


Sentiment Analysis in Literary Studies. A Critical Survey

Simone Rebora <simone_dot_rebora_at_uni-mainz_dot_de>, Johannes Gutenberg University Mainz 
<https://orcid.org/0000-0002-1501-3774>

Abstract

The article sets up a critique of Sentiment Analysis (SA) tools in literary studies, both from a theoretical and a computational point of view. In the first section, a possible use of SA in narratology and reader response studies is discussed, highlighting the gaps between literary theories and computational models, and suggesting possible solutions to fill them. In the second section, a stratified taxonomy of SA tools is proposed, which distinguishes: (1) the emotion theory adopted by the tool; (2) the method used to build the emotion resources; (3) the technique adopted to accomplish the analysis. A critical survey of six representative SA tools for literary studies (Syuzhet, Vader, SentiArt, SEANCE, Stanford SA, and Transformers Pipelines) closes the article.

1. Introduction

After years of disinterest and neglect, Sentiment Analysis (SA) has recently become one of the most discussed topics in computational literary studies. Generally known as the field of study that analyzes “people’s opinions, sentiments, appraisals, attitudes, and emotions towards entities and their attributes” [Liu 2015, 1], SA aims at accomplishing an automated extraction of the emotional content of text by converting it into machine-readable information, such as numbers and discrete labels (e.g., “positive” vs. “negative”), which can be then analyzed statistically or visualized via plots and graphs. It is precisely after June 2014, when Matthew Jockers published a series of thought-provoking posts on his personal blog [Jockers 2014] [Jockers 2015a] [Jockers 2015b], that this computational technique started to acquire a growing relevance in DH research. To date, SA has been used in multiple studies, ranging from the identification of the “basic shapes” of stories cf. [Jockers 2015b] [Reagan et al. 2016], to the large-scale investigation of children’s literature [Jacobs et al. 2020] and online reading communities [Pianzola et al. 2020], with applications on both contemporary and historical languages [Sprugnoli et al. 2020]. This success is paralleled — even if without a causal relation — by the so-called *affective turn* [Clough and Halley 2007] in literary studies [Keen 2011], through which emotions regained a key role in narrative theory and reader response studies, after the neglect — and even the patent opposition — of both structuralist and post-structuralist studies.

However, as it happens with all vanguards and new trends, SA is currently experiencing its peak of theoretical and methodological instability as well. While research in a field like stylometry (the computational analysis of style) has already reached some of the highest levels in terms of scientificity and theoretical awareness,^[1] the same cannot be said with SA, where most tools and methods still lack a validation, connections with literary theory are frail or disputable, and an organizational effort of the research community (such as that of the Special Interest Group “Digital Literary Stylistics”, or of the Federation of Stylometry Labs) is still lacking.

A first, extensive survey of SA for computational literary studies was proposed by [Kim and Klinger 2018b], who distinguished five main areas of application:

1. Classification of literary texts in terms of the emotions they convey
2. Genre and story-type classification

1

2

3

3. Modeling sentiments and emotions in texts from previous centuries
4. Character network analysis based on emotional relations
5. Miscellaneous and more general applications

However, at the end of their survey, Kim and Klinger notice how “the methods of sentiment analysis used by some of the DH scholars nowadays have gone or are almost extinct among computational linguists” [Kim and Klinger 2018b, 33]. There is a clear gap in terms of methodology that needs to be filled (or, at least, directly addressed), if research wants to move further. This necessity is also accompanied by a more general need to understand the internal logics of tools that are frequently used in a purely goal-oriented research. In other terms, a criticism of the tools and methods currently adopted in SA is as necessary as a free exploration of its potential. Without denying the usefulness of exploratory — or even “deformative” [Buurma and Gold 2018] — research, a critical awareness of shortfalls and limitations in distant reading should constitute the necessary groundwork for its fruitful application.

With this article, I will attempt a comprehensive criticism of SA tools in literary criticism by combining two approaches. First, as already done by [Herrmann et al. 2015] for stylometry and by [Ciotti 2017b] for topic modeling, I will conduct a theoretical inquiry on the possible correspondences and inconsistencies between the concepts of literary theory and the computational techniques of SA. This, in order to provide a solution to a basic question of modeling (cf. [Flanders and Jannidis 2019] [Underwood 2019b] [Piper 2018]): to what extent can SA model literary phenomena and thus make possible an *operationalization* (cf. [Moretti 2013] [Salgaro 2018]) of the fundamental questions of literary studies? Main distinction will be in where the emotions reside: if in the text itself (a perspective generally preferred by narratological approaches) or in its readers (main focus of reader-response studies). Second, by narrowing the perspective on some of the most widely used SA tools in DH, this article will perform an analysis of the technical limitations of such tools, in order to heighten awareness of the risks implicit in any research that bases itself uncritically on their outcomes.

2. Theoretical criticism

2.1 Towards an affective narratology

Examining the taxonomy by [Kim and Klinger 2018b], it becomes immediately evident that the strongest connection between literary theory and SA takes place in the field of narratology. The classification of genres and story types (application 2), in particular, depends primarily on the structure of narratives. [Moretti 2011] has also demonstrated how network analysis (application 4) can prove a powerful approach for the study of plot, while [Zehe et al. 2016] have chosen the narratological device of the “happy ending” to classify their selection of literary texts (application 1). While this does not exclude that SA can also be fruitfully applied to the study of different subjects with different approaches (such as poetry in a neurocognitive perspective, as done by [Papp-Zipernovszky et al. 2021]), the field of narratology and the subject of the novel are clearly the most dominant,^[2] with a distinctive interest towards the classification of literary texts (shared by both applications 1 and 2). It is not by chance, then, that Jockers decided to call his SA tool *Syuzhet* [Jockers 2015a], with a direct reference to the distinction made by Russian formalists between the *fabula* and the plot (also known as *syuzhet*, *récit*, or *intreccio*).

However, when looking for the actual location where this connection takes place, the model reveals itself as apparently faulty. Both [Jockers 2015b] and [Reagan et al. 2016] used SA for demonstrating that the story arcs of the entire (Western) literary production are dominated by six basic shapes.^[3] This conclusion sparked a lively debate — and not a few criticisms (see [Hammond 2017]) — , but it was based on a very problematic assumption. In fact, these basic shapes were obtained by tracking the evolution of positive vs. negative emotions throughout thousands of stories. A precedent for such an approach was found in a lecture by Kurt Vonnegut, based on a (rejected) Master's thesis and circulated in the form of a mock-academic pièce (now available on YouTube). Some connections in narratology might be found, like in the various proposals for a “generative grammar of narrative” [Prince 1973], which already stimulated applications in the field of artificial intelligence — but just for creative purposes [Bringsjord and Ferrucci 2000] — , or in the structuralist assumption that all stories can be reconducted to a universal model [Bremond 1973] — but just when focusing on the *fabula*. Antecedents of the concept of *story arc* can even be found in Gustav Freytag's pyramid of the dramatic action [Freytag 1863] and in Northrop Frye's “U” and an “inverted U” shapes [Frye 1982], representing the

archetypical structures of comedy and tragedy. However, when we look into the most established theorizations of narrative form, such as the still-fundamental proposal by [Genette 1972] or the most recent reformulations by [Bal 2017], almost no trace of emotions can be found. Aspects like focalization, space, time, and characters constitute a story through their manipulation and interaction, while affect is nothing but a device used to engage the reader, with just secondary and indirect consequences on the structure of the narrative. In conclusion, the question becomes unavoidable: is the SA approach actually able to develop a formal model of the phenomenon which, traditionally, has been known by literary scholars as “the plot”?

A series of recent studies, produced in the wake of the already-mentioned *affective turn*, help mitigating a straightforwardly negative answer. Patrick Colm Hogan was the first to introduce the concept of “affective narratology” [Hogan 2011]. His goal was clear and apparently in line with the approaches by Jockers and Reagan: starting from the acknowledgement that “narratological treatments of emotion have on the whole been relatively undeveloped” [Hogan 2011, 1], he aimed at highlighting how “emotion systems define the standard features of all stories, as well as cross-culturally recurring clusters of features in universal genres” [Hogan 2011, 2]. This statement could easily be adopted as an epigraph for many SA literary studies. However, when reaching the core of Hogan's proposal, the correspondences become much more problematic. In particular, Hogan defines a nested system, where works are composed by stories (i.e., by many stories intertwined together), stories by episodes, and episodes by events and incidents, all governed by emotional principles. One of the most important principles in this system is normalcy: in fact, “an episode begins with a shift away from normalcy [and] ends when there is a return to normalcy. [...] If normalcy is restored in a more enduring way, we have not just an episode, but a story” [Hogan 2011, 65]. This poses a fundamental problem for computational modeling, because the basic unit of measurement is not uniform (the pages of a book or the words in a sentence, generally taken as reference points by SA tools), but depends on the development of the story itself (episodes can be closed in a few sentences, or develop through multiple pages). In addition, normalcy is not determined by a simple positive vs. negative emotional status, but by more complex emotion systems (e.g., attachment and sexual desire in the romantic genre), which take shape not in the more general context of the story, but with reference to the goals of single characters. The combination of these elements alone makes the emotional analysis of stories much more complex and sophisticated than a simple tracking of the emotional levels through the pages of a book, involving for example the issues of focalization (as the nature of a sentiment inevitably depends on the chosen perspective), style (if we follow its affective-aesthetic interpretation, cf. [Herrmann et al. 2015, 42]), symbolism, and many others.

