit et intelligence artificielle", *op. cit.*, p. 2362. I. Glenn Cohen dá um exemplo semelhante relativo ao diagnóstico de cancro da mama, referindo-se à sub-representação das mulheres de origem africana que apresentam uma densidade diferente das caucasianas. I. Glenn Cohen, "Informed Consent and Medical Artificial Intelligence: What to Tell the Patient?", *The Georgetown Law Journal*, Volume 108, 2020, pp. 1464 e ss. (pp. 1425-1469).

[22] Antonio Oliva *et alii*, "Management of Medico-Legal Risks in Digital Health Era: A Scoping Review", *op. cit.*, p. 4.

[23] *Idem*, *ibidem*.

máticos da interseção desse instituto jurídico com este domínio de atividade e dificuldades várias que, tradicionalmente, se identificavam na sua aplicação nesta área. Importa, pois, conhecer essas características.

Sem nos demorarmos nas problemáticas atinentes à definição de cada um desses conceitos, consideraremos que um *robot* é uma máquina construída tendo por base o paradigma de atuação da IA[24]. A utilização do *robot* congrega, por isso, potencialidade de concretização de riscos do mundo físico e do mundo digital e da conexão entre ambos[25]. Tradicionalmente, era posta a ênfase na componente física do *robot* (que pode ter forma humanoide ou não), mas, mais recentemente, denota-se o relevo, desde logo para a perspetivação jurídica desta espécie de dispositivo, da inteligência (nomeadamente IA) que ele incorpora[26]. Quanto a esta componente, tomaremos como ponto de partida para a reflexão que levamos a cabo a ideia de que um dispositivo dotado de IA é aquele que é "capaz de, tendo em vista um determinado conjunto de objetivos definidos por seres humanos, criar resultados, tais como conteúdos, previsões, recomendações ou decisões, que influenciam os ambientes com os quais interage"[27].

Não nos concentrando na componente física dos *robots* que se encontra também noutros maquinismos – não sendo, por isso, a fonte das questões mais

[24] Servimo-nos, aqui, de uma das definições que é apresentada por Ugo Pagallo ("Despite the multiplicity of robotic applications, some argue that we are dealing with machines built basically upon the mainstream «sense-think-act» paradigm of AI research (Bekey 2005)"), ainda que este Autor entenda que nenhuma das definições fornecidas dissipa todas as dúvidas sobre a noção de *robot*. Ugo Pagallo, *The law of robots*, Springer, Heidelberg, London, New York, 2013, p. 2. Para maiores desenvolvimentos sobre a definição de robot, veja-se, entre nós, Nuno Sousa e silva, "Direito e Robótica: uma primeira aproximação", *Revista da Ordem dos Advogados*, 2017, pp. 499 e ss. (pp. 487-553).

[25] Erica Palmerini, «Robotica e diritto: suggestioni, intersezioni, sviluppi a margine di una ricerca europea», *op. cit.*, p. 1826. Podem distinguir-se várias espécies de *robots*, nomeadamente *robots* teleoperados, *robots* autónomos (no sentido em que não precisam de intervenção humana durante a fase executiva) e os *robots* cognitivos (que são os que estão aptos a uma atuação inteligente). Para uma síntese, Giovanni Di Rosa, "Quali regole per i sistemi automatizzati «intelligenti»", *Rivista di Diritto Civile*, Ano LXVII, nº 5, setembro-outubro de 2021, pp. 832 a 835 (pp. 823-853).

[26] Amedeo Santosuosso e Barbara Bottalico, "Autonomous systems and the law: why intelligence matters", in Eric Hilgendorf e Uwe Seidel (eds.), *Robotics, Autonomics and the Law*, Baden-Baden, Nomos, 2017, pp. 27 e 30.

[27] Reproduzimos a definição (1) de "sistema de inteligência artificial" constante do art. 3º da Proposta de Regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial (Regulamento Inteligência Artificial) e altera determinados atos legislativos da União de 21 de abril de 2021, disponível em https://eur-lex.europa.eu/legal-content/PT/TXT/HTML/?uri=CELEX:52021PC0206&from=EN.

inovadoras que ora emergem –, cumpre-nos, aqui, destacar, as características que a IA (de que estes dispositivos são *animados*) apresenta e que geram especiais dificuldades no âmbito da responsabilidade civil. Falamos, aqui, de uma tríplice ordem de atributos, entre outros que poderiam ser enunciados[28]: i) por um lado, a sua complexidade que congrega a intervenção de várias pessoas (nomeadamente, o programador, o fabricante, o vendedor), para além dos seus utilizadores ou manuseadores; ii) por outro lado, a sua interconectividade com outros dispositivos e com o meio ambiente de onde colhem informações empíricas e através das quais desenvolvem todas as suas potencialidades e iii) finalmente, e muito especialmente, a autonomia[29] e capacidade de produzir decisões assentes num processo de autoaprendizagem (*self-learning*), partindo dos dados recolhidos (aprendidos) através daquelas interações.

Ora, tendo estas características presentes, vamos destacar 4 núcleos problemáticos. Consideraremos, sucessivamente, as especificidades que podem surgir quanto ao cumprimento dos deveres de informação e de obtenção do consentimento prévio e esclarecido do doente (3.1.), os títulos de imputação que podem ser convocados, considerando também a densificação dos deveres que devem considerar-se existentes na relação de prestação de cuidados de saúde com recurso a *robots* (3.2.), a multiplicidade de intervenientes e a con-

[28] Sobre as particulares características da IA, da internet das coisas e da robótica, considere-se o Relatório da Comissão ao Parlamento Europeu, ao Conselho e ao Comité económico e social europeu: Relatório sobre as implicações em matéria de segurança e de responsabilidade decorrentes da inteligência artificial, da Internet das coisas e da robótica, COM (2020) 64 final, de 16 de fevereiro de 2020, disponível em https://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=CELEX:52020DC0064&from=en (p. 2). Como características das tecnologias digitais emergentes que têm potencial disruptivo com impacto na responsabilidade civil, podem apresentar-se, para além da complexidade e da crescente autonomia referidas agora em texto, a opacidade, a vulnerabilidade tecnológica (aos dados utilizados e aos ciberataques), a abertura dos ecossistemas tecnológicos a outros sistemas e fontes de informação. Sobre estas características veja-se também a "Response of the European Law Institute" quanto à consulta pública da Comissão Europeia sobre "Civil liability – adapting liability rules to the digital age and artificial intelligence", European Law Institute, 2022, pp. 23 e ss., e também Ugo Salanitro, "Intelligenza artificiale e responsabilità: la strategia della Commissione Europea", *Rivista di Diritto Civile*, Ano LXVI, nº 6, novembro-dezembro de 2020, pp. 1246 a 1248 (pp. 1246-1276).

[29] Tal como noutros dispositivos dotados de IA (por exemplo, os veículos automóveis), os graus de autonomia são diversos, indo dos meros *robots* colaborativos ou assistenciais (que carecem de intervenção humana) até aos *robots* totalmente autónomos. Para a indicação dos 5 níveis de autonomia no âmbito dos dispositivos de saúde, veja-se Danielle S. Bitterman, Hugo J. W. L. Aerts e Raymond H. Mak, "Approaching autonomy in medical artificial intelligence", *The Lancet Digital Health*, Vol. 2, nº 9, setembro de 2020, pp. e447 e ss, disponível em www.thelancet.com/digital-health [acedido, pela última vez, em 30-03-2023].

sequente dificuldade de identificação da(s) pessoa(s) a que se pode imputar responsabilidade pelos danos ocorridos (3.3.) e, finalmente, as particularidade que se detetam na operação de apuramento do nexo causal, considerando o contexto de multicausalidade que se acentua, potenciando múltiplas interferências causais (3.4.).

3.1. Do consentimento da pessoa que recorre aos cuidados de saúde com recurso a robots

Como resulta hoje assente, a intervenção prestadora de cuidados de saúde pressupõe a prévia decisão da pessoa que à mesma vai ser submetida e que autoriza a sua consecução[30]. A autodeterminação do paciente ocupa um lugar central no correto desempenho das atividades da área referida, sendo, portanto, exigida, em princípio[31], a obtenção prévia do seu consentimento para o ato que vai ser praticado[32]. O consentimento do paciente deve ser prestado antes de qualquer intervenção e tratamento médico-cirúrgico, como requisito legitimador dessa intervenção, de modo livre e esclarecido[33].

Por isso, como dever instrumental do dever de obtenção do consentimento do doente, é necessário fornecer um conjunto de informações, que sejam relevantes para a ponderação e tomada de decisão por parte do paciente, sobre o ato a realizar, tendo em conta as concretas circunstâncias do caso. Na verdade, o fornecimento de informação visa o objetivo de esclarecimento do paciente, pelo que "deve ser prestado pelo médico com palavras adequadas, em termos compreensíveis, adaptados a cada doente, realçando o que tem importância ou o que, sendo menos importante, preocupa o doente"[34]. Consequentemente, ao teor dos deveres relativos quer à informação a prestar, quer ao consentimento

[30] Várias questões se colocam quando essa pessoa for incapaz *de iure* ou de facto. Não vamos, no entanto, considerar essas hipóteses.

[31] Não nos debruçaremos também sobre os desvios a esta exigência.

[32] Esta exigência extrai-se, entre nós, dos arts. 70º, 81º e 340º do CC, dos arts. 156º e 157º do Código Penal, do art. 135º, nº 11, do Estatuto da Ordem dos Médicos e dos arts. 19º e ss. do CDOM. No mesmo sentido, podem convocar-se diversos instrumentos internacionais que vinculam o Estado Português, de que destacamos o art. 5º da Convenção sobre os Direitos do Homem e a Biomedicina – Convenção de Oviedo.

[33] Sobre o consentimento que deve ser prestado pelo doente, veja-se André Dias Pereira, *O consentimento informado na relação médico-paciente. Estudo de Direito Civil*, Coimbra, Coimbra Editora, 2004, especialmente, para o dever de esclarecimento, pp. 349 e ss., e para a telemedicina, pp. 550 e ss., e, do mesmo Autor, *Direitos dos pacientes e responsabilidade médica*, Coimbra, Faculdade de Direito da Universidade de Coimbra, 2012, pp. 345 e ss, acessível em https://estudogeral.sib.uc.pt/bitstream/10316/31524/1/Direitos%20dos%20pacientes%20e%20responsabilidade%20m%C3%A9dica.pdf.

[34] Art. 19º, nº 3, do CDOM.

a obter tem de ser reconhecido um caráter "elástico", devendo os mesmos "ser aferidos à luz das especificidades de cada caso concreto"[35].

De qualquer modo, pode afirmar-se que o âmbito do dever de esclarecimento tem crescido nas últimas décadas num sentido de promoção da tutela conferida ao direito de autodeterminação da pessoa, detetando-se, aliás, um aumento de litígios relativos à reparação de danos decorrentes da violação desse direito[36]. Entende-se, pois, que o esclarecimento deve abranger aspetos relativos aos atos a praticar, às finalidades que com eles se pretendem alcançar e às "consequências funcionais" que dos mesmos podem resultar[37]. Considera-se ademais que, quando existam, também devem ser fornecidas informações sobre as alternativas terapêuticas, cirúrgicas ou farmacológicas e o grau comparativo de sucesso, apresentando vantagens e desvantagens de cada uma delas[38]. De qualquer modo, ainda que o *quantum* de informação varie em função de um conjunto de critérios (nomeadamente a natureza do ato, a finalidade curativa ou não curativa prosseguida com ele, a frequência e probabilidade[39] de ocorrência de um determinado risco), não parece poder entender-se cum-

[35] Acórdão do STJ de 14 de dezembro de 2021 (Processo nº 711/10.2TVPRT.P1.S1). Sobre o conteúdo da informação, ver, por todos, André Dias Pereira, *O consentimento informado na relação médico-paciente...*, *op. cit.*, pp. 369 e ss.

[36] Este fenómeno aparece, aliás, num contexto já identificado de especial dificuldade para o doente de conseguir obter uma reparação pelos danos decorrentes de um erro técnico, à luz das regras jurídicas vigentes. A dificuldade de demonstração do erro e do seu caráter ilícito e culposo parece ter ditado uma acentuação da fundamentação dos pedidos ressarcitórios em falhas dos profissionais relativas à necessária obtenção do consentimento esclarecido dos pacientes. Este caminho facilita a tarefa do doente, já que, à luz das regras de distribuição do ónus probatório, caberá ao profissional médico provar a existência de prestação do consentimento informado do paciente. Trata-se de um facto impeditivo do direito do doente a ser ressarcido pelos danos, cuja responsabilidade imputa ao médico (art. 342º, nº 2, do CC). Nesse sentido, vejam-se os Acórdãos do STJ de 2 de junho de 2015 (Processo nº 1263/06.3TVPRT.P1.S1), de 14 de dezembro de 2021 (Processo nº 711/10.2TVPRT.P1.S1) e de 18 de janeiro de 2022 (Processo nº 19473/17.6T8LSB.L1.S1), disponível na base de dados www.dgsi.pt, e, na doutrina, André Dias Pereira, *O consentimento informado na relação médico-paciente...*, *op. cit.*, pp. 187 e ss.

[37] O nº 2 do art. 19º do CDOM prevê que "o esclarecimento deve ser prestado previamente e incidir sobre os aspetos relevantes de atos e práticas, dos seus objetivos e consequências funcionais, permitindo que o doente possa consentir em consciência".

38 "A autodeterminação nos cuidados de saúde implica, não só que o paciente consinta ou recuse uma intervenção determinada heteronomamente, mas também que disponha de toda a informação relativa às diversas possibilidades de tratamento", como se afirma no Acórdão do STJ de setembro de 2020 (Processo 148/14.4TVLSB.L1.S1), disponível na base de dados www.dgsi.pt.

[39] Aliás, segundo o art. 19º, nº 5, do CDOM, o esclarecimento "deve ser feito, sempre que possível, em função dos dados probabilísticos e facultando ao doente as informações necessárias para que possa ter uma visão clara da situação clínica e tomar uma decisão consciente".

prido o dever de esclarecimento quando há uma referência genérica, não especificada e não atualizada, aos aspetos que se considerem relevantes *in casu*. Ora, não nos suscitam dúvidas de que, pelo menos no momento presente[40], a utilização de *robots* constituirá, precisamente, um dos elementos que deve ser comunicado ao paciente. Assim, tratando-se, por exemplo, de uma intervenção cirúrgica com assistência de *robot*, essa informação deve ser fornecida, assim como devem ser prestadas informações sobre o procedimento que vai ser efetuado, os riscos que ele acarreta e as vantagens que daí se derivarão (incluindo os riscos e as vantagens que resultam especificamente da utilização do *robot* cirúrgico, nomeadamente por comparação aos riscos e vantagens associadas à não utilização de tal *robot*[41]). *Não nos parece que possa exigir-se, sempre, ao profissional médico que preste uma informação completa sobre o funcionamento do robot*, seja na sua componente física ou de *hardware*, seja na sua componente de sistemas operativos ou de *software*, nomeadamente dotados de IA, que nele estão incorporados e que determinam o seu funcionamento. Na verdade, com frequência, não conhecerá sequer o profissional, nem lhe será exigível que conheça em detalhe todos os pormenores de *design* e funcionamento dos dispositivos e da sua componente algorítmica – a complexidade dos dispositivos sobre que refletimos é, muitas vezes, grande, exigindo conhecimentos de outras áreas do saber que ultrapassam aquelas que devem ser dominadas pelos profissionais de saúde[42] –, nem essa informação será determinante da decisão de um paciente que também não dominará o saber correspondente[43].

[40] A evolução tecnológica e a omnipresença destes dispositivos num futuro próximo poderão vir a descaracterizar a relevância qualificada dessa informação para a tomada de uma decisão.

[41] A possibilidade de intervenção sem *robot* cirúrgico deve ser apresentada, quando se perfile como alternativa, para que o paciente possa decidir se prefere que a intervenção se dê com utilização de *robot* ou sem ela. A proliferação de *robots* cirúrgicos, nomeadamente em certas intervenções, poderá levar a que, a prazo, a alternativa não exista, restando ao paciente submeter-se à intervenção (com *robots*) ou não se submeter de todo a ela.

[42] Fala-se em "*black-box*" a propósito dos dispositivos em análise, com um fundamento diversificado. Uma das razões para o emprego desta expressão contende com a opacidade dos mesmos, desde logo da componente algorítmica e de IA que neles está incorporada, nomeadamente quando se trata de "unsupervised algorithms". Antonio Oliva *et alii*, "Management of Medico-Legal Risks in Digital Health Era: A Scoping Review", *op. cit.*, p. 4. Os dispositivos podem precisamente ser classificados em função dessa complexidade e da opacidade que apresentam. I. Glenn Cohen, "Informed Consent and Medical Artificial Intelligence: What to Tell the Patient?", *op. cit.*, p. 1431. Lembre-se que a explicabilidade dos sistemas é "limited, coslty, and not always fully feasible in the whole extent", como se sintetiza na "Response of the European Law Institute" quanto à consulta pública da Comissão Europeia sobre "Civil liability – adapting liability rules to the digital age and artificial intelligence", *op. cit.*, p. 26.

[43] Ora, os deveres de informação e de obtenção de consentimento informado por parte dos profissionais médicos apareciam, precisamente, como deveres que serviam o intuito de "combater

Como sabemos, o consentimento devidamente prestado, para além de ser uma exigência que afasta a ilicitude da intervenção do prestador de cuidados de saúde, constitui também um elemento demarcador dos riscos associados à prestação consentida que, em caso de concretização, levarão a que o doente arque com as consequências danosas que àqueles riscos se liguem causalmente, sem que possa demandar a reparação dos danos sofridos aos prestadores dos cuidados de saúde. Ora, não nos parece que se possa derivar do consentimento prestado pelo paciente à intervenção de um *robot* a assunção de outros riscos que não aqueles que resultem da utilização correta de um *robot* e de que o doente tenha sido instruído. O consentimento prestado pelo doente não pode servir para que ele arque, definitivamente e sem possibilidade de reparação, com as consequências desvantajosas decorrentes da utilização de um *robot* de cuja atuação tenha resultado uma prestação de cuidados de saúde *defeituosa* como consequência de o algoritmo nele incorporado ter sido *erradamente* desenhado (nomeadamente dirigindo-se à consecução, a título principal, do objetivo de poupar recursos em detrimento do objetivo de potenciar o sucesso terapêutico da patologia que o paciente apresente), nem para que o doente suporte as consequências danosas decorrentes da circunstância de a informação com que a IA integrada no *robot* opera não ser representativa de determinadas características presentes no caso e que, segundo a ciência médica, são relevantes para o diagnóstico de uma determinada patologia. De igual modo, o consentimento prestado pelo doente não pode servir para que ele suporte, sem possibilidade de reparação, as consequências lesivas decorrentes da utilização de um *robot* de cuja atuação tenha resultado uma prestação de cuidados de saúde *defeituosa* como consequência de uma falha de atualização do *software*, ou de uma falha humana na introdução de dados sobre o paciente, ou de uma falta de carregamento/fornecimento de energia no decurso da intervenção, ou de um *bug* informático ou de um ataque de um *hacker*.

Assim, sem prejuízo da prestação de todas aquelas informações que sejam específicas da utilização de *robots* e que possam ser relevantes para a decisão do doente (nomeadamente, quando haja a alternativa de uso ou não uso de *robot*, a informação comparativa das vantagens e desvantagens das duas opções em alternativa deverá considerar-se pertinente), a proteção do doente

– sem aniquilar – a já assinalada característica de desequilíbrio que é apanágio das relações médico-paciente", Rute Teixeira Pedro, *A Responsabilidade Civil do Médico...*, *op. cit.*, pp. 77-78. No mesmo sentido, Acórdão do STJ de 18 de janeiro de 2022 (Processo nº 19473/17.6T8LSB.L1.S1), disponível na base de dados www.dgsi.pt.

dar-se-á através de um conjunto de outros comportamentos que constituirão objeto de outros tantos deveres dos profissionais de saúde, das entidades que prestam empresarialmente esses serviços recorrendo a esses profissionais, dos técnicos que desenharam o *software* ou dos fabricantes, por exemplo. No que respeita aos profissionais médicos, como explicitaremos brevemente na próxima subsecção, parece-nos que as *leges artis* da profissão médica devem ser entendidas de forma atualizada[44] e haverá, portanto, que se exigir do profissional que se mantenha atualizado, que conheça os novos dispositivos, que saiba como deve interagir com eles, que tenha conhecimento dos dados que neles deve introduzir, dos momentos em que deve fazer atualizações e daqueles em que deve abster-se de utilizar determinados dispositivos, por exemplo.

Acresce que a intervenção tuteladora do direito deve situar-se, desde logo, a montante, prevendo-se um conjunto de *standards* de segurança que os dispositivos devem respeitar e sem os quais os mesmos não serão licenciados para a utilização na área de prestação de cuidados de saúde. Nessa medida, o profissional deste setor tem o dever (como uma regra que compõe as suas *legis artis* do tempo moderno) de só recorrer a dispositivos licenciados. Assim, por exemplo, por um lado, deve garantir-se que quem vai interagir com o dispositivo conhece as variáveis que estão a ser tidas em conta (idade, sexo, estatura, peso, raça) no seu processamento para averiguar se elas se adequam à especificidade do caso do paciente; por outro lado, deve também garantir-se que ele se assegura que, no que respeita aos algoritmos empregues em dispositivos a ser utilizados na área da saúde, o objetivo primacial é o da promoção do interesse do paciente[45], nomeadamente o do aproveitamento e potenciação

[44] Veja-se o novo art. L. 4001-3 do Code de la Santé Publique francês que prevê várias obrigações (de informação, de prestação de contas) quando um profissional de saúde pretende usar um sistema de "traitement de données algorithmiques dont l'apprentissage a été réalisé a partir de données massives". Os deveres que impendem sobre aqueles que conceberam o sistema têm como beneficiários os doentes e os médicos. Gaëlle Marti, Lucie Cluzel-Métayer e Samir Merabet, "Droit et intelligence artificielle", *op. cit.*, p. 2362.

[45] Vejam-se as garantias de qualidade e segurança previstas no art. 48º do CDOM relativamente à telemedicina e os correspondentes deveres que impendem sobre os profissionais médicos. Aí se prevê que: "1. O médico só deve utilizar a telemedicina depois de se certificar que a equipa encarregue da sua realização garante um nível de qualidade suficientemente alto, funciona de forma adequada e cumpre com as normas estipuladas. 2. O médico deve dispor de sistemas de suporte e utilizar controlos de qualidade e procedimentos de avaliação para vigiar a precisão e a qualidade da informação recebida e transmitida. 3. O médico só deve utilizar a telemedicina depois de se certificar que o sistema utilizado e os seus utilizadores garantem o segredo médico, nomeadamente através da encriptação de nomes e outros dados identificadores".

de todas as possibilidades de cura, de sobrevivência que ele apresente e não outros objetivos, como o de otimização dos custos[46].

3.2. Do título de imputação de responsabilidade civil ao prestador de cuidados de saúde por danos provocados pela utilização de robots

No âmbito da responsabilidade por prestação de cuidados de saúde, vigora o princípio geral de responsabilidade subjetiva previsto no art. 483º, nº 2, do CC. Na verdade, "só existe obrigação de indemnizar independentemente de culpa nos casos especificados na lei" e, não existindo uma previsão normativa que determine a responsabilidade objetiva na área em análise[47], a responsabilidade dependerá da prática de um ato ilícito e culposo, quer se trate de responsabilidade civil extracontratual (art. 483º do CC), quer se trate de responsabilidade contratual (art. 798º do CC). É conhecida a tradicional dificuldade de prova dos pressupostos de responsabilidade civil (subjetiva) por danos causados na prestação de cuidados de saúde, sabendo-se que ela onera, em princípio, o (doente) lesado, que tem o encargo de demonstrar a verificação dos referidos pressupostos (art. 342º, nº 1, do CC) para que o seu direito à reparação dos danos sofridos seja reconhecido. Vários mecanismos de intervenção *pro damnato* têm sido desenvolvidos, aliás, para responder a este problema que pode redundar num défice de proteção dos pacientes[48]. É assim para a prestação de cuidados de saúde nos moldes tradicionais, sem recurso a *robots* e à IA, e é também assim quando esse recurso exista.

Na verdade, segundo a regra do nº 1 do art. 342º do CC, o doente terá, em regra, de provar que ocorreu a prática, pelo prestador de cuidados de saúde, de um facto ilícito e culposo do qual derivaram, num nexo causal juridicamente, os danos que sofreu e que também tem o ónus de provar. As características – entre outras, a sua complexidade, a sua interconectividade com outros instrumentos e a sua autonomia – dos dispositivos robóticos que incorporem IA podem elevar a dificuldade da tarefa probatória que recai sobre o doente lesado a um nível superior.

[46] No exemplo dado por I. Glenn Cohen, "Informed Consent and Medical Artificial Intelligence: What to Tell the Patient?", *op. cit.*, p. 1445.

[47] Sublinhe-se que estamos a refletir sobre a responsabilidade do prestador de cuidados de saúde, não nos debruçando, por isso, sobre a eventual verificação de responsabilidade do produtor do *robot* (ou das várias componentes que o integram).

[48] Veja-se o nosso trabalho (Rute Teixeira Pedro, *A Responsabilidade Civil do Médico...*, *op. cit.*), em que fazemos o diagnóstico do problema, apresentamos vários remédios de índole processual para o debelar e analisamos criticamente a figura da perda de chance que serve o propósito de lhe responder no que concerne à dificuldade probatória do nexo causal.

Concentremo-nos, neste ponto[49], nos requisitos da ilicitude e da culpa[50], na atuação do prestador de cuidados de saúde. Ora, tratando-se de dispositivos dotados de algum grau de autonomia, cuja atuação é, em larga medida, imprevisível e que o utilizador não controla integralmente, podemos falar, nestes casos, de comportamentos ilícitos e culposos do prestador de cuidados de saúde?

Sabemos que, na responsabilidade contratual[51], a ilicitude da atuação do devedor desses serviços traduzir-se-á no incumprimento de alguma das obrigações, que, para ele, emergem do contrato que celebrou com o doente. Na aferição do desvalor da conduta obrigacional, não se pode olvidar que a relação contratual de prestação de cuidados de saúde deve ser entendida como uma relação obrigacional complexa[52], dela derivando, para além de deveres principais de prestação, também múltiplos deveres laterais ou acessórios de conduta – assentes, desde logo, no princípio da boa-fé (art. 762º, nº 2, do CC) – cuja violação consubstanciará também um incumprimento que poderá fundar uma obrigação indemnizatória pelos danos que dele decorram. Pensamos que a utilização de *robots* para a prestação de cuidados de saúde importará a emergência de um conjunto de deveres laterais[53], destacando-se deveres de informação e esclarecimento do doente (a que já nos referimos na subsecção anterior), e também deveres de proteção do mesmo[54]. Parece-nos, pois, que, dependendo das circunstâncias do caso, pode ser afirmada a existência para

[49] Na subsecção 3.4., deter-nos-emos brevemente no requisito do nexo causal.

[50] Consideraremos aqui estes dois requisitos, sabendo, como já o referimos no nosso trabalho *A Responsabilidade Civil do Médico...*, *op. cit.* (pp. 103 e ss.), que os requisitos da culpa e da ilicitude na área em estudo são difíceis de separar.

[51] Que, em regra, é convocada, como se referiu *supra*, nos casos que estamos a considerar.

[52] Sobre a relação obrigacional complexa, veja-se Jorge Ribeiro de Faria, *Direito das Obrigações*, Vol. I, com Miguel Pestana de Vasconcelos e Rute Teixeira Pedro, atualização e ampliação da 2ª ed., Coimbra, Almedina, 2020, pp. 153 e ss.

[53] Uma tarefa central será precisamente a de definir "relevant standards of care", o que não é uma tarefa "qualitatively new" para os juristas, já que "the adaptation of requirements of care to changed technical, economic or even social circumstances has in fact long since formed part of the jurist's craft". Assim, Ernst Karner, «Liability for robotics: current rules, challenges, and the need for innovative concepts», in Sebastian Lohsee, Reiner Schulze e Dirk Satudenmayer (eds.), *Liability for artificial intelligence and the Internet of things*, Baden-Baden, Nomos, 2019, p. 118. Nesse sentido também, Expert Group on Liability and New Technologies – New Technologies Formation, *Liability for artificial intelligence and other emerging digital technologies*, *op. cit.*, pp. 44 e ss.

[54] Sobre os deveres de proteção, ver, por todos, Manuel Carneiro da Frada, *Contrato e Deveres de Protecção*, Separata do Volume XXXVIII do Suplemento ao Boletim da Faculdade de Direito da Universidade de Coimbra, Coimbra, 1994.

os devedores da prestação de cuidados de saúde de deveres de cuidado na escolha e vigilância do *robot*[55], *de obtenção de informação sobre o funcionamento do mesmo, de atualização do software*, de (verificação da) introdução correta dos dados do paciente, de manutenção da componente física do *robot*, de carregamento/fornecimento energético ininterrupto (durante a intervenção) do aparelho, de preservação da incolumidade do acesso ao domínio eletrónico (por exemplo, mantendo o sigilo de palavras-passe e usando aplicações antivírus e *fire-wall*), de utilização de *robots* certificados para o efeito[56], de recusar a utilização de instrumentos dotados de IA quando se saiba (ou se deva saber) que o algoritmo utilizado não serve o propósito primacial de promoção da saúde ou que esse algoritmo não considera, devidamente, no seu juízo, as características do doente a tratar[57] ou foi atingido por um ciberataque que pode ter afetado o seu funcionamento. Alguns desses deveres podem até merecer a qualificação como obrigações de resultado. Difícil, como sabemos, será, desde logo, provar que houve violação culposa desses deveres e que ela se liga causalmente aos danos sofridos. No que respeita à responsabilidade extracontratual, nos termos do nº 1 do art. 483º do CC, a ilicitude traduzir-se-á, em princípio, na violação de um direito absoluto (em regra, nestes casos, um dos direitos de personalidade), ou na violação de uma norma destinada a proteger interesses do doente, sem que, para tal, lhe seja conferido um correspondente direito subjetivo (norma de proteção)[58].

[55] António Pinto Monteiro dá conta dos deveres que recaem sobre o devedor quanto às suas condutas na escolha, vigilância ou instrução do *robot* e que, quando violados, poderão conduzir à responsabilidade do devedor por danos decorrentes de uma deficiente atuação do *robot*. António Pinto Monteiro, "«Qui facit per alium, facit per se» – será ainda assim na era da robótica?", *Direito e Robótica*, *Número especial de Estudos de Direito do Consumidor*, agosto de 2020, pp. 19 a 20.

[56] A regulamentação do setor tem, necessariamente, de passar pela atuação das entidades reguladoras a quem caiba aprovar a utilização de dispositivos dotados de IA ou de outras tecnologias emergentes. Para um estudo comparativo dos regimes aplicáveis a esse propósito nos EUA e na Europa, veja-se Urs J. Muehlematter, Paola Daniore e Kerstin N. Vokinger, "Approval of Artificial intelligence and machine learning-based medical devices in the USA and Europe (2015-20): a comparative analysis", *The Lancet digital health*, Vol. 3, janeiro de 2021, pp. e195 e ss, disponível em www.thelancet.com/digital-health [acedido, pela última vez, em 30-03-2023].

[57] Não podendo parece-nos exigir-se ao profissional de saúde que conheça em detalhe o funcionamento dos algoritmos. De qualquer modo, parece-nos que se deve afirmar um dever de conhecimento das bases de funcionamento do *robot* e, consequentemente, um dever de rejeição da utilização desses instrumentos quando "lacks an epistemic warrant that the AI/ML is reaching good decisions". I. Glenn Cohen, "Informed Consent and Medical Artificial Intelligence: What to Tell the Patient?", *op. cit.*, p. 1461.

[58] Considere-se o art. 4º, nº 8, do CDOM, que impõe um dever de permanente atualização científica e preparação técnica por parte do médico. Esta previsão normativa, que nos parece dever abranger

Consideremos, agora, o pressuposto da culpa. Também, nesse plano, podem surgir dificuldades para demonstrar a omissão da diligência e competência exigíveis ao prestador dos cuidados de saúde, segundo as circunstâncias do tráfico. A negligência traduzirá, portanto, o desvio da atuação adotada pelo prestador de cuidados, em relação a um modelo de comportamento – em termos de conhecimentos científicos e tecnológicos, de competência técnica, de prudência e de atenção – que ele podia e devia ter observado. A definição do padrão de exigibilidade não pode ignorar o contexto de robotização e o grau de autonomia que os *robots* apresentam, a opacidade do seu sistema operativo e a imprevisibilidade da sua atuação[59].

Não se olvida, aqui, o préstimo que representam para o doente lesado as presunções que o desoneram do encargo de demonstração de factos que sejam constitutivos do seu direito à reparação, cabendo ao prestador de cuidados de saúde ilidi-las. Para além da presunção vigente no âmbito obrigacional (a do art. 799º)[60], podemos, no âmbito extracontratual, convocar, para aqui, as presunções previstas nos dois números do art. 493º e o raciocínio que sustenta a teoria dos deveres de segurança no tráfego[61]. Falamos, por um lado, da presunção que recai sobre quem tiver em seu poder coisa móvel ou animal com o dever de a/o vigiar (art. 493º, nº 1). Pensamos que esta previsão norma-

também os instrumentos dotados de IA e os *robots* de que se sirva, deve também ser, de qualquer modo, considerada para a definição dos deveres de conduta dos médicos e para a interpretação e integração do contrato de prestação de serviços médicos. Considerem-se também os deveres previstos nos arts. 46º a 49º do CDOM no âmbito da telemedicina.

[59] As dificuldades de demonstração de um erro técnico que deva ser qualificado como ilícito e possa ser imputado, a título de culpa, ao profissional podem acentuar a tendência (*supra* referida) de se optar pela via de demonstrar que houve um incumprimento culposo dos deveres de informação. Daí decorrerá que, se os médicos "have adequately discharged all their informations duties", torna-se muito difícil a afirmação da responsabilidade desses profissionais num sistema de responsabilidade "negligence-based", como alerta Jean-Sébastien Borghetti, "How can Artificial Intelligence be defective", in Sebastian Lohsee, Reiner Schulze e Dirk Satudenmayer (eds.), *Liability for artificial intelligence and the Internet of things*, Baden-Baden, Nomos, 2019, p. 75.

[60] Segundo o entendimento que nos parece melhor e defendemos desde 2005, a presunção do art. 799º pode aplicar-se no âmbito da prestação de cuidados de saúde e mesmo relativamente a obrigações que se qualifiquem como obrigações de meios. Rute Teixeira Pedro, *A Responsabilidade Civil do Médico..., op. cit.*, p. 87.

[61] Esta teoria assenta no princípio geral, desenvolvido na Alemanha e aceite no nosso ordenamento, segundo o qual "aquele que no tráfego cria ou mantém uma fonte de perigo é obrigado a tomar as medidas necessárias para afastar esse perigo". Ribeiro de Faria, *Direito das Obrigações*, *op. cit*. p. 425. Para um maior desenvolvimento sobre a figura, Sinde Monteiro, *Responsabilidade por Conselhos, Recomendações ou Informações*, Coimbra, Almedina, 1989, pp. 307 e ss, e Menezes Cordeiro, *Da Boa-fé no Direito Civil*, Coimbra, Almedina, 2001, pp. 832 e ss.

tiva poderá, aqui, ser aplicada[62], na medida em que se aceite, como nos parece defensável, que quem toma a decisão de empregar um *robot* médico ou usar um dipositivo dotado de IA deverá assumir o encargo de vigiar o seu funcionamento. Assim sendo, o vigilante responderá pelos danos que a coisa dotada de IA causar, salvo se provar que nenhuma culpa houve da sua parte ou que os danos se teriam igualmente produzido ainda que não houvesse culpa sua. O sucesso da utilização desta presunção pelo doente pode ser denegado, a montante, porque se rejeite com base na autonomia da coisa (e opacidade do seu funcionamento) a afirmação do dever de vigilância, ou, aceitando-se a existência de um dever de vigilância (como nos parece que deva afirmar-se), ser comprometido, a jusante, pela facilidade de demonstração de que o prestador de saúde agiu sem culpa[63].

Por outro lado, falamos também da presunção que recai sobre quem cause danos a outrem no exercício de uma atividade perigosa, por sua própria natureza ou pela natureza dos meios utilizados. Sabemos que, em geral, a atividade de prestação de cuidados de saúde não é qualificada como perigosa. No entanto, em certos casos, por força da especial perigosidade dos instrumentos utilizados, entende-se que está presente um perigo qualificado que justifica a subsunção da situação ao nº 2 do art. 493º. Como exemplos paradigmáticos apareciam, tradicionalmente, os da utilização de um bisturi elétrico ou de uma incubadora com termóstato estabilizador da temperatura e o do tratamento médico com ondas curtas ou raios-x ou de um tratamento dentário

[62] Admitindo a aplicação, veja-se Filipe Albuquerque Matos, destacando o esforço de "interpretação atualista" (p. 201), "Responsabilidade por danos causados a terceiros por robôs", *op. cit.*, pp. 199 e ss., e Mafalda Miranda Barbosa, *Inteligência Artificial. Entre a Utopia e a distopia, alguns problemas jurídicos*, *op. cit.*, pp. 78 e ss e Nuno Sousa e silva, "Direito e Robótica: uma primeira aproximação", *op. cit.*, p. 522. Também Henrique Sousa Antunes, rejeitando a aplicação do art. 491º, admite a aplicação do nº 1 do art. 493º, na medida em que "enquanto ao robô faltar personalidade jurídica, a sua natureza confundir-se-á com uma das duas realidades enunciadas no artigo (coisa ou animal).", "Inteligência Artificial e Responsabilidade Civil: Enquadramento", *op. cit.*, p. 147.

[63] Em sentido que nos parece oposto, Filipe Albuquerque Matos, "Responsabilidade por danos causados a terceiros por robôs", *op. cit.*, pp. 201 e 202. A nossa afirmação feita em texto explica-se pelo facto de entendermos que, se o nível de controlo do utilizador do *robot* é baixo (porque a sua autonomia é muito elevada), os deveres que sobre esse utilizador impenderão serão menos numerosos e mais ténues, podendo, por esse facto, ser mais simples o seu cumprimento e a demonstração desse cumprimento. Como afirma Ernst Karner, provando-se a falta de controlo sobre o dispositivo, perde-se o fundamento para a afirmação da culpa do utilizador do mesmo – Ernst Karner, "Liability for robotics: current rules, challenges, and the need for innovative concepts", *op. cit.*, p. 120.

com broca[64]. Ora, parece-nos que a utilização de *robots* ou outros dispositivos dotados de IA, pelos riscos que se lhes podem associar, até por força da imprevisibilidade que a autoaprendizagem do sistema operativo pode gerar, poderá permitir afirmar a presença de uma perigosidade de nível superior[65] que justifica a aplicação do nº 2 do art. 493º[66]. O prestador de cuidados de saúde será, então, obrigado a reparar os danos produzidos no exercício dessa atividade, exceto se mostrar que empregou todas as providências exigidas pelas circunstâncias com o fim de prevenir a produção desses danos. Ora, mais uma vez, cremos que a responsabilidade do prestador de cuidados de saúde pode ser afastada com alguma facilidade, demonstrando este sujeito que cumpriu os deveres que deixámos acima enunciados[67].

Ora, parece-nos que o cenário que procurámos descrever aponta para uma acentuada dificuldade do paciente lesado em obter proteção ressarcitória. Pensamos, pois, que se justificará uma intervenção legislativa que esclarecesse, se não mesmo sectorialmente, pelo menos, genericamente, que, em caso de utilização de *robots* ou outros dispositivos dotados de IA, o utilizador

[64] O primeiro exemplo era dado por Figueiredo Dias e Sinde Monteiro, "Responsabilidade Médica em Portugal, *Boletim do Ministério da Justiça*, nº 332, janeiro de 1984, p. 53. Os dois últimos exemplos são apresentados por Antunes Varela, *Das Obrigações em Geral*, Vol. I, 10ª ed., Coimbra, Almedina, 2004, p. 595.

[65] O que, naturalmente, pressuporá uma análise casuística para averiguar se concretiza a exigência qualificada de perigosidade especial demandada pelo art. 493º, nº 2, do CC. Na verdade, não se poderá afirmar que todos os dispositivos dotados de IA aumentem, necessariamente, em todos os domínios o grau de perigo de exercício de uma atividade. Nesse sentido, Filipe Albuquerque Matos, "Responsabilidade por danos causados a terceiros por robôs", op. cit., p. 198, Mafalda Miranda Barbosa, *Inteligência Artificial. Entre a Utopia e a distopia, alguns problemas jurídicos*, *op. cit.*, p. 89 e Nuno Sousa e Silva, "Direito e Robótica: uma primeira aproximação", *op. cit.*, p. 521. Depois de considerar essa objeção, Henrique Sousa Antunes conclui que "a perigosidade deve, também, ser aferida pelo grau de envolvimento da atividade com os bens pessoais que serve. Quanto maior for a proximidade da conduta, nomeadamente pela sua reiteração, a bens existenciais, maior a probabilidade de um dano grave. E isso determina a sua perigosidade.", Henrique Sousa Antunes, "Inteligência Artificial e Responsabilidade Civil: Enquadramento", *op. cit.*, p. 146. Note-se também que a interconexão entre produtos e serviços informáticos pode ser fonte de riscos sistémicos que não podem ser desconsiderados. Gerald Spindler, "User liability and strict liability in the internet of things and for robots", in Sebastian Lohsee, Reiner Schulze e Dirk Satudenmayer (eds.), *Liability for artificial intelligence and the Internet of things*, Baden-Baden, Nomos, 2019, p. 127 (pp. 125-143).

[66] Admitindo a aplicação, veja-se Filipe Albuquerque Matos, "Responsabilidade por danos causados a terceiros por robôs", op. cit., p. 198.

[67] No sentido em que um maior grau de autonomia pode redundar num menor número e intensidade de deveres que recaem sobre o utilizador e que, nessa medida, ele mais facilmente cumprirá e provará que cumpriu.

responderia objetivamente independentemente de culpa pelos danos produzidos[68]. Essa é, aliás, a solução proposta na Resolução do Parlamento Europeu, de 20 de outubro de 2020[69], que contém recomendações à Comissão sobre o regime de responsabilidade civil aplicável à inteligência artificial, quando se trate de sistema de inteligência artificial de alto risco, como serão os sistemas usados na área que nos ocupa[70].

3.3. Da multiplicidade de intervenientes e da dificuldade de identificação da(s) pessoa(s) a que se possa imputar a responsabilidade pela atuação robótica na área da saúde

Um dos maiores problemas colocados ao funcionamento da responsabilidade civil pela utilização de dispositivos dotados de inteligência artificial, nomeadamente *robots*, é o de haver uma multiplicidade de pessoas cuja intervenção é necessária para que a operação de tais dispositivos seja levada a cabo e que se apresentam como potenciais responsáveis pelos danos que resultem da sua atuação.

Para além do utilizador e manuseador do *robot* – cuja posição jurídica e termos de eventual responsabilização considerámos como elemento central desta nossa reflexão –, não pode esquecer-se a intervenção de muitos outros sujeitos que se apresentam em posições jurídicas variadas. Pensemos, por exemplo e só para nomear alguns, naqueles que fabricam o objeto físico que constitui o corpo do *robot*, ou naqueles que desenham o *software* incorporado

[68] Defendendo a previsão de um regime de responsabilidade objetiva, veja-se também Filipe Albuquerque Matos, "Responsabilidade por danos causados a terceiros por robôs", *op. cit.*, pp. 203 e ss., e Mafalda Miranda Barbosa, *Inteligência Artificial. Entre a Utopia e a distopia, alguns problemas jurídicos*, *op. cit.*, pp. 97 e ss.

[69] Disponível em https://www.europarl.europa.eu/doceo/document/TA-9-2020-0276_PT.html#title3.

[70] A responsabilidade objetiva do operador de sistemas de inteligência artificial de alto risco encontra-se prevista e regulada nos arts. 4º a 7º da proposta. Esta proposta adota um modelo dualista, prevendo para os outros sistemas de inteligência artificial uma responsabilidade subjetiva com presunção de culpa, como resulta dos arts. 8º e 9º. A Comissão Europeia, nas soluções jurídicas que incluiu na proposta de Diretiva relativa à adaptação das regras de responsabilidade civil extracontratual apresentada no dia 28 de setembro de 2022 (disponível no endereço https://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=CELEX:52022PC0496&from=EN), trilhou um caminho diverso, não contemplando, nessa sua proposta, a consagração de responsabilidade objetiva no contexto de utilização de IA. A Comissão optou por uma intervenção "minimamente invasiva", à luz do princípio da proporcionalidade, propondo apenas "medidas relativas ao ónus da prova", pelo menos numa primeira fase. Admite, no entanto, que, numa segunda fase, depois de se avaliar o efeito da solução inicialmente proposta, a previsão de responsabilidade objetiva se possa revelar "mais adequada". (Exposição de motivos, Proposta de Diretiva, p. 7).

no dispositivo, ou naqueles que configuram as atualizações que são introduzidas no sistema (e que podem não ser os mesmos que desenharam a versão inicial do *software*), ou naqueles a quem o utilizador comprou esses bens, ou também naqueles que transmitem as informações que alimentam o sistema e a partir de cujos *inputs* o dispositivo autoaprende e se desenvolve, ou ainda naqueles terceiros (*hackers*) que podem, indevidamente, ter obtido acesso ao sistema operativo, nele introduzindo alterações, ou, finalmente, naqueles que desempenharam algumas dessas funções ou praticaram alguns desses atos relativamente a outros dispositivos com os quais o *robot* está interconectado e interage e que podem constituir a causa da sua intervenção lesiva. Nesta enunciação, que não pretende ser exaustiva, deixámos propositadamente até agora no silêncio um outro interveniente que pode ser identificado: falamos do próprio *robot*, se ele atingir um grau total de autonomia e se, consequentemente, se tomar a decisão de lhe reconhecer personalidade (eletrónica).

A verdade é que a profusão de pessoas a quem pode ser assacada, potencialmente, responsabilidade – desconsiderando agora a autonomização, para esse efeito, do robot –, a complexidade do dispositivo e a opacidade que o seu funcionamento pode revestir dificultam a identificação das pessoas a quem, em concreto, pode ser atribuída responsabilidade civil pelos danos ocorridos. Pensamos, desde logo, na dificuldade de afirmação – *rectius*, de demonstração – da prática de um ato ilícito e culposo por parte de alguma (ou várias) daquelas pessoas, ou mesmo, se considerarmos o regime de responsabilização do produtor, na dificuldade de sequer afirmar – *rectius*, demonstrar – a existência de um defeito do dispositivo. Na verdade, não poderá derivar-se da verificação de um resultado danoso – e, portanto, indesejado (por exemplo, a morte do paciente, a não verificação da cura do mesmo) – a afirmação da existência de um defeito[71] ou a atuação ilícita e culposa de alguma daquelas pessoas acima referidas. Ora, faltando a demonstração dos requisitos de que depende a afirmação da responsabilidade de alguma daquelas pessoas, o resultado será o de o paciente lesado não conseguir obter a reparação dos danos sofridos. A multiplicidade de intervenientes e a dificuldade de fazer imputar juridicamente responsabilidade a (pelo menos) um deles podem redundar na desproteção indevida do paciente, que é um resultado que deve ser

[71] Se um algoritmo desenhado para fazer diagnósticos erra no diagnóstico, não pode concluir-se que se está perante "*a result of a defective design*", como destaca Jean-Sébastien Borghetti, "How can Artificial Intelligence be defective", *op. cit.*, p. 67 (pp. 63-76). O Autor apresenta, depois, padrões alternativos para aferir do defeito do algoritmo (pp. 68 a 71).

combatido[72]. Os riscos da utilização dos dispositivos em análise não deverão recair, necessariamente, apenas sobre aquele que recorreu aos serviços de outrem que usa, para o exercício da sua atividade – potenciando a sua capacidade de atuação e de obtenção de vantagens económicas –, *robots* e outros dispositivos dotados de IA concebidos e construídos por outros agentes económicos que também se aproveitam economicamente dessa atividade[73]. Tudo agentes que podem até estar mais bem posicionados para promover a segurança e fiabilidade dos *robots* e a minimização dos riscos que a eles se associam. Parece, aliás, para que se encontre uma solução equilibrada – sobretudo na área da saúde em que o aproveitamento das vantagens dos *robots* e outros instrumento com IA a todos beneficiará –, que se deverá buscar um regime em que todos suportem um quinhão das desvantagens associadas aos eventos desvantajosos em que um desses instrumentos é utilizado, seja sob a forma de assunção de responsabilidade civil (independentemente de culpa) pelos danos verificados, pagando uma indemnização, seja sob a forma de assunção do encargo de contratação (obrigatória) de um seguro, pagando o respetivo prémio[74].

As dificuldades que vimos assinalando de fazer funcionar de forma efetiva, eficiente e juridicamente justa (distribuindo devidamente os encargos inerentes aos eventos lesivos) é, aliás, o argumento mais forte para se defender o artifício do reconhecimento de personalidade eletrónica aos *robots* ou a outros dispositivos com IA, dotando-os de um lastro financeiro que alimentaria o pagamento de indemnizações quando a responsabilidade dos mesmos fosse afirmada[75].

[72] O que se pode denominar como o problema da "responsibilty gap". Erica Palmerini, "Robotica e diritto: suggestioni, intersezioni, sviluppi a margine di una ricerca europea", *op. cit.*, p. 1835.

[73] As soluções jurídicas devem, aliás, promover soluções que impeçam a diluição de responsabilidade, como destaca a OMS: "To avoid diffusion of responsibility, in which «everybody's problem becomes nobody's responsibility», a faultless responsibility model («collective responsibility»), in which all the agents involved in the development and deployment of an AI technology are held responsible, can encourage all actors to act with integrity and minimize harm. In such a model, the actual intentions of each agent (or actor) or their ability to control an outcome are not considered.", in *Ethics and governance of artificial intelligence for health: WHO guidance*, *op. cit.*, p. 28.

[74] A determinação da obrigatoriedade de contratação de um seguro de responsabilidade civil encontra-se prevista na solução proposta na Resolução do Parlamento Europeu, de 20 de outubro de 2020 (veja-se art. 4º, nº 4).

[75] Nesse sentido também, entre outros, Ugo Salanitro, "Intelligenza artificiale e responsabilità: la strategia della Commissione Europea", *op. cit.*, p. 1250. Gerhard Wagner, quanto à opção de atribuição de personalidade eletrónica, ainda que reconheça que pode resolver problemas

De qualquer modo, como procurámos deixar explicado *supra* (na subsecção 3.2), a autonomia que seja reconhecida aos *robots* não esvazia a responsabilidade que pode ser imputada a quem os introduz no tráfego jurídico, nomeadamente no exercício de atividades de prestação de cuidados de saúde. E parece-nos que poderá ser assim mesmo num cenário (eventual e de verificação incerta) de existência de autonomia plena desses dispositivos e de (eventual) consequente atribuição de reconhecimento de personalidade (eletrónica) aos mesmos.

