How do Decision Support Systems Nudge?

F. Pedrazzoli, F. A. D'Asaro^{*}, M. Badino

Department of Human Sciences, University of Verona, Verona, (Italy) International Conference on Computer Ethics: Philosophical Enquiry (CEPE) 2023, Chicago, IL

Keywords: nudging, decision support systems, ethical AI

Extended Abstract

In this paper, we explore Decision Support Systems and their ability to nudge users toward particular decisions. DSSs are computer-based systems designed to assist decision-makers in solving (un)structured problems by utilizing data and models. Over time, these systems have evolved with the integration of Artificial Intelligence, and we categorize them into Traditional DSSs and AI-Powered DSSs. We then examine the concept of nudging and its digital counterpart, defining it as a value-embedding practice that shapes environments to encourage individual behaviors aligned with foundational values without imposing restrictive measures or mandatory treatments. We identify three types of digital nudges: (i) nudging in digital environments through data collection and exploitation, and (iii) nudging through digital artifacts such as AI-powered artifacts and services.

The effectiveness of these nudging techniques depends on the specific application context and objectives. Further research is needed to investigate translation strategies from visualization-based nudges to dialogue-based nudges and the potential of language-based nudging. The paper concludes by acknowledging the limitation of the unclear distinction between Traditional and AI-Powered DSSs due to the ambiguity inherent to the concept of AI, and the need for further investigation into the relationship between nudging and boosting.

1 Introduction

The prevalence of Artificial Intelligence (AI) is on a rapid rise, and it's gradually penetrating into various aspects of our society and everyday routines. One of the most promising applications in this field is the possibility of exploiting and involving AI in decision-making pipelines. Tools that facilitate and support decisions have been classically dubbed Decision Support Systems (DSSs), and their scope ranges, e.g., from simple interfaces to extract information from databases to search engines and more sophisticated software that uses Machine Learning or Knowledge Representation methods to advice their users [Tariq and Rafi, 2012]. Given their involvement in the decision-making process, DSSs can operate, by their very nature, some form of influence on the user. For instance, a search engine can lead the user to access different websites according to the order in which results are presented. The type of influence operated by a DSS can dramatically vary from unharmful forms of persuasion

[Orji and Moffatt, 2018] to nudging [Caraban et al., 2019] and manipulation [Susser et al., 2019].

The focus of the present paper is to relate DSS and nudging in order to understand better how a DSS can nudge in light of the most recent advances in AI. The original definition of nudging [Thaler and Sunstein, 2008] assigns it a neutral meaning: it is "any aspect of the choice architecture that alters people's behavior in a predictable way without forbidding any option or significantly changing their economic incentive." On the other hand, other authors provide and use a different meaning of this very concept. For instance [Bovens, 2009] assumes nudging can take a negative meaning when it "works in the dark", i.e., it bypasses the user.

In this paper, we aim to analyze DSSs and their relation to nudging. We propose that a successful analysis of this problem should start off by providing working definitions of DSSs and nudging, with the intention of highlighting how new trends in AI are reshaping the concept of nudging itself. The paper is structured as follows: in Section 2, we propose a taxonomy of DSSs, in Section 3, we define nudging and its characteristics, in Section 4, we analyze the main differences between the classical definition of nudging and its digital counterpart, in Section 5 we analyze how AI-powered interactions enable DSSs to nudge users in new ways, and finally we draw some conclusions and discuss the next steps of our research.

2 What is a Decision Support System?

To the best of our knowledge, Decision Support Systems (DSSs) were first introduced and articulated in [Gorry and Morton, 1971]. DSS are defined there as "interactive computer-based systems, which help decision makers utilize data and models to solve unstructured problems". Since then, other proposals were put forward (see, e.g., [Keen, 1980, Turban, 1995]). Only recently, however, we have started seeing DSSs that exploit Artificial Intelligence, possibly making the interaction with the end-user more complex and articulated. Therefore, previous definitions and taxonomies appear to be too old-fashioned if one wants to deal with this emerging class of systems that go beyond simple interfaces to query unstructured data. Some more recent proposals advance the idea of dub bing these systems "intelligent" (see, e.g., [Tariq and Rafi, 2012]). However, we take a less ambitious stance, and distinguish between Traditional Decision Support Systems (T-DSSs) and AI-Powered Decision Support Systems (AI-DSSs). The primary differences between the two lie in their levels of transparency and the complexity of their decision-making processes, rather than on their "intelligence". As DSSs become more sophisticated and AI-driven, epistemic concerns about opacity, accountability, as well as ethical concerns may arise. One notable issue is the agency tradeoff between the human and the AI-DSS, which can cause problems related to the amount of decision power shared between the parties. This calls for careful consideration and regulation to address these challenges effectively. Before tackling these issues, we move on to defining T-DSSs and AI-DSSs in greater detail.

