SentText: A Tool for Lexicon-based Sentiment Analysis in Digital Humanities

Thomas Schmidt
Media Informatics Group, University of Regensburg, Germany
thomas.schmidt@ur.de

Johanna Dangel
Media Informatics Group, University of Regensburg, Germany
johannadangel97@web.de

Christian Wolff
Media Informatics Group, University of Regensburg, Germany
christian.wolff@ur.de

Abstract

We present *SentText*, a web-based tool to perform and explore lexicon-based sentiment analysis on texts, specifically developed for the Digital Humanities (DH) community. The tool was developed integrating ideas of the user-centered design process and we gathered requirements via semi-structured interviews. The tool offers the functionality to perform sentiment analysis with predefined sentiment lexicons or self-adjusted lexicons. Users can explore results of sentiment analysis via various visualizations like bar or pie charts and word clouds. It is also possible to analyze and compare collections of documents. Furthermore, we have added a close reading function enabling researchers to examine the applicability of sentiment lexicons for specific text sorts. We report upon the first usability tests with positive results. We argue that the tool is beneficial to explore lexicon-based sentiment analysis in the DH but can also be integrated in DH-teaching.

Keywords: sentiment analysis; dictionary-based approaches; Digital Humanities; usability; tool; user-centered design; lexicon-based approaches

1 Introduction

Sentiment analysis (or opinion mining) is a term used to describe computational methods for predicting and analyzing sentiment, predominantly in written text (Liu, 2016, p. 1). Sentiment analysis is especially popular for social media content (Moßburger et al., 2020; Schmidt, Hartl, Ramsauer, Fischer, Hilzenthaler, & Wolff, 2020; Schmidt, Kaindl, & Wolff, 2020) and any

other form of user generated content (cf. Mäntylä et al., 2018). Sentiment analysis offers various benefits for industry and is used, for example, to investigate the popularity of enterprises or political parties on social media (cf. ibid.), predict the mood of users in usability testing (Schmidt, Schlindwein, Lichtner, & Wolff, 2020), in health informatics (Hartl et al., 2019) and gaming (Halbhuber et al., 2019). Sentiment Analysis is focused on the prediction of the attitude towards an entity on a bipolar scale or nominal classes consisting of positive, neutral and negative (Vinodhini & Chandrasekaran, 2012). The neighboring research area of emotion analysis is similar in its approaches and main goals but focused on the prediction of more differentiated emotional classes like anger, sadness and happiness (cf. Binali et al., 2010).

Both research fields have gained increasing popularity in the field of Digital Humanities (DH), especially in the area of Computational Literary Studies (cf. Kim & Klinger, 2018). Sentiment and emotion analysis have been used to investigate various types of literary texts like novels (Jannidis et al., 2016; Kakkonen & Kakkonen, 2011; Reagan et al., 2016), fairy tales (Alm & Sproat, 2005; Mohammad, 2011), plays (Mohammad, 2011; Nalisnick & Baird, 2013; Schmidt & Burghardt, 2018a, 2018b; Schmidt, Burghardt, & Dennerlein, 2018b; Schmidt, 2019; Yavuz, 2021), fan fictions (Kim & Klinger, 2019a, 2019b), online writings (Pianzola et al., 2020), historical political texts (Sprugnoli et al., 2016) or pop song lyrics (Napier & Shamir, 2018; Schmidt, Bauer, Habler, Heuberger, Pilsl, & Wolff, 2020). Research goals vary for the application of sentiment and emotion analysis on these text sorts. Most research tries to explore general purpose applications on these texts to analyze descriptive results, e.g., with a focus on sophisticated visualizations of sentiment and emotion distributions and progression or comparisons of different works (Kakkonen & Kakkonen, 2011; Mohammad, 2011; Reagan et al., 2016, Napier & Shamir, 2018; Schmidt, Burghardt, & Dennerlein, 2018b; Schmidt, 2019). Others evaluate different methodological approaches for this challenging text sort comparing the performance on annotated text units (Schmidt & Burghardt, 2018a; Kim & Klinger, 2019a). Examples of other projects and research goals are the analysis and prediction of plot developments (Reagan et al., 2016), character relations (Nalisnick & Baird, 2013; Schmidt & Burghardt, 2018b; Yavuz, 2021) or "happy endings" (Jannidis et al., 2016) via sentiment and emotion analysis. One branch of research focuses on the annotation of texts with sentiment or emotion information to create well-curated corpora for evaluation and machine learning and to investigate annotation behavior and agreement statistics (Alm & Sproat, 2005; Sprugnoli et al.,

2016; Schmidt, Burghardt, & Dennerlein, 2018a; Schmidt, Burghardt, Dennerlein, & Wolff, 2019a; Schmidt, Jakob, & Wolff, 2019; Schmidt, Winterl, Maul, Schark, Vlad, & Wolff, 2019). Modern approaches explore multimodal methods to analyze sentiment in cultural artefacts (Schmidt, Burghardt, & Wolff, 2019; Ortloff et al., 2019).

