Sentiment Analysis

Rachele Sprugnoli

https://orcid.org/0000-0001-6861-5595

Abstract This chapter presents an overview of Sentiment Analysis with a focus on how it is typically applied in the Digital Humanities field. More specifically, we discuss linguistic issues, such as irony and the use of emoji, that make sentiment analysis challenging and we provide a brief description of several tasks and sub-tasks, all related to subjective texts but seen from different angles: i.e., subjectivity classification, document- and sentence-level polarity classification, aspect-based sentiment analysis, stance detection, irony detection and emotion analysis. In addition, we introduce lexicon-based and machine learning approaches to sentiment analysis. Open issues and best practices for the application of sentiment analysis methods in Digital Humanities are also discussed and the chapter closes with a list of emergent trends in the field.

Keywords Sentiment Analysis, Opinion Mining, Emotion Analysis

1. Introduction

Sentiment Analysis (SA) is a field of research, within the area of Natural Language Processing (NLP),[1] that aim to identify and classify opinions, feelings, personal evaluations towards entities (e.g., people, places, products), events, topics as expressed in written texts (Liu 2022). In its simplest form, SA distinguishes texts according to their polarity (or sentiment orientation): "I love detective stories" has a positive polarity, "I don't like romance books" has a negative polarity and "Agatha Christie was an English writer" has a neutral polarity.[2]

There are many alternative expressions used in the literature to refer to this multifaceted problem: we find, among others, opinion mining, opinion extraction, sentiment mining, affect analysis, polarity detection. In this context the words *sentiment* and *opinion* are often considered synonyms; although they are not, their distinction is very subtle, and they are closely connected. The sentence "I get bored reading

1 A distinction is traditionally made between Computational Linguistics, seen as a branch of linguistics, and Natural Language Processing, seen as a branch of engineering or computer science (Bender 2016). However, in this chapter, we will adopt an integrated view of these two areas since both have the goal of carrying out linguistic analysis and use linguistic data as input.

2 Unless otherwise specified, the examples in this chapter were created by the author.

Rachele Sprugnoli: Sentiment Analysis, in: Christopher A. Nunn and Frederike van Oorschot (eds.): Compendium of Computational Theology 1. Introducing Digital Humanities to Theology, Heidelberg: heiBOOKS, 2024, S. 267–282. DOI: https://doi.org/10.11588/heibooks.1521.c21953

romantic books" expresses a negative sentiment prompted by the feeling of boredom, whereas "I think romantic books all have the same plot" expresses a negative opinion; this example shows that a negative sentiment implies a negative opinion, and a negative opinion is due to a negative sentiment. SA also includes other areas of research and applications that require more granular distinctions which will be addressed in a specific section; for example, aspect-based SA identifies the sentiment of specific attributes or components of an entity.

The growth of interest in SA goes hand in hand with the increasing diffusion of online reviews, forums, microblogs, and social networks which produce an enormous volume of subjective texts, in which users express their opinions and evaluations. SA is also considered a valid tool in the corporate, communicative, and social science fields: in fact, there are many applications which monitor the opinion of customers towards a service or product, or which study the attitude of users on social networks. There are also works that adapt methods and techniques of SA to the humanities with applications to historical, literary, or classical language texts.

In this chapter we introduce the basic definitions and concepts related to SA research with the aim of making the reader aware of the challenges related to SA, especially in the field of Digital Humanities (DH).

2. Why Sentiment Analysis is Challenging – Some Linguistic Peculiarities

The examples given in the previous section are extremely simple from a linguistic point of view but the language we use to express our subjective evaluations is complex, made up of many components that make SA an interesting challenge both for humans and computers.

First of all, the same word in different contexts can have different meanings that encode different sentiments. For example, the adjective *sharp* can be associated with a negative sentiment when it means "keenly and painfully felt" but has a positive sentiment when it means "having or demonstrating ability to recognize or draw fine distinctions."[3]

Furthermore, opinions are not always expressed explicitly and directly but often have an implicit or comparative form. Implicit opinions are those referring to facts or effects related to the object of the opinion: for example, the sentence "this book just makes me yawn" describes a side effect of reading a boring book. Comparative opinions, on the other hand, juxtapose different elements based on the same aspect as in

3 Definitions taken from WordNet 3.1: http://wordnetweb.princeton.edu/perl/webwn (Accessed: 23 June 2024).

"I think Agatha Christie's novels have a more linear plot than those of Arthur Conan Doyle:" it is interesting to note that understanding the sentiment of this last example is difficult because it depends on the reader's personal taste.

Implicit opinions often require world (extra-linguistic) knowledge to be correctly interpreted. The sentence "She looks like a Botticelli madonna!" expresses a positive sentiment by referring to the harmony and beauty of the faces painted by the Renaissance artist; on the contrary, "He looks like a Picasso painting!" makes a reference to the unstructured faces of cubism and therefore to a face with disproportionate features.