2.1 Traditional DSSs

T-DSSs utilize straightforward technologies, allowing users to grasp and follow the decision-making process. They typically avoid complex machine learning methods and serve as simple decision aids with clear processes. T-DSSs are designed to help users make data-driven decisions through data analysis techniques and information

visualization. We regard, for example, analytics of various web platforms as T-DSSs.

The nature of these systems is mainly instrumental, and therefore it is up to their developers to design their interfaces. The effectiveness of T-DSSs ultimately depends on these design choices (e.g., what information is displayed or hidden, how information is presented, etc.) Thus, T-DSSs are tools that mainly rests on Human-Computer Interaction (HCI) techniques for what concerns their effectiveness.

2.2 AI-powered DSSs

AI-powered DSSs represent a step forward compared to T-DSSs, as they are capable of analyzing large amounts of data more quickly and efficiently and providing more accurate and relevant information, suggestions, and recommendations. To exemplify, this category of DSS includes recommendation systems, forecasting systems, simulation systems, image recognition systems, chatbots, etc. AI DSSs may be further classified as Strong DSSs (S-DSSs) and Dynamic-Interface DSSs (DI-DSSs). They usually incorporate machine learning techniques, which can be challenging to interpret. As we have already highlighted above, these systems often offer more advanced and accurate recommendations but their lack of transparency can lead to concerns about accountability and trustworthiness. Dynamic-Interface DSSs not only (possibly) use machine learning for decision making but also implement AI-powered technologies in their interfaces (e.g., conversational agents). This enhances user interaction but increases opacity in both the decision-making process and interface, making it difficult to understand the system's operation and recommendations.

In both S-DSSs and DI-DSSs, the AI component augments these tools with some degree of agency, i.e., autonomy, interactivity, and adaptability. These differences make S-DSSs and DI-DSSs belong to Human-AI Interaction (HAII) rather than HCI as in the case of T-DSSs. This aspect has an impact on the role of these systems on the whole decision process. In fact, AI-powered DSSs may be perceived as collaborators rather than simple support tools due to the increased agency level. It is in this sense that sometimes these systems are said to belong to the Human-AI Collaboration field.

3

3 What is a Nudge?

According to Thaler and Sunstein [Thaler and Sunstein, 2008], "to nudge someone" means to intervene in the environment or choice architecture in which someone operates, with the aim of making their behavior predictable and, therefore in a certain sense more likely. Their proposal, therefore, starts from a rather simple and intuitive idea: the environment in which someone operates is not neutral with respect to the decisions and actions that can be taken within it. It is not new that altering the form of something can produce effects on the behavior of its users. Let us consider, for instance, how the design of an object can influence both the interaction with the object itself and the instrumental use that can be made of it. For example, in the case of a pair of scissors, the two rings at the base of the blade suggest that the handle of the tool is for one hand and that two fingers are the necessary actuators to operate it. Furthermore, the profile of the tips of a pair of scissors can significantly modify the tool's