From a methodological standpoint, the vast majority of the aforementioned research projects in DH often rely on the general purpose, rule-based method of lexicon- (also often called dictionary-) based sentiment analysis. The main idea of lexicon-based methods in Natural Language Processing (NLP) is purely descriptive and is to count words according to large lists of words of a specific category (e.g., positive or negative sentiment), which is called the dictionary or lexicon, to derive insights about the analyzed texts. In the context of sentiment and emotion analysis, some of the most famous English sentiment lexicons are the *Bing Lexicon* (Hu & Liu, 2004) and the *NRC Emotion Lexicon* (Mohammad & Turney, 2013). They consist of large lists of words annotated with their most likely sentiment connotation (e.g., positive, negative, neutral). These words are referred to as sentiment bearing words (SBWs). They are either assigned with a sentiment class or by values with a linear metric, e.g., between −3 (negative) and 3 (positive), to represent to what degree a word is connoted with a specific sentiment. In the case of sentiments consisting of nominal assignments for sentiment classes the value for a positive word is regarded as 1 and for a negative word as −1. To gain an overall value of the expressed sentiment of a text, all values of the detected words are summed up. In the same way, one can perform emotion analysis (Mohammad & Turney, 2013) with lexicons that consist of list of words for specific emotion categories. For more information about the creation of sentiment or emotions lexicons and an overview of research cf. Mohammad (2020).

In general, the lexicon-based approach has been proven inferior compared to advanced machine learning approaches (cf. Kim & Klinger, 2018) for multiple use cases. Current state-of-the-art machine learning approaches in sentiment and emotion analysis involve large word embeddings and deep neural networks (cf. Zhang et al., 2018; Shmueli et al., 2019). Lexicon-based approaches are currently only recommended for areas in which not enough large-scale annotated corpora exist which would be necessary for modern methods. This is mostly the case for under-resourced languages or text sorts which are uncommon in the NLP-community which is the case for literary and historic texts. However, even outside of this use case, libraries and APIs

applying lexicon-based text analysis have regained popularity in the NLP-community to calculate benchmarks or perform first explorations (e.g., VADER; Hutto & Gilbert, 2014).

While rare compared to other text sorts, one can find first projects in DH exploring state-of-the-art deep learning techniques, e.g., on literary texts (Kim & Klinger, 2019a). Nevertheless, lexicon-based methods are still widely popular and the established sentiment analysis method in DH. The vast of majority of recent and current projects use some sort of lexicon-based techniques (Alm & Sproat, 2005; Mohammad, 2011; Nalisnick & Baird, 2013; Reagan et al., 2016; Sprugnoli et al., 2016; Jannidis et al., 2016; Schmidt & Burghardt, 2018a, 2018b; Schmidt, 2019; Pianzola et al., 2020; Schmidt, Bauer, Habler, Heuberger, Pilsl, & Wolff, 2020; Yavuz, 2021). Reasons for this are certainly the lack of large-scale and well-annotated training corpora that would be necessary for state-of-the-art-methods but also that it is a rather transparent and easy method enabling researchers to perform fast and comprehensible explorations of textual sentiment. Schmidt, Burghardt, and Wolff (2018) discuss this dominance of lexicon-based methods in DH in more detail.

Nevertheless, to our knowledge, a usable and accessible tool to perform lexicon-based sentiment analysis specifically for DH researchers that is adaptable to the various use cases of DH and that allows the easy introspection of results does not exist so far. While this is certainly no problem for DH researchers familiar with coding and computational methods, humanities scholars with limited IT skills are quickly discouraged. As Burghardt and Wolff (2014) argue, it is especially important for tools in DH to be accessible and of high usability to lower the participation threshold and help establish digital methods in the humanities community. Furthermore, it is important to note that humanities scholars are a special user group with very specific needs and skills that have to be taken into account during the development phase. While one can identify a growing interest in projects developing such accessible tools (e.g., Schmidt, Burghardt, Dennerlein, & Wolff, 2019b; Schmidt, Jakob, & Wolff, 2019), there is still a lack of easy-to-use and accessible tools for various methods in DH and the tool *SentText* was developed to close this gap for lexicon-based sentiment analysis.