It should be noted, however, that research in computational literary studies is currently paving the way for solving most of these problems. For example, the identification of scenes in narratives (and thus of Hogan's episodes) is at the center of the effort by [Gius et al. 2019], who already highlighted its relevance for narratological studies. [Kim et al. 2017] produced their story arcs by distinguishing Plutchik's eight basic emotions (joy, trust, fear, surprise, sadness, anticipation, anger, and disgust) and thus moved beyond the flattening distinction between positive and negative sentiments. Finally, all the studies collected by [Kim and Klinger 2018b] under the “character network analysis” label focus on the emotions of characters. What is currently missing, is a fruitful integration of all these approaches in a coherent framework.

In her recent contribution on the subject, [Elkins 2022] compares the results of more than thirty-six different computational methods in the creation of *emotional arcs* (a term which she prefers to *plot arcs*, as it indicates “an underlying sentiment structure that occurs even when very little happens plot-wise” [Elkins 2022, 5]). While strengthening the approach introduced by Jockers and providing multiple arguments to confirm its usefulness in literary studies, Elkins consciously decides not to face the more complex narratological issues described here, thus still leaving unrealized a true *operationalization* of Hogan's narrative theory.

2.2 From narratology to reader response

Hogan's is not the only proposal for an inclusion of affect into narrative science. In her sophisticated contribution, [Breger 2017] takes a distance from Hogan, by distinguishing affect from emotions. Following the Deleuzian interpretation, affect can be intended as an “(asubjective, asymbolic) ‘intensity’” [Breger 2017, 229], which resists any formalization or reduction to universal features. The interesting aspect of Breger's proposal is that the narratological

8

9

10

11

function of affect is consequently limited to the process of *worldmaking* (the mental creation of fictional worlds), which happens only through an active collaboration between author, text, and readers. While Hogan tried to ground his model uniquely in the inherent features of narratives, excluding — or at least putting aside — the readers, Breger seems to follow a growing tendency in literary studies, which gives a new relevance to readers (be they real or implied).^[4]

Such a tendency is also evident in [Oatley 2012], who devises a series of experiments with his own readers. By intermixing the chapters of an original short story with a taxonomy of emotions in literature, Oatley shows how such emotional systems sustain all narratives. A powerful support is found in the Sanskrit theory of *rasas*, intended as “essences of the everyday emotions [...] that seem not so much individual [...], but universal: aspects of all humankind” [Oatley 2012, 34].

12

In a similar effort, [Pirlet and Wirag 2017] refer to Monika Fludernik's theory of a “natural” narratology [Fludernik 1996], which “foregrounds the reader and focuses on the cognitive mechanisms underlying reader's construction and interpretation of narrative” [Pirlet and Wirag 2017, 48]. Fludernik's *naturalism* depends on the idea that narratives are built in readers' minds through a re-shaping and variation of everyday human experience. By including the affective component, such a narratology may “expand its purview and become even ‘more natural’” [Pirlet and Wirag 2017, 49].

13

These contributions are just a sample of a currently growing trend in literary studies. The widest body of research on reader's affects and emotions, in fact, can be found in the field of reader response theory, whose origins can be traced back to Aristotle's concept of catharsis [Miall 2018], but which distinguished itself more recently for its “scientific” approach to the study of reading. Through the adoption of empirical methods [Peer et al. 2012], in fact, reader's experiences are analyzed and “measured” via questionnaires and interviews, but also using technologies like eye tracking and fMRI scans. In a recent series of papers, SA also has been introduced in the field.

14

[Jacobs et al. 2017] adopted SA on Shakespeare's sonnets not to visualize their (frequently improbable) plot arcs, but to measure the “emotion potential” of their verses, with the goal of predicting possible readers' reactions and thus devising new experiments on selected texts. [Pianzola et al. 2020], then, used SA on a social reading platform [Cordón-García et al. 2013], Wattpad, with a goal that connects even more strongly narratology and reader response theory. Given that on Wattpad readers can write comments for each paragraph in a novel (reaching even millions of comments), Jocker's technique was adapted to compare two emotional arcs: that of the text and that of readers' reactions. Results were used to isolate the passages that showed the highest levels of harmony or discrepancy (i.e., where the correlations between the two emotional arcs were the highest or lowest), thus identifying the textual features that support or hinder narrativization.

15

These studies prove how fruitful the integration between SA tools and literary studies can be. And if, on the one hand, they risk moving too deeply into the realm of social science thus losing contact with literature *stricto sensu*, on the other hand, they confirm how such a tendency is inscribed in the very practice of distant reading [Underwood 2017]. However, extensive work and reasonable carefulness are still required: main risk is that of an oversimplification of phenomena that necessarily escape any reductionism, while the current results of computational analyses are just scraping their surface. In any case, it seems more and more evident that, while a “theory” for distant reading is still lacking [Ciotti 2018], it is precisely through theoretical reasoning that SA (and other computational methods with it) can actually meet the needs of literary scholars. The possibility of testing proposals like those by Hogan and Oatley can prove extremely valuable for the development of literary studies, whatever the result (a confirmation or a denial) will be. And to the skeptics who — sometimes with good reasons — fear the imposition of quantitative methods over the irreducible subjectivity of literary criticism, it can be answered with [McCarty 2005] that the process of modeling in computational literary studies is never the simple reduction of the phenomenon to a formula, but rather a continuous dynamics between the construction of a model and the confrontation with a reality that always escapes it — the same dynamics that, in the end, sustains any theorization about literature.

16

3. Tools criticism

A direct confrontation with literary theory proves fundamental when setting up a criticism of SA tools. However, it is not

17

sufficient when trying to reach a full awareness of the potential and limitations in their use for literary studies. Once the context in which a tool can be employed has been identified, an equal attention should be dedicated to the specific method it adopts. Indeed, each method implies a model — and thus it also implies a theory. SA, in fact, can be performed by selecting or combining an ample variety of approaches, ranging from simple wordcount to the most complex deep learning techniques, and connecting with multiple psycholinguistics theories. Choosing one approach over the other means also defining the very nature of the object under examination.

3.1 A stratified taxonomy

When trying to propose a taxonomy of SA tools, at least three main distinctions should be made, based on three interconnected aspects: (1) the emotion theory adopted by the tool (T); (2) the technique used to build the emotion resources (ER); (3) the method adopted to accomplish the analysis (M).

18

3.1.1. Emotion theories

As for the emotion theory, an ample selection of competing frameworks is currently available, with their advantages/disadvantages and a lively dispute on which one is the best. However, they can be divided into two main families:

19

- T1. Dimensional representations of emotions
- T2. Discrete (or systemic) representations of emotions

Dimensional representations are generally connected to [Russell 1980], who proposed a bi-dimensional system able to chart all emotional states. By combining the two dimensions of valence (positive vs. negative, e.g., “good” vs. “bad”) and arousal (calm vs. intense, e.g., “pleasurable” vs. “exciting”) any human emotion could be logically represented. Many SA tools adopt this theory by simplifying it further, i.e., by reducing it to valence alone, on a continuous scale that ranges between two extremes (e.g., -1 and +1). This solution, chosen in studies such as [Jockers 2014], [Reagan et al. 2016], and [Elkins 2022], offers an efficient simplification for the analysis, but it also implies the loss of relevant information, when for example aesthetic appreciation (e.g., beautifulness or ugliness) needs to be distinguished from embodied response (e.g., pleasure or pain). It should be noted, incidentally, that this interpretation is also at the basis of the very idea of SA. Especially in its commercial applications, SA aims at mining opinions (by stressing this specific meaning of the word *sentiment*), thus a distinction in terms of positive/negative valence becomes sufficient for accomplishing the task.