Na verdade, se considerarmos o âmbito da responsabilidade contratual, a responsabilidade do devedor ultrapassa a responsabilidade por atos próprios. Assim, como resulta do art. 800º, o devedor é responsável perante o credor pelos atos das pessoas que utilize para o cumprimento da obrigação, como se tais atos fossem praticados pelo próprio devedor[76]. Ora decisivo é, portanto, saber quem é o devedor da prestação em cujo cumprimento se deu a produção dos danos para se identificar a pessoa sobre a qual recairá o dever de indemnizar[77]. A operação de identificação da pessoa com quem o doente celebrou

probatórios, considera duvidoso que se justifique a criação de uma nova entidade jurídica para o efeito, Gerhard Wagner, «Robot Liability», in Sebastian Lohsee, Reiner Schulze e Dirk Satudenmayer (eds.), *Liability for artificial intelligence and the Internet of things*, Baden-Baden, Nomos, 2019, pp. 60 e 61 (pp. 27-62). No sentido da preferência pela previsão de responsabilidade objetiva do "system's keeper" e da obrigatoriedade de contratação de um seguro de responsabilidade civil em detrimento do reconhecimento de personalidade eletrónica aos *robots*, Ernst Karner, "Liability for robotics: current rules, challenges, and the need for innovative concepts", *op. cit.*, p. 123. Contra este expediente técnico se pronunciam Filipe Albuquerque Matos, "Responsabilidade por danos causados a terceiros por robôs", *op. cit.*, em especial, pp. 169 e ss., e Mafalda Miranda Barbosa, "Inteligência artificial, e-persons e direito: desafios e perspetivas", *Direito e Robótica*, *Número especial de Estudos de Direito do Consumidor*, agosto de 2020, em especial, pp. 58 e ss.

[76] Note-se que, não sendo uma questão que mereça uma resposta pacífica, acompanhamos o entendimento segundo o qual a responsabilidade do devedor nos termos do art. 800º abrange não só o incumprimento de deveres de prestação, mas também a violação de deveres laterais de conduta, nomeadamente os deveres de proteção. Julgamos que assim deve ser pois é a própria "ocasião" de cumprimento que propicia a produção do dano. Se não fosse a relação obrigacional, os auxiliares não teriam tido a possibilidade de atuar negativamente sobre a esfera jurídica do credor ou pelo menos não teriam naquelas circunstâncias (o doente numa ocasião de vulnerabilidade).

[77] É, portanto, tarefa decisiva esta de identificação da pessoa com quem o doente celebrou o contrato ou as partes com quem celebrou os contratos, pois podem ser celebrados vários. Pode o doente celebrar vários contratos, desde logo com vários médicos (a contratação do cirurgião e do anestesista pode fazer-se por figuras contratuais autónomas). Pode celebrar contrato com um médico e outro relativo ao internamento com uma clínica em que vai ser feita a intervenção cirúrgica. Pode celebrar um único contrato (denominado contrato total), por exemplo, com a clínica, em cujo objeto caberá a prestação de cuidados de saúde, nomeadamente de serviços médicos.

o contrato e a delimitação do programa obrigacional permitirá saber por que atos o devedor responde nos termos do art. 800º do CC. Cabendo na previsão normativa deste preceito todas as "pessoas que o devedor utilize para o cumprimento da obrigação", parece que serão abrangidos todos aqueles que tenham sido introduzidos pelo devedor na realização da prestação debitória, não sendo determinante, para o efeito, nem a existência de uma relação de dependência do auxiliar perante o devedor[78], nem da existência de autonomia desse auxiliar face ao mesmo devedor. O devedor responde pelos atos dos seus auxiliares quer eles sejam dependentes, quer eles sejam independentes[79]. Ora, parece-nos que o resultado a que se deveria chegar seria aquele segundo o qual o devedor responderia também por uma falha no cumprimento que se ficasse a dever à atuação do *robot*[80]. *Contra essa solução parece-nos, no entanto, depor a dificuldade – que se nos afigura no momento presente insuperável – de não se conseguir afirmar que, se fosse o devedor a praticar o ato adotado pelo robot*, o devedor também responderia, na medida em que essa projeção se faz com base num juízo de culpa[81] – que se apresenta impossível de formular quanto a um *robot*.

[78] Como parece ser exigido pelo art. 500º ao demandar-se a existência de uma relação de comissão entre comitente e comissário.

[79] Note-se que o art. 800º não pressupõe uma "prévia imputação do facto (danoso)" ao auxiliar do devedor, antes se operando uma *projeção* do "comportamento do auxiliar na pessoa do devedor", o que "possibilita um alargamento da zona de responsabilidade e da tutela do lesado", nas palavras Manuel Carneiro da Frada, *Contrato e Deveres de Protecção*, *op. cit.*, pp. 209 e 211. O Autor acrescenta que "A mobilização de terceiros e a sua introdução no programa obrigacional onerá-lo-á [ao devedor] deste modo como risco da sua actividade", *idem*, p. 212.

[80] Apesar de afirmar que "Cada um é que sabe com que meios pode ou deve cumprir (...). Uma deficiente actuação do robô corre por conta e risco de quem o utiliza, como sucederia se essa deficiente actuação ficasse a dever-se a qualquer problema do sue sistema informático ou de outros meios utilizados por esse contraente", António Pinto Monteiro encontra um obstáculo (entre outros) à aplicação do art. 800º na utilização da palavra "pessoa", o que impediria, segundo o Autor (que ressalva apenas uma "analogia, muito generosa"), a aplicação do art. 800º quando o auxiliar fosse um *robot* autónomo, a menos que se lhe reconhecesse personalidade eletrónica – António Pinto Monteiro, "«Qui facit per alium, facit per se» – será ainda assim na era da robótica?", *op. cit.*, pp. 16 a 19. Não nos parece que este obstáculo seja decisivo para a não aplicação do art. 800º aos *robots* autónomos.

[81] Como se poderia afirmar a culpa de um *robot*? Este problema é assinalado por António Pinto Monteiro, "«Qui facit per alium, facit per se» – será ainda assim na era da robótica?", *op. cit.*, p. 20. Henrique Sousa Antunes admite a possibilidade de adaptação do conceito de imputabilidade à atuação dos *robots*, considerando, para esse efeito, "justificada a revisão da referência ao bom pai de família", como "juízo técnico que determinará o comportamento exigível". Afirma, pois que, "em razão da sofisticação das capacidades identificadas e da subordinação da inteligência artificial aos interesses humanos, ao padrão da conduta exigível ao homem médio deve substituir-se a referência

Mais difícil ainda será uma resposta no âmbito delitual em que a imputação a alguém (comitente) de atos de terceiro (seu comissário) depende da verificação dos requisitos do art. 500º do CC. Ora, as dificuldades advirão de a aplicação do art. 500º pressupor a existência de um poder de direção sobre a outra pessoa (o que será difícil de afirmar no caso dos *robots* totalmente autónomos, em que o prestador de cuidados de saúde não tem possibilidade de os dirigir ou controlar) e de demandar uma imputação autónoma de responsabilidade ao comissário, no caso, ao *robot*[82]. *Mais uma vez somos confrontados com a pergunta: como se configura a prática de um ato ilícito e culposo pelo robot*, não se vislumbrando que possa incorrer na prática de atos geradores de responsabilidade objetiva, na falta de previsão normativa nesse sentido?

Detetadas as dificuldades no plano do direito constituído, acompanhamos, no entanto, o entendimento de que a solução a aplicar, quando se recorre à atuação de *robots* e o auxílio prestado por estes dispositivos é "functionally comparable to human labour"[83], deve proporcionar um nível de proteção idêntica à que existe quando há recurso aos auxiliares humanos.

3.4. Da potenciação das interferências causais: a robotização e a acentuação da multicausalidade na prestação de cuidados de saúde

Finalmente, neste périplo muito rápido sobre os desafios colocados à afirmação da responsabilidade civil do prestador de cuidados de saúde quando há utilização de *robots*, não podemos deixar de fazer breves observações sobre as dificuldades que se encontrarão na operação de aferição do nexo causal entre o ato que se considera fundante da responsabilidade civil e os danos verificados[84]. Para além de a afirmação da existência dessa conexão causal,

ao melhor comportamento possível do robô nas circunstâncias consideradas, Henrique Sousa Antunes, "Inteligência Artificial e Responsabilidade Civil: Enquadramento", *op. cit.*, p. 153.

[82] Referindo-se à exigência de dupla imputação, Manuel Carneiro da Frada, *Contrato e Deveres de Protecção*, *op. cit.*, pp. 205 e ss.

[83] "No one should be able to exclude the attribution provided for by vicarious liability provisions simply by employing technical means of support instead of human helpers", como afirma Ernst Karner, "Liability for robotics: current rules, challenges, and the need for innovative concepts", *op. cit.*, p. 120. No direito positivo austríaco, já se entende defensável essa solução *de iure constituto* – *Idem*, *ibidem*. No sentido de que essa deve ser a solução para o problema referido, também o Expert Group on Liability and New Technologies – New Technologies Formation, *Liability for artificial intelligence and other emerging digital technologies*, *op. cit.*, pp. 45 e 46.

[84] Esta dificuldade é também assinalada no Expert Group on Liability and New Technologies – New Technologies Formation, *Liability for artificial intelligence and other emerging digital technologies*, *op. cit.*, p. 20. Igualmente, a Comissão destaca a possível dificuldade acrescida que pode verificar-se quanto à prova do nexo causal (e da culpa), quando descreve as

em termos juridicamente relevantes, ser necessária para que se reconheça o direito à reparação do lesado, será também por referência ao nexo causal afirmado que se delimitará o âmbito de danos ressarcíveis e, portanto, o *quantum* reparatório[85].

A dificuldade na afirmação do nexo causal não é um problema novo no âmbito do funcionamento da responsabilidade civil por danos causados na prestação de cuidados de saúde[86]. Como já tivemos oportunidade de analisar, nesta área de atividade, acentua-se o fenómeno de "concorrência de interferências causais"[87]. A atuação do prestador de cuidados de saúde dá-se num contexto em que intervêm outros elementos potenciais causadores do dano – a própria patologia e a sua evolução natural, outras caraterísticas do paciente pré-existentes ou adquiridas com impacto no desenrolar dos acontecimentos, a atuação de outras pessoas que integram a equipa de profissionais e técnicos que intervêm na prestação do cuidado. Do concurso de fatores que se apresentam com potencialidade geradora dos danos em causa resulta a dificuldade da afirmação de que algum deles constitui condição necessária da produção dos danos verificados.

A utilização de um dispositivo de IA, nomeadamente de um *robot*, acrescenta mais um segmento causal à multicausalidade do desenrolar dos acontecimentos[88]. Dessa circunstância se pode concluir que a complexidade tradicional pode acentuar-se, sobretudo quando estão em causa dispositivos cujo comportamento é caracterizado pela opacidade. Será o caso daqueles que estão dotados de capacidade de *deep learning*, que não permite, pois, a recons-

razões e objetivos da proposta de Diretiva relativa à adaptação das regras de responsabilidade civil extracontratual à inteligência artificial apresentada a 28 de setembro de 2022. In *Exposição de Motivos da Proposta*, p. 2.

[85] Pelo menos num primeiro momento, já que podem intervir outras causas limitadoras do *quantum* reparatório, desde a que resulta da aplicação do art. 494º ou do art. 570º, até limites decorrentes da fixação de montantes máximos para o montante indemnizatório, como se prevê no art. 5º da Proposta de Regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial.

[86] Jean-Sébastien Borghetti destaca a dificuldade tradicional de aferição do nexo causal "when medical treatments or pharmaceutical are at stake", não sendo, portanto, um problema específico da utilização de *robots*. Jean-Sébastien Borghetti, "How can Artificial Intelligence be defective", *op. cit.*, p. 75.

[87] Sobre a problemática, veja-se nosso Rute Teixeira Pedro, *A Responsabilidade Civil do Médico...*, *op. cit.*, pp. 160 e ss.

[88] A transformação da sociedade "de riscos monocausais a uma sociedade de riscos multicausais" e a mutação "da explicabilidade à inevitabilidade da opacidade" são duas das dez mudanças produzidas pelas tecnologias digitais emergentes, no elenco apresentado por Henrique Sousa Antunes, *Direito e Inteligência Artificial*, *op. cit.*, pp. 32 e ss. e pp. 38 e ss., respetivamente.

trução do *iter* da sua atuação (nem sequer pelo próprio prestador de cuidados de saúde a quem faltarão conhecimentos sobre os exatos termos de funcionamento do dispositivo e a previsão do modo como evoluirá o comportamento que o *robot* assumirá). Tratar-se-á, então, de mais uma das situações em que a afirmação do nexo causal com o grau de certeza que legalmente é exigido constitui uma tarefa de dificuldade inultrapassável para o doente lesado e que conduziram a que se ensaiassem mecanismos de intervenção *pro damnato*: quer numa perspetiva de direito constituído – onde avultam mecanismos que operam no âmbito processual, no plano da atividade probatória ou da apreciação judicial do resultado dessa atividade, e mecanismos que operam no plano substancial, nomeadamente o que reconhece relevância ressarcitória à perda de chance[89] –, quer numa perspetiva de direito a constituir – defendendo-se a revisão do regime aplicável, ultrapassando o modelo tradicional de funcionamento de responsabilidade civil[90]. Não nos podemos demorar mais neste ponto, mas ele consubstancia uma das dificuldades que no mundo de hoje, tecnologicamente muito avançado, se intensificarão.

Não pode, no entanto, deixar de se dizer que, paradoxalmente com o que vimos de afirmar, a circunstância de certos *robots* ou dispositivos apresentarem uma *black box* (quando ela não se caracteriza pela opacidade) poderá facilitar a tarefa probatória ao doente lesado, nomeadamente quanto ao nexo casual. Na verdade, podendo aceder-se à informação aí registada, torna-se possível a reconstrução do conjunto de acontecimentos ocorridos, permitindo estabelecer, com clareza, em alguns casos, nexos causais juridicamente relevantes e, noutros casos, excluir, com igual clareza, o liame causal[91].

4. Observações conclusivas

O recurso a dispositivos que incorporam as mais recentes inovações tecnológicas, nomeadamente as que respeitam à inteligência artificial e à robótica, manifesta-se, já há algum tempo, na área dos cuidados de saúde, tendo-se intensificado no contexto pandémico. Trata-se de uma tendência que apresenta vantagens inequívocas para os pacientes que carecem desses cuidados

[89] Veja-se o nosso Rute Teixeira Pedro, *A Responsabilidade Civil do Médico...*, *op. cit.*, pp. 327 e ss.

[90] Sobre sistemas alternativos de compensação dos danos causados pela atividade médica, com a apresentação de propostas de reforma para a área, André Gonçalo Dias Pereira, *Direitos dos pacientes e responsabilidade médica*, *op. cit.*, pp. 727 e ss. e pp. 763 e ss.

[91] Por força deste maior acesso ao conhecimento do desenrolar de acontecimentos, também se atenuará a tradicional assimetria que se considera existir na relação entre o prestador de cuidados de saúde e o doente leigo na matéria. Antonio Oliva *et alii*, "Management of Medico-Legal Risks in Digital Health Era: A Scoping Review", *op. cit.*, p. 4.

e ganhos evidentes na qualidade e eficiência do desempenho da atividade nesse setor, o que representa também um significativo benefício social para a comunidade.

A evolução em curso tende a intensificar-se e, com essa intensificação, multiplicar-se-ão, inevitavelmente, os eventos lesivos no âmbito de uma prestação de cuidados de saúde em que se verificou a utilização de um desses dispositivos. Trata-se de constelações fácticas que, como vimos, o regime jurídico vigente da responsabilidade civil não deixa sem resposta, nomeadamente no que respeita à aferição da eventual responsabilidade do prestador de cuidados de saúde. Como deixámos descrito nas páginas precedentes, encontramos no direito constituído soluções que podem ser aplicadas aos problemas que emergem desta nova fenomenologia e que, em larga medida, coincidem com questões que já se colocam no mesmo âmbito relativamente à prestação de cuidados de saúde nos moldes tradicionais. As características dos *robots* e em particular da IA de que os mesmos são *animados* (a sua complexidade, a sua interconectividade com outros dispositivos e com o meio ambiente e, muito especialmente, a sua crescente autonomia e capacidade de produzir decisões assentes num processo de autoaprendizagem) intensificam as já conhecidas dificuldades de funcionamento da responsabilidade civil no setor da área de prestação de cuidados de saúde.

Procurámos demonstrar que as respostas que podem equacionar-se no plano de direito constituído concitam dúvidas e são de difícil efetivação prática, podendo conduzir a um défice de proteção do paciente lesado. Por isso, tal como aconteceu por ocasião de outras revoluções científicas, o regime de responsabilidade civil pode vir a sofrer uma transformação. Importa, pois, uma revisão legislativa clarificadora[92] que ofereça uma solução integrada que, simultaneamente, dê guarida ao objetivo de assegurar uma suficiente proteção dos doentes e ao objetivo de promoção da evolução tecnológica em curso[93],

[92] A promoção da certeza e segurança jurídicas deverá ser um objetivo a ser prosseguido com as intervenções legislativas que venham a ocorrer, como reconhece a Comissão na Exposição de motivos da Diretiva relativa à adaptação das regras de responsabilidade civil extracontratual à inteligência artificial apresentada a 28 de setembro de 2022. Precisamente por os ordenamentos jurídicos predisporem mecanismos que os Tribunais podem utilizar ("adaptar pontualmente o modo como aplicam as regras em vigor"), quando confrontados com pedidos em que as características específicas da IA tenham um relevo significativo, antecipa-se a acentuação da insegurança jurídica. In *Exposição de Motivos da Proposta de Diretiva*, p. 2.

[93] O que está em linha com a Proposta de Regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial apresentada em 21 de abril de 2021 (é "do interesse da União preservar a liderança tecnológica da UE e assegurar que novas tecnologias, desenvolvidas e exploradas respeitando os valores, os direitos fundamentais e os

distribuindo equilibradamente as consequências desvantajosas da utilização dos *robots* pelos vários sujeitos que intervêm e beneficiam da sua colocação no tráfego jurídico (nomeadamente, os produtores, os distribuidores e os utilizadores)[94]. Antevê-se, pois, que as alterações que se avizinham passem, por um lado, pela consagração de novas áreas de funcionamento de responsabilidade objetiva[95], por outro lado, pela previsão de presunções que invertam o ónus da prova dos requisitos da responsabilidade civil (nomeadamente da culpa e também de nexo causal[96]) precipuamente dirigidas à aplicação quanto à atuação dos operadores dos dispositivos em análise e, finalmente, pela determinação da obrigatoriedade de contratação de seguros de responsabilidade civil para cobrir os danos produzidos por eles.

Porto, 4 de janeiro de 2023.

Bibliografia

Antunes, Henrique Sousa, "Inteligência Artificial e Responsabilidade Civil: Enquadramento", *Revista de Direito da Responsabilidade Civil*, Ano I, 2019, pp. 139 e 140 (pp. 139-154), acessível em https://revistadireitoresponsabilidade.pt/2019/inteligencia-artificial-e-responsabilidade-civil-enquadramento/. [acedido, pela última vez, em 30/03/2023]

Antunes, Henrique Sousa, *Direito e Inteligência Artificial*, Lisboa, Universidade Católica Portuguesa, 2020.

princípios da União, estejam ao serviço dos cidadãos europeus", in *Exposição de motivos da Proposta de Regulamento*, p. 1). Também as soluções vertidas na Proposta de Diretiva relativa à adaptação das regras de responsabilidade civil extracontratual à inteligência artificial apresentada pela Comissão Europeia visam "assegurar que as pessoas que pedem uma indemnização por danos que lhes sejam causados por um sistema de IA gozam de um nível de proteção equivalente ao nível de que gozam as pessoas que pedem uma indemnização por danos causados sem o envolvimento de um sistema de IA", como se lê no considerando nº 7 da proposta, p. 19.

[94] Importante é maximizar as vantagens comunitárias, minimizando, do mesmo passo, o risco de produção de danos, na síntese de Gerhard Wagner, "Robot Liability", *op. cit.*, p. 30.

[95] Como se avança na Resolução do Parlamento Europeu, de 20 de outubro de 2020, mas que a Comissão entendeu não seguir na sua Proposta de Diretiva. Veja-se *supra* nota 70.

[96] Esse é o modelo de intervenção seguido na Proposta de Diretiva relativa à adaptação das regras de responsabilidade civil extracontratual à inteligência artificial apresentada pela Comissão Europeia em que consagra uma presunção (ilidível) de culpa e uma presunção (ilidível) de nexo causal aplicável em determinadas situações previstas, respetivamente, no art. 3º, nº 5 e no art. 4º da Proposta de Diretiva. Para uma primeira apreciação das soluções acolhidas na Proposta de Diretiva, vejam-se as observações que tecemos no nosso "Responsabilidade civil (extracontratual) e inteligência artificial (IA): dos desafios e das possíveis respostas jurídicas aos mesmos" a ser integrado na obra em homenagem ao Senhor Prof. Doutor António Pinto Monteiro (no prelo).

Barbosa, Mafalda Miranda, "Inteligência artificial, e-persons e direito: desafios e perspetivas", *Direito e Robótica*, *Número especial de Estudos de Direito do Consumidor*, agosto de 2020, pp. 57-92.

Barbosa, Mafalda Miranda, *Inteligência Artificial. Entre a Utopia e a distopia, alguns problemas jurídicos*, Gestlegal, 2021.

Bitterman, Danielle S; Aerts, Hugo J. W. L. e Mak, Raymond H, "Approaching autonomy in medical artificial intelligence", *The Lancet Digital Health*, Vol. 2, nº 9, setembro de 2020, pp. e447-e449, disponível em www.thelancet.com/digital-health [acedido, pela última vez, em 30-03-2023].

Borghetti, Jean-Sébastien, "How can Artificial Intelligence be defective", in Sebastian Lohsee, Reiner Schulze e Dirk Satudenmayer (eds.), *Liability for artificial intelligence and the Internet of things*, Baden-Baden, Nomos, 2019, pp. 63-76.

Cohen, I. Glenn, "Informed Consent and Medical Artificial Intelligence: What to Tell the Patient?", *The Georgetown Law Journal*, Volume 108, 2020, pp. 1425-1469.

Cordeiro, Menezes, *Da Boa-fé no Direito Civil*, Coimbra, Almedina, 2001.

Cunha, Carolina, "O doente sem horário: breve anatomia dos problemas jurídicos suscitados pelas aplicações móveis na área da saúde", *Direito e Robótica*, *Número especial de Estudos de Direito do Consumidor*, agosto de 2020, pp. 43-56.

Devillé, Rembrandt; Sergeyssels, Nico e Middag, Catherine, «Basic Concepts of AI for legal scholars», in Jan De Bruyne e Cedric Vanleenhove (eds.), *Artificial Intelligence and the Law*, Cambridge, Intersentia, 2021, pp. 1-22.

Dias, Figueiredo e Monteiro, Sinde, "Responsabilidade Médica em Portugal", *Boletim do Ministério da Justiça*, nº 332, janeiro de 1984, pp. 21-79.

European Law Institute, "Response of the European Law Institute" quanto à consulta pública da Comissão Europeia sobre «Civil liability – adapting liability rules to the digital age and artificial intelligence», European Law Institute, 2022.

Expert Group on Liability and New Technologies – New Technologies Formation, *Liability for artificial intelligence and other emerging digital technologies*, 2019, disponível em https://www.europarl.europa.eu/meetdocs/2014_2019/plmrep/COMMITTEES/JURI/DV/2020/01-09/AI-report_EN.pdf. [acedido, pela última vez, em 30/03/2023].

Faria, Jorge Ribeiro de com Vasconcelos, Miguel Pestana de e Pedro, Rute Teixeira, *Direito das Obrigações*, Vol. I, 2ª ed., Coimbra, Almedina, 2020.

Frada, Manuel Carneiro da, *Contrato e Deveres de Protecção*, Separata do Volume XXXVIII do Suplemento ao Boletim da Faculdade de Direito da Universidade de Coimbra, Coimbra, 1994.

Ibrahim, Hussein; Liu, Xiaoxuan; Zariffa, Nevine; Morris Andrew D. e Denniston, Alastair K. "Health data poverty: an assailable barrier to equitable digital health care", *The Lancet Digital Health*, Vol. 3, nº 4, março de 2021, pp. e260-e265, disponível em www.thelancet.com/digital-health [acedido, pela última vez, em 30/03/2023].

Karner, Ernst, "Liability for robotics: current rules, challenges, and the need for innovative concepts", in Sebastian Lohsee, Reiner Schulze e Dirk Satudenmayer (eds.), *Liability for artificial intelligence and the Internet of things*, Baden-Baden, Nomos, 2019, pp. 117-124.

KOCH, BERNHARD, "Product liability 2.0 – mere update or new version?", in SEBASTIAN LOHSEE, REINER SCHULZE e DIRK SATUDENMAYER (eds.), *Liability for artificial intelligence and the Internet of things*, Baden-Baden, Nomos, 2019, pp. 99-116.

MARTI, GAËLLE ; CLUZEL-MÉTAYER, LUCIE e MERABET, SAMIR, "Droit et intelligence artificielle", *La semaine Juridique. Édition Générale*, n.os 51 e 52, 20 de dezembro de 2021, pp. 2359-2364.

MATOS, FILIPE ALBUQUERQUE, Responsabilidade por danos causados a terceiros por robôs", *Direito e Robótica*, *Número especial de Estudos de Direito do Consumidor*, agosto de 2020, pp. 155-212.

MONTEIRO, ANTÓNIO PINTO, "«Qui facit per alium, facit per se» – será ainda assim na era da robótica?", *Direito e Robótica*, *Número especial de Estudos de Direito do Consumidor*, agosto de 2020, pp. 11-32.

MONTEIRO, SINDE, *Responsabilidade por Conselhos, Recomendações ou Informações*, Coimbra, Almedina, 1989.

MUEHLEMATTER, URS J.; DANIORE, PAOLA e VOKINGER, KERSTIN N., "Approval of Artificial intelligence and machine learning-based medical devices in the USA and Europe (2015-20): a comparative analysis", *The Lancet digital health*, Vol. 3, janeiro de 2021, pp. e195-e203, disponível em www.thelancet.com/digital-health [acedido, pela última vez, em 30-03-2023]

OLIVA, ANTONIO *et alii*, «Management of Medico-Legal Risks in Digital Health Era: A Scoping Review», *Frontiers in medicine*, publicado em janeiro de 2022, https://doi.org/10.3389/fmed.2021.821756 [acedido, pela última vez, em 30/03/2023]

ORGANIZAÇÃO MUNDIAL DE SAÚDE (OMS), *Ethics and governance of artificial intelligence for health: WHO guidance*, World Health Organization, 2021, disponível em https://www.who.int/publications/i/item/9789240029200. [acedido em 30/03/2023]

PAGALLO, UGO, *The law of robots*, Springer, Heidelberg, London, New York, 2013, p. 2.

PALMERINI, ERICA, Robotica e diritto: suggestioni, intersezioni, sviluppi a margine di una ricerca europea», *Responsabilità civile e previdenza*, Volume 81, Fascículo 6, 2016, pp. 1816-1850.

PEDRO, RUTE TEIXEIRA, *A Responsabilidade Civil do Médico. Reflexões sobre a noção da perda de chance e a tutela do doente lesado*, Colecção do Centro de Direito Biomédico da Faculdade de Direito da Universidade de Coimbra nº 15, Coimbra, Coimbra Editora, 2008.

PEREIRA, ANDRÉ DIAS, *O consentimento informado na relação médico-paciente. Estudo de Direito Civil*, Coimbra, Coimbra Editora, 2004.

PEREIRA, ANDRÉ DIAS, *Direitos dos pacientes e responsabilidade médica*, Coimbra, Faculdade de Direito da Universidade de Coimbra, 2012, acessível em https://estudogeral.sib.uc.pt/bitstream/10316/31524/1/Direitos%20dos%20pacientes%20e%20responsabilidade%20m%C3%A9dica.pdf. [acedido pela última vez em 30-03-2023]

ROSA, GIOVANNI DI, "Quali regole per i sistemi automatizzati «intelligenti»", *Rivista di Diritto Civile*, Ano LXVII, nº 5, setembro-outubro de 2021, pp. 823-853.

SALANITRO, UGO, "Intelligenza artificiale e responsabilità: la strategia della Commissione Europea", *Rivista di Diritto Civile*, Ano LXVI, nº 6, novembro-dezembro de 2020, pp. 1246-1276.

Santosuosso Amedeo e Bottalico, Barbara, «Autonomous systems and the law: why intelligence matters», in Eric Hilgendorf e Uwe Seidel (eds.), *Robotics, Autonomics and the Law*, Baden-Baden, Nomos, 2017, pp. 27-58.

Schwab, Klaus, *The Fourth Industrial Revolution*, New York, Crown Business, 2017.

Silva, Nuno Sousa e, "Direito e Robótica: uma primeira aproximação", *Revista da Ordem dos Advogados*, 2017, pp. 487-553.

Spindler, Gerald, "User liability and strict liability in the internet of things and for robots", in Sebastian Lohsee, Reiner Schulze e Dirk Satudenmayer (eds.), *Liability for artificial intelligence and the Internet of things*, Baden-Baden, Nomos, 2019, pp. 125-143.

Varela, Antunes, *Das Obrigações em Geral*, Vol. I, 10ª ed., Coimbra, Almedina, 2004.

Wagner, Gerhard, "Robot Liability", in Sebastian Lohsee, Reiner Schulze e Dirk Satudenmayer (eds.), *Liability for artificial intelligence and the Internet of things*, Baden-Baden, Nomos, 2019, pp. 27-62.

V
IA, CONTRATOS E CONSUMO

Inteligencia artificial y tecnología *blockchain*: transparencia e información como pilares de la protección del consumidor*

Artificial intelligence and blockchain technology: transparency and information as pillars of consumer protection

Beatriz Sáenz de Jubera Higuero**

resumo: La evolución y transformación digital del mercado mundial, y particularmente el europeo, se desarrolla muy rápidamente, con grandes avances vinculados al *Big Data,* la inteligencia artificial y la tecnología *blockchain*. La transformación digital que se está viviendo es también una prioridad legislativa y política en la Unión Europea. Ante estas realidades digitales el Derecho debe dar respuesta y, sobre todo, seguridad. El régimen jurídico debe atender a la protec-

* Trabajo desarrollado en el marco del proyecto de investigación "Consentimiento, abusividad y transparencia en los contratos de contenidos y servicios digitales con consumidores", Ministerio de Ciencia e Innovación, Programa Estatal para Impulsar la Investigación Científico-Técnica y su Transferencia (dentro del Plan Estatal de Investigación Científica, Técnica y de Innovación 2021-2023), referencia PID2021-124191OB-I00, cuyo investigador principal es el Prof. Dr. D. Sergio Cámara Lapuente, Catedrático de Derecho Civil de la Universidad de La Rioja.

** Profesora Contratada Doctora de Derecho Civil (acreditada a Profesora Titular de Universidad) Universidad de La Rioja (España). Investigadora colaboradora del CIJ – Faculdade de Direito da Universidade do Porto. Investigadora asociada de la Cátedra euroamericana de protección jurídica de los consumidores. beatriz.saenz-jubera@unirioja.es. https://orcid.org/0000-0002-0838-4534

ción de los usuarios, la economía y la competencia y, sobre todo, proporcionar seguridad jurídica y seguridad en general, pues esta es esencial en este ámbito. Y concretamente en el marco de las relaciones con los consumidores y usuarios la información y la transparencia en torno a los procesos, el uso de datos y algoritmos y los efectos de las decisiones automatizadas debe presidir el régimen jurídico de estas nuevas realidades, particularmente en inteligencia artificial y en la tecnología *blockchain*, base de los *smart contracts*.

PALAVRAS-CHAVE: Consumidores. Información. Inteligencia artificial. *Smart contracts*. Tecnología *blockchain*. Transparencia. Usuarios.

ABSTRACT: The evolution and digital transformation of the world market, and particularly the European one, is developing very quickly, with great advances linked to Big Data, artificial intelligence and blockchain technology. The digital transformation that is taking place is also a legislative and political priority in the European Union. Faced with these digital realities, the Law must respond and, above all, security. The legal regime must attend to the protection of users, the economy and competition and, above all, provide legal certainty and security in general, as this is essential in this area. And specifically in the framework of relations with consumers and users, information and transparency regarding processes, the use of data and algorithms and the effects of automated decisions must preside over the legal regime of these new realities, particularly in the intelligence. technology and in the blockchain technology base of smart contracts.

KEYWORDS: Consumers. Information. Artificial intelligence. Smart contracts. Blockchain technology. Transparency. Users.

1. Planteamiento: la transformación digital en Europa

Internet y el comercio electrónico ya es una realidad consolidada (más aún con la pandemia del COVID-19), con los distintos cambios que ha implicado en las relaciones humanas, sociales y económicas o comerciales, en los contratos y sus formas y soportes y por los nuevos problemas y conflictos que se han planteado y que han exigido la adaptación y actualización de las instituciones jurídicas tradicionales, así como la reinterpretación o reforma de las normas existentes o la aprobación de nuevas normas.

Diariamente son muy numerosas las operaciones y transacciones electrónicas que se hacen a nivel nacional e internacional, de intercambio de bienes, por operaciones empresariales y bancarias, de contratos de suministro y prestación de servicios...; además de la consolidación del uso de la firma electrónica y el auge de las plataformas virtuales y la influencia de Internet en las relaciones personales a través de las redes sociales que ha puesto de manifiesto en el ámbito contractual una vía especial para el desarrollo de publicidad y de ofertas comerciales (*influencers, youtubers, instagramers,* envío directo de ofertas comerciales por *Facebook*...).

El desarrollo de las nuevas tecnologías es imparable y se produce muy rápidamente, de modo que esa realidad introducida con el comercio electrónico e Internet ha evolucionado enormemente hacia una realidad digital presidida por los grandes avances tecnológicos que han llevado a una gran transformación digital de la sociedad y la economía, ante lo que el Derecho no puede quedar impasible: el Metaverso, el *Big Data* y el análisis masivo de datos, el empleo de algoritmos e inteligencia artificial y robótica en las operaciones contractuales con automatización de procesos y decisiones, la tecnología *blockchain*, el llamado "Internet de las cosas" (*Internet of Things – IoT*), el almacenamiento en la nube, la digitalización de bienes y servicios para ser suministrados *online*... todo esto tiene gran impacto económico y social y plantea, además de nuevas oportunidades de avance y mejora social y económica, nuevos retos y desafíos que deben identificarse y afrontarse para que, además del desarrollo de la economía (digital en este caso) con respeto a una competencia leal, se garanticen los intereses de todos los intervinientes en las operaciones y transacciones contractuales, especialmente los más débiles o vulnerables (como los consumidores y, dentro de estos, las personas con discapacidad).

Es necesario contar con una regulación adecuada; revisar y reformar lo necesario y aprobar las normas precisas para atender a esa realidad, sus problemas, controversias y retos, protegiendo a los usuarios, la economía y la competencia y, sobre todo, proporcionando seguridad jurídica y seguridad

en general. La seguridad en esta realidad digital es esencial: sin ella, aparte de poder implicar una posible vulneración de derechos e intereses, no habrá confianza en el mercado digital y en las operaciones que en él se desarrollen, y con ello también se perderá la confianza en la innovación y el desarrollo y avance tecnológicos, que, por lo general, tienen como finalidad una mejora en las relaciones humanas y económicas, en el consumo y en el bienestar general de la sociedad. Y junto con la regulación, pilares esenciales en la práctica y, particularmente, en la protección del consumidor y usuario, son la información y la transparencia en todo el proceso contractual y en el uso de las nuevas tecnologías.

A este respecto, además de una realidad social y económica, la transformación digital que se está viviendo es también una prioridad legislativa y política en la Unión Europea, que ha mostrado desde hace años cómo la transformación digital y el mercado único digital son una de sus principales prioridades[1] y para ello ha ido marcando desde 2015 una estrategia digital y ha ido adoptando y aprobando distintos informes, comunicaciones o normas al respecto; atendiendo a los últimos años cabe mencionar los siguientes en el ámbito de las distintas instituciones comunitarias:

a) La "Estrategia del Mercado Único Digital", planteada por la Comunicación de 6 de mayo de 2015 de la Comisión Europea al Parlamento Europeo, al Consejo, al Comité Social y Económico Europeo y al Comité de las Regiones "Una Estrategia para el Mercado Único Digital de Europa" (COM(2015) 192 final)[2] y revisada dos años después a través de otra Comunicación de la Comisión Europea de 10 de mayo de 2017 relativa a la revisión intermedia de la aplicación de la Estrategia para el Mercado Único Digital, "Un mercado único digital conectado para todos" (COM(2017) 228 final)[3], fue sustituida por la Comunicación de 19 de febrero de 2020 de la Comisión Europea al Parlamento Europeo, al Consejo, al Comité Social y Económico Europeo y al Comité de las Regiones "Configurar el futuro digital de Europa" (COM(2020) 67

[1] *Vid*. https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age_es (fecha de última consulta el 15 de noviembre de 2022).

[2] Disponible en https://eur-lex.europa.eu/legal-content/ES/ALL/?uri=celex:52015DC0192 (fecha de última consulta el 15 de noviembre de 2022).

[3] Disponible en https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=COM:2017:228:FIN (fecha de última consulta el 15 de noviembre de 2022). A través de esta Comunicación se evalúan los avances en la realización del mercado único digital, y se señalan los aspectos en los que es necesario hacer más esfuerzos y en los que la evolución del panorama digital exige nuevas actuaciones a nivel de la Unión Europea.

final)[4], incidiendo en la necesidad de contar con un marco normativo fiable: "es necesario disponer de un marco claro que fomente unas interacciones digitales fiables de toda la sociedad, tanto de los ciudadanos como de las empresas. Sin este énfasis en la fiabilidad, el proceso vital de transformación digital no puede prosperar". Esta nueva estrategia se basa en tres pilares fundamentales:

i) La tecnología al servicio de las personas: en este sentido se invertirá en la educación y formación en competencias digitales de todos los ciudadanos europeos; se centra también en la necesidad de protección del usuario digital frente a las distintas amenazas digitales; y también se advierte de la importancia del desarrollo fiable de la inteligencia artificial (IA) respetando los derechos de las personas.
ii) Una economía digital justa y competitiva: se advierte de la gran importancia de que en la era digital se garanticen unas condiciones de competencia equitativas para todas las empresas con independencia de su tamaño; asimismo, se insiste en la necesidad de que los consumidores puedan confiar en los productos, contenidos y servicios digitales, además de en la importancia del acceso, uso y tratamiento de datos incidiendo en la necesidad de construir un mercado único o espacio europeo de datos basado en normas y valores europeos.
iii) Una sociedad abierta, democrática y sostenible: se insiste en el derecho a una tecnología fiable y en que los valores europeos y las normas éticas, sociales y medioambientales deben aplicarse también en el espacio digital; en este sentido se incide en la necesidad de reforzar y modernizar las normas aplicables a los servicios digitales en toda la Unión Europea, aclarando las funciones y responsabilidades de las plataformas en línea y luchando contra la desinformación y la venta de bienes ilícitos y la difusión de contenidos ilícitos, además, se incide en el control de los datos personales, necesitando normas más claras sobre transparencia y rendición de cuentas de quienes actúan como guardianes de la información y los flujos de datos.

[4] Puede consultarse en https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX%3A52020DC0067; sobre su contenido *vid*., asimismo, https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/shaping-europe-digital-future_es y en https://ec.europa.eu/info/sites/default/files/communication-shaping-europes-digital-future-feb2020_en_4.pdf (fecha de última consulta el 15 de noviembre de 2022). Asimismo, a este respecto, cfr., entre otros, Moisés Barrio Andrés, "La nueva Estrategia digital de la Comisión Europea: primeras impresiones", *Diario La Ley*, núm. 37, sección Ciberderecho, 20 de febrero de 2020.

b) El 6 de junio de 2018 la Comisión aprobó una propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establece el "Programa Europa Digital para el período 2021-2027" (COM(2018) 434 final)[5], aprobado finalmente por el Parlamento Europeo y el Consejo el 29 de abril de 2021: Reglamento (UE) 2021/694 del Parlamento Europeo y del Consejo de 29 de abril de 2021 por el que se establece el Programa Europa Digital y por el que se deroga la Decisión (UE) 2015/2240[6]. Fundamentalmente es un programa económico financiero con el fin de apoyar la transformación digital de la industria y favorecer un mejor aprovechamiento del potencial industrial de las políticas de innovación, investigación y desarrollo tecnológico, en beneficio de los ciudadanos y las empresas en toda la Unión; se estructura en torno a cinco objetivos específicos que reflejen los ámbitos de actuación clave: informática de alto rendimiento; inteligencia artificial; ciberseguridad y confianza; capacidades digitales avanzadas, y despliegue y mejor uso de la capacidad digital e interoperabilidad (considerando 14).
c) Por su parte, el Consejo de la Unión Europea adoptó el 7 de junio de 2019 unas Conclusiones sobre "El futuro de una Europa altamente digitalizada más allá de 2020: Impulsar la competitividad digital y económica en toda la Unión y la cohesión digital" (10102/19)[7], que inciden en la importancia de apoyar la innovación y fomentar las tecnologías digitales europeas clave y el desarrollo del 5G, respetar los principios y valores éticos en la inteligencia artificial, fortalecer la seguridad digital y la ciberseguridad de Europa y mejorar las competencias digitales, insistiendo en que todos los europeos y todas las empresas europeas independientemente de su tamaño o ubicación deben beneficiarse de la digitalización, promoviendo por ello un acceso sin obstáculos burocráticos[8]. Un año después, el 9 de junio de 2020 el Consejo adoptó unas Conclusiones sobre "La configuración del futuro digital de Europa"

[5] Disponible en https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX%3A52018PC0434&qid=1661445830525 (fecha de última consulta el 15 de noviembre de 2022).

[6] DOUE L 166/1, de 11 de mayo de 2021. Puede consultarse el texto de este Reglamento en https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX%3A32021R0694 (fecha de última consulta el 15 de noviembre de 2022).

[7] Puede consultarse en https://www.consilium.europa.eu/media/39667/st10102-en19.pdf (fecha de última consulta el 15 de noviembre de 2022).

[8] *Vid*. comunicado de prensa del Consejo al respecto de estas conclusiones: disponible en https://www.consilium.europa.eu/es/press/press-releases/2019/06/07/post-2020-digital-policy-council-adopts-conclusions/ (fecha de última consulta el 15 de noviembre de 2022).

(8711/20)[9], en el marco de una realidad marcada por la pandemia del COVID-19, destacando precisamente el impacto de la transformación digital en la lucha contra la pandemia y en la recuperación tras ella; en estas conclusiones se refiere a la conectividad, las cadenas de valor digitales, la sanidad electrónica, la economía de datos, la inteligencia artificial y las plataformas digitales[10].

d) Dentro de este marco de estrategia generalista de la Unión Europea respecto a la economía y realidad digital, cabe mencionar también la Comunicación de 9 de marzo de 2021 de la Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones "Brújula Digital 2030: el enfoque de Europa para el Decenio Digital" (COM(2021) 118 final)[11], que se basa en la estrategia digital aprobada en febrero de 2020 y que responde al llamamiento del Consejo Europeo a través de sus Conclusiones de 2 de octubre de 2020 sobre el ámbito digital europeo y la necesidad de acelerar la transición digital en Europa[12]. Con esta Comunicación la Comisión Europea presentó una visión, objetivos y vías para una transformación digital exitosa de la Unión Europea hasta 2030. Tal y como se señala en dicha comunicación, la Comisión propone establecer una "Brújula Digital" para traducir las ambiciones digitales de la Unión Europea para 2030 en objetivos concretos y alcanzar estos objetivos; dicha Brújula se basará en un sistema de seguimiento mejorado para verificar la trayectoria de la Unión en relación con el ritmo de la transformación digital, las lagunas en las capacidades digitales estra-

[9] Disponible en https://data.consilium.europa.eu/doc/document/ST-8711-2020-INIT/es/pdf (fecha de última consulta el 15 de noviembre de 2022).

[10] *Vid.* el comunicado de prensa del Consejo respecto de estas conclusiones en https://www.consilium.europa.eu/es/press/press-releases/2020/06/09/shaping-europe-s-digital-future-council-adopts-conclusions/ (fecha de última consulta el 15 de noviembre de 2022).

[11] Disponible en https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX%3A52021DC0118 (fecha de última consulta el 15 de noviembre de 2022).

[12] Conclusiones del Consejo de 2 de octubre de 2020 (EUCO/ 13/20) disponibles en https://www.consilium.europa.eu/media/45932/021020-euco-final-conclusions-es.pdf (fecha de última consulta el 15 de noviembre de 2022). El Consejo en estas Conclusiones, concretamente en la sexta, referida al ámbito digital, señaló: "El Consejo Europeo espera con interés la propuesta de la Comisión de una norma sobre servicios digitales antes de que termine este año e invita a la Comisión a presentar, a más tardar en marzo de 2021, una Brújula Digital global que establezca las ambiciones digitales concretas de la UE para 2030. Dicha Brújula debe crear un sistema de seguimiento de las capacidades y aptitudes digitales estratégicas europeas y perfilar los medios y las principales etapas para alcanzar nuestras ambiciones".

tégicas europeas y la aplicación de los principios digitales. Se articula en torno a cuatro puntos cardinales: a) Ciudadanos con capacidades digitales y profesionales del sector digital muy cualificados; b) Infraestructuras digitales sostenibles que sean seguras y eficaces; c) Transformación digital de las empresas; y d) Digitalización de los servicios públicos.

e) Posteriormente, el 15 de septiembre de 2021 la Comisión presentó la propuesta de Decisión del Parlamento Europeo y del Consejo por la que se establece el "programa de política «Itinerario hacia la Década Digital» para 2030" (COM(2021) 574 final)[13], que insiste en los objetivos marcados en la Comunicación sobre la "Brújula Digital": persigue garantizar que la Unión Europea alcance sus objetivos y metas de transformación digital de nuestra sociedad y economía en consonancia con los valores de la Unión, que refuerce el liderazgo digital de la Unión Europea y promueva políticas centradas en el ser humano, inclusivas y sostenibles que capaciten a los ciudadanos y las empresas; el objetivo es lograr la transformación digital de la Unión en consonancia con esta visión mediante el establecimiento de un proceso claro, estructurado y colaborativo que permita lograr este resultado. El 14 de julio de 2022 el Consejo y el Parlamento Europeo alcanzaron un acuerdo provisional sobre este programa "Itinerario hacia la Década Digital" para 2030[14].

f) Aparte de estos documentos y normas que sientan las directrices de una estrategia general de la Unión Europea en el ámbito digital, cabe mencionar asimismo ciertas normas y documentos aprobados con respecto a realidades digitales concretas:
   - En el marco de la Inteligencia Artificial (IA): la Comunicación de 25 de abril de 2018 de la Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones sobre "Inteligencia artificial para Europa" (COM(2018) 237 final)[15]; la Resolución de 12 de febrero de 2019 del Parlamento Europeo "Una política industrial global europea en materia de inteligencia artifi-

[13] Texto disponible en https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX%3A52021PC0574 (fecha de última consulta el 15 de noviembre de 2022).

[14] Puede consultarse el comunicado de prensa del Consejo al respecto de este acuerdo provisional en https://www.consilium.europa.eu/es/press/press-releases/2022/07/14/policy-programme-path-to-the-digital-decade-the-council-and-the-european-parliament-reach-a-provisional-agreement/ (fecha de última consulta el 15 de noviembre de 2022).

[15] Disponible en https://ec.europa.eu/transparency/documents-register/detail?ref=COM(2018)237&lang=es (fecha de última consulta el 15 de noviembre de 2022).

cial y robótica" (2018/2088(INI))[16]; la Comunicación de la Comisión de 8 de abril de 2019 "Generar confianza en la Inteligencia artificial centrada en el ser humano" (COM(2019) 168 final)[17]; el "Libro Blanco sobre la inteligencia artificial: un enfoque europeo orientado a la excelencia y la confianza europea" (COM(2020) 65 final)[18], presentado por la Comisión Europea el 19 de febrero de 2020 junto con el "Informe sobre las repercusiones en materia de seguridad y responsabilidad civil de la inteligencia artificial, el internet de las cosas y la robótica" (COM(2020) 64 final)[19]; documentos estos últimos que sientan las bases de la propuesta que la Comisión Europea presentó el 21 de abril de 2021 de "Reglamento sobre Inteligencia Artificial" (COM (2021) 2016 final)[20]. Posteriormente, el Consejo a finales de 2022 planteó su posición y propuesta transaccional al

[16] Puede consultarse en https://www.europarl.europa.eu/doceo/document/TA-8-2019-0081_ES.html y en https://www.europarl.europa.eu/doceo/document/TA-8-2019-0081_ES.pdf (fecha de última consulta el 15 de noviembre de 2022).

[17] Disponible en https://ec.europa.eu/transparency/documents-register/detail?ref=COM(2019)168&lang=es (fecha de última consulta el 15 de noviembre de 2022).

[18] Libro Blanco sobre IA que puede consultarse en https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX%3A52020DC0065&qid=1661087591880 y en https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_es.pdf (fecha de última consulta el 15 de noviembre de 2022). Cfr. a este respecto, entre otros, Isabel Zurita Martín, "Las propuestas de reforma legislativa del Libro Blanco europeo sobre inteligencia artificial en materia de seguridad y responsabilidad civil", *Actualidad Jurídica Iberoamericana*, núm. 14, febrero, 2021, pp. 443 y ss. (disponible en https://idibe.org/wp-content/uploads/2021/03/11._Isabel_Zurita_pp._438-487.pdf; fecha de última consulta el 15 de noviembre de 2022).

[19] Informe disponible en https://eur-lex.europa.eu/legal-content/ES/TXT/HTML/?uri=CELEX:52020DC0064&from=ES (fecha de última consulta el 15 de noviembre de 2022). Este informe viene a concluir la existencia de lagunas legales que deben subsanarse en la legislación vigente sobre seguridad de los productos, especialmente en la normativa sobre máquinas. Cfr. Isabel Zurita Martín, "Las propuestas...", *op. cit.*, pp. 452 y ss.

[20] Propuesta de la Comisión Europea de 21 de abril de 2021 (COM (2021) 2016 final), de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en material de inteligencia artificial (Ley de Inteligencia Artificial) y se modifican determinados actos legislativos de la Unión (disponible en https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=celex:52021PC0206, fecha de última consulta el 15 de noviembre de 2022). Esta propuesta de reglamentación se plantea con los siguientes objetivos específicos: "a) garantizar que los sistemas de IA introducidos y usados en el mercado de la UE sean seguros y respeten la legislación vigente en materia de derechos fundamentales y valores de la Unión; b) garantizar la seguridad jurídica para facilitar la inversión e innovación en IA; c) mejorar la gobernanza y la aplicación efectiva de la legislación vigente en materia de derechos fundamentales y los requisitos de seguridad aplicables a los sistemas de IA; y d) facilitar el desarrollo de un mercado único para hacer un uso legal, seguro y fiable de las aplicaciones de IA y evitar la fragmentación del mercado".

respecto y, más tarde, el Parlamento Europeo el 14 de junio de 2023 aprobó formalmente su postura negociadora y enmiendas en relación con esa propuesta de Reglamento de Inteligencia Artificial, de modo que a partir de tal momento se han iniciado ya las negociaciones entre las distintas instituciones comunitarias europeas para así lograr que vea la luz esta norma que está llamada a ser el primer texto legal global regulador de la IA.[21]

– En el ámbito de la contratación digital y las plataformas *online* cabe citar las siguientes normas: la Directiva Europea (UE) 2019/770 del Parlamento Europeo y del Consejo, de 20 de mayo de 2019, relativa a determinados aspectos de los contratos de suministro de contenidos y servicios digitales[22], y la Directiva (UE) 2019/771, del Parlamento europeo y del Consejo, de 20 de mayo de 2019, relativa a determinados aspectos de los contratos de compraventa de bienes[23] (ambas normas fueron transpuestas al ordenamiento jurídico español a través del Real Decreto Ley 7/2021, de 27 de abril[24], que, para ello, modi-

[21] Estos últimos textos fueron aprobados muy posteriormente a la entrega de la versión final de este trabajo para su publicación. Ahora con este breve apunte simplemente se quiere actualizar la información sobre el estado en que se encuentra la tramitación y negociación de esta nueva norma. Se siguen clasificando los sistemas de IA en distintos niveles atendiendo al riesgo que de ellos se derivan respecto a diversos aspectos, pero cabe destacar que existen algunas discrepancias y diferencias en cuanto a la definición de "sistemas de IA", lo que deberá concretarse finalmente en la norma definitiva que se apruebe pues determinará su ámbito de aplicación.
La propuesta transaccional o posición adoptada por el Consejo se encuentra disponible en https://data.consilium.europa.eu/doc/document/ST-14954-2022-INIT/es/pdf (fecha de última consulta el 23 de julio de 2023) y el comunicado de prensa del Consejo de 6 de diciembre de 2022 está disponible en https://www.consilium.europa.eu/es/press/press-releases/2022/12/06/artificial-intelligence-act-council-calls-for-promoting-safe-ai-that-respects-fundamental-rights/ (fecha de última consulta el 23 de julio de 2023).
Por lo que se refiere al Parlamento Europeo, el texto de su posición negociadora sobre este Reglamento o Ley de Inteligencia Artificial aprobado en su reunión de 14 de junio de 2023 puede consultarse en https://www.europarl.europa.eu/doceo/document/TA-9-2023-0236_ES.html (fecha de última consulta el 23 de julio de 2023); y la nota de prensa emitida al respecto está disponible en https://www.europarl.europa.eu/news/es/press-room/20230609IPR96212/la-eurocamara-lista-para-negociar-la-primera-ley-sobre-inteligencia-artificial (fecha de última consulta el 23 de julio de 2023).