World knowledge is also needed to discriminate literal from ironic content. Irony is a type of figurative language that is intentionally used to give a sentence an opposite meaning to the literal one. As defined by Utsumi (2000) in his *Implicit Display Theory,* verbal irony is an utterance or a statement that implicitly displays an ironic environment in which the speaker has a negative emotional attitude toward the incongruity between what he/she expects and what actually is. The term irony is often used as a hypernym for sarcasm (Grice 1975) that indeed is a particular form of irony used to mock or insult in a scornful or caustic way. Both irony and sarcasm are particularly interesting in SA because they are sentiment shifters, i.e., they change the polarity: a sentence like "the wifi connection is great – it's fast as a sloth" means the exact opposite of what it seems (the wifi connection is slow) but its apparent linguistic form would lead to assign it a positive polarity.

Another issue to consider is the presence of emoticons and emojis that play an important role when dealing with informal texts such as posts on social networks and forums. Comparing "Rome :)" to "Rome :(," the opinion on the city is expressed by the emoticon; without it, the name Rome alone would have no polarity. In other words, these elements enhance the expressiveness of a text and convey their own specific sentiment even if not always easily identifiable. For example, the fire emoji is mostly used with the meaning of excellent or attractive (therefore with a positive sentiment), but can also signal anger (thus a negative sentiment) or a fact, such as the presence of fires or excessive heat (having, in this case, a neutral sentiment).

3. One Name, Many Tasks

As already stated by Liu in 2010, SA is a multifaceted problem: it is not a single monolithic linguistic task, it does not have a single solution but can be tackled by considering various levels of analysis.

The first level is addressed by the task called subjectivity classification which aims to distinguish objective texts, containing factual information, from subjective, opinionated texts that express feelings, points of view or personal beliefs. This is the first step towards more in-depth analyses: in fact, in objective texts it is not possible to identify a polarity (they are neutral) while subjective texts can be classified according to their sentiment orientation.

Polarity classification is the next step and consists of assigning to an information unit a value that indicates whether it expresses a positive, negative, or neutral sentiment. This value can be categorical or numeric and the range of possible values can vary considerably depending on the degree of detail we want to achieve. For example, there are binary classifications (with only two values, such as *positive* and *negative*), 3-value classifications (e.g., *positive, neutral, negative* or *+1, 0, -1*), 5-value classifications (e.g., *very positive, somewhat positive, neutral, somewhat negative, very negative* or *+1, +0.5, 0, -0.5, -1*) but also decimal scores in a continuous range (typically between *+1* and *-1*).

Polarity classification can be performed at different granularities, i.e., taking into consideration different types of information units: the whole document, a single sentence at a time, or one specific aspect. Document level SA assigns a polarity score to an entire document (e.g., a book review) by assuming such document as a single information unit expressing the opinion of a single person (the author of the review) on a single entity (a book). The same type of classification can be applied at sentence level. Sentence-level SA is useful because the same document can contain different or even opposite opinions in different sentences. For example, a book review may be made up of neutral sentences, describing the plot without making personal judgments, together with other sentences expressing appreciation or disapproval. An even more granular level of analysis is provided by the entity-based or aspect-based SA,[4] which has the purpose of extracting the opinions expressed on individual entities or on entities' features. In the case of the aforementioned review, the book is the entity object of the evaluation while two relevant features can be the plot and the price; the sentiment can be different for each of these elements, for example it can be positive for the book itself and for the plot but negative for the price as in "I enjoyed reading the book because the storyline is compelling, but the price is too high: not everyone can spend 25 euros on a book!" Therefore, the task has two main phases: the extraction of entities and/or features and then the classification of the sentiment for each of them. It is important to note that the relevant features are entity type specific: if price is important to any commercial product or service, plot is specific to books and movies. A mobile phone, on the other hand, can have battery life and ease of configuration as features to identify, while for hotels the location is particularly important.

4 Aspect-based SA is also known as feature-based SA.

4. Other Related Tasks

In this section, we provide an overview of other tasks that are considered sub-problems of SA, all related to subjective texts but seen from different angles.