We present the tool *SentText* (see Fig. 1 for the logo), a web-based tool for performing lexicon-based sentiment analysis on texts via a user interface without the need of coding skills. We argue that the tool is beneficial for first explorations in DH research, e.g., to compare the applicability of various

sentiment lexicons for a specific text sort but also in DH teaching to introduce students into the possibilities but also challenges of lexicon-based sentiment analysis. While the tool is currently focused on sentiment analysis in its default settings and overall presentation it can easily be applied to any sort of lexicon-based analysis. To adapt to the critical user group of humanities scholars we apply the user-centered design process (cf. Vredenburg et al., 2002) and integrate the feedback of this user group early in the development process. Furthermore, we integrate usability tests at various steps of the development to evaluate and improve the usability of the tool.

Fig. 1 Logo of the tool *SentText*

2 Development and requirements analysis

The tool was developed with the Flask framework[1] in Python 3.7, JavaScript, CSS3 and HTML5. We use multiple libraries for natural language processing like the NLTK[2]. Our development process uses ideas of the user-centered design process (cf. Vredenburg et al., 2002). We develop the tool in multiple iterations integrating the feedback of the final user group at various steps via methods of requirements engineering or usability testing.

We acquired first requirements via interviews at the beginning of development. We conducted semi-structured interviews with six researchers with backgrounds either in DH, literary studies or computational text analysis to gather an understanding of potential text sorts, their workflow and potential functionalities for a lexicon-based sentiment analysis tool. All researchers either have performed research in sentiment analysis and DH or plan to do so. The interviews were conducted either personally or via video call and the audio was transcribed afterward. The interview style was semi-structured

1 https://flask.palletsprojects.com/en/1.1.x/

2 https://www.nltk.org/

with loose guidelines consisting of questions and topics to talk about. We summarize some of the main requirements that had implications on the development of *SentText*. Some of the gathered requirements have also implications on the tool development in other research areas:

- Corpora come in different forms and shapes, the most important file formats that should be supported are XML, TEI and TXT.
- Preferred method to import text is a simple upload possibility.
- Important preprocessing steps before the sentiment analysis that should be integrated are: Lemmatization, stop words filtering, lower casing (However we noticed the higher the technical expertise of our participants the more likely they argued to perform these steps themselves and that they would eventually not trust a web tool.).
- Sentiment lexicons have to be adjusted (e.g., adding and removing words) and it should be possible to use own self-created lexicons; some standard sentiment lexicons should be offered by default.
- the most desired visualization types: a progression curve throughout the text, word clouds, visualizations enabling the comparison between texts or larger text groups
- Results should be traceable and transparent; it should be possible to follow to calculation process to the smallest unit.
- Web is the preferred platform since it is more accessible (no long installation process, no problems with different operating systems).
- All results should be downloadable in standard formats, e.g., CSV, JPEG.
- The tool should consist of a graphical user interface and should have a documentation.

Furthermore, we conducted a market analysis comparing various established online sentiment analysis tools to reflect upon useful functions, design elements and what we may offer with our own tool; to name a few: *SentiStrength*[3], *Sentiment Analyzer*[4] or the *Stanford Sentiment Analysis Tool*[5]. However, none of these tools is directed towards DH, works with transparent lexicon-based methods and oftentimes lacks important functionality like the upload of own material. Nevertheless, we systematically analyzed functional-

3 http://sentistrength.wlv.ac.uk/

4 https://www.danielsoper.com/sentimentanalysis/default.aspx

5 http://nlp.stanford.edu:8080/sentiment/rntnDemo.html

ity, layout, advantages and disadvantages of these tools to derive potential features for our future tool and how to implement them.

3 Functionality

SentText is web-based and can therefore be used without any sort of installation. The tool is available online:
https://thomasschmidtur.pythonanywhere.com/.
We integrated a detailed documentation and "about" page but also offer information at various steps of the tool usage via text or tooltips.

Users can upload texts in UTF-8-encoded TXT- or XML-format and perform the sentiment analysis. Using advanced options users can choose to perform stop words removal, lemmatization (only for German: via textblob[6]) or the integration of negations into the sentiment calculation (negations before a sentiment bearing word reverse the sentiment value).

