

# Herausforderungen für Sentiment Analysis bei literarischen Texten

Thomas Schmidt<sup>1</sup>, Manuel Burghardt<sup>2</sup> & Christian Wolff<sup>1</sup>

Medieninformatik, Universität Regensburg<sup>1</sup>

Computational Humanities, Universität Leipzig<sup>2</sup>

## Zusammenfassung

In diesem Beitrag wird über die Ergebnisse eines laufenden Digital Humanities-Projekt zur Sentiment Analysis in literarischen Texten berichtet und die Implikation von diesem diskutiert. In dem Projekt werden verschiedene Methoden der Sentiment Analysis auf Texte historischer Dramen des 18. Jahrhunderts von G. E. Lessing implementiert und gegeneinander evaluiert. Zur Evaluation wurde ein von Menschen bezüglich des Sentiments annotiertes Testkorpus erstellt. Basierend auf den ersten Erfahrungen des Projekts diskutieren wir über Probleme und Herausforderungen, die sich aus der Perspektive der Informatik zur Sentiment Analysis historischer Dramen ergaben. Es wird deutlich, dass bestehende Standardlösungen der Sentiment Analysis für dieses spezifische Szenario nicht ohne Weiteres anwendbar sind. Vielmehr ist die Informatik gefordert, die bestehenden Methoden anzupassen, weiterzuentwickeln und sich mit besonderen Eigenheiten der Textform historischer literarischer Texte auseinanderzusetzen.

## 1 Einleitung

Sentiment Analysis bezeichnet ein Forschungsfeld, bei dem mittels computergestützter Methoden versucht wird, die Meinungen, Gefühle und Emotionen von Menschen zu analysieren und vorauszusagen (Liu, 2016). Bislang befassen sich dabei die meisten Arbeiten mit der Vorhersage und Analyse der *Polarität* eines Textes, also mit der Frage, ob und in welchem Ausmaß ein Text positiv, negativ oder neutral konnotiert ist (Vinodhini & Chandrasekaran, 2012). Zudem finden sich erste Ansätze, bei denen versucht wird, komplexere Emotionen wie Zorn, Überraschung oder Freude zu analysieren (Alm et al., 2005; Mohammad, 2011). In der Informatik ist Sentiment Analysis eines der aktivsten Forschungsfelder der letzten Jahre (Vinodhini & Chandrasekaran, 2012). Die meisten bestehenden Studien und Projekte befassen sich dabei mit Produkt- und Film-Reviews oder Social Media (Liu, 2016).

Methodisch gibt es zahlreiche bestehende Verfahren und Ansätze für die Durchführung von Sentiment Analysis, die sich grob in zwei größere Gruppen klassifizieren lassen (Liu, 2016): In der ersten Kategorie wird die Sentiment Analysis mit Hilfe von Methoden des überwachten maschinellen Lernens (ML) umgesetzt (Liu, 2016, S. 47-57). Für den Einsatz von ML wird zunächst ein Trainingskorpus benötigt, also eine Sammlung adäquater Texteinheiten, die bezüglich des Sentiments korrekt annotiert sind. Die Erstellung solcher Trainingskorpora erfolgt zumeist per Hand (Prabowo & Thelwall, 2009) und ist jeweils maßgeschneidert für die spezifische Anwendungsdomäne. Einen im Vergleich damit eher generischen Ansatz stellt die Nutzung von Sentiment-Lexika dar. Es handelt es sich um Listen von Wörtern, die jeweils mit Sentimentangaben (z. B. *pos / neg*) annotiert sind. Dabei stellen sich domänenspezifische Besonderheiten für lexikonbasierte Ansätze häufig als problematisch heraus, weshalb sie typischerweise weniger leistungsstark als Ansätze aus dem Bereich des maschinellen Lernens sind (Pang et al., 2002).