- *Stance detection* is the task that determines whether the author of a text is in favor or against an entity, event, or topic (AlDayel & Magdy 2021). From a linguistic point of view, stance is an overt expression used to evaluate a certain target element and position oneself with respect to the others by displaying alignment or opposition (Du Bois 2007). For this reason, stance detection requires a given target to measure the author's viewpoint toward it and the output of the classification is one out of the three labels *Favor, Against, Neither,* instead of *Positive, Negative* or *Neutral* as in the simplest case of polarity classification. Stance and polarity are independent of each other: a positive sentiment does not necessarily lead to a supporting stance, just as a negative sentiment is not necessarily associated with an opposing stance. For example, taking the statement "climate change is a real concern" as target, the sentence "It's so sad that too many people don't plan to do anything while our planet is burning!" expresses a negative sentiment but a supportive stance towards the statement. This task is mostly applied to political and social issues to intercept the position of social network users regarding a political figure or proposals considered divisive, such as drug liberalization and same-sex marriage.
- *Irony detection and sarcasm detection* tasks aim to distinguish between ironic or sarcastic and non-ironic or non-sarcastic texts (Maynard & Greenwood 2014). While irony is usually uncritical, sarcasm is more aggressive; however, both these figurative devices create a mismatch between the literal and the intentional meaning of a text. Sometimes a binary classification is made without differentiating between irony and sarcasm, while in other cases a more detailed classification is attempted by recognizing various types of irony, for example, by distinguishing it from sarcasm, satire or parody (Abu Farha et al. 2022).
- *Emotion analysis* consists of determining which emotions are conveyed in a text. The scientific study of emotions has interested psychologists and anthropologists since the publication of Darwin's seminal work *The Expression of the Emotions in Man and Animals* in 1872. Although the theories are numerous, there are two main approaches on which computational techniques are based. According to the first approach, emotions are innate, universal across different cultures and limited in number thus they can be classified using categorical labels. In NLP, these labels are often borrowed from the theories defined by Ekman (1993) or Plutchik (1980). Ekman identifies six emotions (i.e., anger, disgust, sadness, joy, fear, surprise) while Plutchik defines four

pairs of basic emotions (i.e., joy versus sadness, anger versus fear, trust versus disgust, surprise versus anticipation) which combine with each other to form dyads, i.e., complex emotions (e.g., love is a combination of joy and trust). On the other hand, following the second approach, emotions cannot be labeled but represented according to different dimensions using continuous values. In the circumplex model (Russell 1980) there are two fundamental dimensions of emotional experience: valence, i.e., the level of pleasantness, and arousal, i.e., the intensity of the emotion. A third dimension called dominance is often added to the previous two to encode the degree of control the emotion exerts over the person experiencing it. For example, according to this approach known with the acronym VAD *(Valence-Arousal-Dominance),* anger has low valence, high arousal, and high dominance.

5. Methods

From the point of view of the development of SA systems, two main approaches can be distinguished: those based on lexicons, and those using machine learning algorithms, both supervised and unsupervised.

Lexicon-based methods rely on the intuition that the polarity of a text can be obtained on the basis of the polarity of the words that compose it (Taboada 2011). Such polarity is obtained from lexicons made up of lists of tokens, lemmas, or phrases in which each lexical entry is associated with a categorical or numerical value (e.g., *Positive* or *+1*) quantifying its sentiment orientation. Polarity lexicons are available for numerous languages (Mohammed & Balakrishnan 2020): some have been created manually, employing experts (e.g., linguists or psychologists) or crowdsourcing techniques (Mohammad & Turney 2013),[5] but the development of these resources is very time-consuming thus automatic approaches have also been tested for example by exploiting machine translation or available lexicographical resources and corpora. Lexicons typically record the prior polarity of words, i.e., the sentiment they evoke beyond their context of use. Thus, words like *friendship* and *love* are associated with a positive polarity while *murder* and *hate* with a negative one. Rarer are the lexicons that contain sense-based polarities, the best known is *SentiWordNet* (Baccianella et al. 2010) in which each *WordNet synset* (Miller 1995) has a positive, a negative and an objective score. Based on these lexicons, scripts are created which calculate the ratio between positive and negative words within the text to be analyzed: if the text has more positive words it is classified as positive, otherwise it is classified as negative. This approach is very simple to apply but tends to be less accurate than machine learning methods because the lexicon coverage is not unlimited and because the specific

5 The work is carried out by non-expert collaborators recruited on specific web platforms.

SenticNet

POLARITY
positive

NRC-VAD-Lexicon

VALENCE	AROUSAL	DOMINANCE
0.802	0.549	0.647

SentiWordNet 3.0

PoS	Synset ID	PosScore	NegScore	Gloss
n	07186148	0	0	a request (spoken or written) to participate…
n	04689048	0.5	0	a tempting allurement

DepecheMood++

AFRAID	AMUSED	ANGRY	ANNOYED	DONT_CARE	HAPPY	INSPIRED	SAD
0.045481	0.155150	0.130596	0.154374	0.179065	0.135442	0.148765	0.051123

NRC-Emotion-Lexicon

ANGER	ANTICIPATION	DISGUST	FEAR	JOY	SADNESS	SURPRISED	TRUST	POSITIVE	NEGATIVE
0	1	0	0	0	0	0	0	1	0

Fig. 1 Entries for the noun *invitation* in different polarity and emotion lexicons.

context can vary the polarity of a word. It is important to note that the lexicon-based approach can be also applied in the emotion analysis task: in this case emotion lexicons containing word-emotion associations are used. Fig. 1 shows how different 5 English lexicons are from each other in the way they assign polarity or emotional values to the same word i.e., the noun *invitation.* The lexicons taken into consideration for this comparison are: *SenticNet* (Cambria et al. 2022), *NRC-VAD-Lexicon* (Mohammad 2018), *SentiWordNet 3.0, DepecheMood++* (Araque et al. 2019), *NRC-Emotion-Lexicon* (Mohammad & Turney 2013).