[22] DOUE L 136/1, de 22 de mayo de 2019: https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:32019L0770 (fecha de última consulta el 15 de noviembre de 2022).

[23] DOUE L 136/28, de 22 de mayo de 2019: https://eur-lex.europa.eu/legal-content/es/TXT/?uri=CELEX%3A32019L0771 (fecha de última consulta el 15 de noviembre de 2022).

[24] Real Decreto-ley 7/2021, de 27 de abril, de transposición de directivas de la Unión Europea en las materias de competencia, prevención del blanqueo de capitales, entidades de crédito, telecomunicaciones, medidas tributarias, prevención y reparación de daños medioambientales,

ficó el Real Decreto Legislativo 1/2007, de 16 de noviembre, por el que se aprueba el texto refundido de la Ley General para la Defensa de los Consumidores y Usuarios y otras leyes complementarias, en adelante TRLCU); el Reglamento (UE) 2022/1925, del Parlamento Europeo y del Consejo, de 14 de septiembre de 2022, de Mercados Digitales[25] (*Digital Markets Act* – DMA) y el Reglamento (UE) 2022/2065, del Parlamento Europeo y del Consejo, de 19 de octubre de 2022, de Servicios Digitales[26] (*Digital Services Act* – DSA)[27].

2. Inteligencia artificial

2.1. Confianza, información y transparencia como pilares esenciales de su régimen jurídico

Una de las máximas expresiones y reflejo del gran y rápido avance de las nuevas tecnologías es el progreso y desarrollo en los últimos años de la llamada

desplazamiento de trabajadores en la prestación de servicios transnacionales y defensa de los consumidores (BOE núm. 101, de 28 de abril de 2021: https://www.boe.es/buscar/act.php?id=BOE-A-2021-6872; fecha de última consulta el 15 de noviembre de 2022).

[25] DOUE L 265/1, de 12 de octubre de 2022. Este Reglamento (UE) 2022/1925 del Parlamento Europeo y del Consejo de 14 de septiembre de 2022 sobre mercados disputables y equitativos en el sector digital y por el que se modifican las Directivas (UE) 2019/1937 y (UE) 2020/1828 (Reglamento de Mercados Digitales) puede consultarse en https://eur-lex.europa.eu/legal-content/ES/TXT/HTML/?uri=CELEX:32022R1925&from=ES y en https://www.boe.es/buscar/doc.php?id=DOUE-L-2022-81470 (fecha de última consulta el 15 de noviembre de 2022). Conforme a su art. 54, este Reglamento entrará en vigor a los veinte días de su publicación en el DOUE, pero se aplicará a partir del 2 de mayo de 2023 con carácter general, si bien hay previsiones específicas en relación al contenido de ciertos preceptos: el artículo 3, apartados 6 y 7, y los artículos 40, 46, 47, 48, 49 y 50 serán aplicables a partir del 1 de noviembre de 2022 y el artículo 42 y el artículo 43 serán aplicables a partir del 25 de junio de 2023.

[26] DOUE L 277/1, de 27 de octubre de 2022. El texto completo de este Reglamento (UE) 2022/2065 del Parlamento Europeo y del Consejo de 19 de octubre de 2022 relativo a un mercado único de servicios digitales y por el que se modifica la Directiva 2000/31/CE (Reglamento de Servicios Digitales) se encuentra disponible en https://eur-lex.europa.eu/legal-content/ES/TXT/HTML/?uri=CELEX:32022R2065&from=ES y en https://www.boe.es/buscar/doc.php?id=DOUE-L-2022-81573 (fecha de última consulta el 15 de noviembre de 2022). En el art. 93 de este Reglamento se advierte que el mismo entrará en vigor a los veinte días de su publicación en el DOUE, pero se aplicará con carácter general a partir del 17 de febrero de 2024, si bien "el artículo 24, apartados 2, 3 y 6, el artículo 33, apartados 3 a 6, el artículo 37, apartado 7, el artículo 40, apartado 13, el artículo 43 y las secciones 4, 5 y 6 del capítulo IV, serán de aplicación a partir del 16 de noviembre de 2022".

[27] Puede consultarse la nota de prensa del Parlamento Europeo al respecto de la aprobación de ambos Reglamentos en https://www.europarl.europa.eu/news/es/press-room/20220701IPR34364/dos-leyes-historicas-para-unos-servicios-digitales-mas-seguros-y-abiertos (fecha de última consulta el 15 de noviembre de 2022).

inteligencia artificial (en adelante, IA); término con el que se alude a aquellos procesos o sistemas tecnológicos a través de los cuales se trata de que una máquina emule capacidades, habilidades y conductas típicamente humanas, como el razonamiento, la capacidad de decisión o de planificación, la creatividad..., y hacerlo de modo autónomo; tal proceso se lleva a cabo a través de algoritmos y programas informáticos de análisis de datos y procesamiento de información de forma masiva.

Ahora bien, este concepto engloba realidades muy distintas, unas más complejas que otras dependiendo de las capacidades y autonomías con las que se pretenda dotar a la máquina: traductores automáticos y motores de búsqueda en internet; sistemas de reconocimiento de voz y/o facial; asistentes digitales en ordenadores o *smartphones*; *chatbots*; asistentes virtuales como los dispositivos "Alexa"; el "internet de las cosas"; mecanismos de inteligencia artificial en aparatos diversos como aspiradoras, aire acondicionados o relojes *smartwatch*; vehículos autónomos; drones; robots utilizados en procesos de fabricación de distintos bienes; y, en lo que más interesa para el objeto de este trabajo, la aplicación de la IA en los procesos de contratación *online* sobre todo en las ventas (incluida la publicidad)[28].

La inteligencia artificial es una realidad que ha venido para quedarse: son ya muchas las empresas que acuden a esta tecnología para su desarrollo y para optimizar y mejorar su productividad, la toma de decisiones, la automatización y el control de los procesos y el análisis de mayor volumen de datos de forma más rápida.

Sin embargo, esta tecnología también plantea riesgos y aún son varios los retos a los que se enfrenta en su desarrollo. Entre esos retos está principalmente el de conseguir confianza y seguridad para los usuarios en torno a esos procesos automatizados, especialmente en la toma de decisiones y el análisis y uso de los datos. Contar con una IA confiable llevará a promover sus aspec-

[28] Según advierte la Comisión Europea, «el término "inteligencia artificial» (IA) se aplica a los sistemas que manifiestan un comportamiento inteligente, pues son capaces de analizar su entorno y pasar a la acción –con cierto grado de autonomía– con el fin de alcanzar objetivos específicos. Los sistemas basados en la IA pueden consistir simplemente en un programa informático (p. ej. asistentes de voz, programas de análisis de imágenes, motores de búsqueda, sistemas de reconocimiento facial y de voz), pero la IA también puede estar incorporada en dispositivos de hardware (p. ej. robots avanzados, automóviles autónomos, drones o aplicaciones del internet de las cosas)" (Comunicación de 25 de abril de 2018 de la Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones, "Inteligencia artificial para Europa", COM(2018) 237 final; disponible en https://ec.europa.eu/transparency/documents-register/detail?ref=COM(2018)237&lang=es, fecha de última consulta el 15 de noviembre de 2022).

tos positivos y tratar de evitar o mitigar los aspectos negativos y los riesgos que su uso conlleva. Para lograr esa confianza en la IA son pilares esenciales, además de la seguridad que ofrezcan los dispositivos para hacer frente a posibles ataques o ciberataques:

a) El uso de los modelos de IA con criterios de ética y justicia, para evitar resultados discriminatorios por el uso de datos y algoritmos sesgados. Precisamente una de las cuestiones en las que últimamente más se están centrando los estudios sobre IA, incluidas las propuestas e informes que surgen de las instituciones comunitarias europeas, es en la ética, relacionándola también con: la exactitud, explicabilidad e interpretabilidad de los algoritmos, datos y modelos utilizados; con la causalidad y la forma de medir la parcialidad, toxicidad o sesgos de los datos y sistemas; con la equidad en la creación y uso de los algoritmos; y con la privacidad en el uso de datos[29]. Y es que uno de los retos jurídicos civiles más importantes que plantea la IA recae en la obtención de datos personales y su uso, especialmente respecto a la seguridad en su tratamiento y con respecto al posible sesgo discriminatorio de los algoritmos de análisis masivo de datos[30].
b) La explicabilidad de los algoritmos utilizados en esos procesos y los efectos de las decisiones automatizadas con base en ellos (información al respecto y transparencia). Ciertamente hay algunos algoritmos más fáciles de explicar que otros, pero siempre debe buscarse una forma hábil para explicar a los usuarios cómo y porqué se obtienen tales resultados y decisiones por el uso de un algoritmo; la información y el conocimiento de cómo funciona el algoritmo, por tanto, es esencial; y lo es también para poder conocer precisamente por qué en ciertos casos no ha funcionado bien y darle solución.

[29] Cfr., entre otros, Antonio Mozo Seoane, *Los límites de la tecnología. Marco ético y regulación jurídica*, Reus, Madrid, 2021; y Miguel Peguera Poch, "En búsqueda de un marco normativo para la inteligencia artificial", en Agustí Cerrillo i Martínez y Miguel Peguera Poch (coords.), *Retos jurídicos de la inteligencia artificial*, Thomson Reuters Aranzadi, Cizur Menor, 2020, pp. 44 y ss.

[30] A este respecto *vid*., entre otros, Pablo Pascual Huerta, "Algoritmos y protección de datos personales", en Matilde Cuena Casas y Javier Ibáñez Jiménez (dirs.), *Perspectiva legal y económica del fenómeno FinTech*, Wolters Kluwer La Ley, Madrid, 2021, pp. 559 y ss.; y en esta misma obra colectiva, Mª Teresa Bendito Cañizares, "Ética, algoritmos y decisiones individuales y colectivas automatizadas", pp. 641 y ss.

En el ámbito contractual y de protección de los consumidores y usuarios, especialmente ante la prestación del consentimiento, el uso de sistemas de IA puede plantear retos jurídicos relevantes respecto a la identificación de las partes y su capacidad contractual, así como en cuanto a los efectos de las decisiones automatizadas. La IA ha impactado contractualmente sobre todo a través de ofertas comerciales personalizadas y la toma de decisiones respecto a la concesión o no de productos solicitados por el consumidor (créditos, seguros...)[31]; ello a partir de la recopilación y análisis de gran masa de datos relativos a los usuarios, con intención predictiva de preferencias y comportamientos futuros a través de esos datos del pasado del usuario, buscando asimismo orientar la conducta del consumidor en cierto sentido y clasificando a éste bajo ciertos modelos configurados automáticamente (el llamado *profiling* o perfilado); todo lo cual puede llevar a ciertos resultados injustos, desiguales o discriminatorios entre consumidores, especialmente si los datos y algoritmos usados son sesgados.

Por supuesto que esta problemática también se ve vinculada con la otra ya indicada relativa a la forma de obtención de esos datos de los usuarios, su tratamiento y su uso en esos sistemas de inteligencia artificial; el consentimiento del usuario a este respecto es esencial; pero debe ser un consentimiento plenamente informado sobre los datos que serán objeto del tratamiento, la finalidad del mismo y de todas las consecuencias directas o indirectas que su aceptación conllevará.

Pilares esenciales de la regulación protectora de los consumidores y usuarios, tanto a nivel nacional como comunitario europeo, son la información y la transparencia. En ese sentido, el consumidor concreto afectado por el uso de un sistema de IA debe ser informado de que en ese proceso contractual o precontractual están desarrollándose actividades de recomendación predictiva y decisión algorítmica utilizándose un sistema de IA; y, asimismo, deben también presentársele opciones para desactivarlo.

La exigencia de transparencia en estos procesos exige informar sobre qué datos se han utilizado, quién lo ha hecho, cómo se han usado, con qué fin, cómo se han integrado en el sistema de IA, cómo se ha llegado a tomar una decisión concreta en su caso y qué grado de intervención ha tenido esa IA

[31] A este respecto *vid*., entre otros, Isabel Antón Juárez, "Personalización de precios a través de la inteligencia artificial y el Big Data", en Manuel Paniagua Zurera (dir.), *El sistema jurídico ante la digitalización: estudios de Derecho privado*, Tirant lo Blanch, Valencia, 2021, pp. 379 y ss.; y Antonio Robles Martín-Laborda, "Inteligencia artificial y personalización de precios", en Matilde Cuena Casas y Javier Ibáñez Jiménez (dirs.), *Perspectiva legal y económica del fenómeno FinTech*, Wolters Kluwer La Ley, Madrid, 2021, pp. 573 y ss.

en la toma de la decisión (que entroncaría con la cuestión de la atribución de posibles responsabilidades).

Para lograr esa transparencia es necesario que cada decisión adoptada sea trazable y venga fundamentada; no pueden adoptarse decisiones arbitrarias. Y además es preciso que se reconozca el derecho del consumidor a exigir la explicación de las razones y fundamentos de la decisión adoptada en el marco de ese proceso contractual que le afecta[32].

Transparencia e información, por tanto, son principios esenciales en el marco general de la IA y su fiabilidad o confiabilidad: no sólo en relación con los datos y algoritmos, sino también sobre el porqué de su utilización en determinados procesos o negocios y sobre su desarrollo, uso y supervisión; pilares que deben presidir la protección de los usuarios ante estos sistemas de IA, sea cual sea el nivel de riesgo en que se encuentren, atendiendo al criterio y clasificación contenida en la propuesta de Reglamento europeo sobre inteligencia artificial a la que aludiré a continuación, dentro de la estrategia seguida por la Unión Europea sobre esta materia.

2.2. Estrategia de la Unión Europea en el marco de la Inteligencia Artificial (IA)

Cabe poner de relieve cómo la Unión Europea ha considerado la IA como una prioridad dado el papel central que considera tiene y tendrá en la transformación digital de la sociedad.

Ya el 25 de abril de 2018 la Comisión Europea presentó la estrategia europea sobre la IA a través de la Comunicación al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones, "Inteligencia artificial para Europa" (COM(2018) 237 final)[33], donde expone una iniciativa europea sobre IA, buscando trabajar con todos los Estados miembros en torno a un plan coordinado sobre IA y con los siguientes fines: a) potenciar la capacidad tecnológica e industrial de la UE e impulsar la adopción de la IA en todos los ámbitos de la economía, tanto en el sector privado como en el público; b) prepararse para las transformaciones socioeconómicas que

[32] A este respecto, entre otros, *vid*., Elisa de la Nuez Sánchez-Cascado, "Inteligencia artificial y transparencia. Especial referencia a su utilización en el ámbito de las Administraciones Públicas", *El Notario del Siglo XXI*, núm. 93, septiembre-octubre, 2020 (disponible en https://www.elnotario.es/hemeroteca/revista-93/10160-inteligencia-artificial-y-transparencia-especial-referencia-a-su-utilizacion-en-el-ambito-de-las-administraciones-publicas; fecha de última consulta el 15 de noviembre de 2022).

[33] Disponible en https://ec.europa.eu/transparency/documents-register/detail?ref=COM(2018)237&lang=es (fecha de última consulta el 15 de noviembre de 2022).

origina la IA; y c) garantizar el establecimiento de un marco ético y jurídico apropiado, basado en los valores de la Unión y en consonancia con la Carta de los Derechos Fundamentales de la Unión Europea.

El 12 de febrero de 2019 el Parlamento Europeo aprobó la resolución "Una política industrial global europea en materia de inteligencia artificial y robótica" (2018/2088(INI))[34], donde subraya la importancia de una IA segura y transparente, por lo que insta a instituciones comunitarias y Estados miembros a adoptar diversas medidas al respecto, anima al desarrollo de la investigación sobre ello y señala la necesidad de una regulación integral de la IA.

Incidiendo en esa relevancia de la IA y la importancia de su regulación y atención desde las instituciones europeas, posteriormente, siguiendo con la estrategia de la Unión sobre la IA, la Comisión Europea publicó el 8 de abril de 2019 la Comunicación "Generar confianza en la Inteligencia artificial centrada en el ser humano" (COM(2019) 168 final)[35], donde se recogen una serie de directrices en relación con la fiabilidad de la IA; y un año después, el 19 de febrero de 2020, presentó el "Libro Blanco sobre la inteligencia artificial: un enfoque europeo orientado a la excelencia y la confianza europea" (COM(2020) 65 final)[36] junto con el "Informe sobre las repercusiones en materia de seguridad y responsabilidad civil de la inteligencia artificial, el internet de las cosas y la robótica" (COM(2020) 64 final)[37], documentos que sientan las bases de la propuesta que la Comisión Europea presentó el 21 de abril de 2021 de "Reglamento sobre Inteligencia Artificial" (COM (2021) 2016 final)[38], con la que pretende crear el primer marco legal global sobre

[34] Puede consultarse en https://www.europarl.europa.eu/doceo/document/TA-8-2019-0081_ES.html y en https://www.europarl.europa.eu/doceo/document/TA-8-2019-0081_ES.pdf (fecha de última consulta el 15 de noviembre de 2022).

[35] Disponible en https://ec.europa.eu/transparency/documents-register/detail?ref=COM(2019)168&lang=es (fecha de última consulta el 15 de noviembre de 2022).

[36] Libro Blanco sobre IA que puede consultarse en https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX%3A52020DC0065&qid=1661087591880 y en https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_es.pdf (fecha de última consulta el 15 de noviembre de 2022). Cfr. Isabel Zurita Martín, "Las propuestas...", *op. cit.*, pp. 443 y ss.

[37] Informe disponible en https://eur-lex.europa.eu/legal-content/ES/TXT/HTML/?uri=CELEX:52020DC0064&from=ES (fecha de última consulta el 15 de noviembre de 2022). Este informe viene a concluir la existencia de lagunas legales que deben subsanarse en la legislación vigente sobre seguridad de los productos, especialmente en la normativa sobre máquinas. Cfr. Isabel Zurita Martín, "Las propuestas...", *op. cit.*, pp. 452 y ss.

[38] Propuesta de la Comisión Europea de 21 de abril de 2021 (COM (2021) 2016 final), de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en material de inteligencia artificial (Ley de Inteligencia Artificial) y se modifican determinados

IA y convertir a la Unión Europea en el centro mundial de una inteligencia artificial fiable[39].

Esta propuesta de Reglamento europeo de IA incide en las cuestiones de confianza, seguridad, transparencia e información ya indicadas a partir de la clasificación de los sistemas inteligentes en cuatro niveles según el riesgo que se genera (clasificación desarrollada con el fin de concretar prácticas de IA prohibidas en la Unión Europea por su gran riesgo y otras prácticas que, aún permitidas, exigirían la adopción de ciertos requisitos a cumplir por los dispositivos y/o la previsión de ciertas obligaciones por los proveedores, así

actos legislativos de la Unión (disponible en https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=celex:52021PC0206; fecha de última consulta el 15 de noviembre de 2022). Esta propuesta de reglamentación se plantea con los siguientes objetivos específicos: "a) garantizar que los sistemas de IA introducidos y usados en el mercado de la UE sean seguros y respeten la legislación vigente en materia de derechos fundamentales y valores de la Unión; b) garantizar la seguridad jurídica para facilitar la inversión e innovación en IA; c) mejorar la gobernanza y la aplicación efectiva de la legislación vigente en materia de derechos fundamentales y los requisitos de seguridad aplicables a los sistemas de IA; y d) facilitar el desarrollo de un mercado único para hacer un uso legal, seguro y fiable de las aplicaciones de IA y evitar la fragmentación del mercado".

[39] Como se ha indicado anteriormente, con posterioridad a la entrega de la versión definitiva de este trabajo para su publicación, el Consejo y el Parlamento Europeo adoptaron respectivamente en noviembre-diciembre de 2022 y en junio de 2023 sus posiciones y enmiendas en relación con esta propuesta de Reglamento, de modo que a partir de junio de 2023 se han iniciado ya las negociaciones entre las distintas instituciones comunitarias europeas para así lograr que vea la luz esta norma que está llamada a ser el primer texto legal global regulador de la IA.

Ahora con este breve apunte simplemente se quiere actualizar la información sobre el estado en que se encuentra la tramitación y negociación de esta nueva norma. Se siguen clasificando los sistemas de IA en distintos niveles atendiendo al riesgo que de ellos se derivan respecto a diversos aspectos, pero cabe destacar que existen algunas discrepancias y diferencias en cuanto a la definición de "sistemas de IA", lo que deberá concretarse finalmente en la norma definitiva que se apruebe pues determinará su ámbito de aplicación.

La propuesta transaccional o posición adoptada por el Consejo se encuentra disponible en https://data.consilium.europa.eu/doc/document/ST-14954-2022-INIT/es/pdf (fecha de última consulta el 23 de julio de 2023) y el comunicado de prensa del Consejo de 6 de diciembre de 2022 está disponible en https://www.consilium.europa.eu/es/press/press-releases/2022/12/06/artificial-intelligence-act-council-calls-for-promoting-safe-ai-that-respects-fundamental-rights/ (fecha de última consulta el 23 de julio de 2023).

Por lo que se refiere al Parlamento Europeo, el texto de su posición negociadora sobre este Reglamento o Ley de Inteligencia Artificial aprobado en su reunión de 14 de junio de 2023 puede consultarse en https://www.europarl.europa.eu/doceo/document/TA-9-2023-0236_ES.html (fecha de última consulta el 23 de julio de 2023); y la nota de prensa emitida al respecto está disponible en https://www.europarl.europa.eu/news/es/press-room/20230609IPR96212/la-eurocamara-lista-para-negociar-la-primera-ley-sobre-inteligencia-artificial (fecha de última consulta el 23 de julio de 2023).

como la imposición de "transparencia a determinados sistemas de IA": considerando 14 de la propuesta)[40]:

1) En primer lugar, los sistemas prohibidos por suponer un riesgo inaceptable (por amenazar a la seguridad, la vida y derechos de las personas; entre otros casos, serían los sistemas que manipulan el comportamiento de los consumidores subliminalmente o buscan aprovecharse de sus vulnerabilidades).
2) Junto con estos sistemas prohibidos, estarían los sistemas de alto riesgo, entre los que cabe citar aquellos a través de los cuales podría negarse el acceso a la educación o formación profesional, o el acceso al crédito (como sistemas de *scoring* o de evaluación de solvencia financiera), o incluso en el ámbito de los seguros, por la posibilidad de riesgos de exclusión de los mismos que pudiera acontecer. Estos sistemas deben ajustarse a ciertas obligaciones antes de su comercialización, entre las que destacan, en lo que aquí nos interesa, la necesidad de contar con un registro de la actividad para garantizar la trazabilidad de los resultados; asimismo, contar con una documentación detallada sobre el sistema y su finalidad, y proporcionar esa información de modo claro, suficiente y adecuado al usuario.
3) En tercer lugar, estarían los sistemas de riesgo limitado, representados por el ejemplo de los *chatbots* y asistentes virtuales, en los que su empleo conlleva obligaciones específicas de transparencia, de modo que el usuario sea consciente de que está interactuando con una máquina, para así poder adoptar la decisión consciente e informada de continuar o no con el proceso.
4) Y, por último, nos encontraríamos con los sistemas de riesgo mínimo, que para la Comisión Europea serían la mayoría, y en los que apenas se interviene en la propuesta de Reglamento; aunque eso es así sin perjuicio de que sean aplicables algunas recomendaciones para las empresas que acudan a las mismas, entre las que debieran tenerse en cuenta fundamentalmente los principios de información y transparencia.

[40] Tal y como se indica en la exposición de motivos de esta propuesta reglamentaria, la misma "establece normas armonizadas para el desarrollo, la introducción en el mercado y la utilización de sistemas de IA en la Unión a partir de un enfoque proporcionado basado en los riesgos. [...] Asimismo, prohíbe determinadas prácticas particularmente perjudiciales de IA por ir en contra de los valores de la Unión y propone restricciones y salvaguardias específicas en relación con determinados usos de los sistemas de identificación biométrica remota con fines de aplicación de la ley".

El Parlamento Europeo ha estado trabajando sobre esta propuesta y se creó una "Comisión especial sobre Inteligencia Artificial en la Era Digital" (AIDA, por sus siglas en inglés), cuyo informe final se aprobó por el pleno en mayo de 2022. Con este informe se propone la hoja de ruta a seguir por la Unión Europea en esta materia, bajo un enfoque integral, mostrando que la IA será un impulso para la digitalización; esta hoja de ruta se centra en varias cuestiones[41]:

- La necesidad de un entorno normativo favorable respecto a la IA con normas flexibles, dinámicas y modernas, al entender que la actual legislación de la Unión y la nacional de los Estados miembros no ofrece seguridad jurídica al estar tan fragmentada y ser lenta en la adaptación al desarrollo de las nuevas tecnologías.
- La revisión y ampliación de la política y normativa sobre intercambio de datos en la Unión Europea.
- El refuerzo de la infraestructura digital para garantizar el acceso a los servicios digitales por todos los ciudadanos.
- En relación con lo anterior, se advierte de la importancia de apoyar desde la Unión la educación, formación y adquisición de competencias de IA, para que los ciudadanos obtengan habilidades suficientes y necesarias al respecto de cara a distintos aspectos laborales y de la vida en sociedad.
- Y, por supuesto, se muestra una gran preocupación por la seguridad de la IA (incluida, especialmente, la relacionada con aspectos militares)[42].

3. Tecnología *blockchain* y *smart contracts*

3.1. Notas características y definitorias

En el ámbito contractual, sin duda una de las piezas clave de esta revolución digital que se está viviendo actualmente son los *smart contracts* (también llamados contratos inteligentes) que basan su operatividad y funcionamiento especialmente en la tecnología *blockchain* (o cadena de bloques).

Esta tecnología *blockchain* viene a consistir en un registro codificado en cadena de las distintas transacciones o incidencias en el desarrollo de un con-

[41] Puede consultarse en detalle este informe en https://www.europarl.europa.eu/committees/es/aida/home/highlights y en https://www.europarl.europa.eu/cmsdata/246872/A9-0088_2022_EN.pdf (fecha de última consulta el 15 de noviembre de 2022).

[42] Con posterioridad, el 14 de junio de 2023 finalmente el Pleno del Parlamento Europeo aprobó su postura negociadora en el marco de la tramitación del Reglamento europea sobre IA, comenzando así la negociación entre las distintas instituciones europeas para lograr que esta norma vea la luz y sea aprobada definitivamente.

trato o una operación concreta que permiten el seguimiento de determinados activos digitales (fungibles o no fungibles).

Actúa como una gran base de datos a partir de nodos[43] (ordenadores terceros ajenos al contrato distribuidos en esa gran cadena y que comprueban la veracidad y realidad de las operaciones realizadas y su correspondencia con lo que las partes del contrato habían programado desde el inicio que debía suceder) y basándose en redes P2P (*peer to peer*, redes entre iguales o de pares), a través de las cuales se transmite información entre los nodos directa e inmediatamente. Realmente con la tecnología *blockchain* se reemplazan los intermediarios terceros de confianza que tradicionalmente dan fe de las operaciones realizadas por esta solución tecnológica a través de los nodos[44].

Esa cadena de bloques o base de datos está automatizada y descentralizada en cuanto es compartida y "replicada" entre los distintos nodos o participantes de esa red, lo que supone que no puede ser alterada; esa gran cadena de bloques ordenados cronológicamente funcionaría como un registro inmutable en el que se recogen todas las transacciones y operaciones ejecutadas en la red en relación con un concreto activo o contrato; transacciones en *blockchain* que requerirán del uso de claves públicas y privadas que permitan verificar la identidad de ese participante. A medida que se realicen operaciones

[43] En principio todos los nodos son iguales entre sí, aunque cumplen distintas funciones. Siguiendo a Argelich Comelles, pueden distinguirse: a) los nodos mineros, que emiten y propagan transacciones, ponen en funcionamiento el software necesario como BTCMiner, o CGMiner y crean una copia disponible para los mineros en la *blockchain*; b) los nodos *broadcast* o simples, que solamente emiten transacciones y reciben información de la *blockchain*; y c) los nodos maestros o completos, que permiten la descarga de una copia de la *blockchain* para pasar a ser un nodo más de la *blockchain* (Cristina Argelich Comelles, "*Smart contracts* o *Code is Law*: soluciones legales para la robotización contractual", *InDret*, núm. 2, abril, 2020, nota 62: disponible en https://indret.com/wp-content/uploads/2020/04/04-Argelich-numerat.pdf; fecha de última consulta el 15 de noviembre de 2022).

[44] Según señala Ríos López, los principales elementos de la tecnología *blockchain* son: "a) Un nodo: un ordenador personal o un superordenador, dependiendo de la complejidad de la red. Todos los nodos deben tener el mismo software o protocolo para comunicarse entre ellos, independientemente de la capacidad de cómputo; b) Un software o protocolo estándar: se trata de un software informático que ofrece un estándar común para que los nodos puedan comunicarse entre sí; c) Una red entre pares o de usuario-a-usuario o P2P (Peer-to-Peer), de forma que los nodos de la red se conectan directamente a una misma red; d) Un sistema descentralizado, pues no existe una parte intermediaria que ejerza el control en la red, de modo, que todos los ordenadores conectados a la red son los que la controlan, ya que no existe una jerarquía entre los nodos, son todos iguales entre sí" (Yolanda Ríos López, "La tutela del consumidor en la «contratación inteligente». Los «smart contracts» y la «blockchain» como paradigma de la Cuarta Revolución industrial", *Revista Consumo y Empresa*, *vLex*, núm. 9, enero, 2019, nota 6).

o transacciones, la cadena de bloques aumenta; cada bloque se identifica con una concreta operación realizada y verificada por los nodos, los cuales tras esa validación (en cuanto a correspondencia con lo que las partes del contrato habían programado desde el inicio que debía suceder) le atribuyen un código encriptado que vendría a actuar como huella digital única (*hash*), de modo que ese código se incorpora al registro y a ese concreto bloque y el registro ya permanece inalterable, pues el siguiente bloque de la cadena para poder crearse debe partir de ese *hash* o código único del bloque anterior[45].

Resulta a este respecto ilustrativa la explicación del funcionamiento de la tecnología *blockchain* que realiza Feliu Rey: "Imaginemos una mesa de reuniones alrededor de la cual se sienta un número significativo de personas. Cada una de estas personas (ordenadores o nodos conectados) tiene un libro de registro en blanco donde realiza anotaciones (sistema descentralizado). La primera anotación, sigamos con el ejemplo, es que A tiene 50 acciones y se las quiere transmitir a B. Primero se verifica que A tiene 50 acciones que puede transmitir (bloque con información), y se comprueba que todos los miembros de la mesa están de acuerdo con esta anotación inicial (sistema de verificación por consenso descentralizado). Luego se transmite a B. Como todos tienen en su libro que A es el titular y las puede transmitir, proceden a anotar la transmisión a B. Si A quiere volver a transmitir esas acciones, no podría porque ya no consta en el registro como titular y los miembros de la mesa al verificar tal información rechazarían la anotación, por lo que no permitirían esa transacción. Sólo B podría transmitir las acciones ulteriormente. Intentar una alteración de los registros, aunque no es imposible, exigiría un consenso de todos los miembros de la mesa y una modificación en todos los nodos de cadenas de bloques que recogen un tracto sucesivo, lo que resultaría, sin duda, altamente improbable"[46].

[45] *Vid*., entre otros, Cristina Argelich Comelles, "*Smart contracts...*", *op. cit*.; Cecilia Celeste Danesi, "Influencia algorítmica e inmutabilidad de los Smart Contracts: ¿cómo impactan estas tecnologías en la asimetría contractual?", *Actualidad Jurídica Iberoamericana*, núm. 16, febrero, 2022, pp. 1281 y ss. (disponible en https://revista-aji.com/wp-content/uploads/2022/04/54.-Cecilia-Celeste-Danesi-1270-1287.pdf; fecha de última consulta el 15 de noviembre de 2022); Carlos Domínguez Padilla, "La revolución blockchain y los smart contracts en el marco europeo", *Actualidad Jurídica Iberoamericana*, núm. 16, febrero 2022, pp. 1092 y ss. (disponible en https://revista-aji.com/wp-content/uploads/2022/04/45.-Carlos-Dominguez-1088-1109.pdf; fecha de última consulta el 15 de noviembre de 2022); Yolanda Ríos López, "La tutela...", *op. cit*.; y Jorge Feliu Rey, "Smart Contract: concepto, ecosistema y principales cuestiones de Derecho privado", *La Ley Mercantil*, núm. 47, 2018, pp. 14 y ss.

[46] Jorge Feliu Rey, "Smart Contract...", *op. cit*., p. 15.

Esta tecnología busca proporcionar confianza, seguridad e inmutabilidad a los contenidos de los contratos y a las propias transacciones, reduciendo riesgos y costes, como mejora de los contratos electrónicos más simples y tradicionales.

Las principales características de la tecnología *blockchain*, a la vista de lo indicado hasta ahora, son, por tanto:

a) Esta tecnología proporciona datos e información de muy diversa índole (identificación de personas, objeto, condiciones de la operación, lugar y tiempo de ejecución del contrato o de la transacción...) de modo inmediato que se comparte de forma directa y transparente entre los nodos o participantes que tienen acceso a esa *blockchain* y al registro o "libro mayor" que deriva de ella: la información parte de una única fuente fidedigna (ese "libro mayor" o registro en cadena de bloques descentralizado) y los usuarios participantes pueden ver todas las operaciones y transacciones realizadas con todo detalle en relación con ese contrato o activo.
b) Parte de un "libro mayor" a modo de base de datos distribuido (y por ello descentralizado) entre distintos participantes o nodos autorizados que tienen acceso al mismo y a todos los registros que en él se hacen de modo inmutable[47]. Esa inmutabilidad característica de la *blockchain* determina que no se puede alterar o eliminar datos de ese registro ni añadir contenido nuevo sin que haya una validación común de todos los nodos participantes: cada transacción u operación nueva que se realiza en la red debe venir validada por todos los nodos o, de lo contrario, no se incorporará a la cadena de bloques ni, por tanto, a ese registro inmutable; el consenso es, pues, unas de las características del funcionamiento de esta tecnología.

[47] Según señala Domínguez Padilla, en atención a los permisos requeridos para formar parte de una cadena de bloques (*blockchain*) y tener acceso a ese registro, esta *blockchain* puede ser de tres tipos: "A) Públicas. Donde cualquiera puede descargar en su ordenador los programas necesarios y constituir un nodo y participar en el proceso de consenso; cualquiera que sea parte podrá enviar transacciones a través de Internet las cuales se incluirán en la cadena de bloques. B) Federadas o de consorcio. En esta clase no permiten que cualquier persona pueda configurar un nodo en su PC y participar en el proceso de validación de las transacciones ya que se necesita permiso de acceso que suele concederse a los miembros de un determinado colectivo, por ejemplo: el colectivo de entidades financieras. C) Privadas. En estas cadenas de bloques las autorizaciones para poder realizar transacciones son concebidas por organizaciones privadas que determinarán qué condiciones permitirá la lectura de las transacciones realizadas" (Carlos Domínguez Padilla, "La revolución...", *op. cit.*, p. 1093).

c) La *blockchain* está protegida mediante criptografía asimétrica (de dos claves: una pública y otra privada vinculadas entre sí). Cada transacción u operación realizada (para lo que es preciso utilizar la clave pública y la privada) y validada en relación con un activo se registra en un bloque al que se le asigna un código único encriptado (con base en algoritmos), constituyendo su "huella digital" (el llamado *hash*) que es imposible de cambiar; cada bloque se une al anterior formando una cadena irreversible confirmando una secuencia segura de transacciones, a modo de tracto sucesivo, que no puede alterarse ni manipularse fácilmente.
d) Las transacciones sólo se realizan, validan y registran una vez, eliminando el riesgo de duplicidades. Y, además, ningún participante puede por sí sólo cambiar ni falsificar las operaciones una vez que se han grabado en ese "libro mayor" compartido y distribuido entre todos los participantes autorizados para ello (sólo los miembros autorizados para ese acceso). Si se hubiera producido algún error, se deberá hacer una nueva operación para subsanarlo y entonces deberá validarse de nuevo y registrarse como se hace con toda operación; pero ambas operaciones, la primera erróneamente efectuada y la nueva subsanando ese error, serán visibles para todos los participantes.
e) De todo lo anterior se observa que esa inmutabilidad e inalterabilidad de los registros y transacciones, unido a la existencia de códigos encriptados únicos y a la operatividad compartida y descentralizada de esta tecnología, genera mayor seguridad y confianza, y en principio, mayor eficiencia, especialmente si a esa *blockchain* se le une un "contrato inteligente" (*smart contract*) que, como se comentará a continuación, tiene como principal característica el que se ejecuta automáticamente.

Los *smart contracts* se han desarrollado fundamentalmente (aunque no necesariamente) a partir de esta tecnología *blockchain*, que permite que este contrato electrónico se programe bajo un código autoejecutable incorporado a esa cadena de bloques inmutable.

Su naturaleza es algo controvertida. Desde la perspectiva de la tecnología, un *smart contract* viene a ser un programa informático o protocolo de códigos informáticos (*scripts*) a través de los cuales un dispositivo tecnológico podrá ejecutar autónoma y automáticamente esas secuencias programadas previamente. Desde el punto de vista jurídico, como conjunto cabe calificarlos como negocios jurídicos o contratos derivados de la libre autonomía de la voluntad de las partes a partir del consentimiento prestado por ellas en relación con una causa y un objeto o prestación (es decir, serán contratos en la medida que

cumplan con los presupuestos exigidos por el art. 1261 CC)[48]. Pero realmente la especialidad de esta figura contractual se centraría en el hecho de que esa relación contractual y un contrato que podría ser "tradicional" acoge una forma electrónica particular con base en el lenguaje informático y los elementos tecnológicos empleados para lograr que ese contrato sea autoejecutable, sin necesidad de intervención humana, a partir de programas informáticos o *scripts* (incluso ese contrato puede configurarse sólo parcialmente autoejecutable y ser *smart contract* sólo en parte, en relación con ciertas cláusulas y condiciones que sí se han codificado): una vez verificado (en tecnología *blockchain*, con apoyo en los nodos) que se han cumplido ciertas condiciones o eventos programados previamente por las partes en el contrato electrónico (comprobación sin necesidad de intervención humana), el dispositivo tecnológico virtual procede a ejecutar autónomamente el código o *script* programado de la operación y automáticamente ejecuta las prestaciones a las que se obligaron las partes en ese contrato y que dependían del cumplimiento de esas condiciones o eventos (cumplimiento que ya fue verificado); el dispositivo, conforme a las reglas o algoritmos prefijados al inicio en la fase contractual, procesa los datos de modo automatizado según esas órdenes programadas y, dado que los datos le indican que ciertas condiciones contractuales se han cumplido, procede a ejecutar automáticamente la consecuencia de ese hecho, es decir, las prestaciones contractuales vinculadas al cumplimiento de esa condición o hecho; y una vez que ese código se ejecuta, dicha ejecución es

[48] En relación con esto, cabría distinguir entre dos tipos de *smart contracts*: el *smart code contract* y el *smart legal contract*. Siguiendo a Legerén-Molina, "los primeros aluden a las secuencias de código que son todo o parte de un acuerdo existente y que están almacenadas, verificadas y ejecutadas en una cadena de bloques, con las peculiaridades y características que de ello se derivan. Los segundos [...] constituyen secuencias de código que expresan todo o parte de un acuerdo, pero no están registradas ni se ejecutan en una cadena de bloques; serían, por tanto, una alternativa más próxima a los «contratos tradicionales», usualmente recogidos por escrito. [...] Los dos subtipos o categorías de contratos inteligentes mencionados tienen en común la existencia de acuerdos implementados mediante código que se ejecutan de manera automática una vez se cumplan los presupuestos preestablecidos. En tal sentido, ambos son *code* y ambos son *legal*: tanto los que se ejecutan en la cadena de bloques como fuera de ella están conformados por una secuencia de código y también ambos, en tanto expresión o modo de implementar un acuerdo existente entre partes, son *legal*; o de manera más precisa, pueden producir efectos jurídicos. La distinción mencionada reside, entonces, tanto en el soporte donde se verifican –dentro o fuera de la cadena de bloques– como en las características o consecuencias que pueden producir y que, en lo que ahora se alude, se derivan de tal «ubicación» y no de la secuencia en sí misma" (Antonio Legerén-Molina, "Los contratos inteligentes en España. La disciplina de los *smart contracts*", *Revista de Derecho Civil,* vol. V, núm. 2, abril-junio, 2018, pp. 196-198; disponible en https://www.nreg.es/ojs/index.php/RDC/article/view/320/267; fecha de última consulta el 15 de noviembre de 2022). Cfr. a este respecto Carlos Tur Faúndez, *Smart contracts: análisis jurídico*, Reus, Madrid, 2018, pp. 139-141.

irrevocable (o lo es con gran dificultad técnica), no siendo posible detenerla cuando ese código se ha iniciado.

Por tanto, realmente podemos considerar a los *smart contracts* como una especialidad dentro de un contrato que podría configurarse del modo tradicional que todos conocemos, pero que se ha decidido dotarle, total o parcialmente, de cierto contenido codificado informáticamente y de un programa de ejecución automática; puede ser un contrato de venta, de suministro, de depósito, de seguro... que podría configurarse como tradicionalmente se ha hecho hasta ahora, pero las partes han decidido utilizar herramientas informáticas para su ejecución automática, por considerarlo así más seguro y fiable, y de esta forma se ha convertido en un *smart contract* por ese uso de programas y códigos informáticos.

Ahora bien, debe tenerse en cuenta que por su naturaleza ahora mismo estos contratos inteligentes tendrían un ámbito de aplicación algo limitado en atención a la automaticidad que los caracteriza: no todos los acuerdos de voluntades ni relaciones jurídicas contractuales ni prestaciones pueden convertirse a lenguaje codificado; los cumplimientos de las condiciones programadas deben ser totales y objetivamente verificables para que después se autoejecute la prestación prevista (pago del precio al entregarse la cosa; entrega de la cosa al pagarse el precio; devolución de la cosa al pagarse el precio del depósito o, al revés, autoejecución del pago del depósito al verificarse la devolución del bien depositado; pago de una indemnización automáticamente al verificarse el retraso en la llegada del destino del tren o avión o, en general, la producción del siniestro asegurado...).

De hecho, actualmente, y en tanto no se desarrolle más este tipo de herramientas tecnológicas de autoejecución o el "Internet de las cosas", lo más probable es que los contratos sólo sean parcialmente *smart contracts*: es decir, que sólo parte de sus cláusulas y contenido se configuren bajo lenguaje informático codificado y prevean la ejecución automática de alguna prestación (normalmente, una prestación de pago).

Siguiendo a Feliu Rey, "podemos entender la figura de los Smart Contracts como una forma de articular un proceso contractual, de facilitar el desarrollo y consumación del contrato, o de las posibles consecuencias que se derivan del incumplimiento de aquel. De modo que, no sólo dota de una especial forma al acuerdo, la electrónica o digital, sino también, gracias al lenguaje utilizado, permite que las fases, en su caso, de concreción y cumplimiento de las obligaciones se realicen de forma automática, íntegra o parcialmente, sin intervención humana"[49].

[49] Jorge Feliu Rey, "Smart Contract...", *op. cit.*, p. 14.

Un aspecto esencial y a la vez complejo de estos contratos así configurados es la programación inicial; es decir, trasladar a lenguaje informático y códigos las distintas condiciones, cláusulas y contenido del contrato, y crear o programar esos códigos autoejecutables o *scripts*; los errores o fallos en la programación, así como cualquier fallo de seguridad, puede dar lugar a errores de ejecución, pues debe tenerse en cuenta que lo que se ejecute se corresponderá exactamente con lo estrictamente programado. A partir de esa fase de programación, el desarrollo del contrato se prevé sencillo en su operativa, pues está totalmente automatizado, concibiéndose de esta forma como un contrato más seguro dificultando el incumplimiento o las controversias al respecto, además de suponer el ahorro de ciertos costes de ejecución tradicionales por esa automaticidad, especialmente en cuanto a los relacionados con las reclamaciones por cumplimiento o incumplimiento[50].

Según señala Ríos López, existirían seis fases en la creación de un *smart contract*: " (a) Las partes muestran la voluntad de alcanzar un acuerdo; (b) Determinación de las condiciones del contrato, por las partes, o a partir de un hecho externo; (c) Se escribe el código informático que permitirá la ejecución automática cuando sucedan los hechos previstos en el mismo; (d) Cada una de las transacciones se encripta mediante sistemas de autenticación y verificación seguros en la cadena de bloques; (e) Ejecución y procesamiento: cada una de las transacciones registrada en el bloque se verifica por el sistema de consenso, ejecutándose automáticamente la prestación; (f) Ejecutada la prestación, todos los nodos del sistema la reconocen como tal, siendo inalterable en la cadena de bloques"[51].

[50] En relación con los *smart contracts, vid*., entre otros, Carlos Tur Faúndez, *Smart contracts*..., *op. cit*.; Antonio Legerén-Molina, "Los contratos...", *op. cit*., pp. 194 y ss.; Jorge Feliu Rey, "Smart Contract...", *op. cit*., pp. 2 y ss.; Cristina Argelich Comelles, "*Smart contracts*...", *op. cit*.; Marina Echebarría Sáenz, "Contratos electrónicos autoejecutables (Smart contract) y pagos con tecnología Blockchain", *Revista de Estudios Europeos*, núm. 70, julio-diciembre, 2017, pp. 70 y ss. (disponible en https://uvadoc.uva.es/handle/10324/28434; fecha de última consulta el 15 de noviembre de 2022); Yolanda Ríos López, "La tutela...", *op. cit*.; y Aitor Mora Astaburuaga, "*Smart contracts*. Reflexiones sobre su concepto, naturaleza y problemática en el Derecho contractual", *Revista de Derecho UNED*, núm. 27, 2021, pp. 60 y ss. (disponible en https://revistas.uned.es/index.php/RDUNED/article/view/31068/23490; fecha de última consulta el 15 de noviembre de 2022).

[51] Yolanda Ríos López, "La tutela...", *op. cit*., nota 10. La concreción de estas fases las ha realizado esta autora tomando como base el esquema desarrollado en el Libro Blanco "*Smart Contracts: 12 Uses Cases for Business and Beyond. A Technology, Legal and Regulatory Introduction*", elaborado por la *Smart Contracts Alliance* y la *Chamber of Digital Commerce*, diciembre de 2016, p. 12 (documento disponible en https://d3h0qzni6h08fz.cloudfront.net/Smart-Contracts-12-Use-Cases-for-Business-and-Beyond_Chamber-of-Digital-Commerce.pdf; fecha de última consulta el 15 de noviembre de 2022); ahora bien, cabe advertir que la *Chamber of Digital Commerce* en septiembre de 2018 publicó un

3.2. Problemática jurídico civil que se plantea en torno a esta tecnología en el ámbito contractual y con los consumidores y usuarios

Estos contratos con base en la tecnología *blockchain*, en lenguaje informático y con códigos autoejecutables plantean diversos problemas en el ámbito contractual, pudiendo destacar los siguientes[52]:

a) Esa automaticidad propia de los *smart contracts*, como se ha indicado, exige que los cumplimientos de las condiciones programadas sean totales y objetivamente verificables: en estos contratos no caben cláusulas interpretables bajo criterios como la buena fe, la diligencia desarrollada o la concurrencia de caso fortuito o fuerza mayor ni modulaciones de efectos bajo supuestos de circunstancias sobrevenidas imprevisibles o incumplimientos parciales, de modo que resulta un contrato más inflexible por la inmutabilidad de la tecnología *blockchain* y el contenido registrado en ella. Incluso, puede que no se ejecute el contrato pese a que el cumplimiento se ha llevado a cabo, pero éste no fue exactamente del modo en que se programó: la inflexibilidad de la programación puede llevar en tal caso a ir en contra de lo querido por las partes.
b) Los hechos que permitan ejecutar consecuencias y prestaciones automáticas deben ser objetivos, y además, resulta más factible si las prestaciones que se autoejecutan son prestaciones de dar (no de hacer) y específicas (no genéricas): por ejemplo, retraso en la llegada del tren o del avión a su destino, de modo que, verificada esa condición objetiva de retraso en un determinado período o períodos de tiempo, el ordenador ejecuta el programa autónomamente y procede automáticamente a atribuir a los usuarios la prestación de indemnización por ese retraso. En este sentido, pueden surgir problemas jurídicos relevantes si se programan consecuencias automáticas vinculadas a eventos cuyo cumplimiento no es tan fácil de verificar, pues existen circunstancias que exigirían flexibilizar las consecuencias y esas circunstancias no se han previsto en la programación.

segundo Libro Blanco sobre los *smart contracts* preparado a iniciativa de ésta, al igual que el anterior libro blanco citado, por la *Smart Contracts Alliance*; este segundo Libro Blanco "*Smart contracts: is the Law Ready?*" (puede consultarse en https://digitalchamber.s3.amazonaws.com/Smart-Contracts-Whitepaper-WEB.pdf; fecha de última consulta el 15 de noviembre de 2022.