Da Emotionen und Sentiments wichtige Faktoren narrativer Texte und literaturwissenschaftlicher Interpretation sind (siehe z.B. Winko, 2003; Mellmann, 2015), halten Methoden der Sentiment Analysis in zunehmenden Maße in der computergestützten Literaturanalyse Einzug. Anwendungsgebiete sind dabei bislang vor allem Märchen (Alm et al., 2005; Mohammad, 2011) und Romane (Kakkonen & Kakkonen, 2011; Jockers, 2015). Erste Projekte im Bereich der Dramenanalyse explorieren Sentiment-Verteilungen, -Verläufe und Charakter-Beziehungen in Shakespeares Dramen (Mohammad, 2011; Nalisnick & Baird, 2013). Im Bereich des maschinellen Lernens wurden Ergebnisse von Sentiment Analysis als Features z. B. zur Vorhersage von *happy endings* (Jannidis et al., 2016) oder zur Genrebestimmung (Kim, Padó & Klinger, 2017) genutzt.

Die Mehrzahl der bisherigen Sentiment Analysis-Studien im literaturwissenschaftlichen Kontext nutzt lexikonbasierte Methoden (Mohammad, 2011; Kakkonen & Kakkonen, 2011; Nalisnick & Baird, 2013), die Analyse von Verfahren des maschinellen Lernens sind bislang noch selten (z.B. über Bag-of-words-Modelle bei Kim & Klinger, 2018). Der Grund dafür dürfte sein, dass für den literaturwissenschaftlichen Bereich kaum mit Sentiment-Informationen annotierte Korpora existieren, die als Trainingsdaten für maschinelles Lernen genutzt werden können. Tatsächlich ist die Performanz von ML-Ansätzen im Bereich der Sentiment Analysis meist besser als lexikonbasierte Methoden (Pang et al., 2002). Hinzu kommt, dass im Deutschen grundsätzlich weniger und selten ausreichend evaluierte Ressourcen für lexikonbasierte Sentiment Analysis existieren. Der Mangel annotierter Korpora ist auch ein Problem für die Evaluation von Sentiment Analysis-Methoden. Die Leistung von Sentiment Analysis-Algorithmen wird in der Regel durch den Grad der Übereinstimmung zwischen Ergebnissen der Sentiment Analysis und den Sentiment-Annotationen von Menschen gemessen. Im Bereich literarischer Texte wird die Leistung jedoch bislang mangels Evaluationskorpora meist heuristisch und explorativ beurteilt, indem untersucht wird, ob die Ergebnisse der Sentiment Analysis bekannten literaturwissenschaftlichen Interpretationen entsprechen (Mohammad, 2011; Kakkonen & Kakkonen, 2011; Nalisnick & Baird, 2013; Jockers, 2015).

## 2 Fallstudie: Sentiment Analysis für Dramen

Vor diesem Hintergrund präsentieren wir erste Erfahrungen aus einem interdisziplinären und noch laufenden Projekt zur computergestützten Sentiment Analysis deutschsprachiger Dramen. Dabei wurden im Rahmen einer Evaluationsstudie unterschiedliche Lexika und NLP-Tools untersucht und gegen einen manuell erstellten Gold-Standard abgeglichen (Schmidt & Burghardt, 2018a; Schmidt & Burghardt, 2018b; Schmidt, Burghardt & Dennerlein, 2018a; Schmidt, Burghardt & Dennerlein, 2018b).