Machine learning, in the context of NLP, is the process of training a computational system to perform a certain linguistic task. In the supervised approach, the algorithm is trained taking as input a set of annotated data, that is a selection of texts in which the expected classification is provided (for example a collection of sentences each associated with a polarity value). On the contrary, in the unsupervised method training data is not provided but the system tries to autonomously extract generalizations from input texts. Generally, unsupervised learning tends to be less expensive than supervised learning, as it does not require training data, but the results are less accurate. For this reason, there are numerous initiatives that aim to produce annotated data for all the tasks mentioned in the previous sections, covering many languages and various textual genres. Over time, the machine learning algorithms used have evolved and deep learning techniques are now more widely adopted, leading to major improvements in system performance (Yadav & Vishwakarma 2020).

Whichever method is used, system performances tend to vary greatly depending on the task, text types and granularity of analysis. In general, the greater the number of labels used for classification, the greater the complexity of the task and therefore the lower the performance (Wankhade 2022).

6. Sentiment Analysis in Digital Humanities Research

Although the most common resources and tools for SA fall into categories such as social network analysis and customer opinion monitoring, research in DH has been growing in recent years. In general, the interest in the use of NLP methods for the processing of humanistic data is rising, as demonstrated by the large participation in dedicated scientific events.[6] In this increasingly rich panorama of projects and activities at the intersection between DH and NLP, SA is considered a fruitful technique for enriching textual data especially in the fields of history and literary studies.

Most works in the historical domain primarily use digitized newspaper articles as data to understand how important events or famous figures were perceived by their contemporaries. For example, entity-based SA is used in the *Oceanic Exchanges* project to identify the opinion expressed in 19th century German newspapers towards a group of writers of the same period (Keck et al. 2020), while Viola (2023) employs the same method for analyzing the sentiment towards a selection of entities in US newspapers published by Italian immigrants. On the other hand, Mayer et al. (2022) studies the transnational reception of the execution of Maximilian I, emperor of Mexico, in 1867 relying on newspaper from various countries. The case study presented by Sprugnoli et al. (2016) is different because it detects both prior and contextual polarities in Italian political texts of the first half of the 20th century and demonstrates that sentiment orientation is often implicitly expressed, making it particularly difficult to assign a polarity value even for humans.

The spectrum of research in the field of computational literary studies is broader. Starting from the pioneering work of Anderson & McMaster (1982) on the measurement of affective tones in the chapter of a novel and in a set of children's stories, the applications concerns various textual genres (gothic and romantic novels, fairy tales, plays, fan fiction) and various purposes (understanding what makes one plot more intriguing than another, what role emotions play in the interactions between characters, how emotions can help distinguish between different literary genres, what are

6 See, e.g., the annual workshops of the *ACL Special Interest Group on Language Technologies for the Socio-Economic Sciences and Humanities* (LaTeCH, https://sighum.wordpress.com/events/), the *Computational Humanities Research* (CHR, https://2023.computational-humanities-research.org) Conference, the *Workshop on Ancient Language Processing* (ALP, https://www.ancientnlp.com/alp2023) and the *Workshop on Language Technologies for Historical and Ancient Languages* (LT4HALA, https://circse.github.io/LT4HALA). All addresses were accessed on 23 June 2024.

the emotional arcs of stories) as well described in the survey papers by Kim & Klinger (2019) and Rebora (2023) to which we refer for further details.

7. Open Issues and Best Practices

The works cited so far (as well as all those that we have not been able to cite due to space limitations) have had, in one way or another, to address various issues relating to research practices in the humanities and to the characteristics of humanistic texts written in non-contemporary languages. First, literary, and historical texts are often sparse, inconsistent and incomplete, presenting many orthographic variations due to diachrony and diatopy phenomena. Then, to be processed by NLP systems, texts must be available in machine-readable format: the use of OCR *(Optical Character Recognition)* systems in digitization processes, especially when applied to manuscripts or ancient prints, is not free from errors and it is often necessary to intervene to reduce noise and obtain high quality data. Furthermore, humanities scholars work on textual genres (such as poems, plays, philosophical and historical treatises) that are very different from those usually analyzed by NLP systems: this requires that such systems be appropriately adapted or developed from scratch. Finally, final users of DH applications are humanities scholars who are often not tech-savvy users, so it is important to develop simple, intuitive and transparent systems.