20

On the other hand, discrete representations multiply the number of dimensions, while at the same time distinguishing them more strictly into a series of discrete categories (or “basic emotions”). Two main theories dominate this context: [Plutchik 1991], who proposed eight basic emotions (joy, trust, fear, surprise, sadness, anticipation, anger, and disgust) based on differences in human behavior; and [Ekman 1993], who reduced the categories to seven (anger, contempt, disgust, fear, joy, surprise, and sadness), based mainly on differences in facial expressions. However, theories are much more numerous [Tracy and Randles 2011], and many SA approaches even combine them or reduce them to a simple dichotomy (all positive vs. all negative emotions). A unique framework is still far from being defined, while the results of any SA analysis depend heavily on the system that is chosen as a reference.

21

The biggest issue in the applicability of SA for literary studies, however, concerns the very possibility of a unique framework. [Sprugnoli et al. 2016] demonstrated how the simple distinction between positive and negative sentiment in historical texts is an almost impossible task for human annotators. By evaluating inter-annotator agreement scores on a series of excerpts from historical political speeches, [Sprugnoli et al. 2016] noted that, if the performance of humans is below the threshold of acceptability, delegating this task to a computer might make no sense at all. This warrants an extreme carefulness when applying SA to literary texts. However, [Rebora 2020] has also demonstrated how, while inter-annotator agreement remains low, SA is still able to catch significant correlations, especially when comparing the emotional valence of a text with its connected readers' responses.

22

3.1.2. Emotion resources

Once the theoretical framework has been decided upon, a second, fundamental decision concerns the choice of the emotion resources. In fact, the measurement of the overall emotion or sentiment expressed by a text depends primarily on the emotional values assigned to the smaller units that compose it, be they words, clauses, or sentences. Based on categorizations such as [Taboada et al. 2011] and [Seyeditabari et al. 2018], three main approaches can be distinguished:

23

- ER1. Word lists
- ER2. Vector space models
- ER3. Labeled texts

The first two approaches pertain to the more general category of emotion dictionaries, where lists of words are associated to a series of (basic) emotions. Overall, emotion dictionaries are still the most used SA resource in DH, even if the interest towards labeled texts increased in recent years.

24

Word lists are the simplest approach, but they also require extensive preparatory work and frequently prove too rigid for an adaptation to different contexts. For example, the NRC Emotion Lexicon — also known as EmoLex [Mohammad and Turney 2013] — was developed through crowdsourcing: by using the Amazon Turk online service, its developers asked users to annotate a series of words both in terms of sentiment and (Plutchik's) basic emotions. Final values were then assigned via majority consent. Issues were many, however, starting from the limited trustworthiness of Amazon Turk annotators (even if the developers devised methods to avoid errors or cheating) and culminating in the system of values unavoidably inscribed in the annotations. In the end, Emolex might prove to be a good representative of the emotions experienced by present-time Internet users (where e.g., the verb “to cry” clearly expresses negative sentiment), but not of the system of values that sustains a play by Shakespeare or a novel by Austen (where the same verb “to cry” can simply mean “to say out loud”).

25

For this reason, vector space models are frequently used to adapt the dictionary to a specific linguistic and cultural context, through distributional semantics [Harris 1954] and the computational technique more generally known as word embeddings [Mikolov et al. 2013]. Based on co-occurrences in selected corpora, words are transformed (or modeled) into multi-dimensional vectors, which encode information on semantic similarity. Starting from a selection of “seed words” (such as “good” vs. “bad”; or words indisputably related to basic emotions), it becomes thus possible to automatically assign a value to all words in a dictionary. This technique offers the advantage of tailoring the dictionary to a specific context, depending on the corpus that is used to generate the vectors. In this case, limitations depend primarily on the technical issues of word embeddings: for example, large corpora are required for their creation (but they are not always available, especially for historical languages) and the information encoded in the vectors does not necessarily model semantic similarity (e.g., it happens that the vectors of words such as “good” and “bad” are similar because the two words tend to appear frequently together).

26

[Taboada et al. 2011] noted how word lists and vector space models can be combined into hybrid emotion dictionaries, which try to reach an ideal compromise between advantages and shortfalls of the two approaches.

27

However, a more general issue seems at stake with emotion dictionaries. As already noted, in fact, some kinds of emotions (such as the Deleuzian affects) escape linguistic formalization, thus they may be undetectable through SA approaches. This is one of the reasons why [Seyeditabari et al. 2018] foster the use of labeled texts, where the basic unit of the emotion resource are not just words, but clauses and sentences. Ideally, in fact, even the least formalizable affects can be identified when focusing on a text span. Main issue for these kinds of emotion resources is their scarce availability for literary studies. While such material is readily available for e.g., product reviews and social media content, where the text of a review is generally accompanied by a simple rating (e.g., a number of “stars”) and a Tweet is supported by hashtags and emoticons (which frequently synthesize and disambiguate the emotions expressed), extensive annotation work is required for literary texts, with the above-mentioned issues of inter-annotator agreement. In any case, research in this field is rapidly moving forward, and annotated corpora such as [Kim and Klinger 2018a] have been recently made available.

28

It is worth noticing how this issue is widely discussed also in literary studies, where the proposals in favor of a

29

detectability of emotions through lexical cues are numerous: in the field of Medieval studies, for example, [Rosenwein 2008] bases her analysis of ancient emotional systems on words alone, while [Rikhardsdottir 2017] looks at the stylistics features that mark such aspects; the theory of *emotives* by [Reddy 2010], then, leans on the concept of translation (intended in its wider meaning, as a process of connection between separated contexts) to build a virtuous cycle between language and feelings. Numerous original solutions have been developed also in the context of computational literary studies. [Cavender et al. 2016] developed their own emotion dictionary to study affect in *Ulysses*: to correctly operationalize Jameson's theory of affect, they used it not to identify the passages dominated by such words, but those where such words did not appear. The dominance in these passages of words that pertain to the body, thus signaled the presence of (unexpressed) affects.

3.1.3. Computational methods

The final distinction in a taxonomy of SA tools pertains to the method adopted to accomplish the analysis. Here too, three main distinctions have been proposed [Liu 2015]:

- M1. Simple (or advanced) wordcounts
- M2. Syntactic structure analyses
- M3. Machine learning techniques

Wordcount is evidently the easiest approach, which ignores sentence structure and word order to accomplish the most basic *bag of words* analysis. Given a text and an emotion dictionary, the words that appear in both are counted and their values summed to generate a final score. Such approach proves quite ineffective when dealing with short sentences or complex rhetorical structures but shows a surprising efficiency when the dimensions of the analyzed text increase. Unfortunately, no research on determining the minimum length for a reliable SA of literary texts — as done by [Eder 2013] for stylometry — exists yet. Simple wordcount can then also rely on statistics to better balance the relevance of single words (as done e.g., by [Reagan et al. 2016]): for example, if a positive word tends to appear homogeneously in multiple texts, its emotional valence might be relatively lower than that of words which appear just in a few passages. Statistics can support wordcount in even more complex ways, but when the analysis aims at fine-grained results, other approaches need to be employed.

A further step is the analysis of the syntactic structure of sentences to extract their overall meaning. This can be performed through different levels of complexity, which range from the simple identification of emotion shifters (e.g., negations, in sentences such as “he was not a bad person”, or “it was neither sad nor boring”), to a full parsing of sentences, which reconstruct their dependency trees (thus distinguishing principal from subordinate clauses, coordinating from adversative conjunctions, and so on). In theory, this approach should prove the best when aiming at high levels of precision. However, it has to deal with the multiple issues and limitations in natural language processing (NLP), especially when applied to historical languages. One of the most widely used NLP algorithms in DH, UDPipe [Straka 2018], still commits a substantial number of errors in parsing languages like Latin and Ancient Greek.

Machine learning (ML) places itself at the highest level of the taxonomy. It has recently established itself as the most effective approach to artificial intelligence, which adopts a bottom-up strategy to build a model of knowledge through a trial-and-error process [Buduma and Locascio 2017], where examples provided by humans (e.g., the recognition of emotions in texts) constitute the basis for a sophisticated imitation game. Some of its most advanced applications in SA (which even imitate the functioning of the human brain through artificial neural networks, also known as *deep learning*) are presented by [Rojas-Barahona 2016], [Yadav and Vishwakarma 2020] and [Pipalia et al. 2020]. Even the most complex issues in SA, such as the identification of irony [Van Hee et al. 2018] and sarcasm [Di Gangi et al. 2019], can be approached through ML. However, also ML has some fundamental limitations when applied to literary studies: main issue is that ML algorithms need human-annotated material to “learn” their tasks, thus they primarily depend on labeled texts (ER3), with all the related issues that were discussed above.