Als Analyse-Korpus wurden zwölf Dramen aus dem 18. Jahrhundert von Gotthold Ephraim Lessing gewählt. Bekannte Dramen in diesem Korpus sind *Nathan der Weise* und *Emilia Galotti*. Die Dramen wurden als XML-strukturierte Dateien von der Plattform *TextGrid*<sup>1</sup> bezogen. Um eine systematische Evaluation durchführen zu können, wurde ein begrenztes annotiertes Testkorpus erstellt. Als zentrale Texteinheit wurde die *Replik* als kleinste strukturelle Einheit der Gattung Drama gewählt. Dazu wurden 200 zufällige, aber repräsentative Repliken des Gesamtkorpus gewählt und von sechs Annotatoren unter anderem bezüglich der Polarität annotiert (eher positive oder eher negative Konnotation der Replik). Fünf der Annotatoren waren Laien ohne besondere Kenntnisse über Lessings Werk; einer der Annotatoren ist ein ausgewiesener Experte in diesem Bereich. Über Fragebögen und Interviews wurden weitere Daten zur Gesamterfahrung der Teilnehmer mit der Annotation gemacht. Die Sentiment-Annotation binärer Polarität (positiv/negativ) weist mit einem Kappa-Wert von 0.4 und einer prozentuellen durchschnittlichen paarweise Übereinstimmung von 77% auf eine moderate Übereinstimmung. Diese Übereinstimmungsgrade sind konform mit anderen Annotationsstudien von Emotionen im Bereich literarischer Texte (Alm & Sproat, 2005; Alm et al., 2005; Kim & Klinger, 2018). Andere Anwendungsgebiete in der Sentiment Analysis erreichen deutlich höhere Übereinstimmungsgrade z.B. 0.9 für Film-Reviews (Thet et al., 2010) oder 1.0 für Social Media-Kommentare (Prabowo & Thelwall, 2009). Basierend auf der Mehrheitsentscheidung der Annotatoren wurde jeder Replik eine Annotation zugewiesen. Bei gleichmäßiger Verteilung wurde die Annotation des Experten gewählt. Eine detaillierte Beschreibung des Vorgehens und der Ergebnisse der Annotationsstudie und eine Einordnung in die bestehende Forschung findet man bei Schmidt et al. (2018a). Das Evaluations-Korpus ist online frei verfügbar<sup>2</sup>.

Auf Basis dieses Korpus konnten die zentralen deutschsprachigen Sentiment-Lexika sowie einige NLP-Techniken wie die Nutzung von Stoppwörter, Lemmatisierung und die Erweiterung mit historischen linguistischen Varianten sowie die Zusammensetzung dieser Faktoren untersucht werden. Mit der bestmöglichen Konfiguration konnte eine Erkennungsgenauigkeit von 70% erreicht werden. Dies ist größer als die *random* und *majority baseline*, jedoch etwas geringer als der prozentuelle paarweise Übereinstimmungsgrad der Annotatoren unter einander (77%). Im Bereich der Emotionserkennung belegen Kim und Klinger (2018) für Sätze aus

---

<sup>1</sup> <https://textgrid.de/>

<sup>2</sup> Verfügbar unter:  
[https://docs.google.com/spreadsheets/d/1f72hS2WDRBORxzSY\\_tsM\\_igChG2bvxyTYMVZP6kOnuk/edit?usp=sharing](https://docs.google.com/spreadsheets/d/1f72hS2WDRBORxzSY_tsM_igChG2bvxyTYMVZP6kOnuk/edit?usp=sharing)

literarischen Texten mit lexikonbasierten Verfahren Präzisionswerte von 19%. Insgesamt sind die Ergebnisse deutlich geringer als in anderen Anwendungsgebieten der Sentiment Analysis, bei denen Erkennungsraten zwischen 80 und 95% liegen (Vinodhini und Chandrasekaran, 2012). Dies verdeutlicht die Problematik des Einsatzes der bisher gängigen Sentiment Analysis-Methoden für literarische Texte. Mehr Details zu dieser Evaluationsstudie findet man bei Schmidt und Burghardt (2018a).

Um mögliche Anwendungsfälle der Sentiment Analysis in der quantitativen Dramenanalyse zu explorieren, wurde eine Web-Applikation entwickelt, mit der die Ergebnisse der besten Konfiguration visualisiert werden. Mit diesem Werkzeug können Sentiment-Verteilungen und -Entwicklungen für das ganze Drama, Akte, Szenen, Sprecher und Sprecherbeziehungen untersucht werden. Beispielsweise lässt sich für viele von Lessings Dramen aufzeigen, dass die Polarität der Akte im Verlauf des Dramas zunehmend negativer wird. Auch lassen sich besondere Beziehungen von Figuren bezüglich Sentiment analysieren und im Kontext bisheriger hermeneutischer Annahmen interpretieren.