The lack of large amounts of data on which to train machine learning systems, combined with the demand for systems whose results are easily interpretable has led to the widespread adoption of the lexicon-based approach in DH (Ohman 2021). In fact, machine learning algorithms are often criticized because they are difficult to interpret; they are like black boxes and not even the developers are able to explain perfectly why certain choices and, consequently, certain predictions are made. On the contrary, lexicon-based systems make it easier to understand the results, to highlight trends and passages which can be then re-analyzed through a closer reading. Furthermore, the need for intuitive systems has result in the creation of user-friendly graphical interfaces, more suitable for use by non-experts than programming scripts; some examples are *SEANCE* (Crossley et al. 2017), *Lingmotif* (Moreno-Ortiz 2017), and *SentText* (Schmidt et al. 2021). It is important to notice that *Syuzhet,* the first SA system that had a notable resonance, but also numerous criticisms, in the DH community, is lexicon-based and extremely straightforward from a computational point of view since it is based on simple word count;[7] since then, however, lexicon-based approaches have become more refined and the aforementioned tools include preprocessing functionalities (e.g., stop-word removal, lemmatization) and rules to handle negations.

7 See http://www.matthewjockers.net/2015/02/02/syuzhet (Accessed: 23 June 2024).

Whichever method is used and whether a polarity or emotion analysis is to be performed, there are various aspects to consider (Mohammad 2023). In fact, it is necessary to choose lexicon, data, and system to use, or decide to develop new ad hoc resources suitable for the domain of interest because the existing ones are not in line with the objectives of the research. This involves choosing the type of conceptualization to adhere to, i.e., whether to opt for the categorical or dimensional approach, but also whether to use continuous values or discrete labels, as well as the best level of granularity (in other words, how many classes or how many dimensions you want to capture). In addition, when a new lexicon or a new annotated dataset is to be developed from scratch, it's crucial to choose whether to recruit expert or non-expert annotators, through the adoption of crowdsourcing techniques. This second option, although widely employed when dealing with texts from social networks and contemporary languages, is more difficult to apply when dealing with historical and ancient languages. Furthermore, specifically in the case of ancient languages, the problem of the lack of native speakers must also be addressed because it is impossible to rely on the intuition or personal sensitivity of the annotators, thus it is essential to involve language and culture experts (Sprugnoli et al. 2020).

Defining the most correct procedure to follow can be a long interactive process, made up of several experimentation phases. For example, Schmidt et al. (2021) details the choices made to define a new scheme for annotating emotions in German plays written around 1800. Although at first, they considered adopting the categorical approach using basic emotions defined by Ekman or Plutchik, they soon realized that the psychological theories on which these categories were based did not reflect emotion and affect concepts of literary theories. Thanks to a pilot annotation, they noticed that some emotions were particularly relevant even if they did not belong to any psychological theory (e.g., friendship), while others did not have great importance (e.g., disgust) in dramatic texts. In the end, they came up with a new hierarchical scheme made up of 13 emotion concepts. Another interesting example is given by the analysis of emotions in poems which shows how the same textual genre can be addressed by considering different aspects (Sprugnoli et al. 2023) such as the level of expertise of annotators (experts or crowd workers), the textual unit to be annotated (line, sentence, stanza, whole poem), the number of emotions considered (two or more), the general perspective (emotions are annotated as intended by the author or as perceived by the reader). For instance, each poem of PO-EMO is annotated at both line and stanza levels with 9 emotions elicited in the reader employing both trained experts and crowd workers (Haider et al. 2020). On the other hand, only two classes are assigned by experts at each poem in the Kabithaa corpus made of Odia poems (Mohanty et al. 2018).

Summing up, there is no single right way to proceed: all the choices must be weighted based on the research objective, the characteristics of the texts to be analyzed, and the theoretical context of reference.

8. Conclusion

This chapter described the complexity of SA seen as a multifaceted problem: different tasks, methods and applications are introduced by focusing on research in the DH field. To conclude, we mention five promising lines of research that have emerged in recent years attracting increasing attention.

Large Language Models (LLMs). LLMs are deep neural networks (called transformers) trained on massive amounts of unannotated data to predict how a sentence continues or what is missing in a sentence. Prompts, that is instructions in natural language describing the task to perform, are the way humans interact with a LLM.[8] LLMs displays impressive capacities in many NLP tasks but understanding how to optimize prompts to achieve increasingly better results is an open issue also in the case of SA (Mao et al. 2023).

Multimodality. Multimodal SA allows to go beyond text-based SA integrating linguistic information with audio-visual information extracted from images, audio recordings and videos. The first experiments in the DH field concerned the analysis of plays (Schmidt & Wolff, 2021) and oral history interviews (Gref et al., 2022).