[52] Cfr., entre otros, Carlos Tur Faúndez, *Smart contracts...*, *op. cit.*; Antonio Legerén-Molina, "Los contratos...", *op. cit.*, pp. 194 y ss.; Jorge Feliu Rey, "Smart Contract...", *op. cit.*, pp. 2 y ss.; Cristina Argelich Comelles, "*Smart contracts...*", *op. cit.*; Marina Echebarría Sáenz, "Contratos electrónicos...", *op. cit*, pp. 70 y ss.; Yolanda Ríos López, Y., "La tutela...", *op. cit.*; y Aitor Mora Astaburuaga, "*Smart Contracts...*", *op. cit*, pp. 60 y ss.

c) Precisamente los posibles errores en la programación (al no volcar correctamente al lenguaje informático y codificado las cláusulas previstas o las prestaciones a ejecutar) pueden plantear relevantes conflictos jurídicos en relación con el contenido contractual, la conformidad del consentimiento prestado con lo señalado en lenguaje informático y, por ello, también en relación a la ejecución automática del contrato y su cumplimiento que puede no corresponderse con lo efectivamente querido; y es que lo que se ejecute se corresponderá exactamente con lo estrictamente programado. Ello suscitará asimismo la controversia de la concreción de responsabilidades que puedan surgir al respecto, especialmente teniendo en cuenta la automaticidad de la ejecución de las prestaciones y su irreversibilidad.
d) En relación con lo anterior pueden generarse problemas en la verificación del cumplimiento de las condiciones a las que se aúna la consecuencia ejecutable automáticamente. Normalmente para ello se prevé el uso de los llamados oráculos (fuentes de información externa que permiten verificar esos cumplimientos). En la concreción de cuáles deben ser esos oráculos, su inserción en el programa informático y respecto a su fiabilidad pueden generarse ciertos conflictos; y, asimismo, en relación con la información proporcionada por estos oráculos, o su interpretación y aplicación a la concreta relación contractual que determine la ejecución automática de alguna prestación, pueden producirse errores que generen daños y responsabilidad al respecto que deba concretarse.
e) Por otro lado, cabe plantearse problemas en relación con la eficacia y validez del contrato y sus cláusulas y las consecuencias de una posible ineficacia del contrato con retroactividad de las prestaciones efectuadas, estando como están éstas registradas en una cadena de bloques (*blockchain*) cuya principal característica es la inmutabilidad e inalterabilidad de lo registrado. Del mismo modo habrá que atender a cuestiones vinculadas con el ejercicio del derecho de desistimiento por el consumidor, en su caso, o cuestiones relacionadas con el incumplimiento, total o parcial, de las prestaciones y la posible resolución de los contratos o el ejercicio de otros remedios ante la posible ineficacia o invalidez contractual. En estos casos, aparte del ejercicio de las oportunas acciones y sus efectos, cabe destacar la problemática que pueda surgir en atención al destino y tratamiento de los datos objeto de la encriptación y codificación cuando ese contrato ya no produzca

efectos (por resolución o extinción del contrato por cualquier causa), así como hacer efectivo el "derecho al olvido" digital de las partes.

f) Igualmente cabría plantearse la posibilidad de modificación del contrato codificado por cambio de opinión o parecer de las partes o por alteración sobrevenida de las circunstancias que les llevó inicialmente a celebrar ese contrato, o incluso ante la observación de un error que se quiera rectificar en el contenido del contrato o alguna otra causa justificada: la tecnología *blockchain* y la codificación informática del *smart contract* le dota de una inmutabilidad y una inflexibilidad que puede suponer ciertos problemas prácticos.
g) Asimismo, siendo muy probable que el empleo de estos contratos opere bajo la forma de contratos de adhesión, resulta relevante en este ámbito el control del cumplimiento de los oportunos deberes de información precontractual al consumidor en cuanto al funcionamiento y operatividad de estos contratos autoejecutables y sus consecuencias; así como, por supuesto, el control de las condiciones generales de la contratación y de cada cláusula del contrato, buscando la legalidad y la transparencia en el contenido y las consecuencias del contrato y evitando cláusulas oscuras y abusivas (y de haberlas, habrá que atender a su nulidad y demás consecuencias, como la retroactividad de prestaciones, que puede ser más difícil, como se comentó, en un contexto de tecnología *blockchain*).
h) Por último, si en la contratación se han incorporado elementos de inteligencia artificial, cabría traer aquí también los retos jurídicos vinculados a ésta que se han indicado en relación con el uso de datos y algoritmos y la explicabilidad de todo ello, con base en la información y transparencia.

4. Conclusiones

I. El desarrollo de las nuevas tecnologías es imparable y se produce muy rápidamente, habiendo evolucionado enormemente hacia una realidad digital presidida por los grandes avances tecnológicos que han llevado a una gran transformación digital de la sociedad y la economía, ante lo que el Derecho no puede quedar impasible.

Es necesario contar con una regulación adecuada; revisar y reformar lo necesario y aprobar las normas precisas para atender a esa realidad, sus problemas, controversias y retos, protegiendo a los usuarios, la economía y la competencia y, sobre todo, proporcionando seguridad jurídica y seguridad en general. La seguridad en esta realidad digital es esencial.

II. La transformación digital que se está viviendo es también una prioridad legislativa y política en la Unión Europea, que ha mostrado desde hace años cómo la transformación digital y el mercado único digital son una de sus principales prioridades.

III. Una de las máximas expresiones y reflejo del gran y rápido avance de las nuevas tecnologías es el progreso y desarrollo en los últimos años de la llamada inteligencia artificial (IA).

Esta tecnología también plantea riesgos y aún son varios los retos a los que se enfrenta en su desarrollo. Entre esos retos está principalmente el de conseguir confianza y seguridad para los usuarios en torno a esos procesos automatizados, especialmente en la toma de decisiones y el análisis y uso de los datos. Contar con una IA confiable llevará a promover sus aspectos positivos y a tratar de evitar o mitigar los aspectos negativos y los riesgos que su uso conlleva. Para lograr esa confianza en la inteligencia artificial son pilares esenciales, además de la seguridad que ofrezcan los dispositivos para hacer frente a posibles ataques o ciberataques: a) el uso de los modelos de IA con criterios de ética y justicia, para evitar resultados discriminatorios por el uso de datos y algoritmos sesgados; y b) la explicabilidad de los algoritmos utilizados en esos procesos y los efectos de las decisiones automatizadas con base en ellos. Transparencia e información son principios esenciales en el marco general de la IA y su fiabilidad o confiabilidad.

IV. En el ámbito contractual, sin duda una de las piezas clave de esta revolución digital que se está viviendo actualmente son los *smart contracts*, que basan su operatividad y funcionamiento especialmente en la tecnología *blockchain* (o cadena de bloques), caracterizada por la inmutabilidad de los registros encadenados. Los *smart contracts* se han desarrollado fundamentalmente (aunque no necesariamente) a partir de esta tecnología *blockchain*, que permite que este contrato electrónico se programe bajo un código autoejecutable incorporado a esa cadena de bloques inmutable.

Estos contratos con base en la tecnología *blockchain*, en lenguaje informático y con códigos autoejecutables plantean diversos problemas en el ámbito contractual, principalmente derivados de la automaticidad que caracteriza a dichos contratos inteligentes y a la necesaria programación de las condiciones, códigos, etc., que debe hacerse correctamente. Pero, además, y especialmente en la medida que se configuren como contratos de adhesión, resulta muy importante en el marco de la protección de los consumidores que la transparencia del contenido y efectos del contrato y el cumplimiento de los

oportunos deberes de información precontractual sean principios rectores en cuanto al funcionamiento y operatividad de estos contratos y sus consecuencias.

Bibliografía

Antón Juárez, Isabel, "Personalización de precios a través de la inteligencia artificial y el Big Data", en Paniagua Zurera, Manuel (dir.), *El sistema jurídico ante la digitalización: estudios de Derecho privado*, Tirant lo Blanch, Valencia, 2021, pp. 379-416.

Argelich Comelles, Cristina, "*Smart contracts* o *Code is Law*: soluciones legales para la robotización contractual", *InDret*, núm. 2, abril, 2020, nota 62: disponible en https://indret.com/wp-content/uploads/2020/04/04-Argelich-numerat.pdf; fecha de última consulta el 15 de noviembre de 2022).

Barrio Andrés, Moisés, "La nueva Estrategia digital de la Comisión Europea: primeras impresiones", *Diario La Ley*, núm. 37, sección Ciberderecho, 20 de febrero de 2020.

Bendito Cañizares, Mª Teresa, "Ética, algoritmos y decisiones individuales y colectivas automatizadas", en Cuena Casas, Matilde e Ibáñez Jiménez, Javier (dirs.), *Perspectiva legal y económica del fenómeno FinTech*, Wolters Kluwer La Ley, Madrid, 2021, pp. 641-686.

Danesi, Cecilia Celeste, "Influencia algorítmica e inmutabilidad de los Smart Contracts: ¿cómo impactan estas tecnologías en la asimetría contractual?", *Actualidad Jurídica Iberoamericana*, núm. 16, febrero, 2022, pp. 1270-1287 (disponible en https://revista-aji.com/wp-content/uploads/2022/04/54.-Cecilia-Celeste-Danesi-1270-1287.pdf; fecha de última consulta el 15 de noviembre de 2022).

De la Nuez Sánchez-Cascado, Elisa, "Inteligencia artificial y transparencia. Especial referencia a su utilización en el ámbito de las Administraciones Públicas", *El Notario del Siglo XXI*, núm. 93, septiembre-octubre, 2020 (disponible en https://www.elnotario.es/hemeroteca/revista-93/10160-inteligencia-artificial-y-transparencia-especial-referencia-a-su-utilizacion-en-el-ambito-de-las-administraciones-publicas; fecha de última consulta el 15 de noviembre de 2022).

Domínguez Padilla, Carlos, "La revolución blockchain y los smart contracts en el marco europeo", *Actualidad Jurídica Iberoamericana*, núm. 16, febrero 2022, pp. 1088-1109 (disponible en https://revista-aji.com/wp-content/uploads/2022/04/45.-Carlos-Dominguez-1088-1109.pdf; fecha de última consulta el 15 de noviembre de 2022).

Echebarría Sáenz, Marina, "Contratos electrónicos autoejecutables (Smart contract) y pagos con tecnología Blockchain", *Revista de Estudios Europeos*, núm. 70, julio-diciembre, 2017, pp. 69-97 (disponible en https://uvadoc.uva.es/handle/10324/28434; fecha de última consulta el 15 de noviembre de 2022).

Feliu Rey, Jorge, "Smart Contract: concepto, ecosistema y principales cuestiones de Derecho privado", *La Ley Mercantil*, núm. 47, 2018.

Legerén-Molina, Antonio, "Los contratos inteligentes en España. La disciplina de los *smart contracts*", *Revista de Derecho Civil*, vol. V, núm. 2, abril-junio, 2018, pp. 193-241 (disponible en https://www.nreg.es/ojs/index.php/RDC/article/view/320/267; fecha de última consulta el 15 de noviembre de 2022).

Mora Astaburuaga, Aitor, "*Smart contracts*. Reflexiones sobre su concepto, naturaleza y problemática en el Derecho contractual", *Revista de Derecho UNED*, núm. 27, 2021, pp. 57-97. (disponible en https://revistas.uned.es/index.php/RDUNED/article/view/31068/23490; fecha de última consulta el 15 de noviembre de 2022).

Mozo Seoane, Antonio, *Los límites de la tecnología. Marco ético y regulación jurídica*, Reus, Madrid, 2021.

Pascual Huerta, Pablo, "Algoritmos y protección de datos personales", en Cuena Casas, Matilde e Ibáñez Jiménez, Javier (dirs.), *Perspectiva legal y económica del fenómeno FinTech*, Wolters Kluwer La Ley, Madrid, 2021, pp. 559-571.

Peguera Poch, Miguel, "En búsqueda de un marco normativo para la inteligencia artificial", en Cerrillo i Martínez, Augustí y Peguera Poch, Miguel (coords.), *Retos jurídicos de la inteligencia artificial*, Thomson Reuters Aranzadi, Cizur Menor, 2020, pp. 41-56.

Ríos López, Yolanda, "La tutela del consumidor en la «contratación inteligente». Los «smart contracts» y la «blockchain» como paradigma de la Cuarta Revolución industrial", *Revista Consumo y Empresa*, *vLex*, núm. 9, enero, 2019.

Robles Martín-Laborda, Antonio, "Inteligencia artificial y personalización de precios", en Cuena Casas, Matilde e Ibáñez Jiménez, Javier (dirs.), *Perspectiva legal y económica del fenómeno FinTech*, Wolters Kluwer La Ley, Madrid, 2021, pp. 573-598.

Tur Faúndez, Carlos, *Smart contracts: análisis jurídico*, Reus, Madrid, 2018.

Zurita Martín, Isabel, "Las propuestas de reforma legislativa del Libro Blanco europeo sobre inteligencia artificial en materia de seguridad y responsabilidad civil", *Actualidad Jurídica Iberoamericana*, núm. 14, febrero, 2021, pp. 438-487 (disponible en https://idibe.org/wp-content/uploads/2021/03/11._Isabel_Zurita_pp._438-487.pdf; fecha de última consulta el 15 de noviembre de 2022).

Smart Contracts, Inteligência Artificial e proteção do consumidor*

Smart Contracts, Artificial Intelligence and consumer protection

Fernanda de Araujo Meirelles Magalhães**

RESUMO: Os *Smart Contracts*, de facto, são promissores, mas apresentam algumas questões, como aquelas relacionadas com as relações de consumo. Enquanto uns defendem que os *Smart Contracts* poderiam corrigir a disparidade destas relações, outros, em sentido oposto, apontam a redução da proteção das partes mais frágeis, já que os consumidores, efetivamente, possuem pouca interferência na criação do contrato. Nestes casos, fala-se na utilização da IA para garantir a flexibilização e adaptação do contrato, evitando comportamentos desleais e oportunistas, trazendo maior segurança jurídica para as partes. Deste modo, tenciona-se debater de que forma o uso da IA se expressaria nos *Smart Contracts* de consumo para ajudar na proteção dos consumidores.

PALAVRAS-CHAVE: *Smart Contracts*; *Blockchain*; Relações de Consumo; Inteligência Artificial.

* Este trabalho foi elaborado no âmbito da Bolsa de Investigação para Doutoramento na área de Direito, contrato-programa plurianual de unidades de I&D 2020-2023 entre a FCT, a FDUP e o CIJ e do Projeto "It's a wonderful (digital) world": O direito numa sociedade digital e tecnológica, do CIJ – Centro de Investigação Jurídica, da Faculdade de Direito da Universidade do Porto. Texto traduzido e adaptado da versão original inglesa: Fernanda de Araujo Meirelles Magalhães, "Smart Contracts, Artificial Intelligence and Consumer Protection", *in Consumer Protection in the European Union: Challenges and Opportunities*, Cayetana Santaolalla Montoya (Dir. and Coord.), Luxembourg, European Comission, 2023.

** Doutoranda em Direito, Faculdade de Direito da Universidade do Porto (FDUP). Investigadora Bolseira, Centro de Investigação Jurídica (CIJ). fernandaa.magalhaes@hotmail.com

ABSTRACT: Although Smart Contracts are promising, they present some issues, such as those related to consumer relations. While some people defend that Smart Contracts could correct these relationships, others, on the other hand, point out the reduction of the protection of the more vulnerable parties, since consumers, effectively, barely interfere in the creation of the contract. In those cases, the use of Artificial Intelligence is discussed to ensure the flexibility and adaptation of the contract, providing greater legal certainty for the parties. Therefore, this work intends to discuss how the use of AI would be expressed in consumer Smart Contracts to help in the protection of the consumers.

KEYWORDS: Smart Contracts; Blockchain; Consumer relations; Artificial Intelligence.

1. Introdução

Nos últimos anos, principalmente nas últimas duas décadas, as inovações digitais passaram a ter uma presença mais significativa não só na vida quotidiana, mas no Direito. Não só em razão da velocidade com que as tecnologias se desenvolveram, mas, também, pelo desconhecimento em relação às suas verdadeiras funcionalidades e validade, as transformações tecnológicas dão origem a diversas incertezas, relacionadas com os mais variados aspetos jurídicos tradicionais, como uma simples assinatura. Com o surgimento da *internet*, vimo-nos completamente absorvidos pela sua utilização e, aquilo que antes era receio, passou a ser parte comum da vida dos operadores do direito, *v.g.* o uso de e-mails para troca de informações importantes[1]. Esta influência da acelerada evolução tecnológica no Direito também pode ser observada nos processos eletrónicos e na contratação eletrónica que, apesar de já existir há

[1] Manuel Santos Vítor, "Inteligência Artificial e Contratos", *Inteligência Artificial & Direito*, Manuel Lopes Rocha; Rui Soares Pereira (coord.), Coimbra, Almedina, 2020, pp. 222-223.

tempos, devido aos acontecimentos recentes[2], vem crescendo e evoluindo exponencialmente.

Os *Smart Contracts* aparecem dentro deste contexto de Digitalização – que proporcionou também o desenvolvimento de outras tecnologias, como a *blockchain* e a Inteligência Artificial – e do aumento do recurso aos acordos eletrónicos. No fundo, os Contratos Inteligentes recebem esta denominação "Smart", mas, apesar de, usualmente, relacionarmos o uso da palavra às tecnologias que envolvem a utilização Inteligência Artificial (IA), como os *Smartphones* ou *Smartwatches*, não foi este o motivo da utilização do termo. Szabo utilizou-se desta expressão, pois, na altura em que a tecnologia foi criada, acreditava que os *Smart Contracts* eram muito mais eficientes que os contratos tradicionais, já que dispensavam a intervenção humana para sua execução, sendo "inteligentes" o suficiente para sua autoexecução[3].

Em verdade, sob uma perspetiva geral, os *Smart Contracts*, *per se*, também não poderiam ser chamados de contratos, no sentido conhecido por nós, juristas[4]. Isto posto que, na maioria dos casos, os *Smart Contracts* estão relacionados com a execução de tarefas informáticas, que não possuem relação com o cumprimento de obrigações, como, por exemplo, o envio de uma mensagem. Assim sendo, podemos afirmar que os *Smart Contracts* não são, em princípio, contratos, mas ferramentas digitais baseadas em códigos ou protocolos de computador que facilitam o cumprimento de ordens materiais automaticamente, sem a intervenção de terceiros, podendo vir a desempenhar – ou não – a função de contrato[5].

[2] Aqui nos referimos à pandemia do COVID-19. De facto, mesmo antes da pandemia, já vinha sendo notada a presença da tecnologia no nosso *modus vivendi*, mas, sem dúvida, nestes últimos dois anos vemos a importância da tecnologia para executar desde as tarefas mais simples, às tarefas mais complexas através do ambiente eletrónico. Nota-se, portanto, que o uso da tecnologia já não é mais questionável, o que passa a se questionar é a melhor maneira de a aplicar e obter vantagens do seu uso em prol da sociedade.

[3] Rory Unsworth, "Smart Contract This! An Assessment of the Contractual Landscape and the Herculean Challenges it Currently Presents for "Self-executing" Contracts", *Legal Tech, Smart Contracts and Blockchain* Marcelo Corrales Compagnucci; Mark Fenwick; Helena Haapio, Singapore, Springer, 2019, pp. 19-20.

[4] Cf. Fernanda de Araujo Meirelles Magalhães, "Os Smart Contracts e o papel dos tribunais em matéria de Direito dos Contratos", *Estudos Jurídicos Sobre Inteligência Artificial e Tecnologias*, Fábio da Silva Veiga; Cátia Marques Cebola; Susana Sardinha Monteiro (Coords.), Porto, IBEROJUR, 2022, pp. 146 e ss. Manuel Santos Vítor, "Inteligência Artificial" cit., pp. 225 e ss.

[5] Javier W. Ibáñez Jiménez, *Derecho de Blockchain y de la Tecnología de Registros*, Cizur Menor (Navarra), Aranzadi, 2018 p. 96.

A ferramenta chave, grande influenciadora do desenvolvimento dos *Smart Contracts*, foi a tecnologia DLT[6] (*Distributed Ledger Technology*) *blockchain* (cadeia de blocos). A cadeia de blocos teve – e continua a ter – grande impacto e repercussão, pois armazena transações e informações de forma distribuída e tendencialmente imutável, dificultando a possibilidade de fraude de informações dentro da rede, proporcionando maior segurança para a utilização dos *Smart Contracts* e confiança no meio digital.

Ao mesmo tempo que a *blockchain* trouxe a fiabilidade que os *Smart Contracts* precisavam para o seu bom desempenho, trouxe consigo muitas dúvidas e preocupações, principalmente, relacionadas com a sua utilização em determinado tipo de relações contratuais. As principais questões associadas à utilização dos *Smart Contracts* – a maioria delas devido à sua autoexecutabilidade e potencial imutabilidade, proporcionada pela *blockchain* – envolvem o seu uso naqueles acordos jurídicos mais frágeis, nos quais costuma-se haver maior vulnerabilidade de uma das partes do contrato, como é o exemplo das relações de consumo B2C (*Business-to-Consumer*)[7].

Por um lado, os mais otimistas em relação ao tema defendem que o modo de funcionamento dos *Smart Contracts* possui a capacidade de corrigir a disparidade das relações de consumo e ajuda o consumidor a garantir que seus direitos sejam atendidos. Em contrapartida, há quem argumente que, na verdade, há não um reforço, mas uma redução da proteção das partes mais frágeis, já que os consumidores, efetivamente, possuem pouca interferência na criação do contrato e, ainda que pudessem interferir, não possuem o conhecimento necessário para influenciar na construção dos *Smart Contracts*[8].

Com o propósito de responder estes e outros debates que envolvem o uso dos Contratos Inteligentes, estuda-se não só o uso de ferramentas acessórias – *v.g.* os "oráculos", mas também a aplicação da Inteligência Artificial para melhoria do funcionamento e aplicação do raciocínio humano/jurídico. Neste

[6] As tecnologias distribuídas de livro-razão são uma espécie de "registos de transações digitais e distribuídos que armazenam blocos de dados compartilhados em uma rede de nós de computador". Patrick Laurent; Stéphane Hurtaud; Thibault Chollet; Sébastien Genco, "Distributed Ledger Technologies services: using the power of blockchain", *Deloitte*, 2017, pp. 1-8. Disponível em https://www2.deloitte.com/content/dam/Deloitte/lu/Documents/technology/lu-blockchain-services-21092017.pdf (15.01.2022)

[7] Neste sentido ver Stefan Wrbka, Stefan, "A Multilayer Safeguard Mechanism to Optimise the Potential of Smart Contracts in B2C Transactions", *Smart Contracts – Technological, Business and Legal Perspectives*, Marcelo Corrales Compagnucci; Mark Fenwick; Stefan Wrbka, Oxford, Hard Publishing, 2021, pp. 123-144.

[8] Stefan Wrbka, "A Multilayer" cit., pp. 123-124.

sentido, o propósito é tentar garantir a flexibilização e adaptação dos contratos realizados através dos *Smart Contracts* e evitar comportamentos desleais e oportunistas, garantindo maior segurança jurídica para as partes.

Assim sendo, o presente texto ocupa-se em discutir de que forma o uso da Inteligência Artificial se expressaria nos *Smart Contracts* de consumo e seus reflexos no Direito, principalmente, para as relações jurídico-consumeristas.

2. Os *Smart Contracts*, a tecnologia *blockchain* e as relações de Consumo

2.1. Smart Contracts e a tecnologia blockchain

Os *Smart Contracts* foram conceituados pela primeira vez, nos anos 90, pelo jurista e criptógrafo Nick Szabo, que os definia como "um contrato inteligente é um conjunto de promessas, especificadas em formato digital, incluindo protocolos nos quais as partes cumprem essas promessas"[9]. Este conjunto de promessas funciona sob uma lógica objetiva e condicional (*if* "*this*", *then* "*that*")[10] e se executa de forma automatizada, sem a necessidade de um intermediário de confiança para sua formação e conclusão.

Resumidamente, os *Smart Contracts* são protocolos eletrónicos que podem desempenhar a função de contratos ou qualquer outro tipo de tarefa informática, nos quais as partes estipulam os seus termos de forma prévia e, automaticamente, cumpridas estas exigências, sua execução se efetiva. A grande novidade neste sistema é a não intervenção de grande infraestrutura e intermediários centrais[11].

A primeira[12] conceção de Szabo, que nos remetia às *vending machines*[13], tinha como objetivo principal corrigir as fragilidades dos meios convencio-

[9] "A smart contract is a set of promises, specified in digital form, including protocols within which the parties perform on these promises." Nick Szabo, *Smart Contracts: Building Blocks for Digital Markets*, 1996. Disponível em http://www.truevaluemetrics.org/DBpdfs/BlockChain/Nick-Szabo-Smart-Contracts-Building-Blocks-for-Digital-Markets-1996-14591.pdf (12.02.2019).

[10] Manuel A. Gómez, "(In)fallible Smart Legal Contracts*", in Legal challenges in the New Digital Age*, Ana Mercedes Lopez Rodriguez; Michael D. Green; Maria Lubomira Kubica, Leiden, Koninklijke Brill, 2021, p. 29-30.

[11] Anahiby Becerril Gil; Samuel Ortigoza Limón, "Habilitadores tecnológicos y realidades del derecho informático empresarial", *Revista del instituto de Ciencias jurídicas de Puebla*, vol. 12, nº. 41, México, Nueva Época, 2018, p. 29. Disponível em http://www.scielo.org.mx/pdf/rius/v12n41/1870-2147-rius-12-41-11.pdf. (27.09.2018).

[12] Falamos em primeira conceção, pois os Smart Contracts vieram a ser conceituados, anos depois, pelo criador da plataforma Ethereum, Vitalik Buterin. A diferença do conceito de Buterin para o conceito original de Szabo, é que o conceito mais recente se refere aos Smart Contracts criados e utilizados dentro da plataforma Ethereum, através das DLTs, mais especificamente, da blockchain. O Projeto Ethereum foi lançado por Vitalik Buterin no ano de 2015, num contexto onde a tecnologia

nais de contratação, já que ele acreditava que estes últimos eram ineficientes e extremamente onerosos. No entanto, para que toda sua potencial eficácia pudesse ser utilizada, faltava uma atmosfera tecnológica avançada, que permitisse a sua utilização e armazenamento de informações de forma segura.

Em 2008, com o surgimento do criptoativo Bitcoin e da tecnologia *blockchain* – cuja criação fora atribuída ao pseudónimo Satoshi Nakamoto[14], mas até hoje envolve um mistério sobre quem e quantas pessoas estariam por detrás da sua criação – viu-se uma oportunidade para o avanço e desenvolvimento dos *Smart Contracts*. Isto posto que a tecnologia *blockchain*, devido as suas características de potencial imutabilidade e descentralização, produz maior efetividade e menor custo no processo de formação e realização de acordos e tarefas informáticas, facilitando o cumprimento das obrigações e ou transações, junto a um sistema seguro e transparente de armazenamento de informações[15]. Efetivamente, a *blockchain* faz com que a ideia de desnecessidade de autoridade central de confiança para realização de transações possa ser posta em prática, reduzindo a formalidade e os custos associados aos métodos tradicionais, sem comprometer a autenticidade e a credibilidade.

A tecnologia *blockchain* pode ser definida como uma tecnologia de registo de informações que é feita através do consenso *peer to peer*[16], de forma distribuída. O termo *blockchain* é utilizado em razão do seu sistema de funcionamento, que se dá através de blocos que se sobrepõem – o bloco seguinte sempre terá a informação do bloco anterior – formando sua própria impressão digital. Este

blockchain já existia. Vitalik Buterin, "A next generation smart contract & decentralized application platform", *Ethereum White Paper*, 2014. Disponível em https://github.com/ethereum/wiki/wiki/White-Paper#decentralized-autonomous-organizations (23.05.2021).

[13] *Vending machine*: "a machine from which you can buy small things such as cigarettes, drinks, and sweets by putting coins into it". Disponível em https://dictionary.cambridge.org/pt/dicionario/ingles/vending-machine (13.06.2020).

[14] Satoshi Nakamoto, "Bitcoin: A Peer-to-Peer Electronic Cash System", *Bitcoin*, 2008. Disponível em https://bitcoin.org/bitcoin.pdf (20.10.2020).

[15] Fernanda de Araujo Meirelles Magalhães, "'Smart Contracts': o paradigma entre a imutabilidade e a necessidade de flexibilização contratual em tempos de crise", *Actualidad Jurídica Iberoamericana*, nº 16, febrero 2022, pp. 1254-1269.

[16] "São sistemas distribuídos compostos de nós interconectados, aptos a se auto-organizar em topologias de rede, com o intuito de compartilhar recursos, como conteúdo, ciclos de CPU, largura de banda e armazenamento, com a capacidade de adaptação a faltas e acomodação a um número variável de nós, ao mesmo tempo que mantém a conectividade e o desempenho em níveis aceitáveis, sem a necessidade de suporte ou intermediação de um servidor centralizado". Stephanos Androutsellis-Theotokis; Diomidis Spinellis, "A Survey of Peer-to-Peer Content Distribution Technologies", *Athens University of Economics and Business*, 2004, p. 337. Disponível em https://www.spinellis.gr/pubs/jrnl/2004-ACMCS-p2p/html/AS04.pdf (22.02.2019).

modo de desempenho faz com que a validação das transações seja feita, ao mesmo tempo, em vários lugares e em lugar nenhum, dificultando a introdução de informações falsas ou de burla na cadeia de blocos.

Deste modo, apesar de ter sido criada para a melhoria do Sistema Financeiro[17], também serve de força operativa e prática dos *Smart Contracts*, recebendo o *Smart Contract* previamente codificado, escrito e programado, publicitando e expandindo aquelas informações dentro da rede[18].

2.2. A utilização de *Smart Contracts* nas relações de consumo

De facto, a tecnologia blockchain e os *Smart Contracts* foram criados em momentos e com objetivos distintos, porém, conforme vimos, a primeira possui um papel essencial para o uso dos *Smart Contracts*, em particular, nas relações contratuais. A grande questão, contudo, é saber como assegurar a presença dos principais objetivos do Direito dos Contratos nos *Smart Contracts* baseados na *blockchain*, como, por *v.g.*, a proteção das partes mais vulneráveis[19], a garantia da liberdade e do equilíbrio entre os indivíduos – parte e contraparte – e promoção da justiça contratual[20]. Outros pontos discutidos são os deveres de informação, deveres de diligência, dentre outros deveres que não conseguem se basear nos termos pura e simplesmente escritos no código criptografado.

A liberdade de contratação exige reflexão e conhecimento, sobretudo no que diz respeito à celebração de contratos de consumo, especialmente num momento em que os avanços tecnológicos e a abertura dos mercados oferecem ao consumidor uma diversidade de bens e serviços caracterizados pela oferta e pela procura, com informações cada vez mais complexas[21]. Isto coloca

[17] O Banco Central Europeu, inclusive, reconhece a tecnologia como "força transformadora financeira", tendo o presidente destacado sua adoção para melhorar a eficiência dos serviços bancários e mercantis. Inmaculada Sánchez Ruíz de Valdivia, "Blockchain y plataformas de financiación participativa: dos retos del mercado único digital", *in Relaciones contractuales en la economía colaborativa y en la sociedad digital*, Guillermo García González; Maria Regina Gomes Redinha; Maria Raquel Guimarães; Beatriz Sáenz De Jubera Higuero, Madrid, Editorial Dykinson, 2019, pp. 353-373.

[18] Javier W. Ibáñez Jiménez, *Derecho de Blockchain*, cit.

[19] *E.g.* as relações de consumo, nas quais o consumidor, via de regra, possui maior fragilidade em relação às grandes empresas tanto a nível informacional, quanto a nível económico.

[20] Fernanda de Araujo Meirelles Magalhães, "Os Smart Contracts" cit., pp. 143-156.

[21] Luís Miguel Simão da Silva Caldas, "Direito à informação no âmbito do direito do consumo: o caso específico das cláusulas contratuais gerais", *in Julgar*, nº 21, 2013, p. 204. Disponível em http://julgar.pt/wp-content/uploads/2013/09/11-Silva-Caldas-Direito-%C3%A0-informa%C3%A7%C3%A3o-direito-do-consumo.pdf (15.07.2022)

o consumidor em uma posição de fragilidade, uma vez que as empresas vão sempre ter melhor conhecimento daquilo que está sendo vendido ou fornecido, ao passo que o consumidor pode ter pouco – ou nenhum – entendimento sobre o produto ou serviço contratado.

Neste sentido, o Direito do Consumo busca proteger e amparar os consumidores contra possíveis situações injustas, de natureza diversa e que afetem os interesses das partes, que podem resultar de uma assimetria de conhecimentos, ferramentas e poder entre os consumidores e as empresas no contexto das transações de mercado[22]. Na União Europeia, se olharmos, por exemplo, para o artigo 169º[23] do Tratado sobre o Funcionamento da União Europeia, vemos o reconhecimento dos consumidores como parte vulnerável das relações de consumo, evidenciando o empenho para assegurar a sua proteção. A posição de desvantagem do consumidor, neste sentido, também é reconhecida por parte do ELI, de maneira que foi publicado, em setembro de 2022, os *ELI Principles on Blockchain Technology, Smart Contracts and Consumer Protection*[24], com uma parte especial relativa à esta questão.

Deste modo, observa-se que, a aplicação dos *Smart Contracts*[25] nas relações de consumo B2C, para as empresas[26], a "redução dos custos, procedi-

[22] Gabriele Mazzini, "A System of Governance for Artificial Intelligence through the Lens of Emerging Intersections between AI and EU Law", *in Digital Revolution: New Challenges for law*, Alberto de Franceschi; Reiner Schulze, Germany, Verlag C.H. Beck/Nomos, 2019, p. 267.

[23] Article 169

1. In order to promote the interests of consumers and to ensure a high level of consumer protection, the Union shall contribute to protecting the health, safety and economic interests of consumers, as well as to promoting their right to information, education and to organise themselves in order to safeguard their interests.

2. The Union shall contribute to the attainment of the objectives referred to in paragraph 1 through: (a) measures adopted pursuant to Article 114 in the context of the completion of the internal market; (b) measures which support, supplement and monitor the policy pursued by the Member States.

3. The European Parliament and the Council, acting in accordance with the ordinary legislative procedure and after consulting the Economic and Social Committee, shall adopt the measures referred to in paragraph 2(b).

4. Measures adopted pursuant to paragraph 3 shall not prevent any Member State from maintaining or introducing more stringent protective measures. Such measures must be compatible with the Treaties. The Commission shall be notified of them.

[24] The European Law Institute, "ELI Principles on Blockchain Technology, Smart Contracts and Consumer Protection", 2022. Disponivel em https://www.europeanlawinstitute.eu/fileadmin/user_upload/p_eli/Publications/ELI_Principles_on_Blockchain_Technology__Smart_Contracts_and_Consumer_Protection_Council_Draft.pdf (23.10.2022)

[25] A partir de agora, quando utilizarmos o termo *Smart Contracts*, estaremos a nos referir aos *Decentralized Smart Contracts*, isto é, nos *Smart Contracts* baseados na tecnologia *blockchain*.

[26] Produtores, fornecedores ou comerciantes.

mentos de execução simplificados e supervisão mais rápida do desempenho contratual"[27] parece ser, de facto, benéfica e atrativa. A controvérsia existe, no entanto, em relação à proteção do consumidor, nomeadamente, em saber se o uso dos *Smart Contracts* reforça ou perturba o amparo daqueles, conforme discutiremos no tópico a seguir.

3. A proteção do consumidor nos *Smart Contracts*

No tocante a proteção dos consumidores nas relações realizadas via *Smart Contract* vemos que há, conforme mencionamos, uma divisão de dois grupos: os que defendem que os *Smart Contracts* fortalecem a proteção daquelas partes, consideradas mais frágeis; e os que advogam que haveria, na realidade, uma afetação desta proteção, diante do seu modo de funcionamento, isto é, os consumidores usuários dos *Smart Contracts* se veriam ainda mais prejudicados, muitas vezes, de forma não intencional, mas por inobservância.

A primeira corrente, dos que acreditam no reforço da proteção dos consumidores através do uso dos *Smart Contracts*, tem como alicerce o argumento de que os consumidores passariam a dispor de novos recursos na negociação *online*, corrigindo as disparidades existentes entre as partes das relações B2C, restaurando, assim, o desequilíbrio de informações[28]. De acordo com Eduardo Tatit[29], se entende que, nos *Smart Contracts*, os consumidores podem estabelecer as suas condições de contratação, sem a necessidade de intervenção e um terceiro de confiança[30]. Isto conectaria estes consumidores diretamente aos vendedores que estejam de acordo com as exigências pré-determinadas, ou seja, o contrato só se realizaria quando aquelas condições fossem atendidas[31].

Este primeiro grupo também observa os *Smart Contracts* como ferramentas capazes de auxiliar na aplicação direta do direito dos consumidores, quando estes se encontram lesados por algum tipo de incumprimento

[27] Stefan Wrbka, "A Multilayer Safeguard" cit., p. 127.

[28] Roberto Pardolesi; Antonio Davola, "Smart Contract: Lusinghe ed equivoci dell'innovazione purchessia", *Liber Amicorum – Guido Alpa*, Francesco Capriglione, Milano, Wolters Kluwer, Padova, Cedam, 2019, p. 306.

[29] Eduardo Macedo Leme Tatit, *Smart Contracts: A evolução dos contratos tradicionais*, 2018. Disponível em https://www.linkedin.com/pulse/smart-contracts-evolu%C3%A7%C3%A3o-dos-contratos-tradicionais-eduardo/ (17.08.2019).

[30] Neste sentido, ver Matti Rudanko, "Smart Contracts and Traditional Contracts: Views of Contract Law", *Legal Tech, Smart Contracts and Blockchain*, Marcelo Corrales Compagnucci; Mark Fenwick; Stefan Wrbka, Singapore, Springer, 2019, pp. 59-78.

[31] Eduardo Macedo Leme Tatit, *Smart Contracts*, cit.

(total ou parcial) do contrato por parte da empresa ou entidade contratada[32]. Neste sentido, Bernardo Moraes e Gustavo Mello defendem a efetividade dos Smart Contracts nas relações obrigacionais, já que "são automáticas, aqui, a verificação, a execução e a entrada em vigor de termos do negócio jurídico bilateral – contrato – ao mesmo tempo em que se garante sua estabilidade (porque não pode ser revogado), bem como sua transparência e publicidade"[33].

Por outro lado, há as vozes daqueles que acreditam que todo este potencial "pró consumidor" não é, de todo, verdade.

Em primeiro lugar, porque os consumidores, efetivamente, possuem pouca interferência na criação do contrato, permitindo que partes poderosas – empresas, comerciantes, etc. – se beneficiem às custas do consumidor[34]. Além disso, ainda que os consumidores tivessem o poder de decisão, na maioria dos casos, sequer possuem o conhecimento necessário para influenciar na construção dos *Smart Contracts*[35]. Se analisarmos, por exemplo, os contratos de adesão via *Smart Contracts*, as empresas poderiam se aproveitar e beneficiar do nível motivacional geralmente baixo dos consumidores para mover uma ação para inserir cláusulas abusivas[36] dentro destes contratos. Tendo em conta que as empresas podem impor – por meio de uma decisão unilateral – com relativa facilidade qualquer consequência jurídica factual à contraparte, simplesmente, criando ou emitindo decisões automatizadas, o uso dos *Smart Contracts* poderia vir a fomentar este tipo de prática.

Outro argumento utilizado é que existem situações em que as empresas não possuem a efetiva finalidade de tirar vantagem da situação, mas também não tomam cuidados suficientes para evitar que determinada consequência

[32] Neste contexto, utilizemos como exemplo o artigo 17º do Regulamento (EC) nº 1371/2007, que concede aos passageiros dos transportes ferroviários uma compensação equivalente a 25% do preço do bilhete por um atraso de uma hora e a um reembolso de 50% por atrasos de duas horas ou mais. Em caso de contratação através dos Smart Contracts não haveria o problema de as empresas ferroviárias não desejarem fazer o pagamento voluntariamente ou tentarem dificultar o cumprimento do direito do consumidor, impondo diversos procedimentos burocráticos, uma vez que estes direitos estariam garantidos automaticamente pelo contrato, facilitando a proteção dos consumidores.

[33] Bernardo Bissoto Queiroz de Moraes; Gustavo Marchi de Souza Mello, "Smart legal contracts carregam consigo incontáveis benefícios", *Revista Eletrônica Consultor Jurídico*. Disponível em https://www.conjur.com.br/2018-out-29/smart-legal-contracts-contratos (15.08.2019)

[34] Matti Rudanko, "Smart Contracts" cit., p. 75.

[35] Stefan Wrbka, "A Multilayer" cit., p. 124.

[36] No ordenamento português: Decreto-Lei nº 446/85 – regime das cláusulas contratuais gerais.

seja prejudicial ao consumidor, ou não atenda os direitos daqueles[37]. Consequentemente, a utilização dos *Smart Contracts* neste tipo de relação contratual também dificultaria a justiça e a proteção das partes.

Cumpre salientar que, ao longo dos anos, vem sendo feita uma movimentação por parte da União Europeia para, justamente, proteger os consumidores destas práticas abusivas, geralmente presentes nos contratos de adesão[38]. Aliás, se olharmos para o artigo 8-B da Diretiva (UE) 2019/2161, que altera a Diretiva 93/13/CEE do Conselho e as Diretivas 98/6/CE, 2005/29/CE e 2011/83/UE do Parlamento Europeu e do Conselho no que diz respeito à melhor aplicação e modernização das regras de defesa do consumidor da União, vemos sanções mais duras, que reforçam a ideia de combater este tipo de cláusulas.

Assim sendo, ao ponderarmos estas duas posições, parece-nos que, a depender do caso concreto, pode haver, de facto, a desproteção do consumidor médio em relação ao comerciante usuário dos *Smart Contracts*. Isto é, se já para os operadores do Direito, muitas vezes, atingir uma justiça contratual é uma tarefa difícil, imagine para um código informático que – ao contrário de nós, seres humanos – não pode, por si só, gerar soluções ou mesmo tomar decisões que levem em consideração aspetos que inicialmente eram imprevisíveis ou difíceis de definir claramente[39].

A linguagem objetiva, somada à tendencial imutabilidade e autoexecutividade dos *Smart Contracts* pode trazer, deste modo, desvantagens aos consumidores. Contudo, a automatização presente nos *Smart Contracts*, se em conformidade com a legislação, pode ajudar neste processo de proteção do consumidor, aplicando automaticamente as consequências para o descumprimento de determinada cláusula ou desrespeito à algum princípio, por exemplo.

Considerando o potencial de uso dos *Smart Contracts* nas relações de consumo, diante da facilidade de obtenção de dados/informações e transações

[37] Kevin Werbach; Nicolas Cornell, "Contracts: Ex Machina", *Smart Contracts – Technological, Business and Legal Perspectives*, Marcelo Corrales Compagnucci; Mark Fenwick; Stefan Wrbka, Hard Publishing, Oxford, 2021, p. 36.

[38] Neste contexto, graças à Diretiva 93/13/CEE (Unfair Contract Terms Directive) a proteção contra cláusulas contratuais abusivas possui padrões mínimos comuns a toda a UE. Os objetivos da UCDT, se baseiam na "idea [...] that the consumer is in a weak position vis-à-vis the seller or supplier, as regards both his bargaining power and his level of knowledge, which leads to the consumer agreeing to terms drawn up in advance by the seller or supplier without being able to influence the content of those terms[...].". Commission Notice, "Guidance on the interpretation and application of Council Directive 93/13/EEC on unfair terms in consumer contracts", *in Official Journal of the European Union*, 2019. Disponível em https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX%3A52019XC0927%2801%29 (15.07.2022).

[39] Stefan Wrbka, "A Multilayer" cit., p. 133.

através da *blockchain*, e reconhecendo as possíveis dificuldades existentes neste tipo de relação, de acordo com os *ELI Principles on Blockchain Technology, Smart Contracts and Consumer Protection*[40], em relação à linguagem empregada, os Princípios (8), de forma geral, e (15) parte especial, dispõem que os *Smart Contracts*, em relações B2C, devem ser traduzidos para a linguagem natural, para que aqueles possa compreender seus direitos e deveres, salvaguardando a proteção do consumidor.

No entanto, não se exclui a legalidade de um contrato de consumo apenas por ter sido celebrado via *Smart Contract,* pois isso também implicaria dizer que os acordos eletrónicos não poderiam ser juridicamente vinculativos[41]. Deste modo, nos casos em que não há a tradução em linguagem natural, de forma a buscar o equilíbrio da relação de consumo via *Smart Contract*, pode ser inserido um dever de auditar o contrato para a proteção das partes ou ser direcionado para, por exemplo, a *Unfair Contract Terms Directive*, prevalecendo sempre a interpretação mais favorável ao consumidor.

A respeito da proteção do consumidor e, diante dos desenvolvimentos tecnológicos, como no caso dos *Smart Contracts*, é importante haver esta discussão para adaptação da legislação existente. Contudo, Geraint Howells, Christian Twigg-Flesner e Chris Willett[42], trazem uma reflexão interessante: por um lado, as regras jurídicas devem, sim, se adequar às novas condições tecnológicas e de mercado, considerando os riscos e a vulnerabilidade dos consumidores, que podem vir a ser enfatizados pela utilização destas novas tecnologias; por outro, uma regulamentação fortemente protetora pode vir a minar a evolução destas inovações.

No fundo, Howells *et al.* refletem que a proteção do consumidor deve estar mais focada nos valores – contextos por detrás do nível de proteção – do que puramente na lei existente, de forma a haver um equilíbrio entre a proteção dos consumidores e a evolução tecnológica. Diante de dificuldades no uso destas tecnologias, que se encontram em constante processo de desenvolvimento, é importante encontrar um equilíbrio entre a criação de leis, de forma que, ao mesmo tempo, as partes não se vejam lesadas, e a tecnologia não se veja afetada negativamente.

[40] The European Law Institute, "ELI Principles", cit.

[41] The European Law Institute, "ELI Principles" cit., p. 51.

[42] Geraint Howells; Christian Twigg-Flesner; Chris Willett, "Protecting the Values of the Consumer Law in the Digital Economy: The case of 3D-printing", *Digital Revolution: New Challenges for law*, Alberto de Franceschi; Reiner Schulze, Verlag C.H. Beck/Nomos, Germany, 2019, pp. 215 e ss.

De um ponto de vista tecnológico, acredita-se que as novas tecnologias (*e.g.* a IA) e tecnologias acessórias (*v.g.* os "oráculos"), no caso específico dos *Smart Contracts*, podem servir de suporte nas hipóteses em que o Direito das partes se encontra lesado ou não tenha sido observado, auxiliando, portanto, o Direito.

3.1. O uso dos "oráculos" nos *Smart Contracts* para a proteção do consumidor

É verdade, conforme já dissemos, que os *Smart Contracts* e a *blockchain* permitem um ambiente sem confiança, isto é, sem que haja necessidade da intervenção humana para sua execução. No entanto, a grande questão é que a *blockchain* só é capaz de verificar se uma determinada tarefa fora ou não realizada, não podendo captar fatores externos, que dependem de uma interpretação humana e de uma análise de elementos não pertencentes ao contrato, mas que são essenciais para o seu bom funcionamento. Isto não quer dizer que os *Smart Contracts* não devam ser aplicados em absoluto nas relações de consumo, pois, visivelmente, têm suas vantagens. Contudo, deve estar associado a elementos que garantam a proteção das partes mais frágeis, como os consumidores.

Tendo a noção de que o raciocínio revogável e o argumento dialético são importantes no para as relações contratuais e de consumo, a integração dos chamados "oráculos"[43] se torna fundamental para otimizar a qualidade de funcionamento dos *Smart Contracts* nos contratos de consumo. Os "oráculos" são empresas terceiras que conectam o mundo offline com o mundo online, ou seja, estas empresas recolhem as informações externas e trazem para dentro da blockchain para serem validadas e inseridas dentro daquele determinado contrato. Funcionam como uma ponte entre o código criptografado do *Smart Contracts* e a estrutura externa de qualquer conteúdo/informação[44]. Estes oráculos podem ser descentralizados[45], o que torna o sistema mais seguro e

[43] Carlos Tur Faúndez, *Smart contracts: análisis jurídico*, Colección de derecho de las nuevas tecnologías, Madrid, Reus Editorial, 2018, p. 111-115.

[44] Rolf H. Weber, "Smart Contracts: Do we need New Legal Rules?", *in Digital Revolution: New Challenges for law*, Alberto de Franceschi; Reiner Schulze, Germany, Verlag C.H. Beck/ Nomos, 2019, pp. 308-309.

[45] Sobre este ponto, conforme afirma André Feiteiro (e nós concordamos), o "caráter de centralização em um único oráculo é apresentado como um ponto de falha", uma vez que vai contra toda a ideia por detrás dos *Smart Contracts*. André Feiteiro, "The Complementary but not Alternative Utility of Smart Contracts", *in Revista de Direito e Tecnologia*, Vol. 2, 2020, disponível em https://blook.pt/publications/publication/24df53083858/ (23.02.2021).

imparcial, tendo um papel "relevante para evitar o abuso de direito e garantir a segurança jurídica"[46].

No caso de compensação de transportes ferroviários[47], disposta no Regulamento (EC) nº 1371/2007, por exemplo, o "oráculo" transmitiria as informações necessárias para a blockchain e, em caso atraso, o *Smart Contract* automaticamente aplicaria a indemnização correspondente[48]. Neste contexto, os dados recebidos através dos "oráculos" melhorariam o funcionamento do *Smart Contract*, que obteria informações mais precisas e fidedignas, amparando ambas as partes do contrato.

Nota-se que, como a tecnologia *blockchain*, *per se*, não permite reverter ou parar a transação uma vez executada, a presença de um "oráculo" se faz importante para permitir que aquela possa ser contestada, que esteja em conformidade com a legislação pertinente, ou com a situação atualizada daquele determinado contrato[49]. Dificilmente isso poderia ocorrer se não houvesse qualquer conexão do *Smart Contract* com o mundo externo.

A primeira grande questão para a inclusão dos "oráculos", todavia, é a intervenção – de certo modo, humana – de terceiros. A propósito, esta é uma das principais características dos *Smart Contracts*, que o torna diferente de todos os demais meios de contratação eletrónica e, ao mesmo tempo, atrai diversos usuários.

Outro ponto importante a ser considerado é que, quando estamos a falar de eventos elementos numéricos, esta tarefa parece ser fácil de ser realizada pelos "oráculos" e incluída corretamente dentro dos *Smart Contracts*. Entretanto, nos casos em que os conteúdos devem preceder de interpretação, ou estão relacionados a expressões que remetem cláusulas gerais ou conceitos jurídicos genéricos, parece ser mais complicado[50].

Ademais, o uso do "oráculo", apesar de parece-nos ser a alternativa mais acertada para potencializar o uso dos *Smart Contracts* e garantir a proteção dos consumidores, não é totalmente infalível e as suas consequências legais ainda não estão nitidamente definidas[51].

[46] Fernanda de Araujo Meirelles Magalhães, "*Smart Contracts*" cit., p. 1263.

[47] Ver nota 28.

[48] Neste sentido, ver https://etherisc.com/

[49] Rolf H. Weber, "Smart Contracts" cit., p. 307.

[50] Fulvio Sarzana di S. Ippolito; Massimiliano Nicotra, *Diritto della blockchain, Inteliligenza Artificiale e IoT*, Milano, Wolters Kluwer, 2018, p. 105.

[51] Rolf H. Weber, "Smart Contracts" cit., p. 309.

3.2. A aplicação da Inteligência Artificial para aprimoramento dos *Smart Contracts* nas relações de consumo

Respondendo aos receios sobre a utilização dos *Smart Contracts* não só nas relações de consumo, mas nas relações contratuais, em geral, uma possível solução seria a criação de oráculos totalmente automatizados baseados em Inteligência Artificial, aumentando a eficiência das informações obtidas e, consequentemente, diminuindo os custos. Conforme afirma Gabriele Mazzini, "a IA é reconhecida como uma tecnologia com potencial para trazer grandes efeitos transformadores econômicos e sociais, que precisam ser aproveitados para os benefícios das pessoas e da sociedade como um todo, de acordo com uma abordagem europeia"[52].

O uso dos "oráculos" baseados em IA, devido à grande capacidade de recolha de dados em um curto espaço de tempo, serviria para expressar e simular o raciocínio jurídico, podendo avaliar noções fundamentais do direito contratual e de consumo inserindo estas nuances dentro dos *Smart Contracts*. O uso da Inteligência Artificial ajudaria não só numa análise quantitativa de informações, mas também iria auxiliar na aplicação do Direito em questões que exigem interpretação humana[53]. Paralelamente, faria com que o código pudesse se aproximar de forma mais precisa da realidade das partes e das possíveis alterações imprevisíveis.

Não podemos esquecer, no entanto, que a utilização da Inteligência Artificial em transações entre empresas e consumidores precisa ser justa, transparente e compatível com as leis de consumo. A falta de estrutura legal adequada mantém a preocupação em torno do uso destas tecnologias, uma vez que os possíveis impactos negativos de abusos nos *Smart Contracts* se tornam mais evidentes, especialmente, nas relações B2C[54].

A questão do uso da IA, contudo, ainda é muito incerta e não sabemos quais seriam, em concreto, seus efeitos sob os *Smart Contracts* e se a criação de uma legislação específica, neste aspeto, não faria sentido para garantir a proteção das partes mais vulneráveis. Além disso, em caso de falha dos "oráculos", as consequências legais ainda não estão expressamente estabelecidas. Do mesmo modo, adicionar a Inteligência Artificial aos *Smart Contracts* no intuito de proteger o consumidor é interessante, mas traz consigo outras questões relacionadas à tecnologia, que merecem uma atenção especial.