Abbildung 1 zeigt den Polaritätsverlauf pro Akt für das Drama *Der Freigeist*: Auf der x-Achse sind die einzelnen Akte angegeben, die y-Achse gibt die Werte der Polarität normalisiert an der Anzahl der Wörter pro Akt an. Es ist eine Tendenz hin zu steigender Negativität erkennbar. Weitere Visualisierungen für das Lessing-Korpus kann man über das beschriebene Tool einsehen<sup>3</sup>. Weitere Beispiele der Nutzung dieser Visualisierungen für die literaturwissenschaftliche Interpretation im Kontext von Lessings Dramen findet man bei Schmidt et al. (2018b).

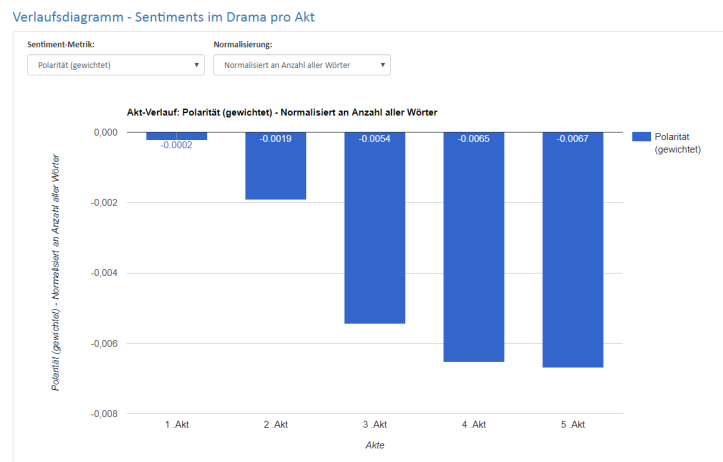


Abbildung 1: Polaritätsverlauf pro Akt für das Drama *Der Freigeist*

<sup>3</sup> [http://lauchblatt.github.io/QuantitativeDramenanalyseDH2015/FrontEnd/sa\\_selection.html](http://lauchblatt.github.io/QuantitativeDramenanalyseDH2015/FrontEnd/sa_selection.html)

### 3 Probleme der Sentiment Analysis in den DH

Die Erkenntnisse des in Kap. 2 beschriebenen Projekts verdeutlichen die Grenzen und Probleme beim Einsatz von Allgemein-Lösungen (hier: bestehende Sentiment-Lexika) aus der Informatik auf Domänen der Geisteswissenschaft (hier: historische Dramen). Neben den technischen Problemen ergeben sich auch nutzer-zentrierte Anpassungsschwierigkeiten aufgrund des Wechsels der bisherigen Hauptdomänen der Sentiment Analysis (Produkt-Reviews, Social Media) auf ein bislang noch weniger erforschtes Anwendungsgebiet.