Please note that at the moment these functionalities are performed for German (as we currently focus on the support of the study of German literary texts), however, users can upload lexicons and stop word lists for other languages and we also plan to include modules for other languages in future work. Users can also choose the sentiment lexicon to be used for the sentiment analysis. We are currently offering per default *SentiWS* (Remus et al., 2010) and *BAWL-R* (Vo et al., 2009), two popular lexicons for German. However, users can also use self-created or adjusted sentiment lexicons if they follow a specific CSV-format, which is defined and explained in the documentation. We plan to add more free lexicons for German (e.g., Waltinger, 2010) as well as for other languages in the future (e.g., Mohammad & Turney, 2013). The adjustment of the lexicons enables users to change lexicons or use resources that are better suited for the specific text sort of interest. Once *SentText* has completed a sentiment analysis run, the results screen is shown (Fig. 2).

On the left side, users can investigate their documents overall but also create groups via "Create Folder" to compare collections of text. By double clicking on a document or a group, users can investigate the results in the

6 https://pypi.org/project/textblob-de/

Fig. 2 Results-Screen – Three groups of documents of different authors are analyzed and compared to each other.

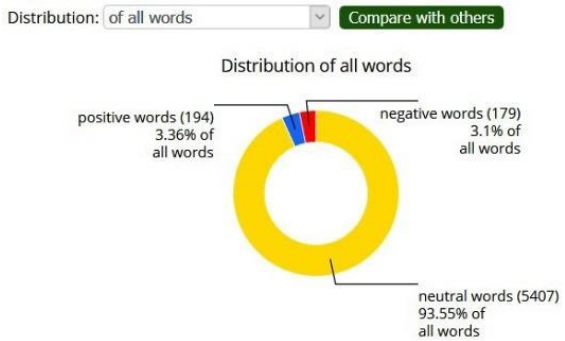

Fig. 3 Pie chart – Distribution of sentiment bearing words

Fig. 4 Word cloud – Strongest negative sentiment bearing words for a German text

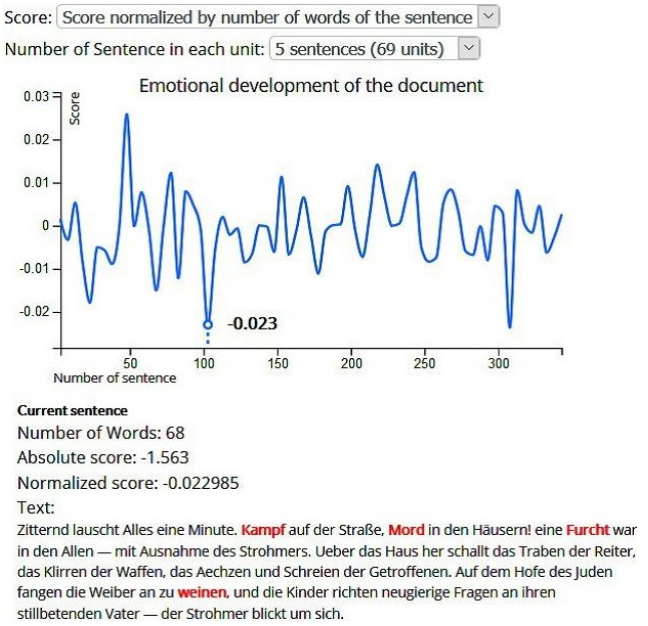

Fig. 5 Sentiment progression throughout a text using normalized values and five sentences per data point unit (text: the novel "Auf Wiedersehen!" by Leo Goldhammer)

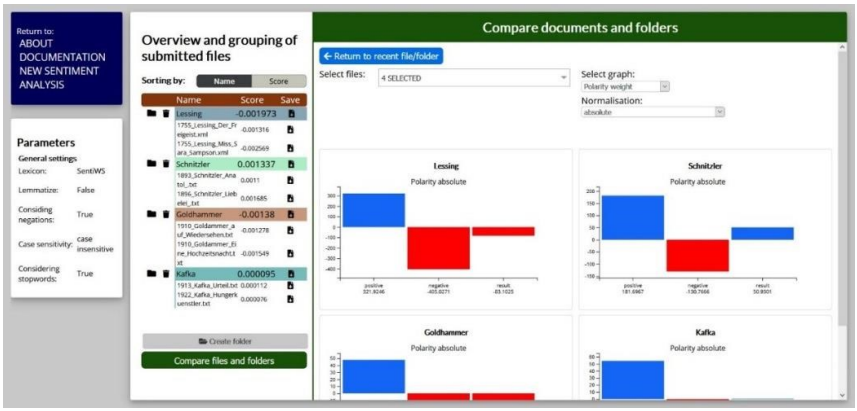

Fig. 6 Comparison of multiple document collections using bar charts

middle area of the page. The tool reports absolute but also normalized values (normalized by the number of tokens of a text) as well as overall values or

SBW-specific values. Users can select to analyze only sentiment bearing words or all words (Fig. 3). The results are shown via various forms of interactive visualization like bar charts, pie charts (Fig. 3) or word clouds (Fig. 4).

