

Aktuelle Herausforderungen der Digitalen Dramenanalyse

Willand, Marcus

marcus.willand@ilw.uni-stuttgart.de
Universität Stuttgart

Trilcke, Peer

trilcke@uni-potsdam.de
Universität Potsdam

Schöch, Christof

christof.schoech@uni-wuerzburg.de
Universität Würzburg

Rißler-Pipka, Nanette

nanette.rissler-pipka@ku.de
Katholische Universität Eichstätt-Ingolstadt

Reiter, Nils

nils.reiter@ims.uni-stuttgart.de
Universität Stuttgart

Fischer, Frank

ffischer@hse.ru
Higher School of Economics, Moskau

Zielstellung und Konzeption

Das hier vorgeschlagene Panel greift mit der Digitalen Dramenanalyse einen sich derzeit dynamisch entwickelnden Bereich der digitalen Literaturwissenschaften auf. Es setzt sich erstens zum Ziel, aktuelle Herausforderungen der Digitalen Dramenanalyse auf verschiedenen Ebenen vorzustellen, wobei insbesondere die Ebenen der dramatischen Gattung, der Netzwerkstrukturen und der dramatischen Figuren im Zentrum stehen werden. Zweitens möchte das Panel mit dem Publikum mögliche Lösungsansätze diskutieren, unter anderem durch Bezug auf vielfältige, vorhandene Erfahrungen mit der Analyse narrativer Texte. In der Summe wird das Panel einerseits eine Zwischenbilanz zum Stand der Forschung anbieten, andererseits auch im Sinne einer Konsolidierung des Forschungsfelds eine Agenda für die weitere Entwicklung formulieren, bei der es nicht zuletzt darum geht, Szenarien einer integrativen, mithin diverse methodische Ansätze synergetisch zusammenführenden Forschung, zu diskutieren.

Dazu wird das Panel eine Art Laborsituation fingieren, in der die Erkenntnisziele, Möglichkeiten

und Grenzen unterschiedlicher methodischer Zugänge zu dem titelgebenden Forschungsbereich der digitalen Dramenanalyse zu Tage treten sollen: *Topic Modeling*, (soziale) Netzwerkanalyse und Analyse der Figurenrede. Diese Gegenüberstellung soll es dem Publikum erlauben, Grundannahmen und Perspektivierungen der jeweiligen Ansätze direkt zu identifizieren und in der Diskussion adressieren zu können. Welche Modellierung des dramatischen Textes liegt der Methode zugrunde? Welche Aspekte eines dramatischen Textes werden durch die jeweilige Methode eigentlich beobachtet, und welche nicht? Und welche Art von Aussagen macht sie möglich?

In den *digital literary studies* wird zwar häufig Methodenkritik geäußert, dies in der Regel aber nur mit Blick auf einzelne Methoden, auch wenn diese auf ganz unterschiedliche Forschungsgegenstände angewandt werden können. Dieses Vorgehen möchte das Panel invertieren, indem nicht eine Methode auf unterschiedliche Objekte, sondern unterschiedliche Methoden auf das gleiche Objekt angewandt werden: Digitalisierte Dramen zwischen 1700 und 1900 aus dem deutsch- und französischsprachigen Raum. Erreicht werden soll durch dieses Vorgehen eine systematische Aufarbeitung der Möglichkeiten einer methodisch reflektierten digitalen Dramenanalyse, die zugleich theoretische und methodologische Grundfragen der digitalen Analyse literarischer Texte im Allgemeinen thematisiert.

In drei Kurzvorstellungen sollen die folgenden Methoden von jeweils einer Forschergruppe des Panels vorgestellt werden, wobei zur besseren Vergleichbarkeit der drei methodisch unterschiedlich aufgestellten Arbeitsgruppen zeitlich und gattungsbezogen vergleichbare Textsammlungen analysiert werden. Zwar werden diese Verfahren jeweils anhand eines individuellen Teilkorpus vorgestellt, es ist jedoch zu berücksichtigen, dass sie alle auf der statistischen Analyse größerer Textmengen basieren.