#### 3.1 Historisches und poetisches Vokabular

Durch die Notwendigkeit des Einsatzes von Sentiment-Lexika ergeben sich einige wortbasierte Probleme für die Sentiment Analysis. Für die korrekte Kalkulation von Sentiment-Werten mittels lexikonbasierter Verfahren ist eine korrekte Identifikation von Sentiment-tragenden Wörtern entscheidend. Im Deutschen sind die existierenden Sentiment-Lexika auf Gegenwartssprache ausgerichtet (z.B. Remus et al., 2010). Dies ist problematisch, da viele ältere literarische Stücke das Vokabular ihrer Zeit verwenden, das sich in vielerlei Hinsicht von der heutigen Standardsprache unterscheiden kann. Ein typisches Problem sind orthographische Veränderungen, durch die Sentiment-tragende Wörter nicht korrekt als solche identifiziert werden können. Einige Beispiele für derartige Wörter aus Lessings Sprachgebrauch sind: *betriegen*, *bös*, *versichren*, *Schwerd*, *verlöre*, *Ehr*, *hämückisch*, *verziehen* (statt verzeihen) oder *sanftres* Gemüt. Ein anderes Problem ist das grundlegende Vokabular, da Sentiment-tragende Wörter verwendet werden, die heutzutage nicht mehr genutzt werden oder archaisch wirken und in zeitgenössischen Sentiment-Lexika nicht enthalten sind z.B. *verdrießlich*, *Nebenbuhler*, *Pöbel*, *Schwachheit*, *Schlingel*, *Schwätzer*, *Pfui* oder *närrisch*. Ein anderes Phänomen sind Wörter, deren Semantik und damit auch ihre Sentiment-Konnotation sich im Verlauf der Zeit verändert haben. Hamilton et al. (2016) konnten bereits für den englischsprachigen Bereich zeigen, dass sich die Bedeutung einer relevanten Anzahl an Wörtern bezüglich ihrer Sentiment-Bedeutung stark ändern kann. Ein Beispiel für ein derartiges Wort im Lessing-Korpus ist das Wort *Freier*, das heutzutage deutlich negativer wahrgenommen wird als zur damaligen Zeit. Auf ähnliche Weise existieren auch Sentiment-Variationen, die eher mit der besonderen Sprache von Lessing zusammenhängen, der gewisse Wörter aus stilistischen und thematischen Gründen so häufig nutzt, dass die heutige Sentiment-Bedeutung nicht mehr zum Tragen kommt. Beispiele dafür sind die Wörter *Herr* und *Fräulein* als höfliche Anreden, mit welchen jedoch zur damaligen Zeit grundsätzlich adlige Personen angesprochen wurden. Ein weiterer Sonderfall sind Sentiment-tragende Wort-Komposita, die mit der vermehrten Verwendung poetischer Sprache zusammenhängen z.B. *Erznarr*, *allerunerträglichst*, *nichtswürdig*, *Bubenstück*, *Galgenschwengel* oder *wohlaffektionierter*.

Wir haben versucht, diesen wortbasierten Problemen über den Einsatz von Lemmatisierung und über die Erweiterung der Lexika mit linguistischen Varianten entgegenzuwirken. Während gängige deutschsprachige Lemmatisierer große Schwierigkeiten mit der historischen Sprache aufzeigen, hat die Lexikon-Erweiterung von allen NLP-Methoden die stärkste Leis-

tungssteigerung erbracht (Schmidt & Burghardt, 2018). Dies zeigt wiederum auf, dass die Verbesserung der korrekten Wort-Erkennung beim Einsatz von lexikonbasierten Verfahren in dieser spezifischen Domäne das zentrale, zu lösende Problem ist. Innerhalb der bisherigen Sentiment Analysis-Forschung gibt es diesbezüglich Versuche, für spezielle Anwendungsgebiete Sentiment Lexika zu erstellen (Hamilton et al., 2016). Auch hier besteht aber weiterhin der Bedarf nach strukturierten textuellen Massendaten der Domäne, die für literarische Texte oft nicht vorliegen. Buechel et al. (2016) greifen die Idee auf, indem sie ein bestehendes Lexikon zur verbesserten Anwendbarkeit für Texte im 18. und 19. Jahrhundert erweitern. Diese Ansätze sind noch nicht vollständig evaluiert und man muss beachten, dass ein großer Unterschied zwischen der historischen Alltagssprache der damaligen Zeit und der poetischen Sprache damaliger Literatur liegt. Des Weiteren ist auch zu beachten, dass die Domäne literarischer Texte breiter und vielseitiger ist als andere Domänen, die bisher in der Sentiment Analysis betrachtet wurden, was ein allgemeines Sentiment-Lexikon für literarische Texte unwahrscheinlich macht. Auch bei der Unterscheidung von Autoren kann man annehmen, dass stilistische Differenzen den effektiven Einsatz domänenspezifischer Lexika erschweren. Neben der Weiterentwicklung domänenspezifischer Sentiment Analysis ist auch der Einsatz hybrider Mischmethoden aus lexikonbasierten Verfahren und ML-Methoden eine vielversprechende noch zu explorierende Technik (z.B. Lalji & Deshmuk, 2015). Dabei werden Ergebnisse von Sentiment-Lexika als Trainingsdaten für ML-Algorithmen genutzt, womit der Mangel an annotierten Korpora teilweise umgangen werden kann.