[52] Gabriele Mazzini, "A System" cit., p. 245.
[53] Matti Rudanko, "Smart Contracts" cit., p. 73.
[54] Stefan Wrbka, "A Multilayer" cit., pp. 143-144.

No ponto atual em que se encontra o desenvolvimento da Inteligência Artificial nos "oráculos", ainda que a tecnologia seja importante para obter informações externas essenciais para o desenvolvimento do *Smart Contract*, pode permanecer a dificuldade da compreensão e termos ambíguos, mantendo-se o problema de interpretação. Havendo esta deficiência interpretativa, consequentemente, irá afetar a execução correta do contrato e, neste caso, a quem se atribuiria a responsabilidade por este erro?

Isto não quer significar, de todo, que o uso da Inteligência Artificial e dos *Smart Contracts* deva ser excluído das relações de consumo. Todavia, é preciso encontrar – ou tentar promover – um equilíbrio entre as partes desta relação que, normalmente, já se encontram em situações assimétricas. Deste modo, antes da sua aplicação, é preciso traçar estratégias que garantam a execução dos termos contratuais de forma automatizada e aproveitem todo o seu potencial, sem desrespeitar as leis de consumo e preservando a lógica que envolve a criação das leis[55].

4. Conclusões

Os *Smart Contracts* apresentam-se como uma inovação tecnológica contratual que modifica expressivamente a forma de contratar e, sobretudo, o modo de execução dos acordos/contratos de consumo. As características dos *Smart Contracts* reduzem a necessidade de um terceiro tanto para administrar os termos negociais quanto para resolver possíveis conflitos entre as partes, o que facilita a contratação e, potencialmente, diminui os custos das transações realizadas por esta via.

O que se observa atualmente é que os consumidores possuem cada vez mais direitos; no entanto, isto não significa nenhuma garantia de aplicação. Principalmente no que diz respeito aos direitos do consumidor já que, por desconhecimento ou desinteresse, quase nunca veem aqueles sendo exercidos. Melhor dizendo, uma minoria mais instruída dos consumidores reivindica os seus direitos, o que, consequentemente, prejudica ainda mais os demais que permanecem inertes, uma vez que há uma dinâmica de preços por parte das empresas para inserir nos valores cobrados o pagamento das reclamações daquela pequena parte da população que postula o seu cumprimento[56].

[55] Stefan Wrbka, "A Multilayer" cit., loc. cit.

[56] Martin Fries, *Law and Autonomous Systems Series: Smart consumer contracts – The end of civil procedure?*, 2018. Disponível em https://www.law.ox.ac.uk/business-law-blog/blog/2018/03/smart-consumer-contracts-end-civil-procedure (19.08.2019)

Resta evidente que, pela sua execução automática, os *Smart Contracts* podem, ser altamente vantajosos e potencialmente aplicáveis em diversos campos. Apesar das críticas a esta tecnologia, principalmente por conta da sua rigidez e automaticidade, a longo prazo e no caminhar do seu desenvolvimento e estudo, a tendência é que esta possa ser mais facilmente – e em maior escala – aplicada em benefício do Direito do Consumo. Entretanto, nota-se uma série de desafios a serem enfrentados para que se possa estabelecer os *Smart Contracts* como um mecanismo confiável de aplicação destes direitos.

A utilização dos "oráculos", sem dúvida, é capaz de responder algumas questões relacionadas com a proteção do consumidor em sede de *Smart Contracts*, mas carrega consigo algumas limitações. A Inteligência Artificial, igualmente, pode auxiliar no processo de interpretação e simulação do raciocínio humano, mas até que ponto e de que maneira isto poderia ocorrer sem gerar novas questões jurídicas – talvez, sem solução. Ademais, havendo necessidade de apoio de regulamentação específica, qual seria o limite para a criação de novas leis para a intervenção sob as novas tecnologias de forma a não impedir o seu desenvolvimento?

Independentemente de qual posição tomar em relação à utilização dos *Smart Contracts* nos contratos de consumo, o que é certo é que, cada vez mais, a economia baseada em papel vem se movendo em direção ao mundo digital. Sendo assim, é preciso não tentar frear a evolução tecnológica, mas entender até que ponto a criatividade tecnológica e as novas formas de comunicação – como no caso dos *Smart Contracts* – se revelam no plano jurídico, para garantir um equilíbrio entre os interesses das partes numa relação de consumo e a proteção de seus direitos.

Bibliografia

A. Gómez, Manuel, "(In)fallible Smart Legal Contracts*", in Legal challenges in the New Digital Age*, LOPEZ RODRIGUEZ, Ana Mercedes; GREEN, Michael D.; KUBICA, Maria Lubomira, Leiden, Koninklijke Brill, 2021, pp. 29-44

Androutsellis-Theotokis, Stephanos; Spinellis, Diomidis, "A Survey of Peer-to-Peer Content Distribution Technologies", *Athens University of Economics and Business*, 2004, p. 335-371. Disponível em https://www.spinellis.gr/pubs/jrnl/2004-ACMCS-p2p/html/AS04.pdf (22.02.2019)

Becerril Gil, Anahiby; Ortigoza Limón, Samuel, "Habilitadores tecnológicos y realidades del derecho informático empresarial", *Revista del instituto de Ciencias jurídicas de Puebla*, vol. 12, nº 41, México, Nueva Época, 2018, p. 11-41. Disponível em http://www.scielo.org.mx/pdf/rius/v12n41/1870-2147-rius-12-41-11.pdf. (27.09.2018)

BUTERIN, VITALIK, "A next generation smart contract & decentralized application platform", Ethereum White Paper, 2014. Disponível em https://github.com/ethereum/wiki/wiki/White-Paper#decentralized-autonomous-organizations (23.05.2021)

CALDAS, LUÍS MIGUEL SIMÃO DA SILVA, "Direito à informação no âmbito do direito do consumo: o caso específico das cláusulas contratuais gerais", *in Julgar*, nº 21, 2013, pp. 203-224. Disponível em http://julgar.pt/wp-content/uploads/2013/09/11-Silva-Caldas--Direito-%C3%A0-informa%C3%A7%C3%A3o-direito-do-consumo.pdf (15.07.2022)

COMMISSION NOTICE, "Guidance on the interpretation and application of Council Directive 93/13/EEC on unfair terms in consumer contracts", *in Official Journal of the European Union*, 2019. Disponível em https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX%3A52019XC0927%2801%29 (15.07.2022)

FEITEIRO, ANDRÉ, "The Complementary but not Alternative Utility of Smart Contracts", *in Revista de Direito e Tecnologia*, Vol. 2, 2020, pp. 71-96. Disponível em https://blook.pt/publications/publication/24df53083858/ (23.02.2021)

FRIES, MARTIN, *Law and Autonomous Systems Series: Smart consumer contracts – The end of civil procedure?*, 2018. Disponível em https://www.law.ox.ac.uk/business-law-blog/blog/2018/03/smart-consumer-contracts-end-civil-procedure (19.08.2019)

HOWELLS, GERAINT; TWIGG-FLESNER, CHRISTIAN; WILLETT, CHRIS, "Protecting the Values of the Consumer Law in the Digital Economy: The case of 3D-printing", *Digital Revolution: New Challenges for law*, FRANCESCHI, ALBERTO DE; SCHULZE, REINER, Verlag C.H. Beck/Nomos, Germany, 2019, pp. 214-244

IBÁÑEZ JIMÉNEZ, JAVIER W., *Derecho de Blockchain y de la Tecnología de Registros*, Cizur Menor (Navarra), Aranzadi, 2018

IPPOLITO, FULVIO SARZANA DI S.; NICOTRA, MASSIMILIANO, *Diritto della blockchain, Inteliligenza Artificiale e IoT*, Milano, Wolters Kluwer, 2018

LAURENT, PATRICK; HURTAUD, STÉPHANE; CHOLLET, THIBAULT; GENCO, SÉBASTIEN, "Distributed Ledger Technologies services: using de power of blockchain", *Deloitte*, 2017. Disponível em https://www2.deloitte.com/lu/en/pages/technology/solutions/blockchain-distributed-ledger-technology-stitch-in-time.html (15.01.2022)

MAGALHÃES, FERNANDA DE ARAUJO MEIRELLES, "'Smart Contracts': o paradigma entre a imutabilidade e a necessidade de flexibilização contratual em tempos de crise", *Actualidad Jurídica Iberoamericana*, nº 16, febrero 2022, pp. 1254-1269

MAGALHÃES, FERNANDA DE ARAUJO MEIRELLES, "Os Smart Contracts e o papel dos tribunais em matéria de Direito dos Contratos", *Estudos Jurídicos Sobre Inteligência Artificial e Tecnologias*, VEIGA, FÁBIO DA SILVA; CEBOLA, CÁTIA MARQUES; MONTEIRO, SUSANA SARDINHA (Coords.), Porto, IBEROJUR, 2022, pp. 143-156

MAZZINI, GABRIELE, "A System of Governance for Artificial Intelligence through the Lens of Emerging Intersections between AI and EU Law", *in Digital Revolution: New Challenges for law*, FRANCESCHI, ALBERTO DE; SCHULZE, REINER, Germany, Verlag C.H. Beck/Nomos, 2019, p. 245-298

MORAES, BERNARDO BISSOTO QUEIROZ DE; MELLO, GUSTAVO MARCHI DE SOUZA, "Smart legal contracts carregam consigo incontáveis benefícios", *Revista Eletrônica Consultor Jurídico*. Disponível em https://www.conjur.com.br/2018-out-29/smart-legal--contracts-contratos (15.08.2019)

NAKAMOTO, SATOSHI, "Bitcoin: A Peer-to-Peer Electronic Cash System", Bitcoin, 2008. Disponível em https://bitcoin.org/bitcoin.pdf (20.10.2020)

PARDOLESI, ROBERTO; DAVOLA, ANTONIO, "Smart Contract: Lusinghe ed equivoci dell'innovazione purchessia", *Liber Amicorum – Guido Alpa*, FRANCESCO CAPRIGLIONE, Milano, Wolters Kluwer, Padova, Cedam, 2019, p. 297 e ss.

RUDANKO, MATTI, "Smart Contracts and Traditional Contracts: Views of Contract Law", *Legal Tech, Smart Contracts and Blockchain*, CORRALES COMPAGNUCCI, MARCELO; FENWICK, MARK; WRBKA, STEFAN, Singapore, Springer, 2019, pp. 59-78

RUÍZ DE VALDIVIA, INMACULADA, "Blockchain y plataformas de financiación participativa: dos retos del mercado único digital", *in Relaciones contractuales en la economía colaborativa y en la sociedad digital*, GARCÍA GONZÁLEZ, GUILLERMO; REDINHA, MARIA REGINA GOMES; GUIMARÃES, MARIA RAQUEL; JUBERA HIGUERO, BEATRIZ SÁENZ DE, Madrid, Editorial Dykinson, 2019, pp. 353-373

TATIT, EDUARDO MACEDO LEME, *Smart Contracts: A evolução dos contratos tradicionais*, 2018. Disponível em https://www.linkedin.com/pulse/smart-contracts-evolu%C3%A7%C3%A3o-dos-contratos-tradicionais-eduardo/ (17.08.2019)

THE EUROPEAN LAW INSTITUTE, "ELI Principles on Blockchain Technology, Smart Contracts and Consumer Protection", 2022. Disponivel em https://www.europeanlawinstitute.eu/fileadmin/user_upload/p_eli/Publications/ELI_Principles_on_Blockchain_Technology__Smart_Contracts_and_Consumer_Protection_Council_Draft.pdf (23.10.2022)

TUR FAÚNDEZ, CARLOS, *Smart contracts: análisis jurídico*, Colección de derecho de las nuevas tecnologías, Madrid, Reus Editorial, 2018

UNSWORTH, RORY, "Smart Contract This! An Assessment of the Contractual Landscape and the Herculean Challenges it Currently Presents for "Self-executing" Contracts", *Legal Tech, Smart Contracts and Blockchain*, CORRALES COMPAGNUCCI, MARCELO; FENWICK, MARK; HAAPIO, HELENA, Singapore, Springer, 2019, pp. 17-61

VÍTOR, MANUEL SANTOS "Inteligência Artificial e Contratos", *Inteligência Artificial & Direito*, ROCHA, MANUEL LOPES; PEREIRA, RUI SOARES (coord.), Coimbra, Almedina, 2020, pp. 221-232

WEBER, ROLF H., "Smart Contracts: Do we need New Legal Rules?", *in Digital Revolution: New Challenges for law*, FRANCESCHI, ALBERTO DE; SCHULZE, REINER, Germany, Verlag C.H. Beck/Nomos, 2019, pp. 299-312

WERBACH, KEVIN; CORNELL, NICOLAS, "Contracts: Ex Machina", *Smart Contracts – Technological, Business and Legal Perspectives*, CORRALES COMPAGNUCCI, Marcelo; FENWICK, MARK; WRBKA, STEFAN, Oxford, Hard Publishing, 2021, pp. 7-37

WRBKA, STEFAN, "A Multilayer Safeguard Mechanism to Optimise the Potential of Smart Contracts in B2C Transactions", *Smart Contracts – Technological, Business and Legal Perspectives*, MARCELO CORRALES; FENWICK, MARK; WRBKA, STEFAN, Oxford, Hard Publishing, 2021, pp. 123-144

Inteligência Artificial e proteção de dados pessoais: a "ditadura" do algoritmo*

Artificial Intelligence and data protection: the algorithms' "dictatorship"

Inês Camarinha Lopes*

RESUMO: Os sistemas de inteligência artificial constituem um recurso crescentemente utilizado por todo o mundo e em diferentes domínios sociais. Os benefícios, a eficiência, celeridade, e a diminuição de custos que esta tecnologia pode representar são incontestáveis. Todavia, a introdução desmesurada no quotidiano dos cidadãos pode significar um perigo para a sua esfera pessoal, entre outros, para o direito à proteção dos seus dados pessoais.

Partindo da noção de IA débil, ainda que se encontre "em construção", ponderam-se as dificuldades de compatibilização desta tecnologia com o regime de proteção de dados pessoais, previsto, *grosso modo*, no Regulamento Geral de Proteção de Dados e discute-se a legitimidade, ética e jurídica, do algoritmo, dotado de um enorme poder decisor e condicionante do desenvolvimento da pessoa em diferentes domínios.

PALAVRAS-CHAVE: Inteligência artificial; transumanismo; pós-humanismo; direito de proteção de dados pessoais; algoritmo; dignidade humana; direitos de personalidade.

* O presente artigo será objeto de publicação e corresponde, com alterações pontuais, ao conteúdo apresentado no I Seminário de Doutoramento em Direito da Universidade Lusófona, no dia 10 de Dezembro de 2021.

** Assistente Convidada na Faculdade de Direito da Universidade do Porto. Mestre em Direito na área das Ciências jurídico-civilistas. Doutoranda em Direito. Investigadora colaboradora do CIJ (Centro de Investigação Jurídica da FDUP).

ABSTRACT: The use of AI systems is increasing all over the world, including in different social domains. The benefits, efficiency, velocity, and the costs reduction that they represent are undeniable. However, the immeasurable introduction into citizens' daily lives can mean a threat to their personal sphere, in particular, to their right of personal data protection.

Starting from the notion of weak Artificial Intelligence, even though that is "under construction", we pounder on this paper the difficulties of making this technology compatible with the legal regime of personal data protection, provided by the General Data Protection Regulation, and either the ethical and legal algorithms legitimacy, which has an enormous decision-making power and conditions the persons' development in different domains.

KEYWORDS: Artificial intelligence; Transumanism; Posthumanism; personal data protection law; algorithm; Human dignity; personal rights.

1. Contextualização do problema

O desenvolvimento tecnológico veio revolucionar a vivência em comunidade, desde a saúde, à celebração de contratos, à economia, à ciência, entre muitas outras áreas económicas e sociais. A rapidez com que este fenómeno se verifica dificulta a pronta resposta da legislação para fazer face às sucessivas ameaças e riscos que a tecnologia coloca aos direitos dos cidadãos[1]. Assim, não obstante as evidentes vantagens e o progresso alcançado com o desenvolvimento tecnológico, na mesma ou em maior medida são potenciados os perigos para a esfera pessoal e patrimonial dos cidadãos.

[1] Os autores apontam ainda o risco de corrosão da democracia e a concentração de poder. Veja-se a este propósito: P. NEMITZ, *Power in times of artificial intelligence. Delphi Interdisciplinary Review of emerging technologies*, 2, 24, 2019, pp. 158 a 160; P. SALES, *Algorithms, Artificial Intelligence, and the Law. Judicature*, 105, 1, 2021, pp. 33; e A. SOLOW-NIEDERMAN, "Administring Artificial Intelligence," *Southern California Law Review*, 93, 4, 2020, pp. 684 e ss.; e ainda J. PRIETO, *Inteligencia artificial, derechos humanos y bienes jurídicos*, Navarra, Aranzadi, 2021, p. 35.

A inteligência artificial (doravante, IA) constituiu um recente e importante avanço tecnológico que, como qualquer outro, deverá ser usado com sabedoria para evitar potenciais efeitos perversos para a humanidade resultantes de uma introdução e uso desmesurado, temerário e à margem da lei destas tecnologias. A IA está presente num vastíssimo leque de domínios sociais: desde os *smart contracts*[2], à saúde com a *E-health*[3], ao setor público e financeiro (no qual temos o exemplo das *bitcoins*)[4]. A IA veio ainda tornar as cidades "inteligentes" (*smart cities*) com aparelhos inovadores de poupança de recursos e eficiência. A sua influência verifica-se também no âmbito da agricultura, pesca, e da mobilidade, com os veículos autónomos. Nem o delicado âmbito do direito sancionatório, inclusive o direito penal adjetivo, escapou aos "tentáculos" desta tecnologia revolucionária, destacando-se para estes efeitos a Resolução do Parlamento Europeu de 6 de outubro de 2021 sobre "[a] inteligência artificial no direito penal e a sua utilização pelas autoridades policiais e judiciárias em casos penais[5] e o caso da testemunha *robot*, intitulada "Alexa", assistente virtual da Amazon, num processo penal nos EUA[6].

Na sequência da apresentação, pela Comissão Europeia, da proposta de Regulamento sobre a IA, a 21 de abril de 2021, foi criada a Comissão Especial sobre a Inteligência Artificial na Era Digital (AIDA). Tratam-se de esforços europeus, e verificam-se, paralelamente, preocupações internacionais, em fomentar a IA e garantir a confiança e segurança destes sistemas para os cidadãos. Note-se que até 2025 estima-se um impacto económico, da auto-

[2] De acordo com Philip Sales os *smart contracts* vieram alterar por completo o paradigma contratual vigente até então: P. Sales, Algorithms, ob. cit., 2021, pp. 24. Sobre o tema, veja-se ainda M. S. Vítor, *Inteligência Artificial e contratos* em: AAVV, *Inteligência*, ob. cit., 2020, pp. 221 a 232, em particular pp. 225 a 228.

[3] Sobre os benefícios e perigos para a privacidade da IA na saúde veja-se: C. Tshider, "The healthcare privacy – artificial intelligence impasse", *Santa Clara High Technology Law Journal*, 36, 4, 2020, pp. 443, na qual a Autora defende o necessário equilíbrio entre a proteção do paciente e o uso da IA. A Autora oferece o exemplo de uma *app* que através de uma foto da área com a patologia consegue detetar com grande probabilidade um problema oncológico (p. 439). Veja-se ainda: J. Prieto, *Inteligencia*, ob. cit., 2020, p. 36.

[4] Veja-se, sobre o tema: A. A. Leal, *Big Data nos setores bancário e financeiro*, em: AAVV, *Inteligência*, ob. cit., 2020, pp. 199 a 220.

[5] Resolução do Parlamento Europeu, de 6 de outubro de 2021, sobre a inteligência artificial no direito penal e a sua utilização pelas autoridades policiais e judiciárias em casos penais (Resolução nº 2020/2016), disponível em: https://www.europarl.europa.eu/doceo/document/TA-9-2021-0405_PT.html .

[6] Na verdade os sistemas de IA são desenvolvidos na sua esmagadora maioria pelos "gigantes digitais": os GAFAM (Google, Apple, Facebook, Amazon e Microsoft): P. Nemitz, *Power*, ob. cit., 2019, p. 158.

matização do conhecimento, *robots* e veículos autónomos entre 6,5 a 12 mil milhões de euros por ano[7].

Remonta ao ano de 1942 a definição, por Isaac Asimov, das três leis da robótica[8] as quais, não obstante se mostrarem desadequadas à atualidade desta tecnologia, são de salientar não apenas porque constituem a genealogia das preocupações legais com os *robots*, intimamente ligados à IA, mas também porque adotam uma perspetiva focada na proteção da pessoa e do seu património em detrimento da máquina. Foi a partir de meados do século XX que a investigação relativa a máquinas inteligentes e o desenvolvimentos de programas com recurso à IA cresceu fortemente[9].

É perante este cenário "sísmico" que o presente artigo nasce e encontra a sua justificação, no qual visamos debater e analisar os perigos para a esfera pessoal dos cidadãos do recurso à IA no seu quotidiano bem como, em particular, a compatibilização de uma futura regulamentação da IA com a da proteção de dados pessoais.

2. A (falta de) conceptualização da IA

A abordagem de qualquer temática carece de uma delimitação prévia, sob pena de não sabermos qual o objeto do estudo. Esta tarefa revela-se, neste caso, rodeada de dificuldades, pois atendendo à novidade da questão e à discordância quanto ao âmbito que carece de regulação, a definição de IA tem sofrido mutações ao longo do tempo e mostra-se variável para a comunidade científica.

Na verdade, o uso da expressão "inteligência artificial" surge para designar várias técnicas, mecanismos e capacidades computacionais que se associam à inteligência humana[10]. Deste modo, a IA surge por contraposição ou comparação à inteligência humana ou biológica[11].

[7] De acordo com a estatística apresentada pela Comissão Europeia em 2019 e disponível em Regular a Inteligência Artificial na UE: as propostas do Parlamento | Atualidade | Parlamento Europeu (europa.eu) (29/11/2021).

[8] As três leis são formuladas por Isaac Asimov são as seguintes: "1) A robot cannot hurt a man or, by inaction, cannot allow a man to be injured; 2) A robot must obey the orders given by humans unless such orders would contravene the first rule; 3) A robot must protect itself as long as this protection does not contravene the first and second rules." Veja-se I. Nicolau, "Human rights and Artificial Inteligence", *Journal of Law and Administrative Sciences*, 12, 2019, p. 64.

[9] R. Martinez, "Artificial Intelligence: Distinguishing between types & definitions", *Nevada Law Journal*, 19, 3, 2019, p. 1023.

[10] P. 4 European Parliamentary Research Service (EPRS).

[11] De acordo com o Congressional Research Service, IA "a computer system capable of human-level cognition". Em: J. G. Martz, "Artificial Intelligence is here, Get ready!", *Catholic University Journal of Law and Technology*, 28, 1, 2019, p. 33.

Esta falta de definição não abona a favor da compreensão holística dos riscos da introdução de sistemas de IA no quotidiano[12]. Esta introdução poderá resultar individualmente considerada ou como ferramenta integrante de um outro produto ou serviço.

Quando se trata de inovações tecnológicas, a existência de uma rígida definição legal traz consigo dificuldades devido à facilidade e rapidez na desatualização, sendo os desenvolvimentos tão imprevisíveis que o conceito não se deve fechar ao "estado da arte" atual.

2.1. A definição legal de IA na UE e no mundo

A Comissão Europeia, na sua proposta de regulamento sobre a IA apresentada em 21 de abril de 2021, define no seu artigo 3º/1 e no considerando 6 "sistema de inteligência artificial", nos seguintes termos: "software that is developed with [specific] tecniques and approaches [listed in Annex 1] and can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with."[13].

Esta noção partiu da homónima definição prevista pela Organização de Cooperação e Desenvolvimento Económico (OCDE) na recomendação sobre a Inteligência Artificial[14], datada de 2019, mas são vários os instrumentos internacionais e europeus que se vão debruçando sobre a matéria[15].

Nos EUA, o "National Artificial Intelligence Iniciative Act", de 2020, visou potenciar e apoiar o investimento, a investigação e desenvolvimento na IA. Tratou-se, por isso, não de uma regulamentação preocupada com retrair esta tecnologia, considerando os seus perigos, mas pelo contrário, que encoraja a sua implementação[16].

[12] No mesmo sentido: H. Malhotra, "Artificial Inteligence: A (semi-inteligent) overview", *International In-house Counsel Journal*, 11, 41, 2017, p. 1.

[13] Sobre a IA na União Europeia veja-se: M. Hildebrandt, "The artificial intelligence of European Union Law", *German Law Journal*, 21, 1, 2020, pp. 74 a 79. Veja-se ainda o documento: Comissão Europeia (2021). AI Watch: AI Standardisation Landscape state of play and link to the EC proposal for na AI reguatory framework. Luxemburgo, publications office of the European Union, pp. 17 e ss..

[14] Trata-se de *soft-law* e por isso, não é vinculativa. O Conselho da Europa encontra-se a diligenciar para ser publicado o enquadramento jurídico para o desenvolvimento, *design* e aplicação da IA.

[15] Sobre regulação da IA ao nível internacional veja-se: D. M. Vicente, A inteligência artificial e iniciativas internacionais, *in* AAVV, *Inteligência artificial & direito*, Coimbra, Almedina, 2020, pp. 93 a 105.

[16] Note-se, inclusive, que a Federal Trade Commission, cujo objetivo é, entre outros, defender a proteção de dados pessoais dos cidadãos, considerou a legislação americana suficiente para proteção do risco de desvios e discriminação. (p. 2 – EPRS).

Do mesmo modo, o Reino Unido adotou, em setembro de 2021, uma "estratégia nacional" para investimento na IA e prevê em 2022 apresentar o regulamento sobre IA.

Para além dos esforços individuais dos Estados, conforme descritos *supra*, destacamos ainda uma iniciativa conjunta, entre os EUA e a UE, no âmbito da criação do "Trade and Technology Council" para um desenvolvimento seguro e responsável da IA.

Estas iniciativas são de aplaudir, considerando a premência de regular uma questão que deixou de pertencer ao futuro e que sem uma cuidada regulamentação poderia redundar ou num desaproveitamento e desencorajamento desta tecnologia ou no seu atentatório e desenfreado desenvolvimento.

No que à noção de IA respeita, a doutrina distingue entre a IA forte e IA débil. A segunda diferencia-se devido à pretensão de replicar a inteligência humana numa máquina. Já a IA débil consiste em tornar a máquina capaz de dar resposta uma tarefa concreta com base nos *inputs* nela introduzidos. Todavia, neste último caso, ao contrário do primeiro, a máquina não tem autonomia cognitiva[17], depende dos contributos humanos previamente inseridos.

A Comissão Europeia, quer no documento de 8 de novembro de 2019 onde define as "diretrizes éticas para uma IA fiável", quer na referida proposta de abril do presente ano, não contempla a IA forte, ainda encarada como "ficção" e não integrante da "realidade". Como é fácil intuir, a IA forte representa um maior risco que se prende com a desumanização que esta tecnologia representa[18], a qual, dizem, poder ter capacidades cognitivas superiores à capacidade humana, por natureza limitada. Esta questão será abordada *infra*, no ponto IV.

A perspetiva adotada pela União Europeia (adiante, UE) que se encontra vertida na dita proposta de regulamento sobre a IA centrar-se-á nos seguintes pilares: a existência de uma definição tecnologicamente neutra, um regime que se baseará no risco do sistema e a adoção de uma perspetiva "antropocêntrica". O primeiro será objeto da nossa análise de seguida e mais abaixo discutiremos os dois últimos postulados.

[17] Sobre a distinção entre IA forte e débil veja-se Marta Albert-Márquez, "Posthumanismo, inteligencia artificial y Derecho", *Persona y Derecho*, 84, 2021, pp. 211 e 212; e R. Martinez, *Artificial*, ob. cit., 2019, pp. 1027 e ss., no qual o Autor defende esta distinção e o seu diferente tratamento legal.

[18] No mesmo sentido, R. Trezza, "Diritto e Intelligenza artificiale", *Etica – privacy – Responsabilitá – decisione*, Pisa, Pancini Giuridica, 2020, pp. 27. Todavia, discordamos do Autor quando defende a "humanização da máquina através do Homem".

a) Uma definição de IA tecnologicamente neutra

Apenas com uma definição tecnologicamente neutra se pode alcançar uma regulamentação prestável para o uso da IA. Caso contrário, rapidamente a norma se tornaria obsoleta e seria fácil contornar o regime jurídico estabelecido. Todavia, definir um conceito puramente tecnológico de modo alheio à tecnologia existente torna o conceito demasiado aberto e acaba por em concreto, não se encontrar definido quais os meios e técnicas que se consideram cabíveis na noção de IA.

Nesta medida, a UE parte da noção descrita *supra* mas no Anexo I à proposta de regulamento vem apresentar uma lista de técnicas e abordagens que, na atualidade, recorrem à IA, a qual será objeto de constantes atualizações.

Esta abordagem proposta pela UE conhece já um precedente, no âmbito da proteção de dados pessoais, pois a noção de "dado pessoal" pretendeu-se ser, igualmente, tecnologicamente neutra. Pode-se por isso afirmar que este entendimento da UE se sucede nas matérias profundamente influenciadas, ou diríamos até imbricadas, pelo domínio tecnológico.

De acordo com o considerando 15 do Regulamento Geral de Proteção de Dados (doravante, RGPD) a proteção de dados pessoais deverá ser tecnologicamente neutral para "evitar ser contornada". Assim, o conceito de "dado pessoal" centra-se no *resultado* – a saber, a identificação ou identificabilidade do titular de dados – e não no meio ou técnica utilizada, contemplando qualquer meio automatizado ou não automatizado.

De facto, esta perspetiva ainda que não se possa concluir ser "imune" ao rápido progresso tecnológico, é aquela que melhor dá resposta às exigências de um regime jurídico adequado e eficaz na proteção da pessoa face à IA e no reforço da confiança nestas tecnologias.

3. Um regime legal baseado no risco (*risk-based approach*)

À semelhança da perspetiva tecnologicamente neutra, a *risk based approach* conhece um precedente recente no âmbito da proteção de dados pessoais, nos termos definidos pelo RGPD.

De acordo com a proposta de regulamento sobre a IA, o seu regime jurídico será diferenciado consoante o risco para os direitos dos cidadãos. Assim, existirão sistemas de IA proibidos porque o risco para os direitos dos cidadãos e sua segurança é intolerável, por exemplo, se constituem práticas que distorcem ou manipulam a realidade (tais como, *deepfakes*, *social scoring*, e certos recursos biométricos); sistemas cuja atividade carece de regulamentação devido ao elevado risco; no caso de o risco ser considerado "limitado", vigorará um regime de transparência, no qual as obrigações para o responsável pelo

sistema serão menores; e, por fim, se o risco é diminuto o uso dos sistemas de IA será permitido, sem restrições, estado prevista a criação de "códigos de conduta" para os fornecedores dos sistemas de IA de baixo risco.

Note-se que no que respeita à identificação do risco e integração de um dado sistema de IA numa das referidas categorias deverá ser apreciada *ex--ante,* depois de efetuada uma cuidadosa avaliação do impacto do concreto sistema e do seu uso.

Todavia, verifica-se a mesma dificuldade que pode ser apontada ao modelo baseado no risco previsto no RGPD: a avaliação do nível de impacto nos direitos dos visados pelos sistemas de IA deverá ser realizada pelos seus fornecedores (no que à proteção de dados respeita, deverá ser realizada pelo responsável pelo tratamento, nos termos dos artigos 35º e seguintes do RGPD[19]).

Ora, sucede que a responsabilidade de definição do risco do sistema de IA que irá determinar a subsunção no regime legal mais ou menos restritivo e o respeito de maiores ou menores obrigações, é daquele que tem interesse na sua difusão e introdução no mercado e não uma entidade imparcial e externa para apreciar de modo isento a existência de riscos do sistema de IA. Numa palavra, o regime jurídico aplicável a um sistema de IA concreto vai ser condicionado pela interpretação de quem tem interesse na aceitação e na facilidade na sua introdução no mercado, aproveitando os espaços "em branco" que possam tornar a sua atividade económica menos onerosa.

Além da dificuldade referida soma-se o facto de poder ser necessário compatibilizar a responsabilidade pelo cumprimento do regime de proteção de dados pessoais previsto no RGPD e o regulamento sobre a IA, cuja responsabilidade, considerando a proposta da Comissão Europeia, recai sobre o fornecedor do sistema de IA. Deste aspeto discorreremos *infra*, no ponto 5. 1..

4. Antropocentrismo, transumanismo e pós-humanismo. O predomínio da máquina sobre o Ser Humano

Conforme explicámos acima a perspetiva adotada em matéria de IA centra-se no Ser Humano, em detrimento da máquina. Trata-se, portanto, de uma perspetiva antropocêntrica, onde a máquina se encontra ao serviço do Homem e não o inverso[20].

[19] Veja-se: A. Barreto Menezes Cordeiro, *Comentário ao Regulamento Geral de Proteção de Dados e à Lei 58/2019,* Coimbra, Almedina, 2021, pp. 279 e ss.; do mesmo Autor: A. Barreto Menezes Cordeiro, *Direito da Proteção de Dados à luz do RGPD e da Lei nº 58/2019,* Coimbra, Almedina, 2020, pp. 317 a 320 e ainda, A. Sousa Pinheiro (coordenação), *Comentário ao Regulamento Geral de Proteção de Dados*, Coimbra, Almedina, 2019, pp. 457 e ss..

[20] No mesmo sentido, enquanto fundamento do ordenamento jurídico, está R. Trezza, Diritto, ob. cit., 2020, p. 19.

O entendimento jurídico de qualquer questão depende e varia consoante os movimentos cultural, social e científico vigente. Assim, estes influenciam profundamente o pensamento jurídico de uma dada época e numa determinada comunidade.

Com o desenvolvimento das tecnologias começa a singrar e ganhar expressividade o movimento "transumanista", nos finais do século XX e inícios do século XXI. Esta filosofia visa utilizar as novas tecnologias no Ser Humano com vista a superar as suas "debilidades", por exemplo a pobreza[21], o envelhecimento, o sofrimento, a doença, ou até a condição de mortal.

Tratar-se-ia de um processo que, ao transcender caraterísticas entendidas como "limites" da condição humana, conduz ao pós-humanismo: isto é, o surgimento de uma nova espécie. Nesta fase deixaria de ter sentido a distinção entre o natural e o artificial, ou seja, entre o humano e a tecnologia, porque o prévio "artificial" passa a fazer parte da condição humana[22].

A IA forte pode conduzir ao pós-humanismo ou não, estando aquela dissociada deste movimento, ainda que o pós-humanismo se possa dizer dependente de IA, o inverso não é verdadeiro. Mas a IA em geral, sobretudo a IA forte, representa o percurso do caminho apontado pelo transumanismo, na medida em que se visa introduzir capacidades intelectuais e de aprendizagem numa máquina, criando inteligência não biológica, ao serviço do Homem. Pensemos num veículo (*robot*) autónomo que é capaz de aprender o caminho até à escola de uma determinada criança e leva-a, na impossibilidade dos pais efetuarem essa deslocação. Neste caso, por exemplo, é superada a limitação da condição Humana de não ser possível estar em dois locais distintos em simultâneo, numa palavra a falta de ubiquidade do Ser Humano é "ultrapassada". De acordo com Ray Kurzweil, em 2045 a inteligência artificial superará a inteligência humana.

Em 1993, Vernor Vinge utiliza a expressão "singularidade tecnológica" para designar a fase "última" de evolução, que conduzia ao fim da condição Humana[23]. A singularidade tecnológica representa um passo "adiante" do

[21] No entanto, a doutrina também destaca a IA como forma de discriminação dos mais pobres, na medida em que apenas as pessoas com condições económicas mais favoráveis terão capacidade para pagar recursos que melhor protegem a sua esfera de privacidade e vida privada. Nesta medida, a privacidade, ao mesmo tempo que é crescentemente ameaçada pela tecnologia, torna-se um privilégio apenas acessível aos mais ricos. Veja-se: J. G. Martz, *Artificial*, ob. cit., 2019, p. 48.

[22] Sobre a definição e o surgimento do transumanismo e do pós-humanismo veja-se: M. Marquez, *Posthumanismo*, ob. cit, 2021, pp. 208 e 214 e R. Trezza, *Diritto*, ob. cit., 2020, pp. 103 e ss..

[23] Vernor Vinge, *The Coming Technological Singularity: How to Survive in the Post-Human Era*, 1993, disponível em: [PDF] The coming technological singularity: How to survive in the post-human era | Semantic Scholar (20/11/2021).

pós-humanismo, sendo que aquela pressupõe a passagem por este mas não o inverso. A singularidade chegaria quando a inteligência humana deixa de controlar a inteligência não biológica, a qual ganha autonomia para evoluir[24].

Estas correntes de pensamento fazem repensar o eminente respeito pela dignidade Humana, o qual deixaria de ser um princípio absoluto. Afinal, se o objetivo último destes movimentos filosóficos é superar, por completo, aquilo que carateriza o Homem (por ser encarado como uma "limitação da condição humana"), estamos a reconhecer que o "pós-humano" é superior à natureza humana[25]. Esta crença na superioridade dos *robots* pode explicar as teses que defendem a "dignidade robótica" ou "antropologia digital" e até a defesa de uma personalidade jurídica robótica.

Apesar de concordarmos com o leitor que está a pensar que algumas destas questões são (ainda) privativas da ficção científica, o certo é que estes movimentos existem e ganham cada vez maior expressividade, e para alguns Autores, o único meio de salvaguardar a ética e o respeito pela dignidade Humana no uso da IA passa por construir uma IA pautada pelos valores Humanos. Todavia, não podemos deixar de nos questionar: poderemos introduzir esta sensibilidade aos valores humanos numa máquina? Se sim, quais os valores humanos que a serem priorizados? De acordo com uma perspetiva minimalista, assegurando o respeito pelos valores suficientes à convivência comunitária, ou maximalista? Deixaremos estas questões na esperança que a sua resposta nunca seja necessária.

5. Os riscos da IA para a esfera pessoal dos destinatários destes sistemas

Conforme resulta da proposta de regulamento sobre a IA apresentada pela Comissão Europeia em abril de 2021, o recurso a sistemas de IA constitui um risco para direitos fundamentais dos visados por esta tecnologia e para a sua segurança, quando os sistemas de IA integram outros produtos e serviços introduzidos no mercado. Em concreto, este documento elenca o direito à não discriminação, liberdade de expressão, dignidade humana, direito à proteção de dados pessoais e direito à privacidade[26].

De facto, os sistemas de IA levantam grandes dificuldades jurídicas e éticas neste domínio, porque, não obstante os benefícios desta tecnologia, a sua

[24] Sobre o tema veja-se: M. Marquez, *Posthumanismo*, ob. cit, 2021, p. 214.

[25] Tal como afirma Malhotra, está a criar o *standard* do "humano ideal" e não da normalidade, em H. Malhotra, *Artificial*, ob. cit., 2017, p. 9. De acordo com o entendimento de Jack Watson, o qual se manifesta contra a IA, a tecnologia torna as questões mais fácil resolução e não necessariamente melhores. Em: J. Watson, "Artificial Intelligence?", *Lawno*, 22, 1, 1997, pp. 36 a 38.

[26] L. N. Santos, *A inteligência artificial e privacidade*, *in* AAVV, *Inteligência*, ob. cit., 2020, pp. 147 a 159.

capacidade, eficiência, rapidez e sobretudo, a opacidade a falta de transparência com que estes sistemas operam atentam de forma direta contra direitos dos visados[27].

Além dos descritos, muitos mais direitos poderiam ser elencados mas, na verdade, apenas no caso concreto poderemos identificar os direitos potencialmente ameaçados pelo sistema de IA. Por um lado, devido à multiplicidade de contextos nos quais o recuso à IA pode surgir (por exemplo, no âmbito sanitário, na saúde, ciência, seguros, crédito, segurança pública, entre outros). Por outro, devido à variedade de sistemas de IA, cuja intrusividade oscila (desde um nível baixo a um nível de intrusividade elevado).

São apontadas várias consequências negativas na comunidade, com repercussões sociais e económicas de relevo, das quais se salienta a extinção de um grande número de empregos cujas funções podem passar a ser realizadas por *robots inteligentes*[28]. Note-se que já não se trata apenas de tarefas mecanizadas e industriais mas inclusive trabalhos que exigem competências cognitivas e intelectivas, tais como a resolução judicial de litígios[29], reduzindo o pensamento jurídico a um mecanismo[30], ou a realização de intervenções cirúrgicas ou o diagnóstico de doenças.

5.1. A proteção de dados pessoais e a IA débil

Devido às limitações de espaço do presente trabalho limitar-nos-emos à abordagem a IA débil e a potencial colisão com a proteção de dados pessoais.

A IA débil depende, na sua existência, de bases de dados. Estas "alimentam" o sistema e permitem, com base na sua análise constante, entre outras competências, que o sistema ofereça uma resposta ao problema. Assim, é impossível dissociar da IA a proteção de dados pessoais.

Os sistemas de IA representam um perigo para o direito à proteção de dados pessoais dos cidadãos[31]. Este é um facto incontornável e incontestado

[27] No mesmo sentido, P. Sales, *Algorithms*, ob. cit., 2021, pp. 25. Sobre os desafios éticos da IA veja-se: P. Rodriguez, "Artificial Intelligence law: applications, risks & opportunities", *Revista Juridica Universidad de Puerto Rico*, 90, 3, 2021, pp. 715 a 719.

[28] Sobre o tema veja-se: I. Nicolau, "Human rights and artificial intelligence", *Journal of law and administrative sciences*, 12, 2019, p. 67 e 69, na qual a Autora afirma que entre 400 e 800 milhões de trabalhadores vão ser substituídos até 2030. E ainda, J. Martz, *Artificial*, ob. cit., 2019, p. 34.

[29] Sobre o "juiz *robot*" veja-se: R. Trezza, *Diritto*, ob. cit., 2020, pp. 17 e ss., 24 e ss. e E. Lashbrooke, "Legal reasoning and Artificial Intelligence", *Loyola Law Review*, 34, 2, 1998, pp. 304 e ss..

[30] M. Marquez, *Posthumanismo*, ob. cit, 2021, p. 225.

[31] Por todos, A. Devia, "La inteligencia artificial, el BigData y la Era digital: una amenaza para los datos personales", *Revista la Propiedad Inmaterial*, 27, 2019, pp. 14 e 15, nas quais se destaca os

quando se pondera o recurso a estes sistemas. O problema, todavia, avoluma-se pelo facto de não podermos falar numa antecipação destes riscos, uma vez que os sistemas de IA que ameaçam os dados pessoais dos cidadãos não vão ser introduzidos, no futuro. Pelo que não se trata de perspetivar e antecipar estas ameaças mas da constatação de uma realidade, infelizmente, presente.

Este juízo negativo que fazemos sobre os riscos da IA para o direito à proteção de dados pessoais, e direitos pessoais em geral, – tais como os direitos à privacidade, honra, imagem[32], não discriminação, livre desenvolvimento da personalidade e liberdades fundamentais como a de expressão e circulação, – não se dirige à tecnologia em si mesma considerada, a qual, sem margem para dúvidas, trouxe e trará muitos benefícios para a sociedade em geral, mas ao lamentável facto de as ameaças aos direitos dos cidadãos serem crescentes e terem já ultrapassado essa fase de iminência, constatando-se que a sua violação não foi evitada e cujos danos causados podem ser irreversíveis.

A compatibilização dos regimes jurídicos sobre os sistemas de IA, de acordo com a proposta de Regulamento da Comissão Europeia, e o que disciplina a proteção de dados pessoais, nos termos previstos, entres outros diplomas, no RGPD, deverá ser compatibilizado, uma vez que aquela tecnologia pressupõe o processamento automatizado de dados, os quais poderão ser informações pessoais.

Sucede que, nos termos da proposta de Regulamento sobre os sistemas de IA, o responsável pelo cumprimento das obrigações nele previstas é o fornecedor do sistema e o responsável pelo cumprimento e respeito pelo direito da proteção de dados pessoais, designadamente os princípios da licitude, lealdade e transparência, minimização de dados, exatidão, limitação da conservação, integridade e confidencialidade, e responsabilidade (artigo 5º do RGPD), é o responsável pelo tratamento[33] (artigo 5º/2 do RGPD, o qual prescreve a res-

perigos da interconexão digital e partilha de dados. A autora afirma ainda que o bom uso dos dados pessoais é também responsabilidade do titular de dados (p. 18). No mesmo sentido, veja-se P. Rodriguez, *Artificial*, ob. cit., 2021, pp. 708 e 709; e ainda: S. A. Jain e S. A. Jain, "Artificial Intelligence: a threat to privacy?", *Nirma University Law Journal*, 8, 2, 2019, pp. 32 e ss..

[32] V. P. Fidalgo, "A Inteligência artificial e direitos de imagem" em: AAVV, *Inteligência*, ob. cit., 2020, pp. 137 a 146.

[33] O conceito de "responsável pelo tratamento" encontra-se definido no artigo 4º/7 do RGPD. Segundo esta disposição, o responsável pelo tratamento é "a pessoa singular ou coletiva, a autoridade pública, a agência ou outro organismo que, individualmente ou em conjunto com outras, determina as finalidades e os meios de tratamento de dados pessoais; sempre que as finalidades e os meios desse tratamento sejam determinados pelo direito da União ou de um Estado-Membro, o responsável pelo tratamento ou os critérios específicos aplicáveis à sua nomeação podem ser previstos pelo direito da União ou de um Estado-Membro." Sobre esta disposição e a sua interpretação jurisprudencial

ponsabilidade do cumprimento e o ónus da prova do cumprimento, destacando-se neste âmbito igualmente o artigo 24º do RGPD[34]).

Repare-se que faz parte da definição de responsável pelo tratamento, ser a pessoa que determina as finalidades e os *meios* do processamento de dados pessoais. Assim, o responsável pelo tratamento pode determinar que o processamento de dados pessoais é realizado com recurso a sistemas de IA. No entanto, o sistema de IA a utilizar para o processamento de dados pessoais poderá não ser transparente para o responsável pelo tratamento. Assim, podemos concluir que o fornecedor dos sistemas de IA que processam dados pessoais, apesar de não ter, formalmente, a responsabilidade de assegurar que o processamento de dados cumpra os *supra* elencados princípios, na prática terá de garantir o seu cumprimento pelo sistema de IA para conseguir introduzi-lo no mercado.

Assim, as capacidades do sistema de IA devem contemplar o respeito pela minimização dos dados pessoais, limitação da conservação, transparência do processamento, exatidão dos dados e a proteção de dados pessoais desde a conceção e por defeito. Por exemplo, um sistema de IA que processe dados pessoais sem permitir ao responsável pelo tratamento que os elimine quando necessário, não cumpre, *ab initio*, o direito da proteção de dados pessoais e poderá fazer incorrer o responsável pelo tratamento em responsabilidade pelo tratamento de dados ilícito.

A compatibilização das responsabilidades pode levantar problemas na prática: imaginemos uma situação onde por defeito do sistema de IA, o processamento de dados realizado não cumpre os princípios descritos, tornando o tratamento de dados ilícito.

Na presença de danos causados a titulares de dados, a responsabilidade civil subjetiva do responsável pelo tratamento pressupõe a existência de culpa do lesante, nos termos gerais. De acordo com o considerando 75 do RGPD, "deverá ser consagrada a responsabilidade do responsável *por qualquer tratamento* de dados pessoais realizado por este ou por sua conta. Em especial, o responsável pelo tratamento deverá ficar obrigado *a executar as medidas* que forem adequadas e eficazes e ser capaz de *comprovar* que as atividades de

ampla, veja-se: Veja-se: Cordeiro, A. Barreto Menezes, *Comentário*, ob. cit. 2021, pp. 88 e 89; do mesmo Autor: A. Barreto Menezes Cordeiro, *Direito*, ob. cit. 2020, pp. 307 e ss. e ainda, A. Sousa Pinheiro (coordenação), *Comentário*, ob. cit., 2019, pp. 137 e ss..

[34] Para uma análise dos artigos 5º e 24º veja-se: A. Barreto Menezes Cordeiro, *Comentário*, ob. cit., 2021, pp. 101 e ss. e 232 e ss., respetivamente; do mesmo Autor: A. Barreto Menezes Cordeiro, *Direito*, ob. cit., 2020, pp. 152 e ss. e 323 e ss., respetivamente; e ainda A. Sousa Pinheiro (coordenação), *Comentário*, ob. cit., 2019, pp. 204 e ss. e 395 e ss., respetivamente.

tratamento são efetuadas em conformidade com o presente regulamento, *incluindo a eficácia das medidas*. Essas medidas deverão ter em conta a natureza, o âmbito, o contexto e as finalidades do tratamento dos dados, bem como o risco que possa implicar para os direitos e liberdades das pessoas singulares."

Note-se ainda que nos termos do considerando 146 do RGPD, "O responsável pelo tratamento ou o subcontratante deverão reparar *quaisquer danos* de que alguém possa ser vítima em virtude de um tratamento que viole o presente regulamento. O responsável pelo tratamento ou o subcontratante pode ser exonerado da responsabilidade *se provar que o facto que causou o dano não lhe é de modo algum imputável*." O direito da União estabelece, portanto, uma presunção de culpa (ilidível, relativa ou *iuris tantum*) do responsável pelo tratamento.

Deste modo, na situação referida, se o defeito do sistema de IA não pode ser imputável ao responsável pelo tratamento, o qual escolheu um meio de processamento de dados pessoais que seria seguro e garantia as medidas técnicas e organizativas adequadas e era um sistema capaz de assegurar o cumprimento do RGPD, perante o titular de dados conseguirá exonerar-se da responsabilidade se provar que o facto *não lhe é de modo algum imputável*. Poderá, nesta situação, o fornecedor do sistema de IA responder pelos danos causados ao titular de dados na medida em que ele não realizou o processamento de dados pessoais (e não foi, por isso, quem violou ilicitamente os direitos do titular de dados e lhe causou danos)? Nos quadros do RGPD cremos não ser possível fundamentar esta responsabilidade do fornecedor pelos danos causados ao titular de dados, restará aguardar que o direito europeu *a constituir* relativo à IA contemple a responsabilidade do fornecedor do sistema, em particular, na legislação europeia sobre o regime de responsabilidade civil na IA, na sequência da Recomendação do Parlamento Europeu à Comissão, de 20 de outubro de 2020, sobre esta matéria[35]. Note-se que no que à proteção do consumidor respeita, a Organização Europeia de Consumo (BEUC) alertou para a necessidade de a proteção do consumidor ser reforçada, designadamente assegurando a lealdade, *responsabilidade*, e transparência.