ability to act and, therefore, more or less indirectly, the likelihood that it will be used for certain purposes and not others. The rounded tips, for example, do not exclude the possibility that a pair of scissors could be used as a cutting weapon, but certainly, reduce the probability of this possible use. This is somewhat isomorphic to the classic nudge example of designing and rearranging objects in a cafeteria: putting healthy food next to the cashier makes customers more likely to buy healthy food. However, what seems to further characterize the practice of nudging, and at the same time places it in a gray area of ethics, is not simply the interest in designing environments or choice architectures but rather in designing them in light of particular structures and cognitive tendencies of human beings. In fact, as pointed out in [Hansen, 2016], "Nudges are called for because of cognitive boundaries, biases, routines, and habits in individual and social decision-making, and work by making use of those boundaries, biases, routines, and habits as integral parts of the choice architecture." Definitions of nudging that include the functioning of nudging itself, or the theory underlying its operation, are more precise than the one analyzed earlier, and encompass the anthropology supported by behavioral economic theory. This, as mentioned, has raised ethical concerns, and in addition to controversy and debate, proposals have also been made to ethically frame the practice of nudging, which include ethical criteria such as transparency, respect for dignity, welfare (justice), and autonomy in the definition. These "ethical boundaries" serve to distinguish nudging, as intended in the ideology of libertarian paternalism, from nudging oriented towards personal interest or manipulation. The ethical aspect of nudging must be evaluated on a case-by-case basis, according to the context and its possible effects.

To conclude with, putting aside ethical considerations, a nudge can be understood as a value-embedding practice that shapes environments to encourage individual behaviors aligned with its foundational values, promoting aspects such as self-care, environmental sustainability, and assistance without imposing restrictive measures or mandatory treatments.

4 What is a Digital Nudge?

The concept of nudging, and its applications, have been quickly adopted by HCI and HAII communities due to the nature of digital environments and artifacts. Indeed, both are designed by human programmers, and both tend to influence their users.

A definition of digital nudging which is closest to its analogical counterpart can be found in [Lembcke et al., 2019]:

"any intended and goal-oriented intervention element (e.g., design, information or interaction elements) in digital or blended environ ments attempting to influence people's judgment, choice, or behavior in a predictable way"

Moreover, just like in analogical nudging, digital nudging is not only concerned with design, but it is also concerned with (i) considering human heuristics and biases, and (ii) improving individual or societal welfare. However, in our view there are at least three meanings according to which the adjective "digital" can be associated with the term "nudging": (i) Nudging in digital (or blended) environments, (ii) Nudging through digital (or blended) environments, and (iii) Nudging through digital artifacts.

In the case of (i), digital nudges are nothing more than standard nudge practices applied in a digital context. For instance, nudge practices such as default or decoy options are frequently employed within digital environments such as e-commerce. In [Bergram et al., 2022], the authors have collected digital nudge practices by analyzing empirical literature on the subject and have categorized them into 10 types (or patterns) of intervention, namely: social, reinforce, disclosure, friction, feedback, default, warning, scarcity, deception, and commitment. It is worth noting that in this case, similarly to their analog counterpart, nudges are mainly based on heuristics and biases. In the case of (ii), on the other hand, nudge practices are specific to digital environments and are possible thank to the properties of these environments themselves. Differently from (i), these nudges are often based on data collection and exploitation, and do not solely rely on heuristics. Through the collection and analysis of user data, it is possible to modify the composition of the choice structure that users navigate, making exposure to certain content more likely than others. In this sense, digital environments become AI-powered digital environments, namely, interactive, autonomous, and adaptable environments that enable personalized nudges. For instance, digital environments such as Social Media Networks are silently animated by Recommender Systems (RS) and Ranking Algorithms (RA), as they proactively and dynamically modify user experience based on the data they collect from users. Therefore, digital nudging practices are not limited to shaping the interface, but also involve the design of RS and RA that dynamically structure the user experience (e.g. [Rieger et al., 2020]).