## 3.2 Akquise annotierter Korpora

Neben den in Kapitel 3.1 beschriebenen Problemen, die spezifisch mit dem Vokabular von literarischen Texten zusammenhängen, werden auch die generellen Probleme lexikonbasierter Verfahren in dieser Domäne nochmal verstärkt: z. B. der Umgang mit Ironie, Sarkasmus, ambige Begriffe und der inhaltliche Kontext. Insgesamt wird bei mangelhaften Ergebnissen mit Sentiment-Lexika der Einsatz von überwachtem maschinellen Lernen empfohlen. Aufgrund der Schwierigkeit der Akquise annotierter Korpora ist dieser jedoch bislang kaum möglich. Dieser grundsätzliche Lösungsweg der Informatik kann in dieser Domäne nicht direkt eingeschlagen werden. Die Gründe dafür sind vielseitig: In bisherigen Annotationsstudien konnte festgestellt werden, dass Sentiment-Annotationen in diesem Bereich sehr subjektiv sind. Die Heterogenität in der Annotation ist deutlich größer, die Übereinstimmungsgrade sind sehr viel geringer. Teilnehmer berichten von einem starken Gefühl der Unsicherheit bei der Annotation. In einer aktuell laufenden Studie mit Germanistikstudenten wurde ein Annotationsschema genutzt, bei dem Teilnehmer die Sicherheit bei der Annotation angeben konnten. Dies hat, so deren Rückmeldung, die Annotation deutlich vereinfacht. Die Modellierung von Vagheit und Unsicherheit steht im Gegensatz zu den bisher vorherrschenden Ansätzen, bei denen eine eindeutige Polarität bestimmt werden soll. Unter anderem durch diesen Bedarf aus den Geisteswissenschaften lassen sich momentane Trends zur Modellierung von Vagheit und Unschärfe in der Informatik beispielsweise über Graph-Datenbanken und *Fuzzy Logic* (z.B. Castelltort & Laurent, 2014) erkennen. Im Bereich der Sentiment Analysis versuchen Hammer et al. (2015) Sentiment-tragende Wörter nicht mehr rein nominal, sondern als Ausdruck einer stochastischen Formel zu modellieren. Weitere Entwicklungen dieser Art, mit der Sentiment Analysis

unsichere Information zu modellieren, scheinen notwendig für den Anwendungsfall literarischer Texte.