In conclusion, it can be stated that ML has become the dominant approach in SA. In a recent *SemEval* task, a total of 89 teams competed for the best SA approach in analyzing multilingual Tweets. ML approaches were the most common and successful among the participants [Patwa et al. 2020].

3.2 SA tools in literary studies

As noted at the beginning, SA approaches in literary studies are frequently many steps behind the most recent advancements in computational linguistics. This depends on the aforementioned intrinsic issues (e.g., the complexity of literary language and the unavailability of annotated corpora), but also on the tendency to adopt already-developed tools, which do not require the expertise of a computer scientist. While, on the one hand, this is a necessity in DH research (which cannot be reduced to a sub-field of computational linguistics), it may also lead to errors and misinterpretations. More subtly, as all tools bring about their implicit theories and biases, any analysis that looks just at the external outcomes without delving into the inner functioning logics, risks unintentionally supporting ideals that are not its own. This is why a critical analysis of SA tools becomes fundamental, focusing at least on the ones that are more extensively used in DH.

35

3.2.1 Syuzhet

Matthew Jocker's *Syuzhet* is the software that originated the most recent wave of interest towards SA in literary studies. *Syuzhet* is probably one of the least advanced software for SA, but it efficiently combines speed and visualization power to produce effective results. When referring to the taxonomy described above, it can be labeled as:

36

- T1, as its default dictionary simply assigns valence to each word (even if it includes also an implementation of the NRC emotion lexicon)
- ER1, because the default dictionary was built through crowdsourcing (or “wisdom-of-the-crowd”)
- M1, because the analysis is run via simple wordcount

Additional feature in *Syuzhet* is a series of visualization algorithms, which apply multiple smoothing functions to generate elegant *plot arcs*. Generally, *Syuzhet* works as follows: (1) the analyzed text is split into sentences, (2) a series of sentiment values is produced for each sentence, and (3) the raw values are processed by the smoothing functions to generate plots. [Swafford 2015] has already shown the main issues related to passages (2) and (3), such as the inability to detect negation and irony, and the distortions generated by smoothing functions like the Fourier transform. In addition, it should be noted that the speed of the algorithm is also determined by the fact that it does not count words in sentences, so the sentiment of the sentence “He had a good hearth and a good mind” (+1.35) is the same of the sentence “He had a good hearth”. Main advantage of *Syuzhet* is its transparency and adaptability: developed as a package for the R programming language, all its functions and resources are freely available and easily modifiable (but basic programming skills are required). It should be noted, however, that more advanced packages, such as *Rsentiment* and *Sentimentr* (the latter, developed as an expansion of *Syuzhet*) are currently available for SA in R [Naldi 2019]. Among the most recent applications of *Syuzhet* in literary studies, [Rybicki 2018] used it to evaluate the trend of sentiments and emotions across three centuries of the English novel, while [Hu et al. 2021] combined it with multifractal theory to analyze narrative coherence and dynamic evolution of the novel *Never Let Me Go* by Kazuo Ishiguro.

37

3.2.2 Vader

Vader [Hutto and Gilbert 2014] is a slightly more advanced SA tool, as it moves beyond the simple word counts carried out by *Syuzhet*. After its first definition, it has been implemented in multiple programming languages, but one of the most used implementations is that in Python. Here, *Vader* has been integrated into the *nltk* library, which provides multiple NLP functions. In the current taxonomy, it can be classified as:

38

- T1, with a focus on valence alone
- ER1, with a dictionary developed through crowdsourcing
- M2, because it identifies valence shifters, intensifiers, et al.

Vader works on a sentence level and produces a numerical output (composed by four values: “positive”, “neutral”, “negative”, and “compound”, which is a normalized sum of the first three). The identification of valence shifters happens through a series of basic rules, which modify the sentiment values assigned to single words if they are preceded or followed by specific particles. For example, the sentence “He had a good hearth” scores a compound value of +0.44 (on

39

a range between -1 and +1), while “He had a good hearth!” scores +0.49 and “He had not a good hearth” scores -0.34. The rules that generate these values are purely mathematical and were defined through an empirical approach. In a series of experiments with multiple annotators (e.g., asking them to evaluate the sentiment of “good” vs. “good!” and “not good”), numerical modifiers were assigned to single particles (e.g., 0.11 for exclamation marks and -0.69 for negations). Such a procedure determines a good balance between accuracy and computing requirements (the software is fast and quite reliable), but it also causes a relative rigidity of the model. It should be noted, in fact, that *Vader* was developed for the analysis of text produced in social media, thus the values of both emotion dictionary and modifiers were tailored for this specific context. In addition, it shows particular issues when dealing with irony and complex syntactic constructions: the sentence “Well, he was like a potato”, cited by [Swafford 2015] when criticizing *Syuzhet*, deceives *Vader* too, with a compound score of +0.55. In literary studies, *Vader* was recently adopted by [Reed 2018] to explore the poetry of the Black Arts Movement and by [Vani and Antonucci 2019] in a software pipeline aimed at generating visual summaries of narratives. However, the second study shows how better results are obtained by adopting a machine learning approach.

3.2.3 SentiArt

SentiArt [Jacobs 2019] tries to cope with the issue of flexibility by adopting vector space models to generate its dictionary. In theory, *SentiArt* can be adapted to all contexts and languages (if enough training material is available). In the current taxonomy, it can be classified as:

40

- T1, with a focus on both valence and arousal
- ER2, with a dictionary developed through word embeddings
- M1, because it processes texts through simple wordcount

The creation of *SentiArt* emotion dictionaries works as follows: given a list of prototypical positive and negative words, a vector space model is used to expand it, by calculating the distance between these words and the entire dictionary. The model can be generated based on a selection of texts (taken from a specific author, genre, or period), or it can be simply downloaded from a repository of pre-trained models — e.g., *FastText* [Joulin et al. 2017]. The first option is advised because it offers the possibility to tailor the dictionary to a specific context; however, it can become problematic when the training material is not enough to produce reliable vectors. Therefore, in the absence of alternatives and when working on contemporary texts, the second option constitutes a valid alternative. *SentiArt* has been developed in Python language, but it can also be used through the graphical interface of *Orange* (by installing the “text” add-on), which offers an easy access to its functionalities (with a limited set of pre-compiled dictionaries). The main advantage of *SentiArt* is that its hit rate reaches almost 100%, thus all the words in a text are given a value (while, in general, sentiment dictionaries cover just 10-20% of a text). This can constitute an issue when also function words (like articles and conjunctions) get a sentiment score: indeed, also such particles might have an impact in terms in valence/arousal (see for example the preposition “Of” at the beginning of *Paradise Lost*); however, both semantic variance (e.g., multiple word senses) and syntactic functions (e.g., intensification and negation) are lost in the process. As a consequence, the analysis should be performed on larger text sets. [Jacobs 2019] tested *SentiArt* on the *Harry Potter* novels, reaching promising results in predicting the emotion potential of text passages and in identifying the personality profile of characters.

41

3.2.4 SEANCE

SEANCE [Crossley 2017] operationalizes the second theoretical framework, that of discrete representations of emotions. By combining basic NLP functions (such as negation detection) with multiple dictionaries, it offers the opportunity to reach a very high granularity in distinguishing discrete emotions. In the current taxonomy, it can be classified as:

42

- T2, with a total of 250 discrete dimensions
- ER1, as all dictionaries are provided through external resources
- M2, because it combines wordcount and basic syntactic rules

SEANCE is distributed as a multi-platform graphical interface, that can be easily used by non-programmers. Its main potential is in the extensiveness of the vocabulary, which combines multiple resources in a single tool.^[5] It includes the already-cited NRC Emotion Lexicon, but also some of the first-ever SA dictionaries, such as the General Inquirer [Stone and Hunt 1963], the Affective Norms for English Words (ANEW, [Bradley and Lang 1999]), and many others. For each sentence or text (users have to prepare them in a plain text or tabular format) a total of 250 dimensions is measured. However, there is a significant overlap in these dimensions, as many of them encode the same phenomena (such as “joy” or “fear”), measured with different dictionaries. In addition, also multiple non-emotional, abstract concepts (such as “causal”, “legal”, and even “aquatic”) are measured. In literary studies, [Thomson 2017] used it to compare the different editions of Wordsworth’s *Prelude*, tracking the change in emotional aspects and providing a quantitative confirmation to already-established critical interpretations.