Finally, in the case of (iii), a nudge is digital because the source accountable for the nudge is a digital technology programmed with the specific purpose of influencing the behavior of the user interacting with it. [Bergram et al., 2022] defines this type of digital nudges as "the event where digital artifacts steer people in particular directions while also allowing them to go their own way." It is worth noting that when we say that a (digital) artifacts nudge someone, we mean something more than simple behavioral influence. Any artifact influences our behavior one way or another. A digital artifact nudges someone when, through its own agency (interactivity, autonomy, and adaptability), it is able to push someone to adopt new behaviors or modify existing ones without manipulation or coercion. These artifacts may widely vary in terms of their complexity. For instance, simple mobile apps to improve users' focus, or physical and psychological wellbeing may be regarded as examples of type (iii) nudges. More complex digital artifacts of this type are AI-powered artifacts or AI-based services (e.g., advanced chatbots and AI-assistants) because of their high degree of interactivity, autonomy, and adaptability. As in the case of AI powered environments, the "AI" element opens up to data, as a resource to design and implement nudges of this type. One of the underlying narratives guiding research in the field of HAII is that the agency of the machines can enhance human agency without replacing it or threatening it [Sundar, 2020]. So, we are trying to design AIs that can help us, as individuals and as a society, in many fields (e.g., healthcare, education, etc.) This could be done in at least two ways. The first consists in thinking of AIs as simple instruments, neutral with respect to the values, with all the responsibilities and ethical concerns attached to its user. The second is the idea that AIs embody values that can be promoted through their own agency, namely, that AIs can perform actions that nudge us to behave consistently with those values.

5 How do Decision Support Systems Nudge?

To answer the question "How does a DSS nudge?" we need to take into account the specific cases we aim to investigate, the application context, and the specific objectives guiding the design of the nudge. Generally, one could respond by listing the ten intervention types we have mentioned above. However, in this section, we will attempt to articulate an answer based on the different types of DSS we have described earlier.

T-DSSs are designed to assist users in making data-driven decisions through data analysis techniques and information visualization. As we already mentioned above, some web platform analytics are T-DSSs. They are tools that allow users to analyze data related to their website or profile traffic. To nudge users, a DSS of this kind can employ graphical data visualization to highlight decision options that align with the user's predetermined goals. However, since these DSSs are comparable to normal, non-intelligent tools, the design of a nudge appears to be a simple design practice oriented towards maximizing the tool's effectiveness.

In the case of AI-powered DSSs, we can identify two ways of nudging. The first concerns the probability that the user accepts the output of the DSS based on the user's perception of the DSS and its performance. Indeed, human perception of these machines is characterized by certain biases and stereotypes. For example, machines are typically considered mechanistic, objective, efficient, unyielding, unemotional, cold, transactional, and prone to being hacked [Sundar, 2020]. Therefore, biases and heuristics of users regarding the capabilities and performance of these machines can be used as a strategic basis on which to operate a nudge. For example, in Europe, by law, every AI will be required to reveal its identity when interacting with a human being. This will expose users to two heuristics that are activated in light of this disclosure. The first is called "automation bias," which is the tendency to overtrust machines and underestimate humans to perform the same tasks [Mosier et al., 1996]; the second is called algorithm aversion [Dietvorst et al., 2015], i.e., the tendency to prefer human judgments over algorithmic decisions even when the latter may be suboptimal. Therefore, where we want the user to follow the advice of the AI, we can try to activate the heuristic of automation bias and counteract that of algorithmic aversion. In this way, we will increase the probability that the user accepts the recommendation of the AI and we will therefore nudge them towards that decision. Conversely, where we want the user to exercise their critical thinking, for example, in a collaborative task, we can try to stimulate algorithmic aversion and counteract automation bias. Therefore, such nudges modulate the credibility of the DSS and the perception of the reliability of its outputs.

The second way in which an AI-powered DSS can nudge concerns its ability to anticipate the user's behavior based on their mental models and epistemic behaviors. This approach is particularly evident in the case of RS embedded in web platforms, due to the large amount of data generated by users. In [Reuver et al., 2021], for example, the authors propose the design of a News Recommendation System to support citizens and public debate, based on what they call the latitude of diversity, that is, the degree of individual acceptance of diversity. In this way, they argue, it is possible to increase the likelihood that users interact with recommendations from points of view, counteracting polarization and fake news through nudging rather than debiasing strategies.