Durch das Feedback der Annotationsstudie konnte auch gezeigt werden, dass bisherige Annotationsschemata der Sentiment Analysis, die fokussiert auf Polarität (positiv/negativ/neutral) und ein klares Ziel (z.B. das Produkt, der Film etc.) sind, nicht passend für Annotation im literarischen Bereich sind. Die Arten von potentiellen Zielen können dabei zu vielseitig sein und sind oft nicht klar erkennbar. Auch hier hängen diese stark von der Interpretation des Annotators ab. Auch die Polarität hat sich als häufig nicht ausreichend herausgestellt, da so komplexere Phänomene wie Emotion, Ironie oder Täuschung nicht behandelt werden können. Um diesen Problemen zu begegnen muss man sich aus literaturwissenschaftlicher Sicht mit Theorien und Ideen zur Modellierung von Emotionen in narrativen Texten auseinandersetzen. Kim und Klinger (2018) beziehen sich auf literaturwissenschaftliche Modelle um ein differenzierteres Annotationsschema zu kreieren und ein somit ein präziser annotiertes Korpus zu erstellen. Die praktische Informatik jedoch kann sich diesem Problem mit Methoden des Usability- und Requirements-Engineering annähern und zusammen mit Literaturwissenschaftlern über Methoden wie *Interviews*, *Fokusgruppen* oder *Contextual Inquiry* weitere Lösungen zur Modellierung erarbeiten. Dabei werden gleichzeitig die speziellen Bedürfnisse dieser Nutzergruppe gegenüber der Sentiment Analysis spezifiziert (welche Fragestellungen sind für Literaturwissenschaftler erkenntnisleitend, welche Informationsbedürfnisse haben sie, die an die Dramen herangetragen werden können). Auf diese Weise kann die Informatik jenseits der literaturwissenschaftlichen Theorie wichtige Erkenntnisse zur erfolgreichen Modellierung und Implementierung der Sentiment Analysis in literarischen Texten akquirieren.

## 4 Fazit

Die Schwierigkeiten bei der Durchführung von Sentiment Analysis auf literarischen Texten und bei der Akquise von Trainingskorpora lassen darauf schließen, dass jenseits der angesprochenen Probleme eine grundlegende Diskrepanz zwischen den bisher vorherrschenden Textformen in der Sentiment Analysis (Wirtschaftskommunikation, soziale Medien) und literarischen Texten vorliegt. Die Informatik sollte beachten, dass die Unterschiede zwischen diesen Texten und Sachtexten auf strukturell unterschiedliche Kommunikationssituationen zurückzuführen sind, die den Kern des Literarischen als künstlerisch gestalteter Sprache berühren und entsprechend berücksichtigt werden müssen. Die offenkundige Schwierigkeit, einen literarischen Text hinsichtlich der Polarität „eindeutig“ zu notieren, ist ein starker Hinweis darauf, dass literarische Texte hinsichtlich des Sentiments keine eindeutige oder einmalige Interpretation benötigen. Insofern ist weitergehend zu fragen, ob nicht zusätzliche Informationsquellen eine interpretationsspezifische Annotation von Sentiment möglich machen. Im Fall des Dramas zum Beispiel fehlt bei einer textfokussierten Analyse der dem Drama innewohnende performative Aspekt gesprochener Sprache. Man könnte beispielsweise der Sentiment Analysis eines Dramentextes die prosodische Analyse einer spezifischen Aufführung des Dramas an die Seite stellen. Literarische Texte sind auch gestaltete Sprache und beim Drama liegt eine andere, vermittelte Kommunikationssituation zugrunde, da der Schauspieler den Text eines Au-

tors (auf ganz unterschiedliche Weise) interpretiert. Damit sind Vermittlungsebenen eingezogen, die der Unmittelbarkeit einer Meinungsäußerung zum Beispiel im Rahmen einer Produktrezension grundsätzlich fehlt. Hier besteht die Chance, poetologische und literaturtheoretische Theorien für die Weiterentwicklung technischer Verfahren der Sentiment Analysis fruchtbar zu machen.