3.2.5 Stanford SA

Even if developed before all the tools presented in this survey, *Stanford SA* [Socher et al. 2013] is still one of the most advanced SA software currently available for digital humanists. Its main distinctive feature is the combination of ML and advanced NLP, with the ideal goal of identifying the sentiment of single sentences. In the current taxonomy, it can be classified as:

- T1, with a focus on valence alone
- ER3, as it works with human-annotated texts
- M2 and M3, because it combines parsing and ML

Stanford SA is written in the Java programming language and is part of *Stanford CoreNLP*, one of the most advanced NLP software suites. It can be used through command line (i.e., by typing a series of pre-formulated commands) and tested on a visually-efficient online demo. In simplified terms, *Stanford SA* works as follows: in a first phase, called “training,” the algorithm is given a series of sentences annotated by human raters. Based on these annotations, the algorithm “learns” how to distinguish five possible sentiments: “very negative”, “negative”, “neutral”, “positive”, and “very positive”. In ML terms, the output of this procedure is also called a *model*, intended as a formal representation of the analyzed phenomenon. At this point, the analysis of new sentences begins: for each sentence, (1) a full dependency tree is automatically built; (2) given that also ML algorithms can be structured as trees, the dependency tree is adapted to provide the structure for the ML algorithm; (3) the algorithm analyses the sentence. As evident, the success of the whole process depends heavily on the quality of the training phase. Here the possible issues are many, because annotation demands a significant amount of time and resources, and *Stanford SA* requires a complex annotation format, which focuses not on single words or sentences, but on all nodes in a dependency tree. When taken out-of-the box, *Stanford SA* performed poorly on nineteenth-century English texts, showing errors also in the reconstruction of dependency trees [Rebora 2018, 232]. This depends on the fact that *Stanford SA* default algorithm is trained on contemporary movie reviews, thus it has substantial issues in adapting to different domains. In conclusion, while *Stanford SA* presents itself as one of the most sophisticated SA algorithms, its complexity in usage and requirements in training have kept digital humanists at a distance. In recent times, however, the interest of digital humanists towards ML approaches for SA — starting from isolated studies such as [Zehe et al. 2017] — has increased substantially.

3.2.6 Transformers Pipelines

Transformers Pipelines represent one of the best compromises between simplicity of usage and complexity of the approach. Based on the *Transformers* architecture, made famous by the success of the BERT language model [Devlin et al. 2019], *Pipelines* allow access to advanced ML functionalities through just a few lines of Python code. In the current taxonomy, they can be classified as:

- T1 and T2, with the possibility to switch between different models
- ER3, as models are created with human-annotated texts
- M3, because they adopt advanced ML

In the simplest implementation, through the “text-classification” pipeline, it is possible to calculate the valence of a

sentence (accompanied by a confidence score). However, by selecting one of the many other text classification models available in the *Hugging Face* repository,^[6] SA can also be accomplished in different languages and by applying many different emotion theories. One problematic aspect here is in the trustworthiness of models, which can prove efficient in accomplishing a task but can also bring about multiple biases [Richardson 2022], even with ethical consequences (e.g., when implicitly modeling racist or sexist biases). While *Transformers Pipelines* constitute the easiest entry way for such a computational technique, it should be noted that most projects in DH try to get the best of it by using more sophisticated implementations. In fact, the possibility of “fine tuning” Transformers models via manual annotation stimulates the development of projects that aim at improving them further.^[7] For example, [Konle et al. 2022] fine-tuned Transformers models to recognize basic emotions in German poetry and categorize poems produced in different periods, while [Grisot et al. 2022] used a similar procedure in a project aimed at evaluating the levels of valence and arousal related to geographical entities in Swiss literature. One possible critical aspect of such an approach has been highlighted by [Underwood 2019c], who noted how, when dealing with traditional literary questions, Transformers do not substantially outperform simpler (wordcount-based) approaches, thus overkilling the problem with a hard-to-implement solution. However, the recent availability of high-standard online resources such as the *Colab Notebooks*, together with the development of research questions which require a fine-grained analysis of texts (see e.g., [Lendvai et al. 2020]), has made the adoption of such a solution more and more advisable in DH.

4. Conclusion

This short survey showed how the gap between state-of-the-art tools and current research in computational literary studies, while still present, seems to be gradually closing itself. And while a community-driven effort like the one in computational linguistics (embodied by phenomena such as the *SemEval* tasks) is still largely absent in DH, the recently growing interest (and criticisms) towards methods like SA suggests that it might be a natural outcome of the current evolution. In fact, among the most relevant acquisitions derived from the debate around [Da 2019], is the importance of validation and reproducibility [Piper 2019], i.e., the construction of a community of practice.

48

Still another, more theoretical issue seems to derive from a matter of modeling. When introducing bleeding edge technology in SA (as well as in all DH tools), a simple, direct connection between the phenomenon and its model seems to get lost: as shown by [Jannidis and Flanders 2019, 93] for vector space models and by [Underwood 2019a] for unsupervised ML, any possible theoretical reasoning risks becoming empty or misleading when we do not know anymore the internal logic of the modeling process, or which phenomenon we are actually modeling. This is especially true for advanced ML approaches, which have been frequently criticized for their lack of transparency. SA adds a further complication to this, because of the ineluctable subjectivity that is inscribed in human emotions. A possible solution to this double conundrum can derive from the practice of annotation. In fact, as ML teaches us, the computational analysis (and prediction) of a phenomenon becomes possible only when humans have found an agreement in identifying it. By asking researchers, students, and literature lovers to annotate texts, testing existing theories and letting more general trends emerge, the dream of building a shared, community-driven “hermeneutic machine” [Ciotti 2017a, 11] might not be that impossible to reach.

49

For the moment, nothing advises against an — informed and critically aware — use of the tools that are currently available, starting perhaps from — but not limiting ourselves to — the tools presented here. Limitations are still many, starting from the fact that resources for the English language substantially outnumber those available for all other languages. However, advantages are equally significant, as in the recognition that all the tools presented here are available in the form of free, open-source, and easily modifiable software. And probably still, in the end, literary studies will continue without the need to include SA tools. In that case, no damage can be done. But if the two will find a way to connect more steadily and learn from each other, their evolution could actually become more than a simple development — and could finally be called “progress”.

50

Notes

[1] As for the scientific validation of stylometric methods in DH, see for example the extensive body of research produced by Maciej Eder (e.g., see [Eder 2012] [Eder 2013] [Eder 2017]) or the detailed inquiry by [Evert et al. 2017]. As for theoretical awareness, see for example

[Kestemont 2014] and [Herrmann et al. 2015]. A full bibliography on stylometry can be consulted on Zotero.

[2] See for example the approach chosen by the most recent monograph on the subject, [Elkins 2022].

[3] Note that the number of shapes is the same, but they do not correspond perfectly. All the plots generated by [Reagan et al. 2016] can be explored interactively through an online Hedonometer.

[4] The concept of the *implied reader*, derived from reception theory and intended as “a textual structure anticipating the presence of a recipient without necessarily defining him” [Iser 1978, 34] has been criticized for its excessive abstraction, through which we risk losing contact with real readers [Salgado 2011]. However, Hogan notices how “there are many cases in which we might wish to say that a given reader’s emotional response is misguided [too]” [Hogan 2016]. Thus, it seems that only a combination between the two (abstract modeling and empirical observation) might actually provide us with a reliable description of the phenomenon of reading.

[5] SEANCE was originally conceived as an expansion of LIWC [Tausczik and Pennebaker 2010], a widely-used (but proprietary) software, which measures more than 100 dimensions in multiple languages (but without including any syntactic rule). In literary studies, LIWC has been used by [Piper 2016] to predict the fictionality of texts.

[6] Note that text classification models can accomplish many different tasks (such as named entity recognition, offensive language recognition, etc.). Still, it is significant that the default “text-classification” pipeline performs SA.

[7] A procedure also known as “transfer learning”, as models which already hold a certain knowledge of human language are adapted to accomplish even more specific tasks.