Finally, the peculiarity of a DI-DSS lies in its interaction with the user through generative or linguistic models. Natural language is traditionally an interactive tool that can be used for a varied set of purposes, ranging from persuasion to manipulation. In other words, language can be used to present rational arguments to convince others and exploit psychological and cognitive mechanisms to manipulate them. In between these two attitudes, we can

legitimately speak of nudging, especially in those cases where the strategic intervention is to the benefit and interest of the user or society. Human shop assistants in physical stores can nudge customers, as they accompany them along the decision-making process and engage in conversations with them. Salespeople are skilled in persuasive and manipulative techniques that involve language. For example, in a clothing store, the shop assistant could push the customer to buy a jacket in different ways, such as immediately proposing the item and providing all the reasons why they believe it is the most suitable for them, or leveraging the customer's vanity and filling them with compliments and flattery. Access to natural language by AIs opens up the use of language as a tool to implement nudges. ChatGPT, for example, has recently been released, and yet it is already being adopted in various digital services. Therefore, in the near future, (i) translation work could be needed for nudging strategies based on visualization into strategies based on dialogue, and (ii) their effectiveness will need to be investigated. For example, a Recommendation System for an e-commerce platform could manifest as a virtual assistant that interacts with the user through natural language, rather than simply ordering products visually in the interface.

Furthermore, the way questions are posed by AI to the human user in a conversation can significantly influence the user's response, similar to what happens with survey or questionnaire questions. Finally, interaction through natural language could open new opportunities for nudging strategies based on the characteristics of the language itself. Language is not just a simple tool for communicating messages but also represents an environment that reflects human experience and culture. Words are not neutral entities, but are emotionally connotated and refer to semantic cores. It might be possible to exploit these features of language to elicit certain behavioral responses from human users.

6 Conclusion

Decision Support Systems (DSSs) are computer-based systems that help decision makers solve unstructured problems by utilizing data and models. Initially proposed in 1971, DSSs have evolved over time with the integration of Artificial Intelligence. The difference between Traditional DSSs and AI-Powered DSSs lies in their level of transparency and complexity of decision-making processes. T-DSSs are simple decision aids that use data analysis and visualization to help users make data-driven decisions. They avoid complex machine learning methods and rely on interface design and informations selection to be effective. AI-powered Decision Support Systems can process large amounts of data quickly and accurately, providing relevant suggestions and recommendations. To better differentiate the use of AI for data analysis versus its use for interface enhancement, we have categorized AI-DSSs into two subgroups: Strong DSSs and Dynamic-Interface DSSs. With the increasing use of AI in DSSs, one notable issue is the tradeoff between human and AI decision power, which can cause problems related to the amount of decision power shared between the parties. In order to articulate an answer to the question "How do DSSs nudge?" we analyzed the concept of nudging and its digital counterpart. We neutrally defined nudging as a value-embedding practice that shapes environments to encourage individual behaviors aligned with its foundational values, promoting aspects such as self-care, environmental sustainability, and assistance without imposing restrictive measures or mandatory treatments. In a similar way, digital nudging refers to intentional interventions in digital or blended environments that attempt to influence people's judgment,

choice, or behavior in a predictable way. We categorized digital nudges into three types: (i) nudging in digital environments using standard nudge practices, (ii) nudging in digital environments through data collection and exploitation, and (iii) nudging through digital artifacts such as AI-powered artifacts and services. Nudging in digital environments involves the application of heuristics and biases, while nudging through digital artifacts on modifying the choice structure based on the data collected from users. Nudging through digital artifacts involves the use of digital technologies programmed to influence the behavior of the user inter acting with it. Therefore, there are many ways in which a DSS can nudge, and all of them depend on the considered application context and abstraction level. T-DSS nudges users through data visualization techniques that align with the user's predetermined goals. AI-powered DSS nudges users in two ways: by modulating the credibility of the DSS based on users' biases and heuristics, and by anticipating users' behavior using mental models and epistemic behaviors. Lastly, DI-DSS nudges users through the use of natural language, opening up opportunities for nudging strategies based on the characteristics of language itself.

The effectiveness of these nudging techniques depends on the specific appli cation context and objectives. Further research is needed to investigate how strategies based on visualization can be translated into strategies based on di alogue and the effectiveness of such translations. Additionally, future studies could explore the potential of language-based nudging by examining the emo tional connotations and semantic cores of words in order to elicit specific be havioral responses from users.

One limitation of this work is that the difference between Traditional and AI-Powered Decision Support Systems is not very sharp. This is due to the ambiguity inherent to the concept of AI itself. As more philosophical work about their distinction is carried out, more aspects of our work may be developed to make our taxonomy more precise and refined in terms of implications. In future work, we also aim to investigate further the relationship between nudging and boosting [Gr⁻une-Yanoff and Hertwig, 2016], with the aim of comparing different anthropological and philosophical assumptions characterizing these two strategies in the field of HAII.