Results can be examined at the word as well as at the sentence level. Users can also visualize the sentiment progression throughout a text as a line graph (Fig. 5). If multiple documents are analyzed, visualizations have a "compare with others"-button (Fig. 6). On the right side of the screen (see Fig. 2 above), there is a "close reading" section (Fig. 7). Users can explore the selected text and what words are marked in what way by the lexicon in detail. This section enables users to investigate what words might be marked wrongly for a specific lexicon and how lexicons should be adjusted. Users can also adjust sentiment values of words during their analysis and download all data as CSV- or XML-files for further analysis as well as all visualizations as PNG-files.

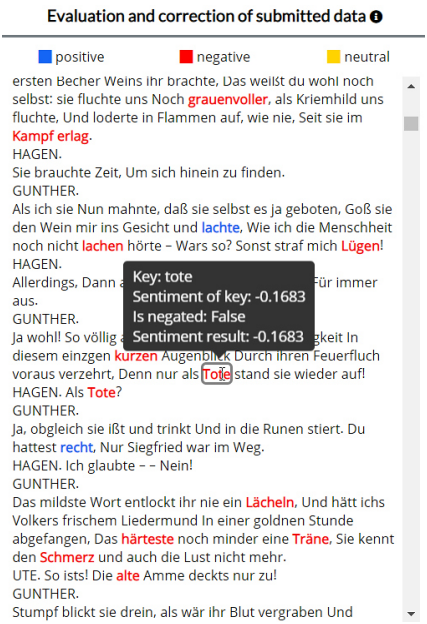

Fig. 7
Close reading section of the tool (text: "Kriemhilds Rache" by Friedrich Hebbel)

4 Evaluation

We have conducted preliminary small-scale usability tests for our tool after the first development iteration to gain some first insights in possible improvements and the overall quality of the tool. The first usability test we performed was focused on the detection of general user type independent usability issues. We conducted a task-based guerilla usability test (Nielsen, 1994) with ten participants (two female, eight male in the age group from 21 to 30 years). The sample consisted almost entirely of students of various degree programs and none of the participants had experience with sentiment analysis. The usability test was lab-based and conducted in a quiet room with one supervisor. The test consisted of nine tasks of which two were explorative ones. The test involved tasks like searching for specific information, analyzing text according to specific parameters, comparing and exporting results. Users were assigned to "think aloud" and give as much Feedback as possible. The tests were recorded, and the supervisor made notes with important feedback during the test.

After analyzing the recordings of the test and the notes taken, we summarized the results by identifying all mentioned usability issues of which we found 21. We ordered these issues according to severity and formulated possible improvements. While we cannot discuss all issues in detail, we want to highlight some of the most important issue groups. A lot of problems can be summarized as missing explanations (e.g., what type of data can be uploaded, how to interpret this score, how to interpret colors). Thus, we integrated more help at fitting usage points and improved upon the documentation. Other issues dealt with an information overload which led us to hide a lot of results at the beginning and make them expandable. Other than that we also identified some basic missing functionality like saving image files by right-click. Despite the usability issues, the average task completion rate (ratio of successfully completed tasks to all tasks) was 89% which points to a tool that is effectively and efficiently to use. We improved upon all identified usability problems in a subsequent development iteration before proceeding to the next usability test.

We conducted the second usability test with four persons with DH or a humanities background (two female, two male) and between 25 and 31 years old. The focus of this test was to gain an overall impression of the tool from the viewpoint of the target group via quantitative parameters. The test

approach was similar to the previous test with the exception that it was performed as remote usability test via Skype. The task success rate was 96% and overall feedback was rather positive. After the test, participants were instructed to fill out a questionnaire: The System Usability Scale (SUS) by Brooke (1996). The SUS is a well-established questionnaire for measuring usability (Bangor et al., 2008) and the tool achieves a score of 92.5 (on a scale from 0 to 100) which is regarded as very good usability (ibid.). While we did not identify major usability issues via this test, we received some feedback for some minor new features.

5 Future work

We argue that *SentText*, in its current state, is a promising and accessible tool for integrating and exploring digital methods in the humanities and the results of the usability tests support this claim. While the sample size for the usability tests is rather small, the results are quite encouraging. We plan to continue our test with more researchers in literary studies to get a better understanding of (1) how and (2) at what steps the tool can best support sentiment analysis projects. Furthermore, we will integrate the usage of this tool in a course about computational literary studies to analyze how this tool can be beneficial for teaching purposes and get feedback from students. Following up, we will continue our next development cycle integrating more lexicons and support for other languages as well as the requirements and features we will acquire with the integration and collaboration of more researchers in literary studies. The application of the user-centered design process, the early integration of the user group as well as the usability tests were highly beneficial for the quality of the current tool and more so reasonable steps due to the special target group.