Works Cited

- Bal 2017** Bal, M. *Narratology: Introduction to the Theory of Narrative*. University of Toronto Press, London (2017).
- Bradley and Lang 1999** Bradley, M. M. and Lang, P. J. “Affective Norms for English Words (ANEW): Instruction Manual and Affective Ratings”. *Technical Report C-1*, University of Florida, NIMH Center for Research in Psychophysiology, Gainesville (1999).
- Breger 2017** Breger, C. “Affects in Configuration: A New Approach to Narrative Worldmaking”, *Narrative*, 25.2 (2017): 227-251. <https://doi.org/10.1353/nar.2017.0012>
- Bremond 1973** Bremond, C. *Logique Du Récit*. Seuil, Paris (1973).
- Bringsjord and Ferrucci 2000** Bringsjord, S. and Ferrucci, D. A. 2000. *Artificial Intelligence and Literary Creativity: Inside the Mind of BRUTUS, a Storytelling Machine*. L. Erlbaum Associates, Mahwah, N.J (2000).
- Buduma and Locascio 2017** Buduma, N. and Locascio, N. *Fundamentals of Deep Learning: Designing next-Generation Machine Intelligence Algorithms*. O’Reilly Media. Sebastopol, CA (2017).
- Buurma and Gold 2018** Buurma, R. S. and Gold, M. K. “Contemporary Proposals about Reading in the Digital Age”. In D. H. Richter (ed), *Companion to Literary Theor*. Wiley, Hoboken (2018), pp. 131-150.
- Cavender et al. 2016** Cavender, K., Graham, J. E., Fox, R. P. Jr., Flynn, R., and Cavender, K. “Body Language: Toward an Affective Formalism of Ulysses”. In Ross, S. and O’Sullivan, J. C. (eds), *Reading Modernism with Machines: Digital Humanities and Modernist Literature*. Palgrave Macmillan, London (2016), pp. 223-242.
- Ciotti 2017a** Ciotti, F. “Modelli e Metodi Computazionali per La Critica Letteraria: Lo Stato Dell’arte”. In Alfonzetti, B., Cancro, T., Di Iasio, V., and Pietrobon, E. (eds), *L’Italianistica Oggi*. Adi Editore, Roma (2017), pp. 1-11.
- Ciotti 2017b** Ciotti, F. “What’s in a Topic Model? Critica Teorica Di Un Metodo Computazionale per l’analisi Del Testo”, *Testo e Senso*, 18 (2017): 1-11.
- Ciotti 2018** Ciotti, F. “What Theory for Distant Reading in Literary Studies?” In *EADH2018*. EADH, Galway (2018), pp. 1-3. https://eadh2018.exordo.com/files/papers/91/final_draft/What_Theory_for_Distant_Reading_in_Literary_Studies-abstract.pdf.

- Clough and Halley 2007** Clough, P. T. and Halley, J. O'M. (eds) *The Affective Turn: Theorizing the Social*. Duke University Press, Durham (2007).
- Cordón-García et al. 2013** Cordón-García, J.-A., Alonso-Arévalo, J., Gómez-Díaz, R., and Linder, D. *Social Reading*. Chandos, Oxford (2013).
- Crossley 2017** Crossley, S. A., Kyle, K., and McNamara, D. S. "Sentiment Analysis and Social Cognition Engine (I): An Automatic Tool for Sentiment, Social Cognition, and Social-Order Analysis", *Behavior Research Methods*, 49.3 (2017): 803-21.
- Da 2019** Da, N. Z. "The Computational Case against Computational Literary Studies", *Critical Inquiry*, 45.3 (2019): 601-639. <https://doi.org/10.1086/702594>
- Devlin et al. 2019** Devlin, J., Chang, M.-W., Lee, K. and Toutanova, K. "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding". ArXiv:1810.04805 (2019). <http://arxiv.org/abs/1810.04805>
- Di Gangi et al. 2019** Di Gangi, M. A., Lo Bosco, G., and Pilato, G. "Effectiveness of Data-Driven Induction of Semantic Spaces and Traditional Classifiers for Sarcasm Detection", *Natural Language Engineering*, 25.2 (2019): 257–85. <https://doi.org/10.1017/S1351324919000019>
- Eder 2012** Eder, M. "Mind Your Corpus: Systematic Errors in Authorship Attribution". In *Digital Humanities 2012: Conference Abstracts*, (Hamburg, Germany). Hamburg Univ. Press, Hamburg (2012), pp. 181-185. https://sites.google.com/site/computationalstylistics/preprints/m-eder_mind_your_corpus.pdf?attredirects=0.
- Eder 2013** Eder, M. "Does Size Matter? Authorship Attribution, Small Samples, Big Problem", *Digital Scholarship in the Humanities*, 30.2 (2013): 167-182. <https://doi.org/10.1093/lc/fqt066>
- Eder 2017** Eder, M. "Visualization in Stylometry: Cluster Analysis Using Networks", *Digital Scholarship in the Humanities*, 32.1 (2017): 50-64. <https://doi.org/10.1093/lc/fqv061>
- Ekman 1993** Ekman, P. "Facial Expression and Emotion", *American Psychologist*, 48.4 (1993): 384-392. <https://doi.org/10.1037/0003-066X.48.4.384>
- Elkins 2022** Elkins, K. *The Shapes of Stories: Sentiment Analysis for Narrative*. Cambridge University Press (2022).
- Evert et al. 2017** Evert, S., Proisl, T., Jannidis, F., Reger, I., Pielström, S., Schöch, C., and Vitt, T. "Understanding and Explaining Delta Measures for Authorship Attribution", *Digital Scholarship in the Humanities*, 32.suppl_2 (2017): ii4–ii16. <https://doi.org/10.1093/lc/fqx023>
- Flanders and Jannidis 2019** Flanders, J. and Jannidis, F. (eds) *The Shape of Data in the Digital Humanities: Modeling Texts and Text-Based Resources*. Routledge, Taylor and Francis Group, London; New York (2019).
- Fludernik 1996** Fludernik, M. *Towards a "natural" Narratology*. Routledge, London; New York (1996).
- Freytag 1863** Freytag, G. *Die Technik des Dramas*. Hirzel, Leipzig (1863).
- Frye 1982** Frye, N. *The great code: the Bible and literature*. Routledge, London (1982).
- Genette 1972** Genette, G. *Figures III*. Éditions du Seuil, Paris (1972).
- Gius et al. 2019** Gius, E., Jannidis, F., Krug, M., Zehe, A., Hotho, A., Puppe, F., Krebs, J., Reiter, N., Wiedmer, N., and Konle, L. "Detection of Scenes in Fiction". In *DH2019 Book of Abstracts*. ADHO, Utrecht (2019). <https://dev.clariah.nl/files/dh2019/boa/0608.html>.
- Grisot et al. 2022** Grisot, G., Reborá, S., and Herrmann, J. B. "Sentiment lexicons or BERT? A comparison of sentiment analysis approaches and their performance". *DH 2022 Conference Abstracts*. ADHO, Tokyo (2022), pp. 469-470
- Hammond 2017** Hammond, A. "The Double Bind of Validation: Distant Reading and the Digital Humanities' 'Trough of Disillusionment'", *Literature Compass*, 14.8 (2017): e12402.
- Harris 1954** Harris, Z. S. 1954. "Distributional Structure", *WORD*, 10.2-3 (1954): 146-162. <https://doi.org/10.1080/00437956.1954.11659520>
- Herrmann et al. 2015** Herrmann, J. B., Schöch, C., and van Dalen-Oskam, K. "Revisiting Style, a Key Concept in Literary Studies", *Journal of Literary Theory*, 9.1 (2015): 25-52.
- Hogan 2011** Hogan, P. C. *Affective Narratology: The Emotional Structure of Stories*. Bison, Lincoln (2011).