References

- [Bergram et al., 2022] Bergram, K., Djokovic, M., Bezen, con, V., and Holzer, A. (2022). The digital landscape of nudging: A systematic literature review of empirical research on digital nudges. In *Proceedings of the* 2022 CHI Con ference on Human Factors in Computing Systems, CHI '22, New York, NY, USA. Association for Computing Machinery.
- [Bovens, 2009] Bovens, L. (2009). The ethics of nudge. *Preference change: Approaches from philosophy, economics and psychology*, pages 207–219.
- [Caraban et al., 2019] Caraban, A., Karapanos, E., Gon, calves, D., and Cam pos, P. (2019). 23 ways to nudge: A review of technology-mediated nudging in human-computer interaction. In *Proceedings of the 2019 CHI conference on human factors in computing systems*, pages 1–15.
- [Dietvorst et al., 2015] Dietvorst, B. J., Simmons, J. P., and Massey, C. (2015). Algorithm aversion: people erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General*,

144(1):114.

[Gorry and Morton, 1971] Gorry, A. G. and Morton, S. S. (1971). Framework for management information systems. *Sloan Management Review*, 13(1).

- [Gr[°]une-Yanoff and Hertwig, 2016] Gr[°]une-Yanoff, T. and Hertwig, R. (2016). Nudge versus boost: How coherent are policy and theory? *Minds and Machines*, 26(1):149–183.
- [Hansen, 2016] Hansen, P. G. (2016). The definition of nudge and libertarian paternalism: Does the hand fit the glove? *European Journal of Risk Regulation*, 7(01):155–174.
- [Keen, 1980] Keen, P. G. (1980). *Decision support systems: a research perspective*. Pergamon Press, Oxford, New York.
- [Lembcke et al., 2019] Lembcke, T.-B., Engelbrecht, N., Brendel, A. B., Her renkind, B., and Kolbe, L. M. (2019). Towards a unified understanding of digital nudging by addressing its analog roots. In *Pacific Asia Conference on Information Systems*.
- [Mosier et al., 1996] Mosier, K. L., Skitka, L. J., Burdick, M. D., and Heers, S. T. (1996). Automation bias, accountability, and verification behaviors. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 40(4):204–208.
- [Orji and Moffatt, 2018] Orji, R. and Moffatt, K. (2018). Persuasive technology for health and wellness: State-of-the-art and emerging trends. *Health informatics journal*, 24(1):66–91.
- [Reuver et al., 2021] Reuver, M., Mattis, N., Sax, M., Verberne, S., Tintarev, N., Helberger, N., Moeller, J., Vrijenhoek, S., Fokkens, A., and van Atteveldt, W. (2021). Are we human, or are we users? the role of natural language processing in human-centric news recommenders that nudge users to diverse content. In *Proceedings of the 1st Workshop on NLP for Positive Impact*, pages 47–59.
- [Rieger et al., 2020] Rieger, A., Theune, M., and Tintarev, N. (2020). Toward natural language mitigation strategies for cognitive biases in recommender systems. In 2nd workshop on interactive natural language technology for explainable artificial intelligence, pages 50–54.
- [Sundar, 2020] Sundar, S. S. (2020). Rise of machine agency: A framework for studying the psychology of human–AI interaction (HAII). *Journal of Computer-Mediated Communication*, 25(1):74–88.
- [Susser et al., 2019] Susser, D., Roessler, B., and Nissenbaum, H. (2019). Technology, autonomy, and manipulation. *Internet Policy Review*, 8(2).
- [Tariq and Rafi, 2012] Tariq, A. and Rafi, K. (2012). Intelligent decision support systems-a framework. In *Information and Knowledge Management*, volume 2, pages 12–20.
- [Thaler and Sunstein, 2008] Thaler, R. and Sunstein, C. (2008). Nudge: Improving Decisions about Health, Wealth, and Happiness. Yale University Press.

[Turban, 1995] Turban, E. (1995). Decision Support and Expert Systems: Management Support Systems.

Englewood Cliffs, N J.