References

Alm, C. O., & Sproat, R. (2005). Emotional sequencing and development in fairy tales. In *International Conference on Affective Computing and Intelligent Interaction* (pp. 668–674). Berlin, Heidelberg: Springer.

Bangor, A., Kortum, P. T., & Miller, J. T. (2008). An empirical evaluation of the system usability scale. *International Journal of Human-Computer Interaction*, 24(6), 574–594.

Binali, H., Wu, C., & Potdar, V. (2010). Computational approaches for emotion detection in text. In *4th IEEE International Conference on Digital Ecosystems and Technologies* (pp. 172–177). Piscataway, NJ: IEEE.

Brooke, J. (1996). SUS – A quick and dirty usability scale. In Jordan, Patrick W. (Ed.): *Usability Evaluation in Industry*. London: Taylor & Francis, pp. 189–194.

Burghardt, M., & Wolff, C. (2015). Humanist-Computer Interaction: Herausforderungen für die Digital Humanities aus Perspektive der Medieninformatik. In *Book of Abstracts. Workshop "Informatik und die Digital Humanities"*, Leipzig.

Halbhuber, D., Fehle, J., Kalus, A., Seitz, K., Kocur, M., Schmidt, T. & Wolff, C., (2019). The Mood Game – How to use the player's affective state in a shoot'em up avoiding frustration and boredom. In: Alt, F., Bulling, A., & Döring, T. (Eds.), *Mensch und Computer 2019 – Tagungsband.* New York, NY: ACM Press. DOI: 10.1145/3340764.3345369

Hartl, P., Fischer, T., Hilzenthaler, A., Kocur, M. & Schmidt, T. (2019). AudienceAR – Utilising Augmented Reality and Emotion Tracking to Address Fear of Speech. In Alt, F., Bulling, A., & Döring, T. (Eds.), *Mensch und Computer 2019 – Tagungsband.* New York, NY: ACM Press. DOI: 10.1145/3340764.3345380

Hu, M., & Liu, B. (2004). Mining and summarizing customer reviews. In *Proceedings of the 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 168–177), Berlin, Heidelberg: Springer.

Hutto, C., & Gilbert, E. (2014). Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Proceedings of the International AAAI Conference on Web and Social Media* (pp. 216–225). Reading, UK: Academic Conferences Limited.

Jannidis, F., Reger, I., Zehe, A., Becker, M., Hettinger, L. & Hotho, A. (2016). Analyzing Features for the Detection of Happy Endings in German Novels. *arXiv preprint* arXiv:1611.09028

Kakkonen, T., & Kakkonen, G. G. (2011). SentiProfiler: creating comparable visual profiles of sentimental content in texts. In *Proceedings of Language Technolo-*

gies for Digital Humanities and Cultural Heritage (pp. 62–69). Cham: Springer Nature.

Kim, E., & Klinger, R. (2018). A survey on sentiment and emotion analysis for computational literary studies. *arXiv preprint* arXiv:1808.03137

Kim, E., & Klinger, R. (2019a). Frowning Frodo, Wincing Leia, and a Seriously Great Friendship: Learning to Classify Emotional Relationships of Fictional Characters. *arXiv preprint* arXiv:1903.12453

Kim, E., & Klinger, R. (2019b). An Analysis of Emotion Communication Channels in Fan Fiction: Towards Emotional Storytelling. *arXiv preprint* arXiv:1906.02402

Liu, B. (2016). *Sentiment Analysis. Mining Opinions, Sentiments and Emotions.* New York: Cambridge University Press.

Mäntylä, M. V., Graziotin, D., & Kuutila, M. (2018). The evolution of sentiment analysis—A review of research topics, venues, and top cited papers. *Computer Science Review, 27*(Febr.), 16–32.

Mohammad, S. (2011). From once upon a time to happily ever after: Tracking emotions in novels and fairy tales. In *Proceedings of the 5th ACL-HLT Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities* (pp. 105–114). Association for Computational Linguistics.

Mohammad, S. M. (2020). Practical and Ethical Considerations in the Effective use of Emotion and Sentiment Lexicons. *arXiv preprint* arXiv:2011.03492

Mohammad, S. M., & Turney, P. D. (2013). Crowdsourcing a Word-Emotion Association Lexicon. *Computational Intelligence, 29*(3), 436–465.