## 5 Literaturverzeichnis

- Alm, C. O., Roth, D., & Sproat, R. (2005). Emotions from text: machine learning for text-based emotion prediction. In *Proceedings of the conference on human language technology and empirical methods in natural language processing* (pp. 579-586). Association for Computational Linguistics
- Buechel, S., Hellrich, J., & Hahn, U. (2016). Feelings from the Past – Adapting Affective Lexicons for Historical Emotion Analysis. In *Proceedings of the Workshop on Language Technology Resources and Tools for Digital Humanities (LT4DH)* (pp. 54-61).
- Castellort, A., & Laurent, A. (2014). Fuzzy queries over NoSQL graph databases: perspectives for extending the cypher language. In *International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems* (pp. 384-395). Springer, Cham.
- Hamilton, W. L., Clark, K., Leskovec, J., & Jurafsky, D. (2016). Inducing domain-specific sentiment lexicons from unlabeled corpora. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing* (Vol. 2016, p. 595). NIH Public Access.
- Hammer, H., Yazidi, A., Bai, A., & Engelstad, P. (2015). Building domain specific sentiment lexicons combining information from many sentiment lexicons and a domain specific corpus. In *IFIP International Conference on Computer Science and its Applications* (pp. 205-216). Springer, Cham.
- 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.
- Jockers, M. L. (2015). *Revealing sentiment and plot arcs with the syuzhet package*. Retrieved from <http://www.matthewjockers.net/2015/02/02/syuzhet/>
- Kakkonen, T. & Kakkonen, G. G. (2011). SentiProfiler: creating comparable visual profiles of sentimental content in texts. In *Proceedings of Language Technologies for Digital Humanities and Cultural Heritage* (pp. 62-69).
- Kim, E. & Klinger, R. (2018). 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* (pp. 1345-1359). Association for Computational Linguistics.
- Kim, E., Padó, S., & Klinger, R. (2017). Prototypical Emotion Developments in Literary Genres. In *Proceedings of the Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature* (pp. 17–26).
- Lalji, T., & Deshmukh, S. (2016). Twitter Sentiment Analysis Using Hybrid Approach. *International Research Journal of Engineering and Technology*, 3(6), 2887-2890.
- Liu, B. (2016). *Sentiment Analysis. Mining Opinions, Sentiments and Emotions*. New York: Cambridge University Press.



- Mellmann, K. (2015). Literaturwissenschaftliche Emotionsforschung. In: Rüdiger Zymner (Hg.): *Handbuch Literarische Rhetorik*. Berlin/Boston, 173-192.
- 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.
- 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* (pp. 479-483).
- Pang, B., Lee, L., & Vaithyanathan, S. (2002). Thumbs up?: sentiment classification using machine learning techniques. In *Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10* (pp. 79-86). Association for Computational Linguistics.
- Prabowo, R., & Thelwall, M. (2009). Sentiment analysis: A combined approach. *Journal of Informetrics*, 3(2), 143-157.
- Remus, R., Quasthoff, U. & Heyer, G. (2010). SentiWS-A Publicly Available German-language Resource for Sentiment Analysis. In *LREC* (pp. 1168-1171).
- Schmidt, T. & Burghardt, M. (2018a). An Evaluation of Lexicon-based Sentiment Analysis Techniques for the Plays of Gotthold Ephraim Lessing. In: *SIGHUM Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities (LaTeCH-CLL 2018)* (pp. 139-149). Retrieved from <http://aclweb.org/anthology/W18-4516>
- 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 lausannois d'informatique et statistique textuelle. Retrieved from <https://epub.uni-regensburg.de/37575/>
- Schmidt, T., Burghardt, M. & Dennerlein, K. (2018a). Sentiment Annotation of Historic German Plays: An Empirical Study on Annotation Behavior. In: Sandra Kübler, Heike Zinsmeister (eds.), *Proceedings of the Workshop on Annotation in Digital Humanities (annDH 2018)* (pp. 47-52). Sofia, Bulgaria. Retrieved from <http://ceur-ws.org/Vol-2155/schmidt.pdf>
- 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. Retrieved from <https://epub.uni-regensburg.de/37579/>
- Thet, T. T., Na, J. C., & Khoo, C. S. (2010). Aspect-based sentiment analysis of movie reviews on discussion boards. *Journal of information science*, 36(6), 823-848.
- 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.
- Winko, S. (2003). Über Regeln emotionaler Bedeutung in und von literarischen Texten. In: Fotis Jannidis & Gerhard Lauer & Matias Martinez & SW (eds.): *Regeln der Bedeutung. Zur Theorie der Bedeutung literarischer Texte*. Berlin, New York: de Gruyter, 329-348.