Panelvorträge

Topic Modeling und Gattung (Christof Schöch, Nanette Rißler-Pipka)

Der Einsatz von *Topic Modeling* (Blei 2012) für die im weitesten Sinne inhaltliche Erschließung großer Sammlungen literarischer Texte zeigt zwei Dinge: Erstens sind die erzielten Ergebnisse, insbesondere die jeweils dominanten Typen der *Topics*, textsortenabhängig (Schöch 2016). So sind nicht-fiktionale expositorische Texte (bspw. Pressemitteilungen) durch abstrakte thematische *Topics* geprägt, fiktionale Erzähltexte (bspw. Romane) aber durch *Topics*, die sich auf narrative und deskriptive Motive beziehen. Auch dramatische Texte zeichnen sich durch ein eigenes Profil solcher Typen von *Topics* aus, in dem diskursive und metadiskursive *Topics* eine besondere

Rolle spielen. Dieser Umstand schärft auch den Blick auf die spezifische, textuelle Funktionsweise der jeweiligen Gattung und literarischer Texte insgesamt.

Zweitens zeigt sich, dass sich einzelne dramatische Untergattungen wie Tragödie, Komödie oder Tragikomödie zwar in Bezug auf die jeweils dominanten Einzel *topics* unterscheiden (und beispielsweise jeweils ein unterschiedlich strukturiertes *Liebes-Topic* haben können). Zugleich fördert *Topic Modeling* aber keine scharfen Trennungslinien zu Tage, sondern zeigt auf, wie prototypisch gedachte Untergattungen in der Praxis unscharf ineinander übergehen können (Schöch, im Erscheinen). Beide genannten Phänomene sind bekannt, aber sowohl in methodischer bzw. informatischer als auch in literaturtheoretischer Perspektive derzeit nicht ausreichend klar erfasst und damit auch nicht empirisch überprüfbar.

Netzwerkanalyse (Frank Fischer, Peer Trilcke)

Die in den quantitativen Sozialwissenschaften entwickelten Verfahren der Netzwerkforschung zielen auf eine formale Analyse sozialer Strukturen (Wasserman / Faust 1994). Angewandt auf literarische Texte ermöglichen sie Strukturbeschreibungen, die aus einer signifikant anderen Perspektive erfolgen als traditionelle literaturwissenschaftliche Verfahren der semantikbasierten Strukturanalyse (z.B. Titzmann 1977), insofern sie nicht die semantische Organisation literarischer Texte, sondern die ästhetische Modellierung sozialer Formationen im Medium der Literatur analysieren (Trilcke 2013). Ob ihres stark formalisierten Charakters operieren netzwerkanalytische Konzeptualisierungen dabei zunächst mit epistemischen Objekten, die sich erheblich von den Objekten der »klassischen« Literaturwissenschaft unterscheiden. Gerade deshalb aber bilden solche Konzeptualisierungen ein ebenso attraktives wie kontroverses Experimentierfeld für computerbasierte Zugänge zum Gegenstandsbereich »Literatur«, die nicht nur neue Antworten auf alte Fragen finden, sondern dezidiert *andere* Fragen formulieren wollen. Diese Ausrichtung wird noch unterstützt durch die *distant reading*-Affinität der literaturwissenschaftlichen Netzwerkanalyse: Zwar lässt sich die visuelle Auswertung in Form von statischen oder dynamischen Netzwerkgraphen noch im Sinne des »traditionellen« Paradigmas der Einzeltextanalyse verwenden (vgl. Moretti 2011); insbesondere die Auswertung von Netzwerkdaten mittels statistischer Methoden zielt jedoch auf die vergleichende Analyse größerer Korpora, die im Bereich der digitalen Dramenanalyse etwa mit historiographischem (Fischer et al. 2015) oder typologisierendem (Trilcke et al. 2016) Erkenntnisinteresse betrieben wird. Der hohe Abstraktionsgrad insbesondere der statistischen Ergebnisse von literaturwissenschaftlichen Netzwerkanalysen sowie deren Korpusorientierung führen allerdings zu einer Spannung zu »traditionellen« Analyse- und Interpretationspraktiken der Literaturwissenschaft,

mit denen die Ergebnisse der Netzwerkanalyse auf den ersten Blick schwer zu vermitteln sind. Hier zeigen sich gleichermaßen die Gefahren (ein Transfer der Ergebnisse zwischen digitalen Methoden und »traditioneller« Literaturwissenschaft wird unmöglich) wie die Potenziale (die »andersartigen« Ergebnisse der digitalen Methoden führen zur produktiven Irritationen der »traditionellen« Literaturwissenschaft) der Methode, die in diesem Einzelvortrag anhand der Netzwerkanalyse von Dramen aus dem *dlina*-Korpus exemplarisch diskutiert werden sollen.