A harmonização do regime jurídico de tratamento de dados pessoais com o regime jurídico sobre a IA resultará ainda da previsível alteração da noção de "dados biométricos" prevista no artigo 4º/14) do RGPD. Torna-se, assim,

[35] Resolução do Parlamento Europeu de 20 de Outubro de 2020 na qual recomenda a Comissão Europeia e definir uma proposta de regulamento sobre o regime de responsabilidade civil para a Inteligência Artificial. Resolução nº 2020/2014(INL) disponível em: https://www.europarl.europa.eu/doceo/document/TA-9-2020-0276_EN.pdf (1/12/2021). Sobre o tema da responsabilidade na IA veja-se Y. Bathaee, "Artificial Intelligence Opinion Liability", *Berkeley Technology Law Journal*, 35, 1, 2020, pp. 113 a 170.

visível o impacto desta nova legislação sobre a IA no direito da proteção de dados pessoais. A identificação biométrica constitui um delicado uso da IA, uma vez que os dados biométricos, à data apenas os que permitam identificar o titular (sem incluir os casos onde o titular é apenas identificável), são considerados sensíveis considerando o risco para os direitos fundamentais do titular e o risco de discriminação. Trata-se de tecnologia que a ser usada indiscriminadamente e sem parcimónia, permite a vigilância em massa dos cidadãos, sendo por isso necessários limites à sua aceitação legitima.

a) A "ditadura" do algoritmo

Os sistemas de IA débil dependem de bases de dados. Os dados são o "combustível" que permite ao sistema de IA desempenhar a sua tarefa. Quando estamos na presença de dados pessoais, de acordo com o artigo 22º do RGPD[36], o titular de dados tem o direito de não ficar sujeito a nenhuma decisão individual automatizada (isto é, tomada com base exclusiva num tratamento de dados automatizado), que produza efeitos na sua esfera jurídica ou o afete significativamente. Todavia, o nº 2 desta norma afasta este direito do titular de dados num conjunto de situações[37].

O sistema de IA poderá ter como objetivo último a tomada de uma decisão individual automatizada com base nos dados pessoais objeto de processamento, incluindo a definição de perfis[38]. Assim, o sistema de IA poderá, com base nos dados pessoais dos titulares, analisar esses dados, prever comportamentos, preferências e o desempenho do titular, avaliar aspetos relativos à saúde, profissão, finanças ou outros aspetos pessoais.

Esta tarefa do sistema de IA que consiste na avaliação de aspetos pessoais do titular de dados é realizada de acordo com algoritmo que é introduzido. Este constitui as operações ou instruções a seguir pelo sistema de IA[39].

[36] Veja-se: A. Barreto Menezes Cordeiro, *Comentário*, ob. cit., 2021, pp. 220 e ss.; A. Sousa Pinheiro (coordenação), *Comentário*, ob. cit., 2019, pp. 386 e ss..

[37] Transcrevemos o nº 2 do artigo 22º do RGPD: "O nº 1 não se aplica se a decisão: a) for necessária para a celebração ou a execução de um contrato entre o titular de dados e um responsável pelo tratamento; b) for autorizada pelo direito da União ou do Estado-Membro a que o responsável pelo tratamento estiver sujeito, e na qual estejam igualmente previstas medidas adequadas para salvaguardar os direitos e liberdades e os legítimos interesses do titular de dados; ou c) for baseada no consentimento explícito do titular de dados.".

[38] Sobre a definição de perfis no quadro da IA veja-se: S. A. Jain e S. A. Jain, *Artificial*, ob. cit., 2019, pp. 34 e 35.

[39] O conceito de "algoritmo" é definido por Philip Sales nos seguintes termos: «an automated instruction, or a "coded recipe that gets executed when it encounters a trigger. It can be as simple as a mere "if, then" statement.» e acrescenta que «That capacity to change, adapt and grow based on new data" is AI», em: P. Sales, , *Algorithms*, ob. cit., 2021, pp. 24.

É a capacidade do sistema de "aprender" o algoritmo e com ele processar os dados pessoais que constituem o *input* desta tecnologia que dita estarmos na presença de um sistema que recorre a IA (débil). O algoritmo dita como este *software* irá tomar uma decisão, ou formular uma recomendação ou previsão, numa palavra, gerar o seu *output*, com base nos dados pessoais processados. Já nos sistemas de IA forte, há uma indiferença face ao algoritmo, porque deixa de depender deste contributo humano que foi introduzido.

É neste sentido que utilizamos a metáfora "ditadura" do algoritmo, na medida em que o algoritmo se apresenta como o decisor de um conjunto de aspetos pessoais e com grande impacto na esfera jurídica do titular. É em função do *output* apresentado pelo sistema de IA, determinado pelo algoritmo, que um cidadão pode ver o recurso ao crédito recusado, ou ficar excluído de um processo de recrutamento eletrónico para um determinado emprego, ou ver-lhe negada a adesão a um seguro de saúde em função do algoritmo ter feito uma previsão de determinada doença ao identificar um risco mais elevado no titular de dados.

Note-se que a crítica a este complexo sistema de decisão não se dirige sempre ao algoritmo, em si mesmo considerado. Cremos que dependerá do caso concreto, uma vez que o algoritmo poderá não ser incorreto ou injusto, mas antes os dados pessoais que alimentam o sistema estarem deturpados. Se a base de dados da qual parte o algoritmo é incompleta ou errónea, o algoritmo introduzido com base no qual o sistema vai dar a resposta à tarefa necessariamente vai gerar um *output* também incorreto[40]. Se o quadro for o descrito, a decisão ou avaliação realizada poderá redundar num tratamento discriminatório, injusto e que afeta a esfera jurídica pessoal e/ou patrimonial do titular de dados, constituindo uma violação dos seus direitos.

Referências bibliográficas

AAVV, AI Watch: AI Standardisation Landscape state of play and link to the EC proposal for na AI reguatory framework, Luxemburgo, Publications office of the European Union, 2021

A. A. Leal, Big Data nos setores bancário e financeiro, AAVV, Inteligência artificial & Direito, Coimbra, Almedina, 2020, pp. 199 a 220

Cordeiro, A. Barreto Menezes, Direito da Proteção de Dados à luz do RGPD e da Lei nº 58/2019. Coimbra, Almedina, 2020

Cordeiro, A. Barreto Menezes, Comentário ao Regulamento Geral de Proteção de Dados e à Lei 58/2019. Coimbra, Almedina, 2021

[40] Sobre o tema, veja-se: B. Sobel, "Artificial Intelligence's fair use crisis", *Columbia Journal of law & the arts*, 41, 1, 2017, pp. 95 e 96.

Devia, A., La inteligencia artificial, el BigData y la Era digital: una amenaza para los datos personales, Revista la Propiedad Inmaterial, 27, 2019
Hildebrandt, M., The artificial intelligence of European Union Law. German Law Journal, 21(1), 2020
Jain, S. A. / Jain. S. A., Artificial Intelligence: a threat to privacy?, Nirma University Law Journal, 8(2), 2019
Lashbrooke, E., Legal reasoning and Artificial Intelligence, Loyola Law Review, 34 (2), 1988
Malhotra, H., Artificial Intelligence: A (semi-intelligent) overview, International In--house Counsel Journal, 11(41), 2018
Marquez, Albert M., Posthumanismo, inteligencia artificial y Derecho. Persona y Derecho, 84, 2021
Martinez, R., Artificial Intelligence: Distinguishing between types & definitions. Nevada Law Journal, 19 (3), 2019
Martz, J. G., Artificial Intelligence is here, Get ready!. Catholic University Journal of Law and Technology, 28 (1), 2019
Nemitz, P., Power in times of artificial intelligence. Delphi Interdisciplinary Review of emerging technologies, 2(24), 2019
Nicolau, I., Human rights and Artificial Inteligence. Journal of Law and Administrative Sciences, 12, 2019
Pinheiro, A. Sousa (coordenação), Comentário ao Regulamento Geral de Proteção de Dados, Coimbra, Almedina, 2019
Prieto, J., Inteligencia artificial, derechos humanos y bienes jurídicos, Navarra, Aranzadi, 2021
Rodriguez, P. Artificial Intelligence law: applications, risks & opportunities. Revista Juridica Universidad de Puerto Rico, 90(3), 2021
Sales, P., Algorithms, Artificial Intelligence, and the Law. Judicature, 105(1), 2021
Sobel, B., Artificial Intelligence's fair use crisis, Columbia Journal of law & the arts, 41(1), 2017
Solow-Niederman, A., Administring Artificial Intelligence, Southern California Law Review, 93(4), 2020
Trezza, R., Diritto e Intelligenza artificiale, Etica – *privacy* – Responsabilitá – decisione. Pisa, Pancini Giuridica, 2020
Tshider, C., The healthcare privacy – artificial intelligence impasse, Santa Clara High Technology Law Journal, 36 (4), 2020
Vicente, D. M., A inteligência artificial e iniciativas internacionais, AAVV, Inteligência artificial & Direito, Coimbra, Almedina, 2020, pp. 93 a 105.
Vinge, Vernor, The Coming Technological Singularity: How to Survive in the Post--Human Era, 1993, Disponível em: [PDF] The coming technological singularity: How to survive in the post-human era | Semantic Scholar (20/11/2021)
Vítor, M. S., Inteligência Artificial e contratos, AAVV, Inteligência artificial & Direito, Coimbra, Almedina, 2020, pp. 93 a 105
Watson, J. Artificial Intelligence?, Lawno, 22(1), 1997

Salud Digital a través de plataformas y otros prestadores de servicios: aproximación a un nuevo paradigma de Paciente Digital

Digital Health Platforms: approaching a new paradigm of Digital Patient

Raquel Luquin Bergareche*

resumen: El actual proceso de digitalizacion ha incrementado y diversificado la oferta de servicios privados de salud digital o telemedicina, tanto en contratos negociados individualmente como en los predispuestos en masa con condiciones generales ofertados por asociaciones médicas que se valen de intermediarios que proveen los medios tecnológicos, compañías aseguradoras de salud e incluso operadoras de telecomunicaciones que ofrecen a sus usuarios estos servicios a precios asequibles. La ponencia pretende aproximarse a algunas cuestiones jurídicas que el nuevo paradigma de la Salud Digital plantea desde la doble perspectiva del contratante o consumidor de estos servicios digitales y del paciente digital, con especial referencia a ciertos colectivos vulnerables como personas ancianas o con discapacitad.

palabras clave: Salud Digital. Telemedicina. Servicios de Salud Digital. Discapacidad.

abstract: The current process of digitalisation has generated an increase in and diversification of the private offer of services concerning telemedicine ("Med-Tech") which affects contracts negotiated individually and other predis-

* Profesora Titular de Derecho civil. Universidad Pública de Navarra. INARBE. ORCID 0000-0003-0330-6542. raquel.luquin@unavarra.es

posed mass contracts with general conditions provided by Medical Associations by intermediaries that provide technology, health insurance companies and telecommunications companies that offer these services at very low prices. The work aims to approach some legal issues that the new paradigm of Digital Health raises from the double perspective of digital services consumers and digital patients, with special reference to vulnerable groups such as elderly or disabled people.

KEYWORDS: Med-Tech. Telemedicine. Health Digital services. Disability.

1. Introducción. Telemedicina y servicios de salud

Bajo el concepto de telemedicina se incluyen diversas modalidades de prestación de servicios prestados a través de o haciendo uso de la tecnología digital: La relación médico-paciente transita hoy hacia un nuevo modelo que, por un lado, se vale de lo digital para ofrecer servicios a distancia con indiscutible reducción de costes de todo tipo (teleconsulta, registro y transmisión de datos, telediagnóstico, intervenciones en remoto incluida la cirugía de precisión o robotizada, y, por otro, se acerca a un modelo de "medicina personalizada" que utiliza el perfil genético y molecular de los individuos como guía para la adopción de decisiones médicas individualizadas en la prevención, diagnóstico y tratamiento de enfermedades. una relación entre profesional médico y paciente.

Hoy en día el paciente digital se relaciona telemáticamente con el profesional médico sin listas de espera, desplazamientos innecesarios, de forma eficiente y personalizada, en un entorno de digitalización creciente. Además, el contexto actual de crisis sanitaria y económica asociada a diversos factores satura y sobrecarga los servicios públicos sanitarios, allá donde existen modelos de sanidad pública, e incrementa las listas de espera, en especial en padecimientos y enfermedades relacionadas con la salud física y también mental, acentuadas con la pandemia de la COVID19.

[1] Este trabajo se ha realizado en el marco del Proyecto de Investigación "Contratación de Servicios de Telemedicina: Actualidad y Desafíos Jurídicos" (TED2021-129472B-00/MICINN "Next Generation EU"/PRTR).

Sin duda se produce un salto cualitativo esencial que va más allá del instrumento digital de comunicacion entre médico y paciente: el actual grado de desarrollo tecnológico permite al "paciente digital" el acceso desde determinadas plataformas a su historia cllínica, a la consulta, entre otros, de resultados de analíticas o de radiografías digitalizadas. Se redefinen los contornos del perfil del profesional sanitario, así como sus funciones y deberes y, del otro lado, se reconfigura el derecho a la autonomía del paciente en el proceso de toma de decisiones sobre su salud, de las que cada vez se hace más corresponsable. La relación médico-paciente requiere una adaptación al nuevo contexto tecnológico y de desarrollo de la Inteligencia Artificial: el paciente digital, más allá de su autonomía en la toma de decisiones en los términos de la LBRAP, debe colaborar en este contexto en la obtención y registro de los datos relativos a su estado físico o salud, desempeñando un nuevo rol protagonista y de auténtico "corresponsable" en el marco de este nuevo paradigma[2].

2. Contratación de servicios de Salud Digital a través de plataformas y otros prestadores de servicios

La relación jurídica entre el médico y el paciente que presta servicios de salud, también en el ámbito digital, se incardina bajo los moldes y contenido normativo del contrato de arrendamiento de servicios. Contrato civil regulado en nuestro ordenamiento jurídico en los arts. 1583 a 1587 de un Código civil decimonónico que, en este punto, se muestra insuficiente y obsoleto.

En el ámbito privado, estos servicios pueden prestarse en el seno de relaciones jurídicas "intuitu personae", en los que los particulares contratan a aquellos profesionales que les ofrezcan, generalmente por su prestigio o experiencia, mayores garantías en el cuidado o recuperación de un bien tan preciado como la salud. Modalidad de contratación entre particulares (piénsese en un médico con consulta privada) que en la práctica, cuando se trata de medicina digital, es hoy minoritaria debido al coste inherente a la digitalización y la IA. Frecuentemente son empresas y corporaciones las que publicitan y ofertan una pluralidad de servicios de salud digital en el mercado para lo cual reclutan cuadros de profesionales médicos. El predisponente (v.gr, multinacionales de las comunicaciones, grandes aseguradoras y otros proveedores) utilizan clausulados contractuales en masa con condiciones generales.

Las normas que rigen ambos tipos de contratos (los negociados individualmente y los que se conciertan mediante la adhesión del consumidor a

[2] https://www.fundacionidis.com/folletos/Experiencia_de_paciente_digital_2021.pdf

clausulados predispuestos con condiciones generales) son diferentes: en el primer caso se contrata con un particular, en el segundo con un consumidor.

Deben así distinguirse, por un lado, los servicios de telemedicina ofertados en el marco de un contrato de seguro privado, los que proveen empresas dedicadas a la comercialización online de servicios médicos privados y aquellos ofertados por grandes compañías tecnológicas especializadas en Salud Digital y operadores de las telecomunicaciones que los ofrecen a sus clientes a un coste asequible. Lo cual habría que distinguir de la prestación de este tipo de servicios directamente por los facultativos médicos, individualmente o asociados, generalmente mediante sociedades profesionales que recurren al apoyo de empresas tecnológicas para su organización, publicidad, comercialización, etc.. La distinción es importante, por cuanto va a determinar la responsabilidad en caso de incumplimiento y en el supuesto de materializarse daños personales o de naturaleza patrimonial.

A los contratos de adhesión predispuestos en masa con condiciones generales les será de aplicación en el ordenamiento español la Ley 7/1998, de 13 de abril, sobre Condiciones Generales de la Contratación y el Real Decreto Legislativo 1/2007, de 16 de noviembre por el que se aprueba el Texto Refundido de la Ley General para la Defensa de los Consumidores y Usuarios y normativa complementaria[3], recientemente modificado para adaptarse a la normativa comunitaria.

La Directiva 2000/31/CE, del Parlamento Europeo y el Consejo, de 8 de junio de 2000, relativa a determinados aspectos jurídicos de los servicios de la sociedad de la información conocida como "Directiva sobre el comercio electrónico", incorporada al ordenamiento español por la Ley 34/2002, de 11 de junio, de Servicios de la Sociedad de la Información y del Comercio Electrónico (LSSICE), ha puesto las bases para la creación del marco jurídico necesario para dotar a los operadores económicos de seguridad jurídica en el mercado de servicios digitales[4]. Su Considerando 18 dispone que aquellas actividades que no pueden realizarse a distancia, como, por ejemplo, el asesoramiento médico que requiere del reconocimiento físico de un paciente, no constituyen servicios de la sociedad de la información.

[3] Ni la Ley 7/2010, de 31 de marzo, General de la Comunicación Audiovisual ni la posterior Ley 9/2014, de 9 de mayo, General de Telecomunicaciones contemplan la prestación en masa y a través de medios telemáticos de servicios médicos o relativos a la salud de los pacientes.

[4] Juan José Castelló Pastor, "Nuevo régimen de responsabilidad de los servicios digitales que actúan como intermediarios a la luz de la propuesta de Reglamento relativo a un mercado único de servicios digitales", en Juan José Castelló Pastor (Dir.), Desafíos jurídicos ante la integración digital: aspectos europeos e internacionales, Aranzadi-Thomson Reuters, 2021, pp. 38-77.

El mes de abril de 2021 fueron objeto de transposición al ordenamiento español: las Directivas (UE) 2019/770 del Parlamento Europeo y del Consejo de 20 de mayo de 2019 relativa a determinados aspectos de los contratos de suministro de contenidos y servicios digitales ("Directiva de servicios digitales") y la Directiva (UE) 2019/771 del Parlamento Europeo y del Consejo de 20 de mayo de 2019, relativa a determinados aspectos de los contratos de compraventa de bienes ("Directiva sobre compraventa de bienes"). Por ello, el Título VIII del Real Decreto-ley 7/2021, de 27 de abril[5] lleva por título "Transposición de directivas de la Unión Europea en materia de contratos de compraventa de bienes y de suministro de contenidos o servicios digitales".

Habiendo transcurrido más de dos décadas desde la aprobación de la Directiva de Comercio Electrónico, lapso temporal que implica un auténtico punto de inflexión en aspectos sanitarios, tecnológicos, económicos y sociales, el actual proceso de digitalización ha originado nuevos y cada vez más complejos modelos de negocio y relaciones jurídicas de prestación de servicios en línea (en ámbitos como el comercio, comunicación, salud, trabajo, ocio, cultura, etc), obligando a revisar algunos principios y normas que han devenido obsoletas ante los avances tecnológicos y de la IA. Una de las prioridades de la Comisión Europea (2019- 2024)[6] es alcanzar una "Europa adaptada a la era digital", para lo cual el ejecutivo comunitario se propone "configurar el futuro digital de Europa" sobre tres pilares: la tecnología al servicio de las personas, una economía digital que sea justa a la vez de competitiva y una sociedad abierta, democrática y sostenible. A tal fin, el 15 de diciembre de 2020, inmersa la comunidad global en una crisis sanitaria sin precedentes por la pandemia de la Covid-19, la Comisión presentaba la Propuesta de un Reglamento relativo a un mercado único de servicios digitales conocido como "Ley de Servicios Digitales" o "Digital Services Act" y otra sobre los mercados competitivos y justos en el sector digital conocida como "Ley de Mercados Digitales" o "Digital Market Act", normas que modernizan el régimen aplicable a los prestadores de servicios. El 5 de julio de 2022, el Parlamento Europeo aprobaba finalmente, tras un intenso periodo de consultas y negociaciones, el Reglamento que contiene la "*Digital Markets Act"* (DMA) o "Ley de Mercados Digitales", norma jurídica comunitaria de aplicación directa a todos los estados miembros que sujeta a las grandes plataformas digitales a un régimen regulatorio que establece obligaciones específicas, prohíbe ciertas

[5] https://www.boe.es/diario_boe/txt.php?id=BOE-A-2021-6872

[6] "Una Europa adaptada a la era digital. Capacitar a las personas con una nueva generación de tecnologías" https://ec.europa.eu/info/strategy/priorities-2019- 2024/europe-fit-digital-age.es

conductas y contiene un estricto régimen sancionador que, en cierto modo, es réplica del previsto para las infracciones que constituyen vulneraciones de la normativa sobre competencia. La Digital Services Act, por su parte, ha sido publicada en el DOUE a la fecha de cierre de este trabajo: 27 de octubre 2022. El Reglamento (UE) 2022/2065 del Parlamento Europeo y del Consejo de 19 de octubre de 2022 relativo a un mercado único de servicios digitales y por el que se modifica la Directiva 2000/31/CE ("Reglamento de Servicios Digitales" o DSA)[7] pretende ir más allá[8]: las indudables funcionalidades de los servicios digitales no ocultan que los mismos se han convertido en fuente de riesgos tanto para la sociedad en su conjunto como para las personas que hacen uso de ellos. En su Comunicación "Shaping Europe's Digital Future" ("Configurar el futuro digital de Europa"), la Comisión Europea se comprometió a actualizar las normas horizontales que definen la responsabilidad[9] y obligaciones de los prestadores de servicios digitales, y especialmente de las denominadas "plataformas en línea".

En su Exposición de Motivos, la conocida como "Digital Services Act" (DSA) señala que "sobre la base de los principios esenciales establecidos en

[7] Tras la adopción por el PE en el pasado mes de julio de 2022 del conocido como "paquete de servicios digitales" o "Digital Services Act package" que comprende la denominada "Ley de Servicios Digitales" (en adelante, DSA) y la "Ley de Mercados Digitales" (DMA), ambos textos deben ser adoptados por el Consejo de la UE, tras lo cual serán firmadas por los Presidentes de ambas instituciones y publicadas en el DOUE, entrando en vigor 20 días desde su publicación. Sobre la base de los principios establecidos en la Directiva sobre el comercio electrónico, la DSA se propone como objetivos garantizar las mejores condiciones para la prestación de servicios digitales innovadores en el mercado interior de la UE, contribuir a la seguridad en línea y a la protección de los derechos fundamentales y establecer una estructura de gobernanza robusta y duradera para la supervisión efectiva de los prestadores de servicios intermediarios.

[8] https://www.boe.es/doue/2022/277/L00001-00102.pdf

[9] Merece un trabajo más amplio que el que permite esta ponencia el análisis de los posibles daños, patrimoniales y personales, que puden generarse de la medicina digital o telemedicina en el marco del modelo de Salud Digital ("e-Health"). Paradigma que se fundamenta en una nueva manera de concebir la relación médico-paciente y que que confiere un nuevo protagonismo activo al paciente digital, que parte de su autonomía en la toma de deciciones que afectan a su salud ("derecho de autodeterminación en el ámbito sanitario", reconocido en la LBRAP) si bien reforzando el nivel de autorresponsabilidad (piénsese en la medición y registro de sus propias variables –presión arterial, niveles glucémicos, etc.- en el uso de dispositivos electrónicos de última generación) sobre la base de la prestación de un consentimiento digital de nuevos perfiles que plantea no pocas cuestiones (entre ellas, las relativas a la protección de los datos personales, que en el ámbito sanitario son especialmetne sensiblles y merecedores de singular protección juridica. Raquel Luquin Bergareche, "Capítulo 6. Prestación de Servicios de Salud Digital: Algunas Reflexiones desde el Derecho Civil", en El impacto de la inteligencia artificial en la teoría y la práctica jurídica (Solar Cayón, J.I. y Sánchez Martínez, Mª.O., Madrid, julio 2022, pp. 167 a 194.

la Directiva sobre el comercio electrónico, que mantienen su validez hasta hoy, esta propuesta pretende garantizar las mejores condiciones para la prestación de servicios digitales innovadores en el mercado interior, contribuir a la seguridad en línea y la protección de los derechos fundamentales, y establecer una estructura de gobernanza robusta y duradera para la supervisión efectiva de los prestadores de servicios intermediarios".

En el marco de la UE, los intermediarios o proveedores de servicios de intermediación en línea son agentes fundamentales de la transformación digital. Sin embargo, los servicios de intermediación en línea son frecuentemente utilizados como canales de venta de mercancías ilegales, falsificadas o peligrosas, de prestacion de servicios contrarios a la ley o a los derechos reconocidos en la normativa de prortección de los consumidores y usuarios o de difundir contenido ilícito por Internet: por esta razón, se ha optado por establecer un régimen de responsabilidad de los intermediarios en línea, cuyas primeras bases se contienen en la "Directiva sobre el comercio electrónico".

La DSA se aplica, si bien con distinto grado de obligaciones, a los servicios de Alojamiento de datos, a los motores de búsqueda en línea, a las redes sociales y a los "*marketplaces*". En concreto, se incluye la siguiente clasificación[10]:

Los "Servicios de intermediación" son los que ponen a disposición de los usuarios las infraestructuras de red e incluyen proveedores de acceso a Internet y registradores de nombres de dominio, entre otros;

Los "Servicios de alojamiento de datos" almacenan información proporcionada por los destinatarios del servicio a petición e incluyen, entre otros, los servicios de computación en nube o de alojamiento web;

Las "plataformas en línea" son las redes sociales o *marketplaces*, como los prestadores de servicios de alojamiento de datos.

Distinción que cobra importancia en el campo de la Salud Digital en que algunos servicios se prestan con el apoyo de intermediarios o a través de plataformas. Las aseguradoras de salud disponen de servicios digitales a través del propio seguro médico con marca propia o impulsando nuevos negocios digitales. También los grupos hospitalarios están evolucionando sus canales hacia plataformas más generalistas. Comunidades de profesionales médicos, agrupados o no por especialidades (cirugía plástica, nutrición, servicios de salud mental, etc.), de manera que, a la fecha de este trabajo, es habitual la

[10] Carolina Pina, "Ley de Servicios Digitales (DSA): un nuevo marco legal para las plataformas digitales de servicios intermediarios". Disponible en https://www.garrigues.com/es_ES/garrigues-digital/ley-servicios-digitales-dsa-nuevo-marco-legal-plataformas-digitales-servicios

prestación de servicios digitales específicos, como teleconsultas o consultas médicas por videollamada o pre-diagnósticos o sistemas de pre-evaluación de síntomas del paciente basados en aplicaciones.

3. El consumidor contratante de servicios de salud digital y el nuevo concepto de paciente digital

En el destinatario de servicios de Salud Digital debe distinguirse su doble condición de "parte-contratante" y de "paciente-digital". En lo que respecta al particular que contrata o hace uso de servicios privados de Salud Digital, puede ser "arrendatario-usuario" de servicios médicos digitales (que se rige por la normativa civil general) o bien consumidor si es parte de un contrato (de adhesión) de servicios médicos digitales con condiciones generales predispuestas por el oferente (rigiéndose por la normativa de consumidores y usuarios).

El Título III TRLGDCU regula a los contratos celebrados a distancia y contratos celebrados fuera del establecimiento mercantiles. El art. 93 TRLGDCU excluye del ámbito de aplicación de las normas sobre contratos celebrados a distancia y fuera de establecimiento mercantil a "los contratos de servicios relacionados con la salud, prestados por un profesional sanitario a pacientes para evaluar, mantener o restablecer su estado de salud, incluidos la receta, dispensación y provisión de medicamentos y productos sanitarios, con independencia de que estos servicios se presten en instalaciones sanitarias". Exclusión debida a que la asistencia sanitaria exige una regulación especial debido a su complejidad técnica, su importancia como servicio de interés general y su financiación pública.

Del mismo modo, el Cap I del Título IV TRLGDCU, relativo a las garantías y servicios de postventa, establece en el núm. 2 del art. 114 que

"lo previsto en este título no será de aplicación a:

"e) Los contenidos o servicios digitales relacionados con la salud prescritos o suministrados por un profesional sanitario a pacientes para evaluar, mantener o restablecer su estado de salud, incluidos la receta, dispensación y provisión de medicamentos y productos sanitarios".Exclusión que se encuentra en consonancia con lo prescrito por la Directiva sobre determinados aspectos de los contratos de suministro de contenidos y servicios digitales de 2019[11].

[11] En esta línea de "merma de protección" frente al consumidor, cabe destacar, respecto de las normas de Derecho Internacional privado, que, conforme al Reglamento Roma I, como regla general, sería de aplicación a los contratos celebrados con consumidores la regulación del lugar

En su cualidad de "paciente en sentido estricto" (art. 3 LBRAP), es decir, en cuanto persona sometida a cuidados profesionales para el mantenimiento o recuperación de su salud, el paciente digital decide sobre su propio proceso de salud y cada una de las intervenciones y actuaciones médicas que se realicen, prestando su consentimiento informado de forma escrita o por vía digital pero en cualquier caso, de forma autónoma, libre y consciente (consentimiento informado) tras recibir la oportuna información asistencial, como se deriva del art 5 del Convenio de Oviedo y arts. 8 a 10 LBRAP.

Resulta necesario abordar las especificidades que presenta el consentimiento informado del paciente digital en un análisis específico que, hasta el momento, no se ha realizado. El "Consentimiento Informado Digital" (CID) es una declaración de voluntad en la que el paciente, partiendo de haber conocido y entendido la naturaleza, finalidad, riesgos, efectos y consecuencias de un tratamiento médico propuesto de forma personalizada, manifiesta su aceptación capaz, consciente y libre a todos los efectos, asumiendo personalmente la responsabilidad de las decisiones que tome sobre su salud ("derecho de autodeterminación personal" en el ámbito sanitario). Será necesario para ello articular mecanismos jurídicos (más allá de las posibilidades técnicas) que garanticen una voluntad libre y consciente en la emisión y manifestación de dicho consentimiento. El desarrollo tecnológico permite hoy, más allá de las firmas electrónicas y certificados digitales, sistemas de reconocimiento facial y otras aplicaciones algorítmicas basadas en blockchain y aplicadas a este ámbito, suya virtualidad y funcionalidad práctica resultan innegables. Corresponde al jurista la reflexión sobre los límites, de haberlos, de dichas aplicaciones digitales. La tecnología, regida por criterios de eficiencia, debe estar al servicio de un desarrollo que, respetando, a la vez, principios de justicia y equidad, ponga en su centro a la persona humana y la dignidad y derechos que le son inherentes.

de residencia habitual del consumidor. Sin embargo, según resulta de los arts. 3.5 y 7 de 20 4 del RD 81/2014 sobre asistencia sanitaria transfronteriza, en el caso de la telemedicina se aplicará la normativa del estado donde esté establecido el proveedor. En cualquier caso, se trata de una normativa de carácter muy dinámico, como lo demuestra el Real Decreto-Ley 24/2021, de 2 de noviembre, que transpone diversas directivas en este ámbito: el 1 de enero de 2022 han entrado en vigor las modificaciones al TRLGDCU reguladas en el Real Decreto-Ley 7/2021 de transposición de la Directiva (UE) 2019/770 de 20 de mayo de 2019, y la Directiva (UE) 2019/771 de 20 de mayo de 2019. En esta materia véase el reciente trabajo de J. Alvarez Rubio, "Respuestas jurídicas a la personalización de ofertas mediante tratamientos automatizados de datos", en "El impacto de la inteligencia artificial en la teoría y la práctica jurídica" coord. José Ignacio Solar Cayón y Maria Olga Sánchez Martínez, pp. 74-82.

El status del paciente digital, sus derechos y deberes en el nuevo paradigma de salud digital (que sin duda le confieren un mayor protagonismo y una mayor responsabilidad en la toma de decisiones), se halla en íntima conexión con la "lex artis" del profesional médico que presta servicios bajo este nuevo modelo de telemedicina y asistencia personalizada. "Lex artis ad hoc" que, en el ámbito sanitario, delimita los deberes del profesional de la salud y, el actual marco de digitalización, modaliza algunas obligaciones y, acaso, hace surgir otras. Consiguientemente se demanda una revisión de los tradicionales criterios y normas reguladoras de la responsabilidad civil, tanto contractual como extracontractual, cuando se trata de este tipo de servicios.

Desde el punto de vista jurídico, resulta esencial delimitar tales deberes y analizar los criterios y extensión de los límites de la responsabilidad civil que pudiera derivarse de su incumplimiento. Piénsese en la obligación de recabar un consentimiento informado, libre y consciente del paciente digital anciano carente de las capacidades y habilidades digitales mínimas, o en algunas personas con discapacidad que contratan este tipo de servicios, o en los efectos de la violación de la confidencialidad de las historias clínicas y de los datos sanitarios transmitidos y/o registrados telemáticamente. En un hipotético caso de error en el diagnóstico.

Son muchos los interrogantes[12] que, en un primer acercamiento a la cuestión, podrían plantearse: ¿qué responsabilidad cabe exigir al paciente que no ha "facilitado correctamente los datos por vía digital", tal y como aparece en los clausulados generales de estos contratos en masa que es obligación del paciente-usuario?;¿en qué supuestos, y bajo qué criterios, podría hacerse res-

[12] ¿Quid de la "lex artis" que debe exigirse al médico digital en sus funciones de diagnóstico o tratamiento de enfermedades? ¿y cuando recaba, registra y verifica digitalmente datos de pacientes... o hace uso de aplicaciones y recursos de IA...? ¿Quid del error diagnóstico inducido por fallos o defectuosa utilización de los dispositivos de medición, registro o del traslado de datos por parte del mismo paciente?; ¿Podría exigirse, y en base a qué criterios (–culpa "in eligendo", ¿in organizando?, ¿responsabilidad objetiva...?-) responsabilidad por daños y perjuicios a las empresas que proveen las plataformas de contratación digital de servicios médicos y reclutan profesionales sanitarios?;¿Cómo se abordaría jurídicamente un supuesto de error en el diagnóstico efectuado por el segundo facultativo consultado que se materializa en la causación de daños al paciente...?; ¿cómo influye el distinto grado de "competencia digital" o de "habilidades digitales"?; ¿cómo afecta en la relación médico-paciente en lo que se refiere a la responsabilidad civil contractual (art. 1101 CC)?; ¿es posible la prestación del consentimiento informado por identificación biométrica de la persona?; ¿cómo se articula la "información asistencial" en la telemedicina); ¿qué consecuencias o reproches culpabilísticos pueden colegirse, de la "corresponsabilidad" –mayor protagonismo- del paciente en la medicina digital?; ¿en qué medida afecta a la formulación del juicio de imputabilidad y al nexo causal verificado un daño antijurídico (responsabilidad extracontractual)...?, etc.

ponder al médico que atiende al paciente en teleconsulta por los daños causados por fallos o errores tecnológicos? ¿Y a aquellos proveedores que, más allá del suministro de la herramienta digital, controlan -o participan- en el reclutamiento o selección de los cuadros médicos que prestan los servicios de telemedicina (servicios profesionales, sujetos a normativa específica y a códigos deontológicos), y realizan promociones y ofertas, y participan en la realización y verificación del pago por los usuarios del servicio, constituyendo en estos aspectos la "cara visible" frente al consumidor...?

En el caso de pacientes con problematicidad específica, se plantean igualmente algunas cuestiones en el marco de un ecosistema, el de la Salud Digital, que, como decimos, refuerza la autorresponsabilidad del paciente. Tras la Ley 8/2021, de 2 de junio[13], por la que se reforma la legislación civil y procesal para el apoyo a las personas con *discapacidad* en el ejercicio de su capacidad jurídica, las personas discapacitadas se conducen de forma autónoma en el ejercicio de esta capacidad, sin perjuicio de las medidas de apoyo necesarias como la guarda de hecho o la curatela. El nuevo sistema se basa en el reconocimiento de la dignidad y autonomía de las personas discapacitadas, al margen (con carácter general) de mecanismos de representación legal.

Por otra parte, en sociedades como las actuales de envejecimiento poblacional exponencial, las personas de edad avanzada, lógicamente con mayores padecimientos de salud, representan una proporción considerable de pacientes digitales. Estas personas tienen menos habilidades tecnológicas, cuando no carecen absolutamente de ellas, por lo que deberían articularse a todos los niveles, también en el normativo, mecanismos de control de posibles situaciones abusivas o de indebida influencia en la contratación de servicios médicos digitales o en la prestación del consentimiento informado preciso para el tratamiento de datos personales o previo a la realización de pruebas diagnósticas, de intervención sanitaria, rehabilitadoras o cualesquiera otras que afecten a su vida y salud.

4. Algunas reflexiones a modo de conclusión

Tratándose frecuentemente de contratos de adhesión con condiciones generales, habrá de tenerse en cuenta la aplicación a este tipo de relaciones contractuales de la normativa reguladora de las relaciones de Consumo: el vigente TRLGDCU, aprobado por Real Decreto Legislativo 1/2007, de 16 de noviembre. En el ámbito de la Unión Europea, debe tenerse en cuenta el marco normativo de responsabilidad por la prestación de servicios digitales previsto en la

[13] https://www.boe.es/diario_boe/txt.php?id=BOE-A-2021-9233

reciente Ley de Mercados Digitales (DMA) y en la Ley de Servicios Digitales (DSA)[14]. En lo que se refiere a los contratos de suministro de contenidos y servicios digitales, son dos las directivas comunitarias que afectan al TRLGDCU aprobado mediante Real Decreto Legislativo 1/2007, de 16 de noviembre: la Directiva (UE) 2019/770 o "Directiva de servicios digitales"[15] y la Directiva (UE) 2019/771 o "Directiva sobre compraventa de bienes".

En España es aplicable la Ley de Servicios de la Sociedad de la Información y del Comercio Electrónico (LSSICE). El art. 93 TRLGDCU excluye del ámbito de aplicación de las normas sobre este particular a "los contratos de servicios relacionados con la salud, prestados por un profesional sanitario a pacientes para evaluar, mantener o restablecer su estado de salud, incluidos la receta, dispensación y provisión de medicamentos y productos sanitarios, con independencia de que estos servicios se presten en instalaciones sanitarias": exclusión debida a que la asistencia sanitaria exige una regulación especial debido a su complejidad técnica, su importancia como servicio de interés general y su financiación pública.

En su dimensión de consumidor, el destinatario de estos servicios de telemedicina puede beneficiarse de la aplicación de las normas relativas al mismo y, en particular, del TRLDCU (normas de servicios de atención al cliente, regulación de cláusulas abusivas, responsabilidad por productos o servicios defectuosos, integración de la publicidad en el contrato etc.). Sin embargo, además de las normas de contratación a distancia (inaplicables) hay otras de las que no puede beneficiarse aunque estén expresamente pensadas para servicios digitales: así, el art. 66 bis TRDCU sobre suministro de servicios digitales (que regula la obligación de entrega en cierto plazo y la posibilidad de resolución del contrato) no se aplica a los contratos excluidos de las normas sobre contratación a distancia como los servicios sanitarios. Tampoco cabe aplicar las disposiciones sobre garantías y servicios posventa (remedios frente a la falta de conformidad), pues el art. 114 TRLCU excluye del ámbito de aplicación del título correspondiente "los contenidos o servicios digitales relacionados con la salud prescritos o suministrados por un profesional sanitario a pacientes para evaluar, mantener o restablecer su estado de salud,

[14] https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52020PC0825&from=en

[15] En la DSA se definen las responsabilidades a cargo de los prestadores de servicios intermediarios, especialmente las plataformas en línea como los mercados y las redes sociales, determinando la imposición de obligaciones de "diligencia debida" (concepto jurídico indeterminado no exento de problemas a la hora de su concreción práctica) para determinados intermediarios, incluyendo procedimientos de notificación y acción en relación con los contenidos ilícitos y la posibilidad de impugnar las decisiones de moderación de contenidos de las plataformas.

incluidos la receta, dispensación y provisión de medicamentos y productos sanitarios". Exclusión que se encuentra en consonancia con lo prescrito por la Directiva sobre determinados aspectos de los contratos de suministro de contenidos y servicios digitales de 2019.

En lo que respecta a la protección de datos personales en el ámbito de los servicios de salud digital, la Comisión Europea adoptó el 23 de febrero de 2022 la Propuesta de Reglamento sobre normas armonizadas sobre el acceso justo a los datos y a su uso, conocido como "Data Act"[16], pilar básico de la estrategia europea de protección de datos en la era digital, más allá de las normas del RGPD[17] y las contenidas en la Ley de Mercados Digitales (DMA) y en la Ley de Servicios Digitales (DSA), que se revelan insuficientes para dar respuesta a algunas cuestiones. Como las atinentes al riesgo que esta modalidad de servicios médicos puede implicar para la protección de derechos fundamentales como la imagen y la intimidad de un paciente que no es atendido en espacios cuya configuración garantiza su privacidad (consultas privadas o centros hospitalarios), sino a distancia y haciendo uso de aplicaciones incluidas en dispositivos que pudieran, por fallo tecnológico o causa intencional, vulnerar tales prerrogativas. Del mismo modo debe preservarse la confidencialidad de los datos médicos y las historias clínicas cuando el tratamiento de los datos tiene lugar a través de un medio telemático que pudiera adolecer de fallos de seguridad.

En este ámbito, se plantean algunos interrogantes que no ofrecen respuesta unívoca en el actual marco jurídico[18]. A modo ejemplificativo: ¿debe revisarse la "lex artis" o diligencia que debe exigirse al profesional médico que actúa a través de canales digitales en sus funciones de diagnóstico o tratamiento de enfermedades, o que recaba, registra y verifica datos de pacientes por medios telemáticos, o hace uso de aplicaciones y recursos de IA?; ¿Quid en caso de un posible error en el diagnóstico médico provocado o inducido por fallos o defectuosa utilización de los dispositivos de medición, registro o traslado de datos?; ¿podría exigirse, y en base a qué criterios responsabilidad por daños y perjuicios a las empresas que proveen las plataformas de contratación digital

[16] https://digital-strategy.ec.europa.eu/en/policies/data-act

[17] https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:32016R0679&from=EN

[18] Vid. Susana Navas Navarro. "Sistemas expertos basados en inteligencia artificial y responsabilidad civil"; "Telemedicina y protección de datos sanitarios (aspectos legales y éticos)", Diario la Ley, 2019, pp. 7-8 y Raquel Luquin Bergareche. "Prestación de servicios de salud digital: algunas reflexiones desde el derecho civil" en "El impacto de la inteligencia artificial en la teoría y la práctica jurídica", coord. Solar Cayón, J.I y Sánchez Martínez, M. O., La Ley, 2022, pp. 167-194.

de servicios médicos y reclutan profesionales sanitarios para la prestación de servicios médicos online? (¿"culpa in eligendo", "in organizando", responsabilidad objetiva...?), ¿cómo se abordaría jurídicamente la imputabilidad en un supuesto de error del diagnóstico clínico efectuado por un segundo facultativo consultado cuando del mismo se derivan daños para el paciente...? Cuestiones, entre otras que no ofrecen respuesta en el actual marco jurídico regulador de unos servicios que, desbordan los moldes de la normativa codificada (Códig Civil), y que en la actualidad aparece fragmentado, disperso, poco adaptado a la reaidad digital y, por consiguiente, insuficiente.

La relación médico-paciente requiere una adaptación al nuevo contexto tecnológico y de desarrollo de la Inteligencia Artificial: el paciente digital, más allá de su autonomía en la toma de decisiones, cuando se trata de salud digital está implicado en la obtención y registro de los datos relativos a su estado físico o salud, desempeñando un nuevo rol protagonista y de auténtico "corresponsable"[19]. En particular, el denominado "Consentimiento Informado Digital" (CID) es una declaración de voluntad en la que el paciente, después de haber conocido y entendido la naturaleza, finalidad, riesgos, efectos y consecuencias de un tratamiento médico propuesto de forma personalizada, manifiesta su aceptación capaz, consciente y libre a todos los efectos, asumiendo personalmente la responsabilidad de las decisiones que tome sobre su salud ("derecho de autodeterminación personal" en el ámbito sanitario). Es preciso articular mecanismos jurídicos de salvaguarda de una voluntad libre y consciente en la emisión y manifestación de dicho consentimiento. El desarrollo tecnológico permite firmas electrónicas y certificados digitales, sistemas de reconocimiento facial y otras aplicaciones algorítmicas basadas en blockchain, contratos inteligentes, reuniones en el metaverso. Desarrollos tecnológicos que sin duda suponen un avance también en el ámbito sanitario.

La "lex artis" del médico digital en el ámbito sanitario delimita los deberes del profesional de la salud y, el actual marco de digitalización de servicios, se modalizan algunas obligaciones y, acaso, surgen nuevas. Lo que a nuestro juicio reclama una revisión de los criterios y normas reguladoras de la responsabilidad civil, tanto contractual como extracontractual, cuando se trata de este tipo de servicios.

En el caso de pacientes con problematicidad específica, se plantean igualmente algunas cuestiones en el marco de un ecosistema de Salud Digital que refuerza la autorresponsabilidad del paciente. Tras la Ley 8/2021, de 2 de junio, en España las personas discapacitadas se conducen de forma autónoma en el

[19] https://www.fundacionidis.com/folletos/Experiencia_de_paciente_digital_2021.pdf

ejercicio de esta capacidad, sin perjuicio de las medidas de apoyo necesarias como la guarda de hecho o la curatela. El nuevo sistema se basa en el reconocimiento de la dignidad y autonomía de las personas discapacitadas, al margen (con carácter general) de mecanismos de representación legal.

En sociedades como las occidentales de envejecimiento poblacional, las personas de edad avanzada y ancianas, con mayores padecimientos de salud, representan una proporción considerable de pacientes digitales. Estas personas tienen menos habilidades tecnológicas, cuando no carecen absolutamente de ellas, por lo que deberían articularse a todos los niveles, también normativo, mecanismos de control de posibles situaciones abusivas o de indebida influencia en la contratación de servicios médicos digitales o en la prestación del consentimiento informado preciso para el tratamiento de datos personales o previo a la realización de pruebas diagnósticas, de intervención sanitaria, rehabilitadoras o cualesquiera otras que afecten a su vida y salud.

Por último, no debe olvidarse que este ámbito de la Salud Digital se halla cada vez más vinculado al de la Inteligencia Artificial (IA) y la Robótica aplicadas al ámbito médico y biomédico: el primer marco normativo europeo IA está representado por el Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en materia de Inteligencia Artificial y se modifican determinados actos legislativos de la Unión, de 1 de abril de 2021[20], al que se acompaña la Propuesta de regulación sobre Maquinaria y Robots. Normativa que pretende aunar una perspectiva realista de reconocimiento del avance tecnológico que representa la introducción del algoritmo con un enfoque garantista en la protección de los derechos fundamentales de las personas (v.gr, prohibición del reconocimiento facial como modo de identificación de las personas en determinadas circunstancias)

Referencias bibliográficas

Castelló Pastor, Juan José, "Disponibilidad de productos y servicios en línea en el mercado único digital"", en Castelló Pastor, jj.; Guerrero Pérez, a.; Marínez Pérez, M., (Dirs.), "Derecho de la contratación electrónica y comercio electrónico en la Unión Europea y en España", Tirant lo Blanch, Valencia, 2021;

Castelló Pastor, Juan José, "Motores de búsqueda y derechos de autor: infracción y responsabilidad", Aranzadi, Cizur Menor, 2016

Castelló Pastor, Juan José, "Nuevo régimen de responsabilidad de los servicios digitales que actúan como intermediarios a la luz de la propuesta de Reglamento de un mercado único de servicios digitales", en Castello Pastor, J.J. "Desafíos jurídi-

[20] https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-

cos ante la integración digital: aspectos europeos e internacionales, Aranzadi Thomson Reuters, 2021

Luquin Bergareche, Raquel, "Acerca de la redefinición de la autonomía privada en la sociedad tecnológica", Revista Boliviana de Derecho, Nº. 26, 2018

Luquin Bergareche, Raquel, "Capítulo 6. Prestación de Servicios de Salud Digital: Algunas Reflexiones desde el Derecho Civil", en El impacto de la inteligencia artificial en la teoría y la práctica jurídica (Solar Cayón, J.I. y Sánchez Martínez, Mª.O., Madrid, julio 2022

Luquin Bergareche, Raquel, "Contratación tecnológica en la era del algoritmo: gestión de la COVID-19 y futuro Blockchain" en "Covid19: conflictos jurídicos actuales y desafíos / Luquin Bergareche (dir.), 2020, Wolters Kluwer, Bosch, págs. 177-196

Navas Navarro, Susana "Daños ocasionados por sistemas de Inteligencia Artificial", Ed. Comares, 2022; "Sistemas expertos basados en inteligencia artificial y responsabilidad civil"; "Telemedicina y protección de datos sanitarios (aspectos legales y éticos)", Diario la Ley, Madrid, 2019

Pina, Carolina, "Ley de Servicios Digitales (DSA): un nuevo marco legal para las plataformas digitales de servicios intermediarios". https://www.garrigues.com/es_ES/garrigues-digital/ley-servicios-digitales-dsa-nuevo-marco-legal-plataformas-digitales-servicios

Solar Cayón, José Ignacio, "La Inteligencia Artificial Jurídica. El impacto de la innovación tecnológica en la práctica del Derecho y el mercado de servicios jurídicos", Thomson Reuters Aranzadi, Cizur Menor, 2022

VI
MERCADO E EMPRESA NUM MUNDO COMPUTACIONAL

Concorrência e Inteligência Artificial: *the good, the bad and the ugly*[*]

Competition and Artificial Intelligence: *the good, the bad and the ugly*

Inês Neves[**]

resumo: O objetivo do texto é o de expor a Inteligência Artificial como realidade e desafio de efeitos heterogéneos, também no âmbito do Direito da Concorrência. Em face das dúvidas suscitadas pelo papel disruptivo assumido pela Inteligência Artificial, seja nos modelos de negócio das empresas, seja na própria aplicação das normas da concorrência, importa garantir que a procura de um qualquer novel enquadramento jurídico e a opção de adaptar ou substituir os quadros e instrumentos vigentes por outros considerados mais idóneos a acomodar o *specialis* da Inteligência Artificial, não assentem num preconceito ideológico, num prejuízo e/ou numa leitura absolutizante e parcial de realidade que surge hoje verdadeiramente *imposta* ao intérprete. Pelo contrário, quaisquer vias de solução apenas respeitarão as exigências da proporcionalidade se lograrem evitar um *chilling-effect* à inovação. A exploração dos efeitos da Inteligência Artificial e o seu reconhecimento como fenómeno multifacetado constituem, pois, a missão do escrito, dirigido a garantir que qualquer enquadramento normativo que se procure aplicar ao caso não acabe silenciando as vantagens inegáveis que a Inteligência Artificial aporta, seja para a concorrência (*pelo* mercado e *no* mercado), seja para a aplicação *eficaz* das suas normas e princípios.

[*] Logramos em inspiração em "Il buono, il brutto, il cattivo", de 1966, direção de Sergio Leone.

[**] Assistente Convidada da Faculdade de Direito da Universidade do Porto (FDUP) | Investigadora Colaboradora do Centro de Investigação Jurídica da Faculdade de Direito da Universidade do Porto (CIJ). ineves@direito.up.pt. ORCID ID: 0000-0003-0448-2951

PALAVRAS-CHAVE: Inteligência Artificial, concorrência efetiva, restrição da concorrência, abuso, controlo de operações de concentração, *enforcement*

ABSTRACT: This text aims to expose Artificial Intelligence as a reality and a challenge with heterogeneous effects in Competition Law as well. In view of the doubts raised by Artificial Intelligence's disruptive role, both in business models and in the very application of competition rules, the legal framework, especially when amending or repealing existing rules to ensure that they are 'fit for Artificial Intelligence' must not be based on an ideological bias, a prejudice and/or a partial reading of the current reality. On the contrary, any solutions will only respect the requirements of proportionality if they manage to avoid a chilling-effect on innovation. Exploring Artificial Intelligence's effects as a multifaceted phenomenon is thus the focus of this paper. It is meant to ensure that any normative framework applying to the case does not end up silencing Artificial Intelligence's undeniable advantages, either for competition (for the market and in the market), or for the effective application of its rules and principles.