Moßburger, L., Wende, F., Brinkmann, K., & Schmidt, T. (2020). Exploring Online Depression Forums via Text Mining: A Comparison of Reddit and a Curated Online Forum. In *Proceedings of the Fifth Social Media Mining for Health Applications (#SMM4H). Workshop & Shared Task* (pp. 70–81). Association for Computational Linguistics.

Nalisnick, E. T. & Baird, H. S. (2013). Character-to-character sentiment analysis in shakespeare's plays. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics* (Vol. 2: Short Papers, pp. 479–483). Association for Computational Linguistics.

Napier, K., & Shamir, L. (2018). Quantitative sentiment analysis of lyrics in popular music. *Journal of Popular Music Studies, 30*(4), 161–176.

Nielsen, J. (1994). Guerrilla HCI: Using discount usability engineering to penetrate the intimidation barrier. In Randolph G. Bias (Ed.): *Cost-Justifying Usability* (pp. 245–272). Boston, MA: Academic Press.

Ortloff, A.-M., Güntner, L., Windl, M., Schmidt, T., Kocur, M., & Wolff, C. (2019). SentiBooks: Enhancing Audiobooks via Affective Computing and Smart Light Bulbs. In: Alt, F., Bulling, A., & Döring, T. (Eds.), *Mensch und Computer 2019 – Tagungsband.* New York, NY: ACM Press. DOI: 10.1145/3340764.3345368

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), e0226708.

Reagan, A. J., Mitchell, L., Kiley, D., Danforth, C. M., & Dodds, P. S. (2016). The emotional arcs of stories are dominated by six basic shapes. *EPJ Data Science, 5*(1), 31.

Remus, R., Quasthoff, U., & Heyer, G. (2010, May). SentiWS – A Publicly Available German-language Resource for Sentiment Analysis. In *Proceedings of the Seventh conference on International Language Resources and Evaluation (LREC'10)* (pp. 1168–1171). European Language Resources Association.

Schmidt, T. (2019). Distant Reading Sentiments and Emotions in Historic German Plays. In *Abstract Booklet, DH_Budapest_2019* (pp. 57–60). Budapest, Hungary.

Schmidt, T., Bauer, M., Habler, F., Heuberger, H., Pilsl, F., & Wolff, C. (2020). Der Einsatz von Distant Reading auf einem Korpus deutschsprachiger Songtexte. In *Book of Abstracts, DHd 2020* (pp. 296–299). Paderborn, Deutschland.

Schmidt, T., & Burghardt, M. (2018a). An Evaluation of Lexicon-based Sentiment Analysis Techniques for the Plays of Gotthold Ephraim Lessing. In: *Proceedings of the Second Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature* (pp. 139–149). Santa Fe, NM: Association for Computational Linguistics.

Schmidt, T., & Burghardt, M. (2018b). Toward a Tool for Sentiment Analysis for German Historic Plays. In: Piotrowski, M. (Ed.), *COMHUM 2018: Book of Abstracts for the Workshop on Computational Methods in the Humanities 2018* (pp. 46–48). Lausanne, Switzerland: Laboratoire laussannois d'informatique et statistique textuelle.

Schmidt, T., Burghardt, M., & Dennerlein, K. (2018a). Sentiment Annotation of Historic German Plays: An Empirical Study on Annotation Behavior. In: S. Kübler, & H. Zinsmeister (Eds.), *Proceedings of the Workshop on Annotation in Digital Humanities (annDH 2018)* (pp. 47–52). Sofia, Bulgaria.

Schmidt, T., Burghardt, M., & Dennerlein, K. (2018b). „Kann man denn auch nicht lachend sehr ernsthaft sein?" – Zum Einsatz von Sentiment Analyse-Verfahren für die quantitative Untersuchung von Lessings Dramen. In *Book of Abstracts, DHd* 2018.

Schmidt, T., Burghardt, M., Dennerlein, K., & Wolff, C. (2019a). Sentiment Annotation in Lessing's Plays: Towards a Language Resource for Sentiment Analysis on

German Literary Texts. In *2nd Conference on Language, Data and Knowledge (LDK 2019). LDK Posters*. Leipzig, Germany.

Schmidt, T., Burghardt, M., Dennerlein, K., & Wolff, C. (2019b). Katharsis – A Tool for Computational Drametrics. In: *Book of Abstracts, Digital Humanities Conference 2019 (DH 2019)*. Utrecht, Netherlands.