- Hogan 2016** Hogan, P. C. "Affect Studies and Literary Criticism". In *Oxford Research Encyclopedia of Literature* (2016). <https://doi.org/10.1093/acrefore/9780190201098.013.105>
- Hu et al. 2021** Hu, Q., Liu, B., Thomsen, M. R., Gao, J., Nielbo, K. L. "Dynamic evolution of sentiments in Never Let Me Go: Insights from multifractal theory and its implications for literary analysis". *Digital Scholarship in the Humanities*, 36.2 (2021): 322-332. <https://doi.org/10.1093/llc/fqz092>
- Hutto and Gilbert 2014** Hutto, C. J. and Gilbert, E. "Vader: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text". In *Eighth International AAAI Conference on Weblogs and Social Media* (2014). <https://www.aaai.org/ocs/index.php/ICWSM/ICWSM14/paper/download/8109/8122>.
- Iser 1978** Iser, W. *The Act of Reading: A Theory of Aesthetic Response*. Johns Hopkins University Press, Baltimore (1978).
- Jacobs 2019** Jacobs, A. M. "Sentiment Analysis for Words and Fiction Characters From the Perspective of Computational (Neuro-)Poetics", *Frontiers in Robotics and AI*, 6 (2019). <https://doi.org/10.3389/frobt.2019.00053>
- Jacobs et al. 2017** Jacobs, A. M., Schuster, S., Xue, S., and Lüdtkke, J. "'What's in the Brain That Ink May Character...': A quantitative narrative analysis of Shakespeare's 154 sonnets for use in (Neuro-)cognitive poetics", *Scientific Study of Literature*, 7.1 (2017): 4-51.
- Jacobs et al. 2020** Jacobs, A. M., Herrmann, J. B., Lauer, G., Lüdtkke, J. and Schroeder, S. "Sentiment Analysis of Children and Youth Literature: Is There a Pollyanna Effect?" *Frontiers in Psychology* 11 (2020): 574746. <https://doi.org/10.3389/fpsyg.2020.574746>
- Jannidis and Flanders 2019** Jannidis, F., and Flanders, J. "A Gentle Introduction to Data Modeling". In Jannidis, F., and Flanders, J. (eds), *The Shape of Data in the Digital Humanities: Modeling Texts and Text-Based Resources*. Routledge, Taylor and Francis Group, London; New York (2019), pp. 26-95.
- Jockers 2014** Jockers, M. "A Novel Method for Detecting Plot". (2014). <http://www.matthewjockers.net/2014/06/05/a-novel-method-for-detecting-plot/>.
- Jockers 2015a** Jockers, M. "Revealing Sentiment and Plot Arcs with the Syuzhet Package". (2015). <http://www.matthewjockers.net/2015/02/02/syuzhet/>.
- Jockers 2015b** Jockers, M. "The Rest of the Story". (2015). <http://www.matthewjockers.net/2015/02/25/the-rest-of-the-story/>.
- Joulin et al. 2017** Joulin, A., Grave, E., Bojanowski, P., and Mikolov, T. "Bag of Tricks for Efficient Text Classification". In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers*. Association for Computational Linguistics (2017), pp. 427-431.
- Keen 2011** Keen, S. 2011. "Introduction: Narrative and the Emotions", *Poetics Today*, 32.1 (2011): 1-53. <https://doi.org/10.1215/03335372-1188176>
- Kestemont 2014** Kestemont, M. "Function Words in Authorship Attribution. From Black Magic to Theory?" In *Proceedings of the 3rd Workshop on Computational Linguistics for Literature (CLFL)*. Association for Computational Linguistics, Gothenburg, Sweden (2014), pp. 59-66. <http://aclweb.org/anthology/W/W14/W14-0908.pdf>.
- Kim and Klinger 2018a** Kim, E. and Klinger, R. "Who Feels What and Why? Annotation of a Literature Corpus with Semantic Roles of Emotions". In *Proceedings of the 27th International Conference on Computational Linguistics*. Association for Computational Linguistics, Santa Fe, New Mexico, USA (2018), pp. 1345-1359. <http://aclweb.org/anthology/C18-1114>.
- Kim and Klinger 2018b** Kim, E. and Klinger, R. "A Survey on Sentiment and Emotion Analysis for Computational Literary Studies". ArXiv:1808.03137 (2018). <http://arxiv.org/abs/1808.03137v1>.
- Kim et al. 2017** Kim, E., Padó, S., and Klinger, R. "Investigating the Relationship between Literary Genres and Emotional Plot Development". In *Proceedings of the Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature*. Association for Computational Linguistics, Vancouver, Canada (2017), pp. 17-26. <https://doi.org/10.18653/v1/W17-2203>
- Konle et al. 2022** Konle, L., Kröncke, M., Jannidis, F. and Winko, S. "Emotions and Literary Periods". *DH 2022 Conference Abstracts*. ADHO, Tokyo (2022), pp. 278-281
- Lendvai et al. 2020** Lendvai, P., Darányi, S., Geng, C., Kuijpers, M., Lopez de Lacalle, O., Mensonides, J.-C., Rebora, S. and Reichel, U. (2020). "Detection of Reading Absorption in User-Generated Book Reviews: Resources Creation and Evaluation". In *Proceedings of The 12th Language Resources and Evaluation Conference*. European Language Resources Association, Marseille (2020), pp. 4835-4841. <https://www.aclweb.org/anthology/2020.lrec-1.595>

- Liu 2015** Liu, B. *Sentiment Analysis: Mining Opinions, Sentiments, and Emotions*. Cambridge University Press, New York (2015).
- McCarty 2005** McCarty, W. *Humanities Computing*. Palgrave Macmillan, New York (2005).
- Miall 2018** Miall, D. S. "Reader-Response Theory". In Richter, D. H. (ed), *A Companion to Literary Theory*. John Wiley and Sons, Chichester, UK (2018), pp. 114-125. <https://doi.org/10.1002/9781118958933.ch9>
- Mikolov et al. 2013** Mikolov, T., Sutskever, I., Chen, K., Corrado, G., and Dean, J. "Distributed Representations of Words and Phrases and Their Compositionality". ArXiv:1310.4546 (2013). <http://arxiv.org/abs/1310.4546>.
- Mohammad and Turney 2013** Mohammad, S. and Turney, P. D. "Crowdsourcing a Word-emotion Association Lexicon", *Computational Intelligence*, 29.3 (2013): 436-465. <https://doi.org/10.1111/j.1467-8640.2012.00460.x>
- Moretti 2011** Moretti, F. "Network Theory, Plot Analysis", *New Left Review*, 68 (2011). <http://newleftreview.org/II/68/francomorettnetworktheoryplotanalysis>.
- Moretti 2013** Moretti, F. "'Operationalizing: Or, the Function of Measurement in Modern Literary Theory.'" *Pamphlet of the Stanford Literary Lab* (2013), pp. 1-15. <https://litlab.stanford.edu/LiteraryLabPamphlet6.pdf>.
- Naldi 2019** Naldi, M. "A Review of Sentiment Computation Methods with R Packages". ArXiv:1901.08319 (2019). <http://arxiv.org/abs/1901.08319>.
- Oatley 2012** Oatley, K. *The Passionate Muse: Exploring Emotion in Stories*. Oxford University Press, New York (2012).
- Papp-Zipernovszky et al. 2021** Papp-Zipernovszky, O., Mangen, A., Jacobs, A. M. and Lüdtke, J. "Shakespeare Sonnet Reading: An Empirical Study of Emotional Responses". *Language and Literature: International Journal of Stylistics* (2021): 096394702110546. <https://doi.org/10.1177/0963947021105467>
- Patwa et al. 2020** Patwa, P., Aguilar, G., Kar, S., Pandey, S., PYKL, S., Gambäck, B., Chakraborty, T., Solorio, T. and Das, A. "SemEval-2020 Task 9: Overview of Sentiment Analysis of Code-Mixed Tweets". In *Proceedings of the Fourteenth Workshop on Semantic Evaluation*. International Committee for Computational Linguistics, Barcelona (2020), pp. 774–790. <https://doi.org/10.18653/v1/2020.emeval-1.100>
- Peer et al. 2012** Peer, W. v., Hakemulder, J., and Zyngier, S. *Scientific Methods for the Humanities*. John Benjamins, Amsterdam; Philadelphia (2012).
- Pianzola et al. 2020** Pianzola, F., Reborá, S., and Lauer, G. "Wattpad as a Resource for Literary Studies in the 21st Century. Quantitative and Qualitative Examples of the Importance of Digital Social Reading and Readers' Comments in the Margins", *PLoS ONE*, 15.1 (2020): e0226708. <https://doi.org/10.1371/journal.pone.0226708>
- Pipalia et al. 2020** Pipalia, K., Bhadja, R. and Shukla, M. "Comparative Analysis of Different Transformer Based Architectures Used in Sentiment Analysis". In *Proceedings of the 9th International Conference System Modeling and Advancement in Research Trends (SMART)*. IEEE, Moradabad (2020), pp. 411–415. <https://doi.org/10.1109/SMART50582.2020.9337081>
- Piper 2016** Piper, A. "Fictionality", *Journal of Cultural Analytics*, (2016). <https://doi.org/10.22148/16.011>
- Piper 2018** Piper, A. *Enumerations: Data and Literary Study*. The University of Chicago Press, Chicago; London (2018).
- Piper 2019** Piper, A. "Do We Know What We Are Doing?" *Journal of Cultural Analytics*, (2019). <https://culturalanalytics.org/2019/04/do-we-know-what-we-are-doing/>.
- Pirlet and Wirag 2017** Pirlet, C. and Wirag, A. "Towards a 'Natural' Bond of Cognitive and Affective Narratology". In Burke, M. and Troscianko, E. T. (eds), *Cognitive Literary Science*. Oxford University Press, Oxford (2017), pp. 35–54. <https://doi.org/10.1093/acprof:oso/9780190496869.003.0003>
- Plutchik 1991** Plutchik, R. *The Emotions*. University Press of America, Lanham, Md (1991).
- Prince 1973** Prince, G. J. *A Grammar of Stories: An Introduction*. Mouton, The Hague; Paris (1973).
- Reagan et al. 2016** Reagan, A. J., Mitchell, L., Kiley, D., Danforth, C. M., and Dodds, P. S. "The Emotional Arcs of Stories Are Dominated by Six Basic Shapes", *EPJ Data Science*, 5.1 (2016): 31.
- Reborá 2018** Reborá, S. *History/Histoire e Digital Humanities. La Nascita Della Storiografia Letteraria Italiana Fuori d'Italia*. Firenze University Press, Firenze (2018). <http://www.fupress.com/catalogo/history-histoire-e-digital-humanities/3748>.
- Reborá 2020** Reborá, S. "Shared Emotions in Reading Pirandello. An Experiment with Sentiment Analysis". In Marras, C.,