Linguistic Linked Data. Linguistic resources (lexicons and annotated data) for SA are now very numerous but do not interact with each other: using linked data techniques (Iglesias et al. 2017) would make them interoperable, more visible and reusable. Data models, ontologies and interlinked resources are presented every year during the Sentiment Analysis & Linguistic Linked Data workshop series, also including papers on Classical languages (Sprugnoli et al. 2021).

Perspectivism. To train machine learning systems it is necessary to have high quality data, in which labels are assigned in a consistent way; however, annotation is highly subjective when it comes to sentiment and emotions and it is often difficult to have a consensus on the label to assign because multiple interpretations are possible, especially when dealing with literary texts. In case of doubt or disagreement the assignments are forced towards a single label so that the algorithm can learn and make predictions. To change this paradigm, the so-called perspectivism, a more inclusive framework that aim at preserving the different points of view of the annotators, has been proposed (Cabitza et al. 2023).

Reader Response Studies. If the application of SA to the analysis of literary texts is still subject to criticism because it is not easy to find the right balance between computational approaches and narratological theories, reader response studies are

8 An example of such interaction is given by the ChatGPT interface.

finding greater success also in terms of system performance, showing that SA seems to be more efficient when applied to comments on a literary text than on the text itself (Pianzola et al. 2020).

References

Abu Farha, I., Oprea, S.V., Wilson, S., & Magdy, W. (2022). SemEval-2022 Task 6. iSarcasmEval. Intended Sarcasm Detection in English and Arabic. In *Proceedings of the 16th International Workshop on Semantic Evaluation* (SemEval-2022) (pp. 802–814). Seattle: Association for Computational Linguistics. DOI: https://doi.org/10.18653/v1/2022.semeval-1.111 (Accessed: 23 June 2024).

AlDayel, A., & Magdy, W. (2021). Stance detection on social media. State of the art and trends, *Information Processing & Management,* 58(4), 1–22.

Anderson, Clifford W., & McMaster, G.E. (1982). Computer assisted modeling of affective tone in written documents. *Computers and the Humanities,* 16(1), 1–9.

Araque, O., Gatti, L., Staiano, J., & Guerini, M. (2019). Depechemood++. A bilingual emotion lexicon built through simple yet powerful techniques, *IEEE transactions on affective computing,* 13(1), 496–507. DOI: https://doi.org/10.1109/TAFFC.2019.2934444 (Accessed: 23 June 2024).

Bender, E.M. (2016). Linguistic typology in natural language processing, *Linguistic Typology,* 20(3), 645–660.

Cabitza, F., Campagner, A., & Basile, V. (2023). Toward a perspectivist turn in ground truthing for predictive computing, *Proceedings of the AAAI Conference on Artificial Intelligence,* 37(6), 6860–6868. DOI: https://doi.org/10.1609/aaai.v37i6.25840 (Accessed: 23 June 2024).

Cambria, E., Liu, Q., Decherchi, S., Xing, F., & Kwok, K. (2022). SenticNet 7. A commonsense-based neurosymbolic AI framework for explainable sentiment analysis, *Proceedings of the Thirteenth Language Resources and Evaluation Conference,* 3829–3839. URL: https://aclanthology.org/2022.lrec-1.408 (Accessed: 23 June 2024).

Crossley, S.A., Kyle, K., & McNamara, D.S. (2017). Sentiment Analysis and Social Cognition Engine (SEANCE). An automatic tool for sentiment, social cognition, and social-order analysis, *Behavior research methods,* 49, 803–821.

Darwin, Ch. (1872). *The Expression of the Emotions in Man and Animals.* London: John Murray.

Du Bois, J.W. (2007). The stance triangle. In R. Engebretson (Ed.). *Stancetaking in Discourse. Subjectivity, evaluation, interaction* (pp. 139–182). Amsterdam: John Benjamins.

Ekman, P. (1993). Facial expression and emotion, *American psychologist,* 48(4), 384–392.

Gref, M, Matthiesen, N., Venugopala, S.H., Satheesh, Sh., Vijayananth, A., Ha, D.B., Behnke, S., & Köhler, J. (2022). A Study on the Ambiguity in Human Annotation of German Oral History Interviews for Perceived Emotion Recognition and Sentiment Analysis, *Proceedings of the Thirteenth Language Resources and Evaluation Conference,* 2022–2031. URL: https://aclanthology.org/2022.lrec-1.217 (Accessed: 23 June 2024).

Grice, H.P. (1975). Logic and conversation. In P. Cole & J.L. Morgan (Eds.), *Syntax and semantics,* 3. Speech acts (pp. 41–58). New York/San Francisco/London: Academic Press.

Haider, Th., Eger, S., Kim, E., Klinger, R., & Menninghaus, W. (2020). PO-EMO. Conceptualization, Annotation, and Modeling of Aesthetic Emotions in German and English Poetry, *Proceedings of the Twelfth Language Resources and Evaluation Conference,* 1652–1663. URL: https://aclanthology.org/2020.lrec-1.205 (Accessed: 23 June 2024).