Schmidt, T., Burghardt, M., & Wolff, C. (2018). Herausforderungen für Sentiment Analysis-Verfahren bei literarischen Texten. In: Burghardt, M., & Müller-Birn, C. (Eds.), *INF-DH 2018.* Bonn: Gesellschaft für Informatik e.V.

Schmidt, T., Burghardt, M. & Wolff, C. (2019). Towards Multimodal Sentiment Analysis of Historic Plays: A Case Study with Text and Audio for Lessing's Emilia Galotti. In *Proceedings of the DHN (DH in the Nordic Countries) Conference* (pp. 405–414). Copenhagen, Denmark.

Schmidt, T., Hartl, P., Ramsauer, D., Fischer, T., Hilzenthaler, A., & Wolff, C. (2020). Acquisition and Analysis of a Meme Corpus to Investigate Web Culture. In *Digital Humanities Conference 2020 (DH 2020).* Virtual Conference.

Schmidt, T., Jakob, M., & Wolff, C. (2019). Annotator-Centered Design: Towards a Tool for Sentiment and Emotion Annotation. In Draude, C., Lange, M., & Sick, B. (Eds.), *INFORMATIK 2019: 50 Jahre Gesellschaft für Informatik – Informatik für Gesellschaft (Workshop-Beiträge)* (pp. 77–85). Bonn: Gesellschaft für Informatik. DOI: 10.18420/inf2019_ws08

Schmidt, T., Kaindl, F., & Wolff, C. (2020). Distant Reading of Religious Online Communities: A Case Study for Three Religious Forums on Reddit. In *Proceedings of the Digital Humanities in the Nordic Countries 5th Conference (DHN 2020).* Riga, Latvia.

Schmidt, T., Schlindwein, M., Lichtner, K., & Wolff, C. (2020). Investigating the Relationship Between Emotion Recognition Software and Usability Metrics. *i-com, 19*(2), 139–151.

Schmidt, T., Winterl, B., Maul, M., Schark, A., Vlad, A., & Wolff, C. (2019). Inter-Rater Agreement and Usability: A Comparative Evaluation of Annotation Tools for Sentiment Annotation. In Draude, C., Lange, M. & Sick, B. (Eds.), *INFORMATIK 2019: 50 Jahre Gesellschaft für Informatik – Informatik für Gesellschaft (Workshop-Beiträge)* (pp. 121–133). Bonn: Gesellschaft für Informatik. DOI: 10.18420/inf2019_ws12

Shmueli, B., & Ku, L. W. (2019). SocialNLP EmotionX 2019 Challenge Overview: Predicting Emotions in Spoken Dialogues and Chats. *arXiv preprint* arXiv: 1909.07734

Sprugnoli, R., Tonelli, S., Marchetti, A., & Moretti, G. (2016). Towards sentiment analysis for historical texts. *Digital Scholarship in the Humanities*, 31(4), 762–772.

Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M. (2011). Lexicon-based methods for sentiment analysis. *Computational linguistics, 37*(2), 267–307.

Vinodhini, G., & Chandrasekaran, R. M. (2012). Sentiment analysis and opinion mining: a survey. *International Journal of Advanced Research in Computer Science and Software Engineering*, 2(6), 282–292.

Vo, M. L., Conrad, M., Kuchinke, L., Urton, K., Hofmann, M. J., & Jacobs, A. M. (2009). The Berlin affective word list reloaded (BAWL-R). *Behavior Research Methods, 41*(2), 534–538.

Vredenburg, K., Mao, J. Y., Smith, P. W., & Carey, T. (2002). A survey of user-centered design practice. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 471–478). New York, NY: ACM Press.

Waltinger, U. (2010). GermanPolarityClues: A Lexical Resource for German Sentiment Analysis. In *Proceedings of the Seventh International Conference on Language Resources and Evaluation – LREC 2010* (pp. 1638–1642). Dordecht: Springer.

Yavuz, M. C. (2021). Analyses of Character Emotions in Dramatic Works by Using EmoLex Unigrams. In *Proceedings of the Seventh Italian Conference on Computational Linguistics, CLiC-it'20*. Bologna, Italy.

Zhang, L., Wang, S., & Liu, B. (2018). Deep learning for sentiment analysis: A survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 8*(4), e1253.

In: T. Schmidt, C. Wolff (Eds.): Information between Data and Knowledge. Information Science and its Neighbors from Data Science to Digital Humanities. Proceedings of the 16th International Symposium of Information Science (ISI 2021), Regensburg, Germany, 8th—10th March 2021. Glückstadt: Verlag Werner Hülsbusch, pp. 156—172. DOI: doi.org/10.5283/epub.44943.