KEYWORDS: Artificial Intelligence, effective competition, restriction of competition, abuse, merger control, enforcement

1. Considerações introdutórias

De acordo com a Comissão Europeia, o conceito de Inteligência Artificial ('IA') abrange "*sistemas que apresentam um comportamento inteligente, analisando o seu ambiente e tomando medidas – com um determinado nível de autonomia – para atingir objetivos específicos*"[1]. Por seu turno, a definição avançada pelo High-Level Expert Group on Artificial Intelligence salienta características

[1] *Cf.* Comissão Europeia, *Comunicação da Comissão ao Parlamento Europeu, ao Conselho Europeu, ao Conselho, ao Comité Económico e Social Europeu e ao Comité Das Regiões: Inteligência artificial para a Europa*, COM/2018/237 final, Bruxelas (25.4.2018, p. 1).

como *i)* a capacidade de perceção do meio ambiente (e de adaptação do comportamento àquele meio), *ii)* a interpretação desse mesmo meio ambiente, sob a forma de dados mais ou menos estruturados, *iii)* a elaboração de um raciocínio com base nesses dados e *iv)* a assunção de decisões ou a opção por cursos de ação, de acordo com parâmetros predefinidos, sempre de forma a alcançar um determinado objetivo[2]. Em suma, são as capacidades de atuação e de pensamento de forma *racional*[3] que surgem a caracterizar a IA, enquanto realidade que não (mais) se circunscreve a um simples conjunto de receitas ou linhas de comandos dirigidos a resolver um problema específico[4].

Apesar de coincidirem no seu núcleo essencial, estas são, porém, apenas duas de entre uma miríade infindável de tentativas de definição formal da Inteligência Artificial[5]. Saliente-se que, volvidos cinco anos desde a referida definição pela Comissão, e pelo High-Level Expert Group, a incerteza permanece, agora adensada com a Proposta de Regulamento Inteligência Artificial, ainda que aí se alegue avançar uma "definição inequívoca e preparada para o futuro de «inteligência artificial»[6]. Em face da incerteza, concordamos com os que – atenta a extensão das definições já ensaiadas –, afirmam que o "*mundo não precisa de mais uma*"[7]. E não é, em todo o caso, esse o propósito do escrito.

[2] *Cf.* High-Level Expert Group on Artificial Intelligence, "*A definition of AI: Main capabilities and scientific disciplines*", Comissão Europeia, 2018, *in* https://ec.europa.eu/futurium/en/system/files/ged/ai_hleg_definition_of_ai_18_december_1.pdf (27.11.2022), p. 7.

[3] Ou as de agir e pensar como um humano, mimetizando as funções cognitivas do cérebro humano – *cf.* Charles Kerrigan (com contribuições de Suzanne Rab, Stephen Kenny QC, Charlotte Payne e Jason G. Allen), *Introduction to AI*, *in* Kerrigan, Charles (ed.), *Artificial Intelligence – Law and Regulation*, [S. l.], Elgar, 2022, eISBN: 9781800371729, pp. 30-36, §§3.002 e 3.007.

[4] A propósito dos algoritmos – *cf.* Gönenç Gürkaynak, "*Algorithms and Artificial Intelligence: An Optimist Approach to Efficiencies*", *in Competition Law & Policy Debate Journal*, Volume 5, Issue 3, 2019, pp. 29-34, *in* https://ssrn.com/abstract=3783353 (27.11.2022), p. 29.

[5] Sendo até usual distinguir entre aceções mais ou menos latas ou "fortes" – *cf.* Moritz Hennemann, *Artificial Intelligence and Competition Law*, *in* Wischmeyer, Thomas / Rademacher, Timo (eds.), *Regulating Artificial Intelligence*, [S. l.], Springer Cham, 2020, eBook ISBN: 978-3-030-32361-5, pp. 361-388, §4.

[6] Saliente-se que, volvidos cinco anos desde a referida definição pela Comissão, e pelo High-Level Expert Group, a incerteza permanece, agora adensada com a Proposta de Regulamento Inteligência Artificial, ainda que aí se alegue avançar uma "definição inequívoca e preparada para o futuro de «inteligência artificial»".

[7] *Cf.* Charles Kerrigan, *Introductory Essay*, *in* Kerrigan, Charles (ed.), *Artificial Intelligence – Law and Regulation*, [S. l.], Elgar, 2022, eISBN: 9781800371729, pp. 2-23, §§1.032-1.045 (em particular, §1.041).

Pelo contrário, crê-se porventura mais pertinente – até por força do objeto do presente texto – apelar e sinalizar aquilo que se considera converter a IA num *specialis*, e que, em nosso ver, assenta ou resulta da conjugação de características como a *i)* complexidade, *ii)* a autonomia, *iii)* a imprevisibilidade, *iv)* a opacidade, *v)* a vulnerabilidade[8] e a *vi)* interdisciplinaridade[9], num conjunto idiossincrático, que precisamente converte a Inteligência Artificial em algo – senão único –, certamente diferente. Assim, e sem ignorar a questão litigiosa de saber se uma definição prévia da IA verdadeiramente se impõe, para efeitos da respetiva regulação (adequada)[10], optar-se-á, nesta sede, por a perspetivar – independentemente da forma como definida ou encarada[11] –, simultaneamente como realidade e desafio. Com efeito, em ambas as dimensões, a IA surge-nos como *algo* carecido – se de nada mais – certamente da atenção e da reflexão típicas das diferentes ciências e domínios do saber, entre os quais o Direito da Concorrência. O ponto de partida acabado de referir – a IA como realidade e desafio – merece, porém, uma breve densificação.

Vejamos.

A referência à IA como uma realidade permite fazer jus à circunstância de ser ela, hoje, instrumento inegável da satisfação de necessidades e interesses vários. Com efeito, a Inteligência Artificial surge no nosso quotidiano (não raras vezes de forma invisível, encapotada ou pressuposta), mobilizada na execução de tarefas[12] de tal forma distintas entre si, que o único elemento identitário ou agregador é mesmo a ligação umbilical que todas estabelecem

[8] *Cf.* Charles Kerrigan, *Introductory Essay...*, *cit.*, §1.050.

[9] *Cf.* Wolfgang Hoffmann-Riem, *Artificial Intelligence as a Challenge for Law and Regulation*, *in* Wischmeyer Thomas / Rademacher, Timo (eds.), *Regulating Artificial Intelligence*, [S. l.], Springer Cham, 2020, eBook ISBN: 978-3-030-32361-5, pp. 1-29, §1.

[10] *Cf.* Charles Kerrigan, *Introductory Essay...*, *cit.*, §1.077.

[11] Mas aderindo aos que alertam para a necessidade de evitar a respetiva humanização com recurso a analogias exageradas – *cf.* Charles Kerrigan (com contribuições de Suzanne Rab, Stephen Kenny QC, Charlotte Payne e Jason G. Allen), *Introduction to AI...*, *cit.*, §3.022. Considera-se que não se está perante lugares paralelos, nem seria este domínio em que a analogia pudesse vingar. Pelo contrário, se a Inteligência Artificial como fenómeno representa a execução de tarefas em termos que se aproximam das capacidades humanas e porventura as ultrapassam – *cf.* Wolfgang Hoffmann-Riem, *Artificial Intelligence as a Challenge...*, *cit.*, §1 – considera-se que a tentativa de paralelo com a pessoa humana, não só redundará em respostas incompatíveis com princípios basilares do Ordenamento Jurídico, ferindo a *pessoa*, como retirará ou silenciará as particularidades da Inteligência Artificial, refreando as suas possibilidades. Em suma, havendo vários tipos de inteligência, importa manter separadas as águas do (in)orgânico.

[12] Como a identificação de padrões, a avaliação e classificação de imagens, a tradução de texto, entre outras mais avançadas – *cf.* Wolfgang Hoffmann-Riem, *Artificial Intelligence as a Challenge...*, *cit.*, §3.

com processos e/ou técnicas de IA[13]. Note-se que se trata de atividades que – se é certo se consideram básicas ou 'assumidas' por e para quem nasceu, viveu e morrerá apoiado pelas ferramentas de IA –, se mantêm objeto de alguma admiração e de verdadeiro fascínio por todos os que se viram, durante muito tempo, sujeitos a encarar realidades como a tradução de um texto, a introdução de legendas num vídeo, ou, ainda, o bloqueio de correio eletrónico não solicitado[14] como ações dispendiosas (de tempo e de recursos), desde logo porquanto não mecanizadas nem automatizadas.

Nos mais variados domínios, a IA apresenta-se, hoje, portanto, como uma realidade intersticial da sociedade.

No entanto, é também e sobretudo enquanto realidade, que a Inteligência Artificial nos surge, em contrapartida (ou por inerência), como um desafio premente e atual. Já aqui se aludiu ao próprio debate em torno da sua definição. Em todo o caso, enquanto realidade heterogénea, com aplicações várias e efeitos de sinal antagónico a tentativa da sua *captura* por uma fórmula geral e abstrata vem originando a criação de fações cujos extremos são ocupados *i)* ora por aqueles que, perante a inviabilidade de uma definição adequada e *suficiente*, refutam a bondade de uma qualquer noção formal[15], *ii)* ora pelos que, em ordem à garantia de algum nível de segurança jurídica (ainda que mínimo) e à necessidade de definição do "regulado", sustentam a urgência de uma definição ou (pelo menos) de uma delimitação terminológica, para efeitos da compreensão e demarcação do âmbito de aplicação dos quadros normativos que se pretendam a ela aplicar.

A dificuldade terminológica é, porém, apenas uma das questões suscitadas pela IA, reveladoras da adequação do qualificativo que lhe é aqui votado como desafio. Desde logo, a IA surge a desafiar os quadros tradicionais e a *forma mentis* de grande parte dos domínios socioeconómicos. Neles se assiste, *pari passu*, à formulação de *novas* questões (pelo menos *prima facie*) e ao surgimento de interrogações e de debates apaixonados sobre a *suficiência* dos instrumentos, quadros e regimes vigentes para dar resposta àqueles que – ainda que não necessariamente novos – são certamente efeitos *sob uma nova roupagem*,

[13] Incluindo *machine learning*, *machine reasoning* e robótica – *cf.* High-Level Expert Group on Artificial Intelligence, "*A definition of AI: Main capabilities and scientific disciplines*", Comissão Europeia, 2018, *in* https://ec.europa.eu/futurium/en/system/files/ged/ai_hleg_definition_of_ai_18_december_1.pdf (27.11.2022), p. 7.

[14] *Cf.* Comunicação da Comissão ao Parlamento Europeu, ao Conselho Europeu, ao Conselho, ao Comité Económico e Social Europeu e ao Comité Das Regiões, *Inteligência artificial para a Europa*, COM/2018/237 final, Bruxelas, (25.4.2018, p. 1).

[15] Que sempre se afiguraria, mais geradora de problemas, dúvidas e lacunas (de exceção e de regulamentação) do que promotora de vantagens.

geradora de dúvidas sobre o *fit-for-purpose* das disciplinas aplicáveis. Em geral, pode afirmar-se que grandes "problemas" como *i)* a garantia da equidade e da justiça do processo e dos resultados, numa realidade marcada pela rigidez; *ii)* a necessidade (e viabilidade) do controlo democrático das tecnologias; *iii)* a inviabilidade de *outsourcing* da moralidade[16], a que acrescem *iv)* as insuficiências e os riscos alinhados segundo grandes vetores, como a explicabilidade, a robustez e a segurança, a linhagem ou o respetivo enviesamento[17], dão hoje palco a um conjunto de dúvidas sobre o se e o como encarar e regular a Inteligência Artificial.

Ainda relacionada com esta sua natureza bifronte de realidade-desafio está a circunstância de a IA se assumir, também, como um fenómeno ambivalente quanto aos seus efeitos ou impacto. Por outras palavras, ao contrário do que muitas vezes se vem avançando a seu propósito – e que facilmente se explica pelo enfoque tendencial do cérebro humano e das políticas normativas no negativo das realidades, isto é, no que se procura criticar ou combater, sendo menor ou nenhuma a preocupação para com o que nelas se possa enaltecer ou promover – a IA não se resume (ou não se pode ver resumida) a uma *black box* impenetrável, fonte das maiores reservas, perigos e/ou efeitos nefastos. Pelo contrário, está-se perante uma realidade que – ainda que de forma silente ou invisível – se encontra associada a um manancial de possibilidades e de eficiências que não podem ser esquecidas, ignoradas e, muito menos, subalternizadas face aos seus riscos. A título meramente exemplificativo considerem-se eficiências resultantes da IA que vão desde a possibilidade de adoção de processos otimizados com a redução de custos e a diminuição dos preços, à possibilidade de oferta de bens e/ou à prestação de serviços personalizados, customizados e inovadores[18].

Estas e outras *externalidades positivas* serão objeto de referência breve, a propósito do tratamento do *lado bom* da Inteligência Artificial. A sua referência preliminar serve apenas para sublinhar que o que se acaba de salientar sobre a natureza ambivalente ou polimórfica da IA não é facto desconhecido do leitor.

[16] *Cf.* Charles Kerrigan, *Introductory Essay..., cit.*, §§1.054, 1.067, 1.070.

[17] *Cf.* Tirath Virdee, *Understanding AI*, *in* Kerrigan, Charles (ed.), *Artificial Intelligence – Law and Regulation*, [S. l.], Elgar, 2022, eISBN: 9781800371729, pp. 37-55, §§4.038, 4.044-4.048.

[18] Sobre a questão, salientando as várias eficiências associadas aos algoritmos, no plano jusconcorrencial – *cf.* Gönenç Gürkaynak, "*Algorithms and Artificial Intelligence...*", *cit.* No mesmo sentido, sobre a IA em geral – *cf.* Suzanne Rab, *Competition Law*, *in* Kerrigan, Charles (ed.), *Artificial Intelligence – Law and Regulation*, [S. l.], Elgar, 2022, eISBN: 9781800371729, pp. 197-212, §§11.019 e 11.020.

Se é certo que – quando introduzida nos modelos de negócio e nos processos produtivos, ou quando simplesmente aplicada no quotidiano -, subjaz à IA um certo *dark side*, associado a perigos, incertezas e impactos vários, nos âmbitos social, económico e jurídico[19], ninguém se atreverá a negar a sua utilidade e contributo para a prossecução de interesses e a satisfação de necessidades várias, em termos e/ou condições também eles redesenhados pelas potencialidades do digital ou da automação.

Isto mesmo explica também – e desde logo pela missão que lhe cabe –, nenhum domínio (praticamente) existir, imune à problemática em torno da IA. Pelo contrário, em praticamente todos eles, a ambivalência da Inteligência Artificial como fenómeno é causa de controvérsias sobre o melhor caminho, abordagem, estratégia ou enquadramento, à luz de um binómio de todos conhecido: adaptar ou legislar *ex novo*? E que não deixa de se relacionar, ainda, com uma outra dualidade: *hard law* ou flexibilidade? Afinal, está-se perante uma lacuna regulatória justificada pela dinâmica do respetivo desenvolvimento e, bem assim, pela interdisciplinaridade do fenómeno[20].

A ambivalência e o caráter multifacetado da IA conduzem a que uma qualquer *estratégia* que sobre ela se pretenda adotar não possa deixar de assentar num ponto de partida imparcial o suficiente a reconhecer: *i)* que, num mesmo domínio e, bem assim, num mesmo ramo do Direito, a Inteligência Artificial é fonte (potencialmente) promotora de efeitos positivos e nefastos, e bem assim *ii)* que o seu *cross-cutting effect* poderá bem suscitar eventuais rotas de colisão, não já entre efeitos de sinal contrário num mesmo domínio, mas também entre os efeitos positivos produzidos num domínio, umbilicalmente associados a externalidades negativas num outro. Por outras palavras, a ambivalência da IA exige – para que logre ser bem compreendida – diálogo e transversalidade.

[19] Exemplo ilustrativo da ambivalência que se acaba de descrever poderá encontrar-se no domínio laboral e das condições laborais: se é verdade que a IA permite a substituição do humano na execução de tarefas substitutivas e repetitivas, encontrando-se, de igual modo, associada à criação de novos postos de trabalho, sobretudo no domínio da Engenharia e da Informática – *cf.* Gönenç Gürkaynak, "*Algorithms and Artificial Intelligence...*", *cit.*, pp. 32 e 33, e Suzanne Rab, *Competition Law...*, *cit.*, §11.020 – a substituição do trabalho humano, a extinção de postos ditos "tradicionais" e a necessidade de requalificação (em idade nem sempre compatível com a formação em área diferente) são problemas sociais que não podem deixar de ser considerados – *cf.* Comissão Europeia (Crémer, Jacques / de Montjoye, Yves-Alexandre / Schweitzer, Heike), "*Competition policy for the digital era – Final Report*", 2019, ISBN 978-92-76-01946-6, *in* https://ec.europa.eu/competition/publications/reports/kd0419345enn.pdf (27.11.2022), p. 12.

[20] *Cf.* Moritz Hennemann, *Artificial Intelligence...*, *cit.*, §7.

E uma tal constatação – por demais óbvia que possa surgir –, obsta precisamente a uma qualquer decisão, política ou estratégia que prescinda de uma análise holística das esferas sociais e – em particular no que ao Direito respeita –, do sistema jurídico como um todo. A circunstância de se estar perante um fenómeno acima da *summa divisio*, que, porquanto toca diferentes interesses públicos e privados, exige que a análise se não circunscreva aos (mercados dos) produtos e serviços produzidos e/ou prestados com recursos a tecnologias de IA. Pelo contrário, uma análise atinente ao impacto da IA, e, portanto, um diagnóstico adequado do problema, haverá de considerar ou abranger todos os domínios, áreas e interstícios da sociedade, sob pena de enfoque contrário arriscar, não só ignorar as repercussões e as relações de complementaridade que caracterizam a IA como realidade atual, como, e bem assim, menosprezar o seu potencial para se estender a todos os domínios da vida, no futuro[21].

Em face de tudo o exposto, e portanto, à questão de saber como enquadrar e prevenir os riscos e os perigos associados à IA, sem obstar ou refrear as suas potencialidades e eficiências, não poderá deixar de se responder com a exigência de um ponto de partida sólido e imparcial o suficiente a reconhecer a sua multidimensionalidade ou pluralidade de efeitos. Porque se há domínio em que a harmonização de políticas não poderá ocupar a retaguarda é este[22].

1.1. Delimitação

Como resulta da exposição introdutória avançada, o presente texto tem por objetivo chamar a atenção para uma realidade cuja complexidade exige um enquadramento ou uma abordagem *macro* o suficiente para assegurar a coerência das políticas e a efetividade das respostas.

Sem prejuízo da manutenção da bondade desse ponto de partida, crê-se, não só necessário (atento o objeto do texto), como relevante, concretizar o escopo do escrito, num movimento particular-geral.

E recordando o que se referiu sobre o desencontro dos efeitos da IA ou o caráter multifacetado do seu impacto *i)* num mesmo domínio e, bem assim, *ii)* entre domínios, crê-se que a(s) abordagem(ns) holístico-transversal(is)

[21] *Cf.* Wolfgang Hoffmann-Riem, *Artificial Intelligence as a Challenge..., cit.*,§7.

[22] Há, inclusive, quem considere que, de entre os desafios da IA, aquele que se afigura mais importante é precisamente o que respeita à circunstância de aquilo que se afigura positivo em determinado domínio (jurídico) poder chocar com as normas e princípios de um outro. Referindo-se à relação entre as regras da concorrência e as normas em matéria de defesa do consumidor, *cf.* Václav Šmejkal, "*Three Challenges of Artificial Intelligence for Antitrust Policy and Law*", *in Charles University in Prague Faculty of Law Research Paper*, No. 2021/III/3, 2021, pp. 1-17, *in* https://ssrn.com/abstract=3984354 ou http://dx.doi.org/10.2139/ssrn.3984354 (27.11.2022), em particular, pp. 100-104.

encetadas não deverão deixar de ser conduzida(s), *primo*, em cada um dos domínios jurídicos nos quais a IA surge como fonte potenciadora de efeitos divergentes. Trata-se de um ponto de partida, apenas. Não substitui a complementaridade, antes evita a confusão. O enfoque aqui assumido é estritamente (ou em primeira linha) **jurídico**, motivo por que o enquadramento da Inteligência Artificial que aqui se procurará encetar se circunscreve às questões que a mesma coloca ao Direito, naturalmente sem desconhecer que o seu impacto se produz, também, noutras esferas estritamente socioeconómicas, não juridificadas e ajurídicas[23].

Em segundo lugar, e agora já por referência ao domínio jurídico, o texto ocupar-se-á (*apenas*) do tratamento da ambivalência da IA no Direito da Concorrência, domínio onde a sua faceta híbrida ou os respetivos efeitos heterogéneos vêm já sendo objeto de sinalização e de debate (ainda que com cambiantes ou variações entre as diferentes políticas e comportamentos por ele abrangidos). Apesar dessa delimitação prévia, opta-se por não segmentar ou circunscrever a análise, por referência a qualquer um dos blocos de preocupação jusconcorrencial, a saber: *i)* os acordos ou práticas concertadas restritivos da concorrência; *ii)* o abuso de posição dominante; *iii)* o controlo de operações de concentração e *iv)* o regime dos auxílios de Estado. Sobre a não circunscrição do escrito a um destes âmbitos, importa proceder a um esclarecimento apenas.

Com efeito, não obstante possam ser diferentes as questões, a intensidade da dúvida ou o grau de urgência de uma resposta, em cada um dos domínios descritos, não só *i)* o impacto da IA é transversal a todos eles – abrangendo, inclusive, o próprio *enforcement* das normas –, como *ii)* se está perante uma realidade que vai a ponto de fazer questionar o próprio instrumentário dogmático-conceptual desta *divisio* do Direito e, bem assim, alguns dos seus conceitos operatórios, entre os quais os de "mercado relevante", "poder de mercado" e "colusão"[24].

Em resultado, e recordando que *i)* o fito do escrito é, precisamente o de identificar a ambivalência caracterizante da IA, mais atestando como *ii)* enquanto elemento transversal, a IA se logra assumir, em todos aqueles âmbitos, como *algo* potencialmente tão positivo como negativo, ou – recorrendo ao

[23] E, bem assim, sem ignorar que o Direito não poderá ser compartimentado das demais esferas sociais com as quais se relaciona e que poderá, inclusive, ser chamado a regular.

[24] *Cf.* Moritz Hennemann, *Artificial Intelligence..., cit.*, §8. *Vd.*, ainda, Comissão Europeia (Crémer, Jacques / de Montjoye, Yves-Alexandre / Schweitzer, Heike), "*Competition policy for the digital era...*", *cit.*, p. 42 e ss.

jargão da concorrência –, tão pró-concorrencial ou promotor de eficiências, como restritivo da concorrência e diminuidor do bem-estar, a circunscrição do texto a um dos domínios de atuação do Direito da Concorrência deixá-lo-ia necessariamente *manco*. É precisamente a natureza ambivalente, multifacetada ou porventura camaleónica da IA – também no que à concorrência se refere –, que justifica o presente texto. Considera-se, pois, útil e justificado abranger a miríade de comportamentos capturados pelas regras da concorrência, como forma de também assim apelar e comprovar a necessidade de uma análise holística o suficiente a evitar estratégias incoerentes, inclusive numa mesma área do Direito.

Aqui chegados, poderá aceitar-se ou antecipar-se a crítica de que *i)* a sistematização do papel da Inteligência Artificial e do seu impacto nos quadros jurídico-dogmáticos do Direito da Concorrência, e *ii)* o *objetivo* que lhe está associado de confrontar o leitor com aquele que é, atualmente, um *panorama* ou *cenário de dúvida* é missão de *utilidade* ou valor despiciendo.

Importa, por isso, que nos expliquemos.

A assunção expressa de um ponto de partida marcado pela *dúvida* é verdadeiro fundamento de legitimidade das políticas e estratégias que se pretendam adotar sobre a IA nesta área. Isto porque, se é certo que a incerteza não legitima, por si só, a resignação a uma abordagem do tipo "*wait and see*", importa também evitar uma resposta precipitada ou a tentativa de dilatação dos instrumentos jusconcorrenciais, forçando-os a dar resposta a problemas para cuja solução se não veem talhados[25]. Se de nada mais servir, este escrito – ao alertar e fundamentar a necessidade de alguma parcimónia nesta sede –, serve para *desarmar* (porquanto ilegítimas) abordagens inimigas da inovação, da eficiência e, em última linha, da própria concorrência e do bem-estar económico.

Em suma, o que aqui se exporá – sob um prisma *imparcial* – é nada mais do que o impacto da IA no Direito da Concorrência, procurando apelar à sua heterogeneidade ou ambivalência, enquanto verdadeiro pressuposto de legitimidade do caminho (de qualquer caminho) que se opte por seguir. Uma última delimitação negativa importa aqui fazer.

A relação da IA com o Direito da Concorrência tende a ser associada a um outro problema igualmente relevante, e que vem, aliás, sendo objeto de referência mais aturada pela Doutrina. Trata-se da questão de saber se os quadros, instrumentos e normas do Direito da Concorrência se revelam adequados a dar resposta aos desafios suscitados pela IA. Em suma, está em causa do

[25] *Cf.* Suzanne Rab, *Competition Law..., cit.*, §§11.059 e 11.061.

problema da adequação e da suficiência das normas do Direito da Concorrência quando referidas à IA[26]. Esta é, porém, questão autónoma e *distinta* – inclusive do prisma cronológico – daquela que aqui se visa, e que é prévia. Assim, e uma vez mais, importa esclarecer que o objetivo aqui assumido não é o de avançar respostas ou sugestões quanto à política concreta a seguir (nem o de assumir sequer um qualquer diagnóstico relativo a um potencial "*enforcement gap*")[27]. Pelo contrário, o texto cinge-se à tarefa de escalpelizar um (de entre vários) dos pressupostos de legitimidade que importa impor a qualquer política ou estratégia. Trata-se, pois, de um ponto de partida prévio à identificação de propostas de solução.

Naturalmente – admite-se -, a sinalização do caráter multiforme da IA e o reforço da necessidade de um entendimento imparcial a seu respeito certamente confortarão a opção ou a preferência por abordagens mais parcimoniosas, menos extremistas, e alinhadas, porventura, com a lógica do "diálogo regulatório". No entanto, se assim é, estar-se-á sempre, e em qualquer caso, perante uma externalidade ou elemento *acidental* do texto, e que não prejudica o seu foco primeiro, que é outro. Se se quiser, não nos deteremos, nesta sede, sobre a receita – para a cura ou, pelo menos, para o alívio dos sintomas – do *remédio* **x**, **y** ou **z**. Pelo contrário, do que se cuidará, será, antes, da adequada *determinação* do diagnóstico, através da identificação do contexto *sui generis* que circunda a IA como realidade e como desafio, conscientes de que – na ausência de constatação e tratamento expressos desse pano de fundo –, o diagnóstico (qualquer que ele seja) se verá necessariamente inquinado por uma visão apriorística e absolutizante da realidade, e, bem assim, capturado pelos vícios de uma apreensão monolítica da IA: ora exclusivamente positiva ou favorável, ora estritamente crítica, cética, ou completamente agnóstica.

1.2. Sistematização

Em face do objetivo identificado *supra*, e que, como o próprio título do texto indica, implica desnudar a Inteligência Artificial, não apenas como uma opor-

[26] Distinguindo as questões do *i)* impacto (nefasto) da IA no Direito da Concorrência e da *ii)* adequação das suas normas face ao fenómeno, *cf.* Suzanne Rab, *Competition Law...*, *cit.*, §11.004.

[27] E, portanto, não se cuidará aqui da questão da (in)suficiência do Direito da Concorrência ou de saber se será preferível a adoção de uma abordagem mais reformista ou mais conservadora, mais intervencionista ou mais liberal. De igual modo, não se irá a ponto de discutir a questão antecedente relativa *inclusive* ao papel, intervenção ou missão particular – se é que algum(a) – que ao Direito da Concorrência deverá caber, a este nível (e atenta a limitação do seu âmbito). Sobre a questão, *vd.*, *inter alia*, Suzanne Rab, *Competition Law...*, *cit.*, §11.021 e ss.

tunidade e fonte geradora de eficiências (*the good*), mas também, e bem assim, como um perigo e como um desafio (*the bad and the ugly*) para o Direito da Concorrência, o texto encontra-se dividido em quatro grandes capítulos, dirigidos ao tratamento dessa tríplice dimensão.

Num primeiro capítulo, tratar-se-á de um conjunto de generalidades preliminares sobre o Direito da Concorrência como disciplina jurídica e, bem assim, sobre a Inteligência Artificial como fenómeno necessariamente associado à digitalização e aos mercados digitais. Ainda que breve e não inovador, o objetivo é o de firmar um ponto de ordem sobre o âmbito do Direito da Concorrência e a disrupção nele provocada pela IA, para efeitos de legitimar, depois, a análise das respetivas repercussões (positivas e negativas) nesta área do Direito.

Uma vez encerrada a divisão introdutória, abrir-se-á o capítulo dedicado ao elenco das externalidades positivas decorrentes da utilização da Inteligência Artificial. Sobre esta opção sistemática – e ainda que se possa estranhar que um texto dedicado a ilidir a presunção de negatividade associada à IA principie precisamente 'pelo argumento mais forte' (que, segundo aconselha o costume do foro, deveria, pelo contrário, encerrar o discurso). A sistemática compreende-se à luz do propósito de ilidir aquela que é uma verdadeira presunção de efeitos negativos, resultante de uma visão cética ou exclusivamente negativa da IA como fenómeno. O objetivo deste segundo capítulo é, pois, o de apresentar a IA, à luz de uma abordagem otimista, dirigida ao tratamento do respetivo *good* da IA, e que inclui um subcapítulo dedicado às eficiências da sua aplicação no *enforcement* das regras da concorrência.

O terceiro capítulo é, em contraste, voltado aos riscos e aos efeitos negativos associados à IA, ou seja, ao respetivo *bad* e *ugly side(s)*, agora num cenário que, segundo se espera, se apresentará já como *neutro* ou (pelo menos, mais) *equilibrado* para o leitor. À semelhança do capítulo antecedente, também aqui se procederá à autonomização dos riscos associados à introdução e ao recurso a ferramentas de Inteligência Artificial na aplicação e execução das normas da concorrência.

O capítulo final apresentará algumas conclusões para a fundamentação e densificação do pressuposto, do ponto de partida e do critério de legitimidade que aqui se procura(m) avançar e que se crê dever(em) nortear e perpassar quaisquer estratégias ou políticas dirigidas a dar resposta às (novas) questões suscitadas pela IA.

2. Direito da Concorrência e Inteligência Artificial: *friends or foes*?

Como é facto que se considera não desconhecido, o Direito da Concorrência tem um âmbito de aplicação e de intervenção relativamente estreito(s), em torno de três ou quatro blocos normativos, a saber: *i)* os acordos ou práticas concertadas restritivos da concorrência; *ii)* o abuso de posição dominante; *iii)* o controlo de operações de concentração e, ainda, *iv)* o regime dos auxílios de Estado[28]. Isto posto, e sem negar que ainda hoje a busca pela verdadeira *alma* do Direito da Concorrência continua alimentando discussões apaixonadas entre os que o associam a objetivos políticos mais latos e os que, pelo contrário, o acantonam ou agrilhoam à promoção de eficiência(s) (geradora(s), em termos mediatos, de outras externalidades, garantida(s) em termos exclusivos ou mais perfeitos pela *concorrência efetiva*), não se afigura errado afirmar que se está perante um tronco do Ordenamento de balizas relativamente circunscritas.

Além desta sua circunscrição, considera-se também idiossincrasia sua o facto de as respetivas normas (muitas vezes qualificadas como normas *em branco*) se caracterizarem por um preenchimento aberto ao diálogo constante com a jurisprudência e a prática decisória, entre nós, das Autoridades Nacionais da Concorrência ('ANC's), da Comissão Europeia e, em última linha, dos tribunais nacionais e do Tribunal de Justiça da União Europeia. Ou seja, pode sem pejos afirmar-se que a malha regulatória do Direito da Concorrência é composta por normas abertas à densificação e ao preenchimento evolutivos dos seus conceitos e conteúdo, em termos que permitem assegurar a sua adequação às circunstâncias cambiantes da sociedade e dos mercados.

Ora, da conjugação de *i)* um âmbito de aplicação relativamente circunscrito, por um lado, com *ii)* o potencial de preenchimento e atualização das suas normas, através de uma jurisprudência e prática decisórias 'evolutivas', por outro, resulta patente, segundo se crê, a relativa capacidade de conservação dos regimes nacionais e europeu da concorrência, e, bem assim, a circunstância de, ao longo dos anos, as suas normas terem logrado dar resposta ou devido enquadramento a novas questões, práticas, dinâmicas, modelos de negócio e comportamentos.

Acontece que, se até aqui, a abertura e a flexibilidade referidas asseguraram, com relativo brilhantismo, o devido enquadramento a questões e teorias de dano *novas* (ainda que somente *prima facie*), hoje, porém, seja os mercados

[28] Ainda que, quanto a este, a sua recondução ao âmbito do Direito da Concorrência possa ser controvertida, havendo quem o perspetive, antes, como política ancilar da construção e integração do mercado interno, e, portanto, como uma especialidade do Direito da União Europeia, que não caracterizaria o Direito da Concorrência enquanto ramo do Direito.

digitais em geral, seja a Inteligência Artificial em particular, vêm desafiando a *suficiência* dos instrumentos jusconcorrenciais, pelo menos *as they are*. Um tal desafio e os debates que suscita sobre qual a melhor abordagem a adotar perante uma realidade mutável a uma velocidade vertiginosa justificam que nos questionemos: o que terá, afinal, a IA de tão especial que abale a capacidade de adaptação e de maleabilidade das normas e do regime da concorrência, a ponto de estes não mais lograrem *acomodar* essa *diferença*? Serão as potencialidades que confere às empresas (direta ou indiretamente, através de novas *condições* de mercado)? Será o seu caráter opaco, imperscrutável? Será a circunstância de ameaçar algumas das assunções e quadros dogmáticos do Direito da Concorrência?

Algumas destas interrogações lograrão densificação, nos pontos que se seguem.

Antes, porém, de os abordarmos, importa alertar para o seguinte: uma (qualquer) discussão sobre o impacto da Inteligência Artificial no Direito da Concorrência não poderá ser desligada de uma questão mais ampla, atinente aos desafios dos mercados digitais[29], em relação à qual se não crê que a primeira adquira absoluta autonomia dogmática. Pelo contrário, parte dos problemas associados ao recurso e à utilização da Inteligência Artificial (seja pelas empresas, seja pelas autoridades competentes) cruzam-se com os desafios mais latos dos mercados digitais e respetivas idiossincrasias, e que também vêm pondo em xeque a operatividade das regras da concorrência.

Não se poderia encerrar o ponto, sem antes – e retomando a designação dada ao capítulo –, esclarecer que a circunstância de se prefigurar a IA como um desafio (inclusive) para o Direito da Concorrência (cujas normas se autonomizam pela referida capacidade de adaptação e abertura), não é constato necessariamente negativo ou que consinta qualificá-la como *inimiga* daquele. Não só pelas eficiências que se elencarão sumariamente de seguida, mas também, e sobretudo, pela circunstância de se estar, afinal, perante um ramo do Direito onde, segundo se crê, se encontram reunidas as condições para votar e oferecer à IA o enquadramento devido, se há ramo do Direito que se encontra munido dos instrumentos idóneos a, sem rotura, assegurar quadros de *governance* adequados à IA, é o Direito da Concorrência.

Expliquemo-nos.

[29] E a propósito do qual vimos ouvindo falar de *Big Tech*, *gatekeepers* e outros termos afins. Neste capítulo, seguir-se-á de perto a OCDE – Organização para a Cooperação e Desenvolvimento Económico, "*OECD Business and Finance Outlook 2021: AI in Business and Finance*", Paris, OECD Publishing, 2019, *in* https://doi.org/10.1787/ba682899-en (27.11.2022), §4.1.

Ao contrário de outros domínios, em que a análise concreta ou a margem decisória surgem como *corpos estranhos*, no Direito da Concorrência, o casuísmo e o relevo do caso concreto nunca deixaram de estar presentes, considerando-se aqui asseguradas as exigências da segurança jurídica à luz de um entendimento não estritamente formalista (contanto que respeitados, naturalmente, os limites do Estado de direito). Ora, se assim é, é este um domínio em que se afigura possível *testar* e lograr uma resposta adequada à IA, sem que isso implique uma revolução dos *standards* de análise ou dos quadros normativos aplicáveis. Por outras palavras, e sem prejuízo das dúvidas referidas *supra* quanto a uma putativa necessidade de adequação ou de reforma, o Direito da Concorrência reúne características idóneas a garantir o tratamento da Inteligência Artificial enquanto realidade e fenómeno, em termos respeitadores da sua diversidade, e garantes da proporcionalidade do enquadramento.

Assim, e portanto, à questão de saber "Direito da Concorrência e Inteligência Artificial: *friends or foes*?" considera-se adequado responder com esta *vantagem competitiva* do Direito da Concorrência em mente.

3. A Inteligência Artificial como realidade-fenómeno promotor da concorrência: *the good*

Conforme referido aquando da explanação da estrutura do escrito, o presente capítulo procurará alertar para as vantagens (*v.g.* eficiências) da Inteligência Artificial e para o modo como a(s) mesma(s) beneficia(m) a concorrência, não apenas na sua dimensão estática, mas – e quiçá sobretudo –, na sua costela dinâmica, dirigida à promoção da inovação[30]. Porquanto se trata, sobretudo, de eficiências transversais à análise do Direito da Concorrência e ao entendimento e aplicação dos seus princípios e regras, opta-se por proceder, neste capítulo, a uma análise não segmentada por domínios (exceto quando isso resulta expressamente em texto).

Vejamos.

As eficiências resultantes da Inteligência Artificial e, portanto, o seu *lado bom* no domínio jusconcorrencial são, como se referiu já, algo transversal aos diferentes campos de aplicação das regras da concorrência, repercutindo-se

[30] E portanto, determinados comportamentos aparentemente não promotores da concorrência no momento em que analisados, poderão afigurar-se necessários para o desenvolvimento de um produto de melhor qualidade ou para a aquisição de uma escala que garanta a sua oferta a um preço mais baixo, no futuro.

na concorrência *no mercado* e *pelo mercado*[31], e podendo ser perspetivadas, tanto do prisma da procura, como da oferta.

O constato aqui avançado resulta desde logo visível ou evidente se se pensar na possibilidade de identificação de novos parâmetros competitivos, isto é, de novas dimensões com base nas quais as empresas concorrem, desde logo por força da introdução da IA nas cadeias de valor e, portanto, da concorrência com Inteligência Artificial (e que dá lugar ao surgimento de novos ou melhores produtos, serviços inovadores e processos otimizados). A esta dimensão acresce uma outra: a concorrência pela Inteligência Artificial (termo que se emprega como descritivo da circunstância de as empresas concorrerem hoje, e também, pela sua afirmação e pela sua condição de líderes em mercados relacionados com a produção, o desenvolvimento e a comercialização de ferramentas e/ou processos de IA, em si mesma considerada)[32].

A propósito desta distinção: concorrência com e pela Inteligência Artificial importa, no que à segunda respeita, esclarecer que do que se trata é nada mais do que da concorrência que se estabelece entre as empresas, em ordem à obtenção de uma "vantagem competitiva no plano da IA", e que as força a concorrer, agora também por referência aos próprios sistemas e instrumentos de IA. Nela logra descobrir-se uma nova forma de pressão concorrencial exercida sobre as empresas, dirigindo-as ou verdadeiramente as forçando a orientar as suas operações para a digitalização e a introduzir ferramentas, mecanismos e tecnologias de IA nos seus processos[33], em termos reveladores da importância da costela dinâmica da concorrência e da sua ligação umbilical com a inovação.

Fechado o parêntese, e apesar de distintas ou autonomizáveis, cumpre assinalar que, em qualquer uma daquelas dimensões – concorrência com e pela Inteligência Artificial –, e independentemente da sua referência ou pertinência para o acesso ou à atuação no mercado, isto é, para as dimensões da concorrência pelo mercado e no mercado, a Inteligência Artificial surge a impactar de forma positiva nos mercados, enquanto fator desestabilizador.

Vejamos como.

Crê-se hoje relativamente *imediato* o contributo significativo que a IA[34] é capaz de aportar para a ameaça a operadores tradicionais ou incumbentes,

[31] Neste sentido, *cf.* Moritz Hennemann, *Artificial Intelligence...*, *cit.*, §1.

[32] Sobre a distinção, *cf. ibidem*, §8.

[33] Sobre a "vantagem competitiva algorítmica" – *cf.* Gönenç Gürkaynak, "*Algorithms and Artificial Intelligence...*", *cit.*, pp. 33 e 34.

[34] Apesar de os exemplos serem transversalmente aplicáveis à digitalização em geral.

até então protegidos ou relativamente imunes à pressão da concorrência (o que lhes permitia, aliás, lograr resistir à expectativa e aos interesses do consumidor do século XXI). Mercados como o financeiro ou o dos transportes são exemplos de domínios em que a digitalização e, em particular, o emprego de ferramentas de IA (pelo menos, em maior escala ou apelando a uma *aceção forte* de IA) vieram pôr em causa a posição de quase incontestabilidade dos operadores tradicionais, forçando-os a inovar, precisamente para acompanhar (e superar) a oferta de *newcomers*. Novos *players* que agora se apresentam como prestadores de serviços diferentes, inovadores e garantes de um conjunto de possibilidades, até então não abertas à procura (pelo menos, nos termos em que agora oferecidos). O ponto que se acaba de fazer, atinente ao incremento da concorrência no mercado, justifica – mais não seja como pano de fundo –, retomar a missão que a concorrência efetiva (como bem jurídico) desempenha, no plano da promoção (das possibilidades) de escolha do consumidor (em ambas as suas dimensões: qualitativa e quantitativa). E retomá-lo, precisamente para, em conformidade, concluir sobre o papel que a IA poderá desempenhar a este nível.

As suas vantagens no plano da procura não se quedam por aqui. Com efeito, também no que à garantia do caráter informado da escolha e ao problema da assimetria de informação respeita, poderá a IA assumir um importante e insubstituível papel, desde logo em prol do esclarecimento do adquirente, utilizador ou beneficiário último dos bens e dos serviços comercializados e prestados. E isto, enfim, através do aumento da transparência que proporciona e das possibilidades que assim abre à garantia de uma escolha informada e efetiva[35].

Idênticos "positivos" se encontram no lado da oferta. Efetivamente, também neste plano se não poderá ignorar que a introdução da IA nos sistemas produtivos permite às empresas melhorar e otimizar a sua atividade e *outputs*, numa ótica de maior celeridade, menor custo e maior proximidade ou ali-

[35] Como veremos, e este é exemplo evidente de como determinada vantagem poderá também ser perspetivada como problema de raiz, a Inteligência Artificial contribui para aumentar e intensificar o grau de transparência dos mercados, o que, no que ao consumidor respeita, sempre facilita ou potencia as condições para uma escolha mais informada. Segundo Gönenç Gürkaynak, "*Algorithms and Artificial Intelligence...*", *cit.*, p. 30, a atuação dos algoritmos permite reduzir os custos de transação e de pesquisa de informação, permitindo decisões de compra menos dispendiosas e mais céleres, em situações complexas, na medida em que garante aos consumidores a possibilidade de verificação e comparação simultânea de ofertas num só local ou plataforma. Acresce que a possibilidade de elaborar sugestões de compra, baseadas em aquisições prévias, auxilia, também, o consumidor, nesta sede.

nhamento com as preferências do cliente e parceiros comerciais. Tudo isto através da customização dos seus produtos e serviços, e, portanto, da garantia da conquista da preferência de um cliente de necessidades e exigências cambiantes, em ordem às quais, ou em linha com as quais importa dirigir os esforços de *marketing*[36]. Em suma, e em traços gerais, crê-se que, também no plano da oferta, a Inteligência Artificial permite às empresas fazer mais e melhor, com menos: menos custos, menos gastos, menos recursos, e menos tempo.

Este pano de fundo afigura-se revelador de evidentes eficiências, que poderão (e deverão) ser repercutidas *a final*, sob a forma de produtos (ou serviços) a preços mais baixos, de maior qualidade, em maior quantidade, mais inovadores, e disponibilizados de forma mais célere, próxima e alinhada com as expectativas de quem os procura. Crê-se, portanto, legitimada a assunção da IA como algo (também) bom para a concorrência.

Algo mais se pode, porém, avançar nessa sede.

Associada ao incremento da produtividade das empresas, indústrias e processos produtivos, importa também não menosprezar o impacto macro que a IA poderá produzir, no que às economias nacionais e europeia respeita[37]. Saliente-se, apesar de ser esta uma especificidade do regime jurídico europeu que, a 19 de outubro de 2022, a Comissão Europeia adotou uma Comunicação revista sobre as regras em matéria de auxílios estatais à investigação, ao desenvolvimento e à inovação ('Enquadramento IDI de 2022')[38], esclarecendo os princípios e as circunstâncias que terá em conta na apreciação dos auxílios estatais a empresas para atividades de Investigação e Desenvolvimento[39]. Con-

[36] Referindo-se ao impacto pró-concorrencial dos algoritmos e às eficiências associadas numa ótica de custos e qualidade dos produtos e serviços prestados, *cf.* Gönenç Gürkaynak, "*Algorithms and Artificial Intelligence...*", *cit.*, p. 30.

[37] Sobre a questão, *cf. ibidem*, pp. 32 e 33.

[38] *Cf.* Comissão Europeia, *Comunicação da Comissão: Enquadramento dos auxílios estatais à investigação, desenvolvimento e inovação*, C(2022) 7388 final, Bruxelas (19.10.2022).

[39] Recorda-se que o objetivo do controlo dos auxílios de Estado é o de evitar que as subvenções estatais falseiem ou ameacem falsear a concorrência no mercado interno e afetem as trocas comerciais entre Estados-Membros. Por esse motivo, e a essa luz, o artigo 107º, nº 1 do Tratado sobre o Funcionamento da União Europeia estabelece um princípio de proibição dos auxílios estatais, acompanhado da identificação de casos excecionais, em que os mesmos poderão ser considerados compatíveis com o mercado interno – *cf.* artigo 107º, n.os 2 e 3 do Tratado. Em particular, nos termos do artigo 107º, nº 3, alínea c) do Tratado, uma medida de auxílio poderá ser declarada compatível com o mercado interno desde que preenchidas duas condições: *i)* o auxílio facilita o desenvolvimento de uma atividade económica e *ii)* o auxílio não afeta negativamente as condições das trocas comerciais de maneira que contrarie o interesse comum.

sidera-se este um ato relevante, precisamente por sinalizar uma aposta clara da União Europeia na investigação e no desenvolvimento de processos incluindo e envolvendo Inteligência Artificial. E, portanto, a assunção expressa da sua importância e valor acrescentado para as economias nacionais e europeia.

O ponto é, a esta altura, suficientemente demonstrativo de como a concorrência com IA, e pela IA consente(m) e legitima(m) uma visão positiva sobre a Inteligência Artificial.

Assim, seja pelos desafios que ergue aos incumbentes, seja pelas características que lhe associamos atrás (e intimamente relacionadas com as idiossincrasias dos mercados digitais), a Inteligência Artificial cria as condições para mercados mais agressivos ou competitivos, acrescentando novas dimensões ou parâmetros competitivos, sem exigir (pelo menos num primeiro momento) a reunião de meios, em termos configuradores de uma verdadeira barreira à entrada. Pelo contrário, saliente-se que as possibilidades da IA estão, não raras vezes, na dependência de uma *ideia*, que – se é certo carecida, depois, de um investimento necessário à respetiva materialização – adquire um valor autónomo, enquanto *ideia*. Naturalmente, problemas há de disputabilidade que importa resolver. O Regulamento Mercados Digitais é, a esse nível, referência incontornável.[40]

Em face de tudo o exposto, e sem ignorar os riscos que mais à frente se listarão, considera-se seguro afirmar que a introdução e a internalização de ferramentas e mecanismos de IA nos processos produtivos e nos modelos de negócio das empresas levam, não só a que *i)* fatores de diferenciação tradicionais como o preço e/ou a qualidade resultem necessariamente reconfigurados em razão das possibilidades oferecidas pela IA, como, e bem assim, ao *ii)* surgimento ou autonomização de novas dimensões com base nas quais as empresas concorrem.

Assim, (também) em razão da IA, as empresas competem hoje entre si, não já – ou não já apenas -, pela oferta do 'preço mais baixo', mas também, e quiçá de forma preponderante, em parâmetros como *i)* o quão rápido respondem e se adaptam a mudanças no mercado, *ii)* a acuidade com que preveem, interpretam e tratam as informações que alimentam – como *inputs* – os respetivos sistemas e processos, como, e, bem assim, *iii)* a medida em que logram exponenciar a utilização e a aplicação da IA na melhoria da qualidade dos

[40] Regulamento (UE) 2022/1925 do Parlamento Europeu e do Conselho de 14 de setembro de 2022 relativo à disputabilidade e equidade dos mercados no setor digital e que altera as Diretivas (UE) 2019/1937 e (UE) 2020/1828 (Regulamento dos Mercados Digitais) PE/17/2022/REV/1, OJ L 265, 12.10.2022, p. 1-66.

produtos e serviços prestados e na oferta de soluções que, além de (verdadeiramente) mais baratas, se afigurem satisfatórias das necessidades e expectativa dos clientes. Clientes cada vez mais informados e exigentes e, bem assim, cada vez mais dificilmente satisfeitos.

Este é um pano de fundo que urge não ignorar, no momento de sinalizar os vícios e os perigos da Inteligência Artificial. O ruído associado à reconfiguração e/ou multiplicação dos parâmetros competitivos é circunstância que, não só *i)* tem o potencial de dificultar putativas estratégias de coordenação (em razão do aumento dos fatores diferenciadores, nos quais as empresas assentam as suas decisões e baseiam a sua estratégia comercial), como *ii)* aumenta a incerteza ou o risco, exercendo uma maior pressão sobre as empresas, para que se diferenciem e inovem. E, portanto, uma associação estanque entre a IA e a facilitação de estratégias de colusão é, sob este pano de fundo, ligação, no mínimo, precipitada.

Em suma, demonstrado fica um cenário de eficiências que, não só milita em sentido contrário a uma visão cética ou pessimista da Inteligência Artificial, como certamente obsta à sua consideração-redução a fator potenciador de estratégias de colusão ou, em termos mais latos, à sua configuração como *inimigo* da concorrência.

3.1. A Inteligência Artificial como mecanismo indutor de eficiências no *enforcement* das regras da concorrência

Antes de encerrar o capítulo, crê-se relevante sinalizar as possibilidades e as vantagens que a IA também aporta para a aplicação e *enforcement* das regras da concorrência, enquanto missão e incumbência das Autoridades Nacionais da Concorrência e da Comissão Europeia (e, em última linha, dos tribunais nacionais e do Tribunal de Justiça da União Europeia).

Com efeito, o recurso à IA aporta eficiências não despiciendas, também no plano adjetivo ou processual (elemento necessário da garantia da efetividade das regras e normas substantivas ou materiais). As suas potencialidades são sobretudo relevantes, à luz de *i)* investigações e processos cada vez mais complexos[41], *ii)* com um volume crescente de documentação, de dados e de prova associados, cuja gestão *manual* ou não automatizada é, senão impossível, certamente difícil ou insuficiente, e, bem assim, em circunstâncias em que, *iii)* e apesar dessa sua crescente complexidade e *dimensão*, os processos

[41] Complexidade em parte devida, também, à digitalização e às particularidades dos mercados digitais, em ordem aos quais orientadas, agora, as prioridades em matéria de *enforcement*.

se encontram, ainda assim, carecidos de conclusão dentro de limites temporais estritos e inultrapassáveis, ditados pelo princípio da tutela jurisdicional efetiva e do processo equitativo.