Hernández Farias, D.I., & Rosso, P. (2017). Irony, sarcasm, and sentiment analysis. In F.A. Pozzi, E. Fersini, E. Messina & B. Liu (Eds.), *Sentiment Analysis in Social Networks* (pp. 113–128). Amsterdam et al.: Morgan Kaufmann.

Hutto, C., & Gilbert, E. (2024). VADER. A parsimonious rule-based model for sentiment analysis of social media text, *Proceedings of the international AAAI conference on web and social media,* 8(1), 216–225. DOI: https://doi.org/10.1609/icwsm.v8i1.14550 (Accessed: 23 June 2024).

Iglesias, C.A., Sanchez-Rada, F.J., Vulcu, G., & Buitelaar, P. (2017). Linked data models for sentiment and emotion analysis in social networks. In F.A. Pozzi, E. Fersini, E. Messina & B. Liu (Eds.), *Sentiment Analysis in Social Networks* (pp. 49–69). Amsterdam et al.: Morgan Kaufmann.

Keck, J., Knabben, M., & Pado, S. (2020). Who's in the News? Methodological Challenges and Opportunities in Studying 19th-century Writers in Historical Newspapers. *Europeana PRO,* 16. Newspapers, no pag. URL: https://pro.europeana.eu/page/issue-16-newspapers#who-s-in-the-news (Accessed: 23 June 2024).

Kim, E., & Klinger, R. (2019). A Survey on Sentiment and Emotion Analysis for Computational Literary Studies, *Zeitschrift für digitale Geisteswissenschaften,* no pag. DOI: https://doi.org/10.17175/2019_008_v2 (Accessed: 23 June 2024).

Liu, B. (2010). Sentiment analysis. A multi-faceted problem. *IEEE intelligent systems,* 25(3), 76–80.

Id. (2022). *Sentiment analysis and opinion mining.* Cham: Springer. DOI: https://doi.org/10.1007/978-3-031-02145-9 (Accessed: 23 June 2024).

Mayer, A.I.L., Gutierrez-Vasques, X., Saiso, E.P., & Salmi, H. (2022). Underlying Sentiments in 1867. A Study of News Flows on the Execution of Emperor Maximilian of Mexico in Digitized Newspaper Corpora, *Digital Humanities Quarterly,* 16(4), 1–98. URL: http://www.digitalhumanities.org/dhq/vol/16/4/000649/000649.html (Accessed: 23 June 2024).

Maynard, D., & Greenwood, M. (2014). Who cares about Sarcastic Tweets? Investigating the Impact of Sarcasm on Sentiment Analysis, *Proceedings of the Ninth International Conference on Language Resources and Evaluation* (LREC'14), 4238–4243. URL: http://www.lrec-conf.org/proceedings/lrec2014/pdf/67_Paper.pdf (Accessed: 23 June 2024).

Mao, R., Liu, Q., He, K., Li, W., & Cambria, E. (2023). The biases of pre-trained language models. An empirical study on prompt-based sentiment analysis and emotion detection. *IEEE Transactions on Affective Computing*, 14(3), 1743–1753. DOI: https://doi.org/10.1109/TAFFC.2022.3204972 (Accessed: 23 June 2024).

Mohammad, S. (2018). Obtaining reliable human ratings of valence, arousal, and dominance for 20,000 English words, *Proceedings of the 56th annual meeting of the association for computational linguistics,* 1. Long papers, 174–184. DOI: https://doi.org/10.18653/v1/P18-1017 (Accessed: 23 June 2024).

Id. (2023). Best Practices in the Creation and Use of Emotion Lexicons, *Findings of the Association for Computational Linguistics.* EACL 2023, 1825–1836. DOI: https://doi.org/10.18653/v1/2023.findings-eacl.136 (Accessed: 23 June 2024).

Id., & Turney, P. D. (2013). Crowdsourcing a word-emotion association lexicon, *Computational intelligence,* 29(3), 436–465. DOI: https://doi.org/10.1111/j.1467-8640.2012.00460.x (Accessed: 23 June 2024).

Mohanty, G., Mishra, P., & Mamidi, R. (2018). Kabithaa. An annotated corpus of Odia poems with sentiment polarity information, *Proceedings of the Eleventh International Conference on Language Resources and Evaluation* (LREC 2018), 52–57. URL: http://lrec-conf.org/workshops/lrec2018/W11/pdf/15_W11.pdf (Accessed: 23 June 2024).

Moreno-Ortiz, A. (2017). Lingmotif. Sentiment analysis for the digital humanities, *Proceedings of the Software Demonstrations of the 15th Conference of the European Chapter of the Association for Computational Linguistics,* 73–76. URL: https://aclanthology.org/E17-3019 (Accessed: 23 June 2024).