Passarotti, M., Franzini, G., and Litta, E. (eds), *Atti del IX Convegno Annuale AIUCD. La svolta inevitabile: sfide e prospettive per l'Informatica Umanistica*. Università Cattolica del Sacro Cuore, Milano (2020), pp. 216-221. <http://doi.org/10.6092/unibo/amsacta/6316>

- Reddy 2010** Reddy, W. M. *The Navigation of Feeling: A Framework for the History of Emotions*. Cambridge Univ. Press, Cambridge (2010).
- Reed 2018** Reed, E. "Measured Unrest In The Poetry Of The Black Arts Movement". In *DH2018 Book of Abstracts* (2018). <https://dh2018.adho.org/measured-unrest-in-the-poetry-of-the-black-arts-movement/>.
- Richardson 2022** Richardson, S. "Exposing the Many Biases in Machine Learning." *Business Information Review* (2022), 02663821221121024. <https://doi.org/10.1177/02663821221121024>
- Rikhardsdottir 2017** Rikhardsdottir, S. *Emotion in Old Norse Literature: Translations, Voices, Contexts*. D. S. Brewer, Cambridge (2017).
- Rojas-Barahona 2016** Rojas-Barahona, L. M. "Deep Learning for Sentiment Analysis", *Language and Linguistics Compass*, 10.12 (2016): 701-719. <https://doi.org/10.1111/lnc3.12228>
- Rosenwein 2008** Rosenwein, B. H. "Emotion Words." In Nagy, P. and Bouquet, D. (eds), *Le Sujet Des Émotions Au Moyen Âge*. Beauchesne, Paris (2008), pp. 93-106.
- Russell 1980** Russell, J. A. "A Circumplex Model of Affect", *Journal of Personality and Social Psychology*, 39.6 (1980): 1161-1178. <https://doi.org/10.1037/h0077714>
- Rybicki 2018** Rybicki, J. "Sentiment Analysis Across Three Centuries of the English Novel: Towards Negative or Positive Emotions?" In *EADH2018* (2018). <https://eadh2018.exordo.com/programme/presentation/11>.
- Salgaro 2011** Salgaro, M. "La lettura come 'Lezione della base cranica' (Durs Grünbein). Prospettive per l'estetica della ricezione", *Bollettino Dell'associazione Italiana Di Germanistica*, 4 (2011): 49-62.
- Salgaro 2018** Salgaro, M. "The Digital Humanities as a Toolkit for Literary Theory: Three Case Studies of the Operationalization of the Concepts of 'Late Style,' 'Authorship Attribution,' and 'Literary Movement'", *Iperstoria*, 12 (2018): 50-60. http://www.iperstoria.it/joomla/images/PDF/Numero_12/Salgaro_pdf.pdf.
- Seyeditabari et al. 2018** Seyeditabari, A., Tabari, N., and Zadrozny, W. "Emotion Detection in Text: A Review". ArXiv:1806.00674 (2018). <http://arxiv.org/abs/1806.00674>.
- Socher et al. 2013** Socher, R., Perelygin, A., Wu, J. Y., Chuang, J., Manning, C. D., Ng, A. Y., and Potts, C. "Recursive Deep Models for Semantic Compositionality Over a Sentiment Treebank". In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, Seattle (2013), pp. 1631-1642.
- Sprugnoli et al. 2016** Sprugnoli, R., Tonelli, S., Marchetti, A., and Moretti, G. "Towards Sentiment Analysis for Historical Texts", *Digital Scholarship in the Humanities*, 31.4 (2016): 762-772. <https://doi.org/10.1093/lc/fqw027>
- Sprugnoli et al. 2020** Sprugnoli, R., Passarotti, M., Corbetta, D. and Peverelli, A. "Odi et Amo. Creating, Evaluating and Extending Sentiment Lexicons for Latin". In *Proceedings of the 12th Language Resources and Evaluation Conference*. ACM, New York, (2020), pp. 3078-3086. <https://aclanthology.org/2020.lrec-1.376>
- Stone and Hunt 1963** Stone, P. J. and Hunt, E. B. "A Computer Approach to Content Analysis: Studies Using the General Inquirer System". In *Proceedings of the May 21-23, 1963, Spring Joint Computer Conference*. ACM, New York (1963), pp. 241-256. <https://doi.org/10.1145/1461551.1461583>
- Straka 2018** Straka, M. "UDPipe 2.0 Prototype at CoNLL 2018 UD Shared Task". In *Proceedings of the CoNLL 2018 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies*. Association for Computational Linguistics, Brussels (2018), pp. 197-207.
- Swafford 2015** Swafford, A. "Problems with the Syuzhet Package". In *Anglophile in Academia: Annie Swafford's Blog* (2015). <https://annieswafford.wordpress.com/2015/03/02/syuzhet/>.
- Taboada et al. 2011** Taboada, M., Brooke, J., Tofiloski, M., Voll, K., and Stede, M. "Lexicon-Based Methods for Sentiment Analysis", *Computational Linguistics* 37.2 (2011): 267-307. https://doi.org/10.1162/COLI_a_00049
- Tausczik and Pennebaker 2010** Tausczik, Y. R. and Pennebaker, J. W. "The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods", *Journal of Language and Social Psychology*, 29.1 (2010): 24-54. <https://doi.org/10.1177/0261927X09351676>

Thomson 2017 Thomson, D. E. *Prelude as Lifespan Gauge*, *Scientific Study of Literature*, 7.2 (2017): 232-256.

Tracy and Randles 2011 Tracy, J. L. and Randles, D. "Four Models of Basic Emotions: A Review of Ekman and Cordaro, Izard, Levenson, and Panksepp and Watt", *Emotion Review*, 3.4 (2011): 397-405.
<https://doi.org/10.1177/1754073911410747>

Underwood 2017 Underwood, T. "A Genealogy of Distant Reading", *DHQ: Digital Humanities Quarterly*, 11.2 (2017).
<http://www.digitalhumanities.org/dhq/vol/11/2/000317/000317.html>.

Underwood 2019a Underwood, T. "Algorithmic Modeling. Or, Modeling Data We Do Not Yet Understand". In Flanders, J. and Jannidis, F. (eds), *The Shape of Data in the Digital Humanities: Modeling Texts and Text-Based Resources*. Routledge, Taylor and Francis Group, London; New York (2019), pp. 250-263.

Underwood 2019b Underwood, T. *Distant Horizons: Digital Evidence and Literary Change*. The University of Chicago Press, Chicago (2019).

Underwood 2019c Underwood, T. "Do humanists need BERT? Neural models have set a new standard for language understanding. Can they also help us reason about history?" (2019). <https://tedunderwood.com/2019/07/15/do-humanists-need-bert/>

Van Hee et al. 2018 Van Hee, C., Lefever, E., and Hoste, V. "Exploring the Fine-Grained Analysis and Automatic Detection of Irony on Twitter", *Language Resources and Evaluation*, 52.3 (2018): 707-731. <https://doi.org/10.1007/s10579-018-9414-2>

Vani and Antonucci 2019 Vani, K. and Antonucci, A. "NOVEL2GRAPH: Visual Summaries of Narrative Text Enhanced by Machine Learning". In *Text2Story@ ECIR* (2019), pp. 29-37.

Yadav and Vishwakarma 2020 Yadav, A. and Vishwakarma, D. K. "Sentiment analysis using deep learning architectures: A review". *Artificial Intelligence Review*, 53.6 (2020): 4335–4385. <https://doi.org/10.1007/s10462-019-09794-5>

Zehe et al. 2016 Zehe, A., Becker, M., Hettinger, L., Hotho, A., Reger, I., and Jannidis, F. "Prediction of Happy Endings in German Novels Based on Sentiment Information". In *Proceedings of the Workshop on Interactions between Data Mining and Natural Language Processing (2016)*, pp. 9-16.

Zehe et al. 2017 Zehe, A., Becker, M., Jannidis, F., and Hotho, A. "Towards Sentiment Analysis on German Literature". In Kern-Isberner, G., Fürnkranz, J., and Thimm M. (eds), *KI 2017: Advances in Artificial Intelligence*. Springer International Publishing, Cham (2017), pp. 387-394. https://doi.org/10.1007/978-3-319-67190-1_36



This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.