Analyse der Figurenrede (Nils Reiter, Marcus Willand)

Computerlinguistische Methoden wie Named Entity Recognition und Koreferenzresolution (cf. Poesio et al. 2016) erlauben die Erkennung von Figurenreferenzen in der Rede dramatischer Figuren. Die erkannten Referenzen wiederum können genutzt werden, um den Stellenwert einer Figur innerhalb des Gesamttextes zu identifizieren. Neben der direkten Präsenz von Figuren (im Sinne von: Figur spricht; siehe auch das Problem der sog. Konfiguration, hierzu Ilsemann 1995, 2008) lässt sich damit auch die indirekte Präsenz (über eine Figur wird gesprochen) messen.

Im Falle von *Miss Sara Sampson* und *Emilia Galotti* (Lessing 1755, 1772) unterscheiden sich die beiden titelgebenden Figuren – Sara und Emilia – hinsichtlich dieser Dimensionen: Während Sara den größten Redeanteil auf sich vereinigt, spricht Emilia weniger als halb so viel (relativiert für die Länge des Gesamttextes). Im Gegensatz dazu wird *über* Emilia viel öfter gesprochen, so dass sie sozusagen passive Bühnenpräsenz zeigt. Anhand von Figuren wie dem *König* zeigt sich, dass auch passive Figuren die dramatische Handlung beeinflussen können. Dies gilt auch für Figuren und figurenähnliche Entitäten, die nicht in den *Dramatis Personae* genannt werden (*Gott*, das *Volk*).

Unser Beitrag zum Panel diskutiert zum einen die Herausforderungen an die maschinelle Sprachverarbeitung, wenn sie auf Dramentexte angewendet wird (Blessing et al. 2016). Zum anderen wollen wir untersuchen, inwiefern Autorinnen und Autoren sprachliche Eigenheiten der Figuren nutzen, um diese zu charakterisieren und z.B. als bestimmten Figurentypus (zärtlicher Vater, Hanswurst usw.; cf. Sørensen 1984, Aust 1989, Kord 2009) zu kennzeichnen.

Bilanzierung, Konsolidierung, Agenda

Die unterschiedlichen methodischen Zugänge zu dramatischen Texten erlauben zwar eine direkte Gegenüberstellung und Diskussion der drei

Forschungsansätze, ihrer Prämissen, aber auch der Relevanz ihrer Ergebnisse für literaturtheoretische oder -historische Fragestellungen. Die vorgestellten Verfahren sollen letztlich aber nicht als konkurrierend oder unverbunden gedacht werden, sondern als Beiträge zu einem gemeinsamen Ziel: dem differenzierteren literaturwissenschaftlichen Verständnis dramatischer Texte. Vor dem Hintergrund der das Panel leitenden Idee einer Bilanzierung bisheriger und Konsolidierung aktueller Forschung auf dem Gebiet der *Digitalen Dramenanalyse* könnten ausgehend von den Einzelbeiträgen daher folgende Fragen diskutiert werden:

Jede der drei Methoden verfolgt spezifische Fragen und birgt spezifische Herausforderungen. In welchem Maße gibt es gemeinsame Forschungsziele, zu denen jede der Methoden einen Beitrag leisten kann? Können die verschiedenen Methoden beispielsweise einen Beitrag zu einer empirisch gesicherten Gattungsdifferenzierung oder für die literaturgeschichtliche Periodisierung leisten?

Wie können Ergebnisse, die mit unterschiedlichen methodischem Vorgehen gewonnen wurden, in Bezug zueinander gesetzt werden?

Welche Ressourcen (insbesondere Textsammlungen) liegen vor und wie kann die Verfügbarkeit geeigneter Ressourcen für die Digitale Dramenanalyse zukünftig verbessert werden? Wie können die teils unterschiedliche Anforderungen der Methoden an die Formate von Daten und Metadaten aufgefangen werden?

Welche konzeptuellen und datenbezogenen Standards für dokumentbezogene Metadaten und strukturelle oder semantische, lokale Annotationen liegen vor, wie kann die Standardisierung (bspw. durch Annotationsrichtlinien) weiter gefördert werden?

Welche Tools sind für die digitale Dramenanalyse derzeit verfügbar, wie könnte die Tool-Entwicklung zielgerichtet gefördert werden? Welche generischen Tools könnten produktiv eingesetzt werden, wie könnte der Einsatzbereich vorhandener Tools erweitert (Adaptierbarkeit, Übertragbarkeit) und so eine breitere Nutzerbasis geschaffen werden?