Neste plano, portanto, as possibilidades abertas pela IA ao nível do *i)* processamento de dados, do *ii)* reconhecimento de padrões, da *iii)* formulação de decisões de priorização (e alocação de esforços e recursos), e da *iv)* identificação imediata de propostas de solução[42] não deverão ser menosprezadas. A utilização de tecnologias e ferramentas de IA poderá permitir às ANC's logrem, não só *i)* alimentar e engrandecer o acervo de informações e de dados de cuja recolha e tratamento dependem para efeitos da realização de estudos de mercado[43]; *ii)* utilizar essas mesmas informações e respetivo tratamento, para estabelecer domínios de atuação prioritária, fundamentando as decisões de priorização (cada vez mais necessárias) à luz de uma análise certamente mais analítica, completa e menos suscetível a erro(s), como e, bem assim, *iii)* encurtar a duração das investigações, em razão, seja da identificação mais célere e precisa dos elementos probatórios (ir)relevantes e (des)necessários para a prova de uma infração (e a determinação dos seus agentes), seja, ainda, da possibilidade de automatização do processo de identificação de informações confidenciais e sensíveis em documentos cujo acesso deva ser concedido a Co-Visadas e/ou a Terceiros, em termos que desonerem ou, pelo menos, auxiliem as partes nesse labor.

O aumento da efetividade e da celeridade das investigações aporta, por seu turno, vantagens para as próprias Visadas. E isto, não só no que à duração dos processos respeita (bem se sabendo como a respetiva morosidade influi ou faz perigar o seu direito à presunção de inocência, em razão do intervalo dilatado que poderá mediar entre o início e o encerramento do processo), mas, e sobretudo, no que aos pedidos de elementos e tratamento de confi-

[42] *Cf.* Andreas von Bonin / Sharon Malhi, "*The Use of Artificial Intelligence in the Future of Competition Law Enforcement*", *in Journal of European Competition Law & Practice*, Volume 11, Issue 8, 2020, pp. 468-471, *in* https://awards.concurrences.com/IMG/pdf/the_use_of_artificial_intelligence_in_the_future_of_competition_law_enforcement.pdf?67667/695bd2238811a7abb3f664646dfc0ed74d99b8a3975cf6c3331bc7e9efea478c (27.11.2022), p. 468. Os Autores consideram, em particular, a importância das ferramentas de IA na decisão de determinação de quais as empresas e comportamentos a considerar como prioritários, tendo em conta a relativa discricionariedade da Comissão (e das ANC's) na matéria e, bem assim, a circunstância de se estar perante entidades com recursos finitos.

[43] Em particular, a recolha e o tratamento de um largo volume de dados permite a identificação de tendências de mercado de forma mais célere e precisa, além de possibilitar o despiste precoce de indicadores de deficiências ou falhas de mercado, que devam ser objeto de atenção particular – *cf.* Andreas von Bonin / Sharon Malhi, "*The Use of Artificial Intelligence…*", *cit.*, p. 469.

dencialidades se refere, tarefa(s) certamente facilitada(s) pelo emprego de ferramentas de IA, seja pelas autoridades, seja pelas próprias empresas, em termos admitidos pelo contexto normativo aplicável[44].

Enfim, também, portanto, no plano do *enforcement*, o recurso e a utilização de ferramentas de IA se afiguram promissores e garantes da adoção de respostas mais céleres, eficazes, adequadas, e certamente *mais* informadas por parte das autoridades, colocando o processo ao serviço (dos direitos) das partes.

4. Em torno das dúvidas e perigos da Inteligência Artificial e respetivo impacto negativo na concorrência: *the bad and the ugly*

A circunstância de se afastar, no presente texto, uma visão estritamente cética ou negativa da IA não legitima que se caia no extremo oposto de uma sua consideração como realidade puritana ou isenta de crítica, risco ou suspeição. Pelo contrário, praticamente por referência a cada uma das vantagens listadas, é possível acorrentar, não um, mas vários problemas e desafios, os quais oscilando entre a esfera do simples *bad*, e do verdadeiro *ugly*.

Isto dito, e precisamente para efeitos de assegurar a neutralidade e a imparcialidade que se assumem como objetivo do texto, importa agora sistematizar muito brevemente o *dark side* que vem sendo associado à Inteligência Artificial no Direito da Concorrência. No que a ele se refere, importa salientar que – apesar do seu tratamento mais aturado no plano dos acordos e práticas restritivas[45], em particular por referência ao papel desempenhado pelos algoritmos[46] –, não é *suficiente* circunscrever a análise ao plano da colusão algorítmica. Pelo contrário, e como também já se antecipou, a Inteligência Artificial impõe-se como um desafio transversal aos vários domínios jusconcorrenciais, em termos que sempre obstaculizam a discriminação de todos os demais àquele.

[44] *Cf. ibidem*.

[45] Qualificando a "cartelização" como o desafio da IA (para o Direito da Concorrência) mais comentado, *cf.* Václav Šmejkal, "*Three Challenges of Artificial Intelligence...*", *cit.*, em particular, pp. 108-113.

[46] Sobretudo numa ótica de ausência de controlo humano, conjugada com o respetivo potencial para distorcer ou restringir a concorrência, num determinado mercado. Refira-se que a preocupação à qual vem sendo dada maior atenção respeita a cenários de colusão tácita – *cf.* Gönenç Gürkaynak, "*Algorithms and Artificial Intelligence...*", *cit.*, p. 32. Em particular, o foco vem incidindo sobre os algoritmos de preços, alargando-se à coordenação expressa (através de condições de mercado mais adequadas) e à colusão tácita (à qual, recorda-se, não subjaz um qualquer acordo prévio, que se limite a ser facilitado, implementado, concretizado ou monitorizado através dos referidos algoritmos) – *cf.* Suzanne Rab, *Competition Law..., cit.*, §§11.013-11.015.

Sem prejuízo, e precisamente por corresponder ao âmbito em que os riscos e desafios da IA (sobretudo no que à intervenção e atuação de algoritmos respeita) vêm sendo objeto de tratamento mais detido, principiar-se-á pelos domínio dos acordos e práticas restritivas da concorrência.

Sobre eles, e à cabeça, considera-se pertinente retomar a assunção (também aqui preliminarmente avançada) da IA como *innovation-enhancing*. Em particular, aquilo que, crê-se acertado afirmar que aquilo que a mesma consente ou possibilita em termos de inovação não poderá deixar de ser pontuado com um fator que sempre se considerou inimigo dessa inovação, e que poderá estar acorrentado à IA: a redução ou eliminação dos riscos da concorrência e da incerteza estratégica que a mesma permite ou coenvolve. Risco e incerteza esses que, num cenário ideal, são os motores que forçam as empresas a inovar, em razão do *perigo da descompetitividade*, e na ótica de perda da preferência do cliente.

Porquanto associada à diminuição do grau de opacidade do mercado[47], aumentando a transparência e a previsibilidade das condutas (através da facilitação da comunicação e da deteção de desvios)[48], a Inteligência Artificial poderá funcionar como elemento ancilar de uma *conspiração*. É, portanto, natural a atenção votada ao risco de as ferramentas e tecnologias de IA serem a origem da facilitação da coordenação necessária à celebração de um acordo ou prática concertada. A sua utilidade não se cinge sequer ao momento inicial, da *celebração* do acordo (ou garantia da concertação), estendendo-se a fases posteriores, atinentes à respetiva execução, implementação e respetivo (in) cumprimento. Por outras palavras, além de possibilitar ou facilitar a implementação do próprio acordo ou coordenação de comportamentos, o recurso a ferramentas ou elementos de IA (de que são exemplos mais recorrentemente avançados, os algoritmos), tem o potencial de agilizar, também, a deteção de desvios e a aplicação de sanções aos incumpridores de determinado entendimento comum, diminuindo, portanto, os incentivos à *batota*.

Em face do exposto, e em termos sumários, não é precipitado concordar-se com os que afirmam que a Inteligência Artificial surge aqui a resolver grande parte dos problemas associados à estabilidade de um cartel: a saber, *i)* a comunicação limitada, *ii)* a existência de incentivos a incumprir, e, ainda, *iii)* as dificuldades de implementação de um entendimento comum, resultantes do aumento da complexidade do esquema, seja em razão do número

[47] Por exemplo, porque da interação frequente entre algoritmos, sobretudo assentes em *machine learning*, a incerteza via de regra associada à dinâmica da concorrência se vê fortemente diminuída.
[48] *Cf.* OCDE – Organização para a Cooperação e Desenvolvimento Económico, "*OECD Business and Finance Outlook 2021...*", *cit.*, §4.2.1.

de participantes, seja por força da diversidade de produtos envolvidos, seja, ainda, pelas características próprias dos mercados em questão. A esta luz, compreensível é a conclusão de que o protagonismo da IA não se circunscreve, em verdade, aos cenários de colusão expressa, estendendo-se ao domínio da colusão tácita, a propósito da qual, se vem avançando o *grosso* das interrogações, seja pela impenetrabilidade da *machine learning*, seja, ainda, pelos riscos associados à partilha ou aquisição, pelas diferentes empresas, de um mesmo algoritmo, programado para maximizar os lucros[49].

Sem prejuízo de tudo o exposto, e cuja premência não deixa dúvidas, urge – retomando o alerta preliminar do capítulo –, não circunscrever o lado *negro* da IA à matéria dos acordos e práticas concertadas.

Desde logo, também a propósito do ilícito de abuso de posição dominante se vem afirmando que a IA poderá funcionar como uma espécie de *quarto fator*[50], que – somando-se aos desafios inerentes e transversais aos mercados digitais – vem contribuir para a incontestabilidade dos seus operadores[51]. Quer isto significar

[49] *Cf. ibidem*. Sobre os vários cenários de colusão algorítmica, *cf.* Ariel Ezrachi / Maurice E. Stucke, "*Artificial Intelligence & Collusion: When Computers Inhibit Competition*", *in University of Illinois Law Review*, Volume 2017, 2017, *Oxford Legal Studies Research Paper No. 18/2015, University of Tennessee Legal Studies Research Paper No. 267*, pp. 1775-1810, *in* https://ssrn.com/abstract=2591874 (27.11.2022), pp. 1781-1796. Sobre as hipóteses de *hub-and-spoke* em particular, *cf.* OCDE – Organização para a Cooperação e Desenvolvimento Económico, "*Roundtable on Hub-and-Spoke Arrangements: Background Note by the Secretariat*", 2019, *in* https://one.oecd.org/document/DAF/COMP(2019)14/en/pdf (27.11.2022), §83 e ss. Em geral sobre o papel dos algoritmos no plano da colusão, *cf.* OCDE – Organização para a Cooperação e Desenvolvimento Económico, "*Algorithms and Collusion: Competition Policy in the Digital Age*", 2017, *in* https://www.oecd.org/competition/algorithms-collusion-competition-policy-in-the-digital-age.htm (27.11.2022); Competition & Markets Authority, "*Algorithms: How they can reduce competition and harm consumers*", 2021, *in* https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/954331/Algorithms_++.pdf (27.11.2022), e Emilio Calvano / Giacomo Calzolari / Vincenzo Denicolò / Sergio Pastorello, "*Algorithmic Pricing What Implications for Competition Policy?*", *in Review of Industrial Organization*, Volume 55, issue 1, No 9, 2019, pp. 155-171, *in* https://link.springer.com/article/10.1007/s11151-019-09689-3 (27.11.2022). Segundo os últimos Autores, registou-se uma evolução face a um contexto inicial, sendo hoje já possível prefigurar cenários em que os algoritmos aprendem a conluiar-se entre si *i)* ainda quando não desenhados ou instruídos para tal, *ii)* sem que comuniquem entre si, e *iii)* sem conhecimento prévio do ambiente em que operam. Os Autores distinguem, para o efeito, entre "*adaptive algorithms*" e "*learning algorithms*".

[50] Neste sentido, reputando a super-dominância dos controladores de acesso como o desafio mais presente da IA, no plano do Direito da Concorrência – *cf.* Václav Šmejkal, "*Three Challenges of Artificial Intelligence…*", *cit.*, pp. 104-107.

[51] Sobre a questão, e em termos que se acompanham em texto – *cf.* OCDE – Organização para a Cooperação e Desenvolvimento Económico, "*OECD Business and Finance Outlook 2021…*", *cit.*, §4.2.2.

que, se os efeitos de escala e de escopo, as externalidades de rede, os benefícios dos dados e as particularidades das *data-driven strategies*, são já realidades nossas conhecidas, e se, portanto, ninguém nega que a conveniência do uso de determinado serviço ou tecnologia aumentará com o número de utilizadores (o que garante a determinado incumbente uma vantagem inegável), não menos verdade é que a associação destes fatores a ferramentas de IA adensa as características dos mercados do tipo *winner-takes-all*. Esta é uma circunstância que não poderá deixar de se repercutir na determinação do poder de mercado de determinada empresa. Repare-se que, se o aumento do número de utilizadores permite, em consequência ou em círculo, o exponenciar desse número e a produção de mais dados, daí se segue que mais e melhores (porquanto mais precisos) serão os *inputs* utilizados para alimentar as ferramentas de IA, o que, por sua vez, gerará *outputs* que, pelo seu valor acrescentado, captam mais utilizadores, num movimento em espiral que tem por tronco comum a possibilidade de reforço do poder de mercado da empresa dominante. A uma outra luz, a própria *capacidade* da empresa em IA poderá, em si mesma, contribuir decisivamente para a qualificação de uma empresa como dominante em determinado mercado, ainda quando uma análise segundo os critérios tradicionais não permita sustentar a existência de poder de mercado num tal nível[52].

Acresce ao exposto – agora no que aos comportamentos (da empresa dominante) respeita -, que também nesta sede poderá a Inteligência Artificial surgir como potenciadora, senão de novas, pelo menos de formas de abuso agravadas ou facilitadas pelas suas potencialidades. A título meramente exemplificativo, pense-se na recolha de informações sobre empresas concorrentes[53], que permitem à empresa dominante definir, com maior precisão, um nível de preços predatórios tendente ao afastamento (*v.g.* exclusão) de um concorrente seu do mercado[54].

[52] *Cf.* Moritz Hennemann, *Artificial Intelligence...*, *cit.*, §16. Também sobre a questão, *cf.* Alžběta Krausová, "*EU Competition Law and Artificial Intelligence: Reflections on Antitrust and Consumer Protection Issues*", *in The Lawyer Quarterly*, Volume 9, No 1, 2019, pp. 79-84, *in* https://tlq.ilaw.cas.cz/index.php/tlq/article/view/322/321 (27.11.2022), pp. 82-83.

[53] Informações que podem respeitar, seja à sua estrutura de custos, seja aos recursos que têm disponíveis, seja, ainda, à sua capacidade para suportar determinados aumentos de preços.

[54] Em sede de abuso de exploração, são várias e também ambivalentes as preocupações atinentes à personalização dos preços e à discriminação a ela inerente. A este propósito, se é certo que a associação automática da personalização de preços a um dano ao consumidor se afigura imprecisa, pois que possibilita a determinados consumidores a aquisição de produtos que – não fora a diferenciação de preços – não lograriam adquirir, pode a personalização encerrar em si mesma

O potencial negativo da IA não se cinge ao abuso de exclusão. Também no plano dos preços discriminatórios, logra a AI ser mobilizada para potenciar, seja a *i)* aplicação de preços diferentes a clientes em igual condição[55], sem uma qualquer justificação associada, seja a *ii)* prática de preços uniformes ou idênticos por referência a clientes em diferentes condições[56]. Finalmente (no que ao abuso de posição dominante se refere), importa também referir que, quando em cenários de integração vertical[57], a IA poderá, exigir a formulação de novas teorias de dano ou à reconfiguração da noção de "abuso" por referência a comportamentos sobre os quais se não geravam, antes, dúvidas relativamente à sua recondução à "concorrência pelo mérito".

Prosseguindo o périplo pelo âmbito de aplicação do Direito da Concorrência, importa, agora, alertar para o impacto da IA no plano do controlo de operações de concentração, onde a sua utilização pelas empresas envolvidas não deixará de influir na análise prospetiva aí adotada. Em particular, a bitola ou o critério operativo subjacente(s) à decisão de (não) aprovação de uma operação sujeita a notificação prévia – a saber, o perigo de entraves significativos à concorrência efetiva no mercado relevante – não poderá deixar de se ver influenciado pela presença de tecnologias de IA. Isto porque, quando as empresas envolvidas na operação de concentração – concorrentes ou não – logram, através da operação, combinar bases de dados, ou alimentar-exponenciar a sua capacidade de IA, as economias de escala e de escopo daí resultantes[58] deverão ser objeto de consideração detida. Além da atenção a votar aos efeitos conglomerados que poderão resultar da conjugação de várias ou

um comportamento abusivo, desde logo quando, e através da discriminação de consumidores de produtos rivais, por via de uma estratégia de preços seletiva, se busque e logre a exclusão de concorrentes do mercado. Segue-se de perto, OCDE – Organização para a Cooperação e Desenvolvimento Económico, "*OECD Business and Finance Outlook 2021...*", *cit.*, §4.2.2. *Vd.*, ainda, Competition & Markets Authority, "*Algorithms: How they can reduce competition...*", *cit.*, §2.2.3.

[55] Na sua condição de clientes, e apesar de diferenças assinaláveis, nos seus padrões de consumo, incluindo aspetos ou dimensões como a lealdade à marca, hábitos de compra e preferências de consumo.

[56] *Cf.* Suzanne Rab, *Competition Law..., cit.*, §§ 11.029-11.032. Importa, porém, circunscrever o problema dos preços diferenciados ao ilícito de abuso de posição dominante, tendo em conta as eficiências associadas a uma eventual diferenciação, por parte de empresas não dominantes. Esta circunscrição não deixa de exigir, ainda assim, cautela.

[57] Que em si mesma não é um problema, porquanto geradora de eficiências, mas em que a questão do *self-preferencing*, isto é, do *favorecimento* de serviços próprios face aos de concorrentes, assume importância.

[58] E em resultado das quais poderá vir a resultar uma situação de incontestabilidade não imediatamente apreensível.

determinadas bases de dados e tecnologias de IA, importa prestar atenção redobrada aos efeitos coordenados[59].

Ainda no plano do controlo de operações de concentração, uma outra dimensão ou núcleo problemático surge(m) evidente(s), e em relação clara com a regulação dos mercados digitais. Trata-se das assim chamadas *killer* ou *zombie acquisitions*, termo que procura designar as transações pelas quais empresas incumbentes adquirem – não raras vezes por valores astronómicos –, empresas disruptivas (as chamadas *newcomers* ou *mavericks*), *matando-as à nascença*. O objetivo subjacente é o de refrear (ou liquidar mesmo, preventivamente) aquela que é, para já, apenas uma concorrência potencial (ou nem isso), tudo isto num momento em que o volume de negócios ou a quota de mercado do *target* não preenchem, ainda, os limiares exigidos para nascimento do dever de notificação prévia[60], e em que, portanto, a transação muito provavelmente escapará ao controlo *ex ante* das operações de concentração[61]. O Regulamento Mercados Digitais procura, agora, mitigar a questão, através do seu artigo 14º, prevendo uma obrigação de comunicar concentrações projetadas, "sempre que as entidades da concentração ou a empresa objeto da concentração prestem serviços essenciais de plataforma ou qualquer outro serviço no setor digital ou permitam a recolha de dados, independentemente de ser ou não de notificação obrigatória à Comissão nos termos do referido regulamento, ou a uma autoridade da concorrência nacional competente, nos termos das regras nacionais relativas a concentrações de empresas". Trata-se, porém, de obrigação que apenas visa um conjunto restrito de empresas, designadas como controladoras de acesso.

[59] *Cf.* Moritz Hennemann, *Artificial Intelligence...*, *cit.*, §31.

[60] Pelo menos entre nós, o dever de notificação prévia de operações de concentração não se estende a toda e qualquer operação, cingindo-se às operações que preencham os limiares determinados pelas normas aplicáveis.

[61] Sobre a questão, *cf.* OCDE – Organização para a Cooperação e Desenvolvimento Económico, "*OECD Business and Finance Outlook 2021...*", *cit.*, §4.2.3. Saliente-se que o legislador procurou acomodar a lacuna referida em texto, no Regulamento dos Mercados Digitais [Regulamento (UE) 2022/1925 do Parlamento Europeu e do Conselho de 14 de setembro de 2022 relativo à disputabilidade e equidade dos mercados no setor digital e que altera as Diretivas (UE) 2019/1937 e (UE) 2020/1828], cujo nº 1 do artigo 14º dispõe nos seguintes termos: "*O controlador de acesso informa a Comissão de qualquer operação de concentração, na aceção do artigo 3º do Regulamento (CE) nº 139/2004, que esteja projetada, sempre que as entidades da concentração ou a empresa objeto da concentração prestem serviços essenciais de plataforma ou qualquer outro serviço no setor digital ou permitam a recolha de dados, independentemente de ser ou não de notificação obrigatória à Comissão nos termos do referido regulamento, ou a uma autoridade da concorrência nacional competente, nos termos das regras nacionais relativas a concentrações de empresas*".

Este é um périplo muito corrido. Visa-se, aqui, apenas alertar para a dimensão do problema, em termos demonstrativos da sua transversalidade aos diferentes domínios de aplicação substantiva do Direito da Concorrência.

4.1. Dos limites à *efetividade* do *enforcement* das regras da concorrência: em particular, a *equidade* dos métodos e do processo

Também no plano adjetivo-processual se suscitam questões e dúvidas várias sobre o potencial de aplicação da IA.

De facto, e em verdadeiro contraponto com as vantagens, oportunidades ou valor acrescentado da IA no plano do *enforcement* do Direito da Concorrência, problemas há, inerentes à investigação de condutas com recurso a ferramentas de inteligência *inorgânica*[62]. Importa recordar que estão em causa *sistemas*, *tecnologias* ou *meios* que, além de rígidos e desligados de imperativos de *fairness* e equidade, se apresentam relativamente imperscrutáveis, em razão do emprego de linguagem que, pese embora descodificável, se afigura carecida de tradução.

Esta circunstância desdobra-se em duas dificuldades de ordem distinta.

Em primeiro lugar, a introdução da IA nos processos produtivos e o surgimento de modelos de negócio nela assentes, conduz a que o secretismo associado às práticas restritivas da concorrência se veja reforçado pela natureza opaca de realidades como os algoritmos[63]. Por outras palavras, há que não menosprezar a circunstância de se estar perante realidades relativamente opacas que, ao jeito de *black box*, oneram as autoridades com a necessidade de compreender sistemas e tecnologias para cuja compreensão não estão, *a priori*, dotadas de recursos e/ou de meios[64].

[62] *Cf.* Charles Kerrigan, *Introductory Essay...*, *cit.*, §1.032.

[63] Sobre a questão, *cf.* OCDE – Organização para a Cooperação e Desenvolvimento Económico, "*OECD Business and Finance Outlook 2021...*", *cit.*, §4.3.2.

[64] Poderá contrapor-se com a possibilidade de as ANC's, desde logo por força dos poderes que lhes são atribuídos nos termos da lei, e que se veem reforçados com a Diretiva ECN+ [Diretiva (UE) 2019/1 do Parlamento Europeu e do Conselho, de 11 de dezembro de 2018, que visa atribuir às autoridades da concorrência dos Estados-Membros competência para aplicarem a lei de forma mais eficaz e garantir o bom funcionamento do mercado interno], apreender e solicitar informações sobre estas tecnologias. No entanto, importa precisar que os problemas identificados nesta sede se não dirigem ou, pelo menos, se não circunscrevem ao *acesso*, contendendo, antes, com a compreensão de uma realidade não imediatamente apreensível, como o são uma ata, uma mensagem de correio eletrónico ou um clausulado contratual. Em particular, e porque poderá ser necessário compreender e testar a tecnologia em ambiente real, algo mais do que a simples possibilidade de apreensão poderá ser necessária (até para evitar *presunções de ilicitude ou dano* associadas, afinal, a tecnologias amigas da concorrência).

Por seu turno, os mesmos problemas de (in)inteligibilidade poderão colocar-se, também, quanto às empresas Visadas, desde logo quando confrontadas com um processo conduzido com recurso a ferramentas e tecnologias de IA. Em particular, seja numa fase inicial do processo, seja enquanto meio de prova, a utilização de ferramentas e mecanismos de IA pelas autoridades competentes poderá vir associada a entropias processuais, pondo em causa os direitos fundamentais das empresas e, em particular, a possibilidade de compreenderem a acusação que lhes é dirigida e de exercerem cabalmente o seu direito de defesa[65].

Em particular, *i)* a promessa de imparcialidade da IA *vis-à-vis* os preconceitos e pré-juízos humanos ou, ainda, *ii)* o respetivo valor acrescentado face às dificuldades associadas à garantia de uma independência-imparcialidade internas (i.e. referentes ou relativas às convicções pessoais daquele que investiga e decide) não poderão ignorar a enorme vulnerabilidade dos sistemas e tecnologias de IA. Sendo que, por vulnerabilidade, nos não referimos, apenas, à maior suscetibilidade a ataques externos, mas, e sobretudo, à captura, condicionamento e enviesamento pela informação com que alimentadas, resultando num verdadeiro *enforcement bias*[66].

Em geral, seja no que à aceleração, encurtamento e, enfim, efetividade das investigações conduzidas respeita, seja, ainda, por referência à priorização de casos ou de mercados em que a intervenção se afigure mais urgente ou *necessária* – e novamente em contraponto com o que se alegou quanto às *promessas* da IA neste domínio -, é pertinente salientar que a maior eficácia e/ou efetividade dos processos não deixa de poder vir associada a défices notórios em matéria de direitos fundamentais.

Em face desses perigos, quer no que à *i) fairness* dos processos respeita, quer no que à *ii)* inteligibilidade, correção e/ou justeza das decisões se refere, importa garantir que o *design* da tecnologia seja capaz de assegurar a garantia, o respeito e a efetividade dos direitos fundamentais no processo (onde estes se impõem, não como *standards* mínimos, mas, e, antes, como mandatos de

[65] *Cf.* Andreas von Bonin / Sharon Malhi, "*The Use of Artificial Intelligence…*", *cit.*, p. 469.

[66] Na medida em que o recurso a ferramentas de IA diminui a intervenção e as ponderações associadas a uma administração sobre a qual recaem deveres de boa-fé e de imparcialidade (além da necessária vinculação aos direitos fundamentais), e tendo em conta que os *outputs* alcançados se veem conformados pela quantidade e qualidade da informação com que alimentados. Recorde-se, no que às ANC's respeita, que a discricionariedade das autoridades não é ilimitada, apresentando-se, antes, como uma discricionariedade vinculada, e que deve respeito aos direitos fundamentais e aos princípios basilares do Ordenamento Jurídico, *cf. ibidem*. Diferentemente se passam as coisas – pelo menos, na ausência de um quadro de *governance* claro – com a IA.

otimização). Por outra palavras, importa assegurar que as vantagens da IA em matéria de *enforcement* se não logrem obter à custa da erosão das garantias do processo equitativo[67].

5. Sobre o *futuro* do Direito da Concorrência num mundo *monopolizado* pela Inteligência Artificial: principais conclusões

O objetivo do presente texto é, como se avançou em sede preliminar, o de alertar para a ambivalência de efeitos da Inteligência Artificial, não só, mas também (ou sobretudo) no Direito da Concorrência, domínio de que não é privativa, mas onde a sua heterogeneidade justifica tratamento aturado.

Referiu-se, a título preliminar, que um diagnóstico adequado para o problema do impacto da IA nos instrumentos e quadros dogmático-conceptuais do Direito da Concorrência não poderia nem poderá, nunca, laborar sobre uma visão apriorística, enviesada e/ou cega à complexidade da IA e respetivo caráter multifacetado, motivo por que se procurou sistematizar, aqui, uma tríade – *the good, the bad and the ugly* – como pano de fundo demonstrativo da necessidade de refutação da tese da IA como fenómeno necessariamente *negativo, suspeito* ou carecido da intervenção-regulação do Direito (e, em particular, das normas da concorrência).

Pelo contrário, e de forma a evitar um *output* gerador da diminuição dos incentivos à inovação, com dano claro para o bem-estar económico, procurou-se estabelecer um ponto de partida neutro e imparcial, e que se considera necessariamente transversal a quaisquer propostas de solução avançadas.

Sobre estas – e apesar da delimitação negativa com que também se principiou o texto – crê-se relevante deixar, agora e aqui, uma nota final[68]. Em particular, além da adoção de um ponto de partida ou visão do estado das coisas necessariamente imparcial, importa catalogar os desafios e as questões a que se pretende dar resposta, à luz de uma distinção basilar, a saber, entre: *i)* os desafios ou problemas que exigem, de facto, respostas novas ou, pelo menos, adaptadas, e *ii)* aqueles que, ainda que *prima facie* se apresentam

[67] *Cf. ibidem*, pp. 469-471. Os Autores sinalizam, em nosso ver, os principais riscos associados à IA, em sede de *enforcement*, e que se podem reconduzir a dois grandes blocos: *i)* o problema do "*enforcement bias*", sobretudo em razão da existência de indústrias e de mercados em que é maior a abundância de dados e informações e *ii)* as entropias introduzidas no processo derivadas de características particulares da IA (sobretudo, a sua complexidade e opacidade), atentos os perigos que as mesmas representam para a garantia do direito a um processo equitativo, incluindo o direito de defesa e a igualdade de armas.

[68] Ainda recondutível ao objeto do escrito, precisamente por se assumir, também, como verdadeiro pressuposto da formulação de estratégia(s).

nas vestes de problema novo ou diferente, se revelam – uma vez desnudados de adereços – perfeitamente passíveis de resolução e de resposta adequada, através da aplicação das regras da concorrência e da recondução às construções até aqui adotadas.

No que à IA se refere em particular, considera-se premente distinguir *i)* os casos em que a mesma é utilizada apenas e tão só como *instrumento facilitador*[69], entre tantos outros, de *ii)* hipóteses fáticas efetivamente carecidas de um enquadramento, senão novo, pelo menos objeto de fundamentação acrescida. E isto, *i)* seja por inexistir ou ser inviável a associação da ferramenta de IA a uma qualquer vontade, instrução e / ou comando pela empresa-sujeito do Direito da Concorrência, *ii)* seja por não haver sequer evidência de uma qualquer estratégia ou acordo anticoncorrencial, caso em que a IA se circunscreve ou reduz ao papel de mera promotora de condições que facilitam a adaptação ao comportamento esperado de concorrentes[70].

A necessidade de distinção de cenários é, segundo se crê, relevante a vários níveis. No entanto, ela assume particular premência no plano da responsabilidade e da imputação do comportamento restritivo da concorrência[71], e isto,

[69] Isto é, para efeitos de efetivação ou agilização da implementação de um acordo ou prática concertada.

[70] Novamente: ao passo que, no primeiro caso, as regras da concorrência se aplicam da mesma forma como se aplicariam e vêm aplicando a hipóteses de acordos ou práticas concertadas implementados ou efetivados através do recurso a diferentes meios, como o telefone ou uma aplicação de conversação (e em que, portanto, do que se trata é, ainda, de uma conduta humana imputável à empresa), já no cenário, a análise jusconcorrencial terá necessariamente de levar em linha de conta um conjunto de fatores, de entre os quais *i)* os incentivos e a posição das empresas utilizadoras de ferramentas de inteligência artificial, *ii)* o modo de funcionamento e os resultados atingidos pela tecnologia, e, bem assim, *iii)* a justificação, a finalidade e os benefícios logrados com o emprego das ferramentas em questão. De igual modo, e novamente a propósito do primeiro cenário, é útil recordar que, tal como a a conduta dos trabalhadores pode ser imputada à empresa respetiva, ainda quando estes hajam atuado sem a sua chancela, sem o seu conhecimento, ou inclusive contra ordens expressas [*cf.* acórdão do Tribunal de Justiça (Quarta Secção) de 21 de julho de 2016, *SIA "VM Remonts" (anteriormente SIA "DIV un KO") e o. contra Konkurences padome*, Processo C-542/14, ECLI:EU:C:2016:578 ('*VM Remonts*')], também os resultados do emprego e recurso a ferramentas de IA, ainda quando assentes em *machine learning*, poderão vir a ser imputados à empresa que beneficia dos resultados por elas logrados, numa ótica de *follow the flow of money*, e, desde logo, à luz de um princípio de precaução ou numa ótica de responsabilidade objetiva.

[71] Sobre esta, *cf.* Moritz Hennemann, *Artificial Intelligence...*, *cit.*, §33, alertando (§§45-46) para a circunstância de a atribuição de personalidade jurídica a sistemas autónomos não permitir, sequer, resolver o problema, antes o adensando, pela necessidade de revisão dos quadros do Direito da Concorrência. Para um elenco das soluções possíveis, incluindo referências a legislação nacional em vigor, *cf.* Václav Šmejkal, "*Three Challenges of Artificial Intelligence...*", *cit.*, em particular, pp. 108-114. *Vd.*, ainda, sobre a responsabilidade civil inerente à cartelização mediada por algoritmos, Marta

independentemente do modelo de responsabilidade que se adote, e do sujeito último ao qual se pretenda imputar os resultados da IA[72].

Com efeito, no primeiro cenário – em que a IA se resume a instrumento *facilitador* da implementação de acordos ou estratégias anticoncorrenciais desenhadas por humanos -, o caráter anticoncorrencial da conduta (e a sua imputação às empresas) não gera dúvidas, pois que, à semelhança de uma mensagem de correio eletrónico, de uma chamada ou da simples presença numa reunião, a utilização de ferramentas de IA figura como *meio*, bem se sabendo como o caráter anticoncorrencial de uma conduta se não vê dependente do instrumento utilizado[73].

Mais problemáticos, porém, serão os cenários em que, *i)* ora não há uma instrução expressa por humanos (apesar da identificação de uma estratégia anticoncorrencial implementada por determinado mecanismo de IA), *ii)* ora se verifica, em razão da coincidência (efeitos cumulativos) da utilização de IA, uma redução da concorrência no mercado, sem evidências, porém, da adoção de estratégias ou comportamentos anticoncorrenciais[74]. Em nosso ver, apenas (ou sobretudo) estes cenários deverão ser considerados, numa lógica de *Direito a constituir*.

Encerrado este parêntese, permitamo-nos, agora, sedimentar a resposta que se procurava aqui avançar.

Reconhecendo-a como uma solução *humilde*, crê-se que a mesma respeita o domínio do possível e é, acima de tudo, fiel ao desiderato: refutar a clareza de um pressuposto porventura subjacente ao discurso que vem sendo adotado relativamente ao futuro do Direito da Concorrência, numa era marcada pela preponderância da IA.

Reitera-se. Os desafios suscitados pelos mercados digitais e, em particular, pela IA, não se compadecem com respostas fáceis, polares (de sim ou de

Teixeira Pires, "A responsabilidade civil inerente à cartelização mediada por algoritmos", *Yearbook Mestrado Faculdade De Direito*, Volume 3, 2020, Universidade Católica Editora, 2021, pp. 1241-1286, *in* https://www.uceditora.ucp.pt/pt/biblioteca-de-investigacao/3085-yearbook-mestrado-faculdade-de-direito-2020.html (27.11.2022).

[72] Apesar da premência da questão, considera-se que as construções adotadas no Direito da Concorrência manterão a sua importância e valia, desde logo para efeitos da compreensão das noções de "acordo" e de "intenção", permitindo, não só melhor balizar as fronteiras entre a licitude e a colusão ilícita, como, e bem assim, determinar a responsabilidade jusconcorrencial e eventual alocação de responsabilidades entre titulares, *developers*, utilizadores e, ainda – para quem o admita – os próprios mecanismos de IA, como os algoritmos (enquanto *facilitadores*).

[73] Neste sentido, *cf.* OCDE – Organização para a Cooperação e Desenvolvimento Económico, "*OECD Business and Finance Outlook 2021...*", *cit.*, §4.3.1.

[74] *Cf. ibidem*.

não) nem, e muito menos, com soluções precipitadas[75] (negação que não se confunde com o refutar de uma abordagem de tipo preventivo ou intervencionista). Assim, se é certo que a disrupção provocada pela IA não consente que o jurista durma na sombra ou que enverede forçosamente por uma abordagem to tipo *wait-and-see*, também não nos parece que a visão da IA como um *game-changer*, gerador do *fim da concorrência como a conhecemos*, na origem de *uma nova era de governance jurídica* no domínio jusconcorrencial, deva ou possa vingar[76].

Vejamos.

Ainda que se não negue que a IA deve efetivamente ser objeto de debate, a mesma não comporta (forçosamente) a *revogação-substituição* dos instrumentos e conceitos operatórios do Direito da Concorrência. Naturalmente, ela exigirá *i)* a destrinça de casos e *ii)* a adaptação dos princípios e quadros dogmáticos a uma realidade que se autonomiza como verdadeiro *specialis*, pelas características a que se deixou referência. Confia-se, no entanto, na flexibilidade e na abertura que sempre caracterizou as normas da concorrência, assim como se suspeita da bondade de vias disruptivas cuja novidade associada e, portanto, a ponderação que não poderá deixar de lhes estar subjacente, dificilmente se compadecem com a celeridade da mutação. Por outras palavras, sendo a Inteligência Artificial um termo *catch-all*, utilizado para designar realidades muito distintas, caracterizadas (tal como os seus efeitos), pela sua assimetria e heterogeneidade, não poderão as respostas, estratégias ou políticas desenhadas para lhe(s) dar resposta prescindir de um conhecimento sólido, assente em estudos de mercado (necessariamente morosos), sobre a realidade que se pretende regular.

[75] Que, por exemplo, ignorem eficiências ou redundem em problemas de articulação com outras áreas do Direito.

[76] Neste sentido, acompanhamos Nicolas Petit, "*Antitrust and Artificial Intelligence: A Research Agenda*", *in Journal of European Competition Law & Practice*, Volume 8, Issue 6, 2017, pp. 361–362, *in* https://doi.org/10.1093/jeclap/lpx033 (27.11.2022). Segundo o Autor, às afirmações mais recorrentes a propósito da utilização da IA nos mercados – e que respeitam ao aumento ou exponenciar das instâncias de verificação de condutas anticoncorrenciais, às dimensões "não-preço", e, bem assim, ao problema da "fraude" do *design* dos mercados algorítmicos -, é possível contrapor um conjunto de considerações, em nosso ver meritórias da maior atenção. Em primeiro lugar, qualquer foco no papel facilitador dos algoritmos terá de ser compaginado e devidamente ponderado com uma análise do seu efeito desestabilizador, no respeitante aos danos à concorrência. Em segundo lugar, falta, ainda, uma compreensão segura sobre as estratégias compensatórias. Em terceiro lugar, a assimetria e a heterogeneidade dos algoritmos é circunstância que não se pode ignorar. Por fim, e além da necessidade de investigação adicional em matéria de prova, as dificuldades no plano dos próprios objetivos prosseguidos pelo Direito da Concorrência não podem, também, ser menosprezadas.

Em suma, não é, nesta sede, adequada, nem *i)* uma *presunção de restrição da, ou dano à concorrência* pela Inteligência Artificial, nem *ii)* a sua assunção como algo estritamente positivo, gerador de eficiências e, portanto, amigo da concorrência (*pelo* e *no* mercado). Com Gönenç Gürkaynak[77], crê-se que o problema não reside propriamente na IA em si mesma considerada, mas, e antes, nas utilizações que dela são feitas pelas empresas, o que se afigura justificativo de uma abordagem flexível, *tailor-made* e proporcional o suficiente a reconhecer o *specialis* de cada situação concreta. Acresce ao exposto que, além das utilizações e destinações particulares, a assunção da IA ou dos algoritmos como *plus factors* ou garantes da maior probabilidade, estabilidade e/ou duração da colusão (como se viu no capítulo referente ao *bad* e ao *ugly* da IA) não poderá senão ser perspetivada como "cenário alternativo" entre outros tantos, vendo-se a sua verificação *in concreto*[78] dependente de fatores como *i)* o grau de personalização das transações, *ii)* o ruído introduzido pelas dimensões "não-preço" da concorrência, entre as quais a qualidade e o caráter diferenciado do serviço, e, bem assim, *iii)* a eventual adoção de contramedidas ou estratégias compensatórias por parte dos compradores[79].

No que ao plano do *Direito a constituir* respeita, os efeitos e perigos listados por referência à IA poderão ser perfeitamente balizados e contidos, através de processos e quadros de *governance* adequados[80]. Ou seja, crê-se possível assegurar as eficiências da IA, sem uma sujeição inelutável aos seus riscos e desvantagens, tudo passando pela introdução de enquadramentos normativos, atentos a dimensões nem sempre compatíveis entre si: a segurança, a robustez, os direitos, a ética e a flexibilidade. A referência ou subordinação a estes quadros torna-se particularmente necessária e relevante, no que ao *enforcement* das regras da concorrência respeita, onde a gestão dos riscos se afigura concretizável, através da introdução de elementos como *checklists*, *impact assessment tools*, e, bem assim, mediante a manutenção de procedimentos e garantias efetivos, que permitam às Visadas, não só inteligir, como questionar e

[77] *Cf.* Gönenç Gürkaynak, *"Algorithms and Artificial Intelligence..."*, *cit.*, p. 34.

[78] Revelador, no que aos algoritmos respeita, da necessidade de uma análise casuística, à luz de um esquema do tipo *checklist*, *cf.* Michal Gal, *"Algorithmic-facilitated Coordination" – Note for OECD Competition Committee Roundtable on Algorithms and Collusion*, OECD, 2017, *in* https://one.oecd.org/document/DAF/COMP/WD(2017)26/en/pdf (27.11.2022), figura 1 (p. 23).

[79] *Cf.* Suzanne Rab, *Competition Law...*, *cit.*, §11.015.

[80] Sobre o papel que as normas da concorrência desleal poderão desempenhar nesta sede, *vd.* Stefan Scheuerer, *"Artificial Intelligence and Unfair Competition – Unveiling an Underestimated Building Block of the AI Regulation Landscape"*, *in GRUR International*, Volume 70, Issue 9, 2021, pp. 834-845, *in* https://doi.org/10.1093/grurint/ikab021 (27.11.2022).

sindicar a legitimidade das ferramentas, do processo e dos seus resultados[81]. Por fim, também a "transversalidade" que caracteriza a IA e as dificuldades que lhe vão associadas poderão ser colmatadas pela previsão de mecanismos de coordenação com outras entidades e reguladores, sobretudo incumbidas de missões no plano da defesa do consumidor e proteção de dados[82].

Ao invés de uma qualquer alegação fechada quanto àquilo que a IA representa, em termos absolutos, para a concorrência, importa reconhecer e não desconsiderar as hipóteses alternativas, sobretudo quando o objetivo da análise é o de ensaiar propostas de solução[83]. Um ponto de partida necessariamente imparcial permitirá conter os riscos e os perigos da IA (no plano substantivo ou adjetivo), sem refrear a sua potencialidade. Os quadros que se consideram dever ter aplicação ao caso não poderão ser construídos senão em termos que tenham por subjacente um efetivo conhecimento prévio da realidade regulada, o que, tomando o seu tempo, poderá implicar que, no curto prazo – e antes, portanto, da opção pela *hard law* – se considere a aplicação de uma miríade de mecanismos regulatórios que lograrão, no curto prazo, uma resposta adequada ao problema. A título meramente exemplificativo, considere-se o potencial da formulação de recomendações, do diálogo regulatório e da condução de investigações que permitam às autoridades competentes acompanhar de perto os novos modelos de negócio e as potencialidades da IA[84].

E eis, portanto, o fito do presente texto.

É esta uma resposta imparcial, holística e prudente. Uma resposta, para muitos, insuficiente. E para outros, uma "não-resposta" até, porquanto mero pressuposto dela. Não obstante, seja enquanto *resposta*, seja como humilde *ponto de partida*, trata-se e procurou-se aqui fazer um ponto que visa evitar *falsos positivos*, restritivos da liberdade de empresa, enquanto direito fundamental (cf. nº 1 do artigo 61º da Constituição da República Portuguesa e artigo 16º da Carta dos Direitos Fundamentais da União Europeia), e lesivos da concorrência, enquanto bem jurídico.

[81] *Cf.* neste sentido, Tirath Virdee, *Understanding AI...*, *cit.*, §§4.049-4.054.

[82] *Cf.* OCDE – Organização para a Cooperação e Desenvolvimento Económico, "*OECD Business and Finance Outlook 2021...*", *cit.*, §4.3.3.

[83] *Cf.* Suzanne Rab, *Competition Law...*, *cit.*, §11.055.

[84] *Cf.* OCDE – Organização para a Cooperação e Desenvolvimento Económico, "*OECD Business and Finance Outlook 2021...*", *cit.*, §4.3.3.

Bibliografia

CALVANO, Emilio / CALZOLARI, Giacomo / DENICOLÒ, Vincenzo / PASTORELLO, Sergio, "*Algorithmic Pricing What Implications for Competition Policy?*", *in Review of Industrial Organization*, Volume 55, issue 1, No 9, 2019, pp. 155-171, *in* https://link.springer.com/article/10.1007/s11151-019-09689-3 (27.11.2022)

COMISSÃO EUROPEIA, *Comunicação da Comissão ao Parlamento Europeu, ao Conselho Europeu, ao Conselho, ao Comité Económico e Social Europeu e ao Comité Das Regiões*: *Inteligência artificial para a Europa*, COM/2018/237 final, Bruxelas (25.4.2018, p. 1)

COMISSÃO EUROPEIA, *Comunicação da Comissão: Enquadramento dos auxílios estatais à investigação, desenvolvimento e inovação*, C(2022) 7388 final, Bruxelas (19.10.2022).

COMISSÃO EUROPEIA (CRÉMER, Jacques / DE MONTJOYE, Yves-Alexandre / SCHWEITZER, Heike), "*Competition policy for the digital era – Final Report*", 2019, ISBN 978-92-76-01946-6, *in* https://ec.europa.eu/competition/publications/reports/kd0419345enn.pdf (27.11.2022)

COMISSÃO EUROPEIA, *Comunicação da Comissão: Enquadramento dos auxílios estatais à investigação, desenvolvimento e inovação*, C(2022) 7388 final, Bruxelas (19.10.2022)

COMPETITION & MARKETS AUTHORITY, "*Algorithms: How they can reduce competition and harm consumers*", 2021, *in* https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/954331/Algorithms_++.pdf (27.11.2022)

EZRACHI, Ariel / STUCKE, Maurice E., "*Artificial Intelligence & Collusion: When Computers Inhibit Competition*", *in University of Illinois Law Review*, Volume 2017, 2017, *Oxford Legal Studies Research Paper No. 18/2015, University of Tennessee Legal Studies Research Paper No. 267*, pp. 1775-1810, *in* https://ssrn.com/abstract=2591874 (27.11.2022)

GAL, Michal, "*Algorithmic-facilitated Coordination*" – *Note for OECD Competition Committee Roundtable on Algorithms and Collusion*, OECD, 2017, *in* https://one.oecd.org/document/DAF/COMP/WD(2017)26/en/pdf (27.11.2022)

GRÜKAYNAK, Gönenç, "*Algorithms and Artificial Intelligence: An Optimist Approach to Efficiencies*", *in Competition Law & Policy Debate Journal*, Volume 5, Issue 3, 2019, pp. 29-34, *in* https://ssrn.com/abstract=3783353 (27.11.2022)

HENNEMANN, Moritz, *Artificial Intelligence and Competition Law*, *in* WISCHMEYER, Thomas / RADEMACHER, Timo (eds.), *Regulating Artificial Intelligence*, [S. l.], Springer Cham, 2020, eBook ISBN: 978-3-030-32361-5, pp. 361-388

HIGH-LEVEL EXPERT GROUP ON ARTIFICIAL INTELLIGENCE, "*A definition of AI: Main capabilities and scientific disciplines*", Comissão Europeia, 2018, *in* https://ec.europa.eu/futurium/en/system/files/ged/ai_hleg_definition_of_ai_18_december_1.pdf (27.11.2022)

HOFFMANN-RIEM, Wolfgang, *Artificial Intelligence as a Challenge for Law and Regulation*, *in* WISCHMEYER Thomas / RADEMACHER, Timo (eds.), *Regulating Artificial Intelligence*, [S. l.], Springer Cham, 2020, eBook ISBN: 978-3-030-32361-5, pp. 1.29

KRAUSOVÁ, Alžběta, "*EU Competition Law and Artificial Intelligence: Reflections on Antitrust and Consumer Protection Issues*", *in The Lawyer Quarterly*, Volume 9, No 1, 2019, pp. 79-84, *in* https://tlq.ilaw.cas.cz/index.php/tlq/article/view/322/321 (27.11.2022)

KERRIGAN, Charles, *Introductory Essay*, *in* KERRIGAN, Charles (ed.), *Artificial Intelligence – Law and Regulation*, [S. l.], Elgar, 2022, eISBN: 9781800371729, pp. 2-23

KERRIGAN, Charles (com contribuições de Suzanne Rab, Stephen Kenny QC, Charlotte Payne e Jason G. Allen), *Introduction to AI*, *in* KERRIGAN, Charles (ed.), *Artificial Intelligence – Law and Regulation*, [S. l.], Elgar, 2022, eISBN: 9781800371729, pp. 30-36

OCDE – ORGANIZAÇÃO PARA A COOPERAÇÃO E DESENVOLVIMENTO ECONÓMICO, "*Algorithms and Collusion: Competition Policy in the Digital Age*", 2017, *in* https://www.oecd.org/competition/algorithms-collusion-competition-policy-in-the-digital-age.htm (27.11.2022)

OCDE – ORGANIZAÇÃO PARA A COOPERAÇÃO E DESENVOLVIMENTO ECONÓMICO, "*OECD Business and Finance Outlook 2021: AI in Business and Finance*", Paris, OECD Publishing, 2019, *in* https://doi.org/10.1787/ba682899-en (27.11.2022)

OCDE – ORGANIZAÇÃO PARA A COOPERAÇÃO E DESENVOLVIMENTO ECONÓMICO, "*Roundtable on Hub-and-Spoke Arrangements: Background Note by the Secretariat*", 2019, *in* https://one.oecd.org/document/DAF/COMP(2019)14/en/pdf (27.11.2022)

PETIT, Nicolas, "*Antitrust and Artificial Intelligence: A Research Agenda*", *in Journal of European Competition Law & Practice*, Volume 8, Issue 6, 2017, pp. 361–362, *in* https://doi.org/10.1093/jeclap/lpx033 (27.11.2022)

PIRES, Marta Teixeira, "A responsabilidade civil inerente à cartelização mediada por algoritmos", *Yearbook Mestrado Faculdade De Direito*, Volume 3, 2020, Universidade Católica Editora, 2021, pp. 1241-1286, *in* https://www.uceditora.ucp.pt/pt/biblioteca-de-investigacao/3085-yearbook-mestrado-faculdade-de-direito-2020.html (27.11.2022).

RAB, Suzanne, *Competition Law*, *in* KERRIGAN, Charles (ed.), *Artificial Intelligence – Law and Regulation*, [S. l.], Elgar, 2022, eISBN: 9781800371729, pp. 197-212

SCHEUERER, Stefan, "*Artificial Intelligence and Unfair Competition – Unveiling an Underestimated Building Block of the AI Regulation Landscape*", *in GRUR International*, Volume 70, Issue 9, 2021, pp. 834-845, *in* https://doi.org/10.1093/grurint/ikab021 (27.11.2022)

ŠMEJKAL, Václav, "*Three Challenges of Artificial Intelligence for Antitrust Policy and Law*", *in Charles University in Prague Faculty of Law Research Paper*, No. 2021/III/3, 2021, pp. 1-17, *in* https://ssrn.com/abstract=3984354 ou http://dx.doi.org/10.2139/ssrn.3984354 (27.11.2022)

VIRDEE, Tirath, *Understanding AI*, *in* KERRIGAN, Charles (ed.), *Artificial Intelligence – Law and Regulation*, [S. l.], Elgar, 2022, eISBN: 9781800371729, pp. 37-55

VON BONIN, Andreas / MALHI, Sharon, "*The Use of Artificial Intelligence in the Future of Competition Law Enforcement*", *in Journal of European Competition Law & Practice*, Volume 11, Issue 8, 2020, pp. 468-471, *in* https://awards.concurrences.com/IMG/pdf/the_use_of_artificial_intelligence_in_the_future_of_competition_law_enforcement.pdf?67667/695bd2238811a7abb3f664646dfc0ed74d99b8a3975cf6c3331bc7e9efea478c (27.11.2022)