Ohman, E. (2021). The validity of lexicon-based emotion analysis in interdisciplinary research, *Proceedings of the Workshop on Natural Language Processing for Digital Humanities* (NLP4DH), 7–12. URL: https://aclanthology.org/2021.nlp4dh-1.2 (Accessed: 23 June 2024).

Pianzola, F., Rebora, S., & Lauer, G. (2020). Wattpad as a resource for literary studies. Quantitative and qualitative examples of the importance of digital social reading and readers' comments in the margins, *PloS one,* 15(1), 1–46. URL: 10.1371/journal.pone.0226708 (Accessed: 23 June 2024).

Plutchik, R. (1980). A general psychoevolutionary theory of emotion. In id. & H. Kellermann (Eds.), *Theories of emotion* (pp. 3–33). Cambridge, Mass.: Academic press [= *Emotion. Theory, Research, and Experience,* 1]. DOI: https://doi.org/10.1016/B978-0-12-558701-3.50007-7 (Accessed: 23 June 2024).

Rebora, S. (2023). Sentiment Analysis in Literary Studies. A Critical Survey, *Digital Humanities Quarterly,* 17(2), 1–50. URL: http://www.digitalhumanities.org/dhq/vol/17/2/000691/000691.html (Accessed: 23 June 2024).
Russell, J. A. (1980). A circumplex model of affect, *Journal of Personality and Social Psychology,* 39, 1161–1178.
Schmidt, Th, & Wolff, Ch. (2021). Exploring multimodal sentiment analysis in plays. A case study for a theater recording of Emilia Galotti, *Proceedings of CHR 2021. Computational Humanities Research Conference,* 392–404. URL: https://ceur-ws.org/Vol-2989/short_paper45.pdf (Accessed: 23 June 2024).
Schmidt, Th., Dangel, J., & Wolff, Ch. (2021). A Tool for Lexicon-based Sentiment Analysis in Digital Humanities, *Proceedings of the 16th International Symposium of Information Science* (ISI 2021), 156–172. URL: https://epub.uni-regensburg.de/44943/1/isi_schmidt_dangel_wolff.pdf (Accessed: 23 June 2024).
Schmidt, Th., Dennerlein, K., & Wolff, Ch. (2021). Towards a corpus of historical german plays with emotion annotations, *Proceedings of 3rd Conference on Language. Data and Knowledge (LDK 2021). Schloss Dagstuhl. Leibniz-Zentrum für Informatik,* 1–11 [= *Open Access Series in Informatics,* 93]. DOI: https://doi.org/10.4230/OASIcs.LDK.2021.9 (Accessed: 23 June 2024).
Sprugnoli, R., Passarotti, M., Corbetta, D., &. Peverelli, A. (2020). Odi et Amo. Creating, Evaluating and Extending Sentiment Lexicons for Latin, *Proceedings of the Twelfth Language Resources and Evaluation Conference,* 3078–3086. URL: https://aclanthology.org/2020.lrec-1.376 (Accessed: 23 June 2024).
Sprugnoli, R., Passarotti, M., Testori, M., & Moretti, G. (2022). Extending and Using a Sentiment Lexicon for Latin in a Linked Data Framework, *Proceedings of the Workshops and Tutorials. Language Data and Knowledge 2021 (LDK 2021),* 1–14. DOI: https://doi.org/10.5281/zenodo.6303164 (Accessed: 23 June 2024).
Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M. (2011). Lexicon-based methods for sentiment analysis, *Computational linguistics,* 37(2), 267–307. DOI: https://doi.org/10.1162/COLI_a_00049 (Accessed: 23 June 2024).
Utsumiu, A. (2000). Verbal irony as implicit display of ironic environment. Distinguishing ironic utterances from nonirony, *Journal of pragmatics,* 32(12), 1777–1806. URL: http://www.utm.se.uec.ac.jp/~utsumi/paper/jop2000-utsumi.pdf (Accessed: 23 June 2024).
Viola, L. (2023). Networks of migrants' narratives. A post-authentic approach to heritage visualisation, *ACM Journal on Computing and Cultural Heritage,* 16(1), 1–21. DOI: https://doi.org/10.1145/3575863 (Accessed: 23 June 2024).
Wankhade, M., Rao, A. Ch. S., & Kulkarni, Ch. (2022). A survey on sentiment analysis methods, applications, and challenges, *Artificial Intelligence Review,* 55(7), 5731–5780. DOI: https://doi.org/10.1007/s10462-022-10144-1 (Accessed: 23 June 2024).
Yadav, A., & Vishwakarma, D. K. (2020). Sentiment analysis using deep learning architectures. a review, *Artificial Intelligence Review,* 53(6), 4335–4385. DOI: https://doi.org/10.1007/s10462-019-09794-5 (Accessed: 23 June 2024).

Figure Credit

Fig. 1 was created by the author herself and first published here.