Indem das Panel die Vielfalt digitaler Dramenanalysen vorführt und die explorative Kraft methodischer Innovation durch die Digital Humanities für die Literaturwissenschaften betont, möchten wir die fingierte "Laborsituation" im Sinne der theoretischen und wissenschaftspolitischen Implikationen einer auf Überprüfbarkeit und Wiederholbarkeit angelegten Wissenschaft verstanden wissen.

Fußnoten

1. <https://dlina.github.io/Introducing-DLINA-Corpus-15-07-Codename-Sydney/>
2. <https://quadrama.github.io/blog/2016/10/07/ottokar-capulet>

Bibliographie

Aust, Hugo (1989): *Volksstück vom Hanswurstspiel zum sozialen Drama der Gegenwart*. München: Beck.

Blei, David M. (2012): „Probabilistic Topic Models“, in: *Communication of the ACM* 55 (4): 77–84 [10.1145/2133806.2133826](https://doi.org/10.1145/2133806.2133826).

Blessing, Andre / Bockwinkel, Peggy / Reiter, Nils / Willand, Marcus (2016): „Dramenwerkbank: Automatische Sprachverarbeitung zur Analyse von Figurenrede“, in: *DHd 2016: Modellierung - Vernetzung - Visualisierung* 281–284 <http://dhd2016.de/boa.pdf> [letzter Zugriff 24. August 2016].

Fischer, Frank / Göbel, Mathias / Kampkaspar, Dario / Trilcke, Peer (2015): „Digital Network Analysis of Dramatic Texts“, in: *DH2015: Global Digital Humanities* http://dh2015.org/abstracts/xml/FISCHER_Frank_Digital_Network_Analysis_of_Dramati/FISCHER_Frank_Digital_Network_Analysis_of_Dramatic_Text.html [letzter Zugriff 24. August 2016].

Ilseman, Hartmut (1995): „Computerized Drama Analysis“, in: *Literary and Linguistic Computing* 10 (1): 11–21.

Ilseman, Hartmut (2008): „More statistical observations on speech lengths in Shakespeare’s plays“, in: *Literary and Linguistic Computing* 23 (4): 397–407.

Kord, Susanne (2009): „Unmöglichkeiten. Vater-Tochter-Dramen im 18. und 19. Jahrhundert“, in: Martinec, Thomas / Nitschke, Claudia (eds.): *Familie und Identität in der deutschen Literatur*. Frankfurt am Main: Peter Lang 105–126.

Moretti, Franco (2011): „Network Theory, Plot Analysis“, in: *Stanford Literary Lab Pamphlets* 2 <http://litlab.stanford.edu/LiteraryLabPamphlet2.pdf> [letzter Zugriff 24. August 2016].

Poesio, Massimo / Stuckardt, Roland / Versley, Yannick (2016): *Anaphora Resolution: Algorithms, Resources, and Applications*. Berlin / Heidelberg: Springer.

Schöch, Christof (2016): „What Are Literary Topics, Really?“, in: *Digital Humanities Lunch*. Krakau, Institut für Polnische Sprache, 8. April 2016 <http://christofs.github.io/literary-topics/#/> [letzter Zugriff 24. August 2016].

Schöch, Christof (im Erscheinen): „Topic Modeling Genre: An Exploration of French Classical and Enlightenment Drama“, in: *Digital Humanities Quarterly*. (Preprint): <https://zenodo.org/record/166356> [letzter Zugriff 15. November 2016].

Sørensen, Bengt Algot (1984): *Herrschaft und Zärtlichkeit der Patriarchalismus und das Drama im 18. Jahrhundert*. München: C.H. Beck.

Titzmann, Michael (1977): *Strukturelle Textanalyse: Theorie und Praxis der Interpretation*. München: W. Fink.

Trilcke, Peer / Fischer, Frank / Göbel, Mathias / Kampkaspar, Dario (2016): „Theatre Plays as ‚Small Worlds‘? Network Data on the History and Typology of German Drama, 1730–1930“, in: *Digital Humanities 2016*:

Conference Abstracts <http://dh2016.adho.org/abstracts/360>
[letzter Zugriff 24. August 2016].

Trilcke, Peer (2013): „Social Network Analysis (SNA) als Methode einer textempirischen Literaturwissenschaft“, in: Ajouri, Philip / Mellmann, Katja / Rauen, Christoph (eds.): *Empirie in der Literaturwissenschaft*. Münster: mentis 201–247.

Wasserman, Stanley / Faust, Katherine (1994): *Social Network Analysis: Methods and Applications*. New York: Cambridge University Press.