

Textanalyse mit kombinierten Methoden – ein konzeptioneller Rahmen für reflektierte Arbeitspraktiken

Kuhn, Jonas

jonas.kuhn@ims.uni-stuttgart.de
Universität Stuttgart, Deutschland

Pichler, Axel

axel.pichler@ts.uni-stuttgart.de
Universität Stuttgart, Deutschland

Reiter, Nils

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

Viehhauser, Gabriel

gabriel.viehhauser-mery@ilw.uni-stuttgart.de
Universität Stuttgart, Deutschland

Einleitung

Die Zahl und Intensität der Aktivitäten im interdisziplinären Spektrum der *Digital Humanities* (DH) bzw. *Computational Social Science* ist in den vergangenen 5-10 Jahren enorm angewachsen. Nicht zuletzt dank eines undogmatischen Selbstverständnisses der Forschungscommunity basieren die Aktivitäten auf einem Methodenpluralismus, dem eine ebenso facettenreiche Methodenreflexion entspricht. Gleichwohl ist es für die wissenschaftliche Praxis unerlässlich, dass sich Teilcommunities, die ein vergleichbares Erkenntnisinteresse innerhalb des DH-Spektrums verfolgen, über den konzeptuellen Rahmen sowie die aus ihm folgenden Maßstäbe verständigen, die sie an ein methodisch valides Vorgehen anlegen.

Forschung mit kombinierte komputationell/geisteswissenschaftlichen Methoden

Dieser Beitrag fokussiert auf denjenigen Teilbereich der DH, der sich zum Ziel setzt, adaptierbare datenorientierte Computermodelle methodisch adäquat in kombiniert komputationell/geisteswissenschaftliche

Arbeitspraktiken zu integrieren. Methodologisch zielt diese Teildisziplin also darauf ab, Forschungsfragen aus einem geisteswissenschaftlichen Kontext mit kombinierten Methoden (bzw. mit “mixed methods”) adäquat bearbeiten zu können.¹ Basierend auf eigenen Erfahrungen und dem Austausch innerhalb der DH-Community ist unsere Einschätzung, dass Mitglieder von Forschungsteams, die im Spektrum der komputationell/geisteswissenschaftlichen Methodenkombination eingehende Projekterfahrung gesammelt haben, eine ausdifferenzierte Wahrnehmung der zu kombinierenden Arbeitspraktiken entwickelt haben. Ein methodologischer Metadiskurs, der eine enge Verzahnung der kombinierten Methoden thematisiert, findet jedoch nur in engen Zirkeln – häufig projektintern – statt. Für neu etablierte Projektkooperationen ist es daher nach wie vor schwierig, Workflows aufzusetzen, die ein reflektiertes Vorgehen garantieren. Zudem wird für unterschiedliche methodische Komponenten die jeweilige Adäquatheit häufig nicht auf dem gleichen Reflexionsniveau diskutiert: so kann für die “komputationelle” Praxis, das Modellverhalten im Rahmen einer Evaluation und Fehleranalyse zu reflektieren, auf klarere Konventionen aufgebaut werden als etwa für ein Hinterfragen der Annahmen zum Interpretationskontext der untersuchten Texte, die mit der Operationalisierung zentraler Analysekatégorien einhergehen.

Es erscheint uns daher an der Zeit, intensiver über einen geeigneten konzeptionellen Rahmen für Arbeiten aus einem der DH-Teilbereiche zu diskutieren, in denen kombinierte Methoden zum Einsatz kommen – einen Rahmen, der eine gleichermaßen adäquate Reflexion für alle einfließenden Vorannahmen ermöglicht und zudem einfach genug darstellbar ist, dass sich ein methodisch adäquater Workflow mit vertretbarem Aufwand und ohne Brüche konstruieren lässt.

Vier Aspekte des konzeptionellen Rahmens für die Methodenkombination

In diesem Beitrag stellen wir Kernpunkte eines generalisierten arbeitspraktischen Vorgehensmodells dar, das wir aus den Erfahrungen des Stuttgarter *Zentrums für reflektierte Textanalyse* (CRETA²) heraus entwickelt haben, in dem Beteiligte aus einer Reihe von unterschiedlichen Textwissenschaften und Computerwissenschaften kooperieren. Wir beschränken den Forschungsgegenstand auf Texte, öffnen den Raum aber für sehr unterschiedliche Arten von Fragestellungen: eine literaturwissenschaftliche Auseinandersetzung mit (poetischen) Primärtexten soll ebenso abgedeckt werden wie die Analyse von wissenschaftlichen Diskursen (etwa in der Philosophie) oder von Texten als Quellen für historische oder sozialwissenschaftliche Untersuchungen. Eine paradigmatische Auswahl erster Forschungsergebnisse

der am Zentrum angesiedelten Projekte bietet der Band Kuhn/Pichler/Reiter (erscheint).³

Die Motivation für den vorgeschlagenen konzeptionellen Rahmen liegt nicht vordringlich in einer deskriptiven wissenschaftstheoretischen Betrachtung. Vielmehr soll der Rahmen Ansatzpunkte für die konkrete Praxis bieten – etwa für *Best-Practice*-Vorschläge. Wir fokussieren auf vier ineinandergreifende Aspekte, die für die DH zentral sind. Keiner dieser Aspekte ist grundsätzlich neu für die Methodendiskussion – hier geht es jedoch um eine handhabbare Gesamtkonzeption.

Als konkrete Illustration des Vorgehens mögen Arbeiten aus dem QuaDramA-Projekt dienen (Krautter/Pagel 2019, Krautter et al. 2018): ein Korpus von deutschsprachigen Dramen wird mit kombinierten Methoden erschlossen; ein exemplarischer Analyseschritt dabei liegt in der Klassifikation von Figuren nach bestimmten Typen. Einige Figurentypen sind bereits literaturwissenschaftlich etabliert (zärtlicher Vater) oder lassen sich relativ treffsicher aus der Figurentafel (Tochter) oder den Metadaten (Titelfigur) extrahieren. Andere Typen wie z.B. die Protagonistin bzw. der Protagonist entziehen sich einer aus unmittelbar verfügbaren Texteigenschaften oder Metadaten ableitbaren Zuweisung, sind jedoch von Bedeutung für literaturhistorische Betrachtungen (etwa für die Frage, inwieweit Emilia Galotti als Titelfigur aus G. E. Lessings bürgerlichem Trauerspiel (1772) den Status einer Protagonistin hat). Eine Annäherung an derartige Kategorien mit kombinierten Methoden kann ausgehend von klaren Fällen eine vorläufige Operationalisierung ansetzen, darauf aufbauend datenbasierte Computermodelle erzeugen und die Modellvorhersagen auf dem Gesamtkorpus in den Prozess einer Verfeinerung der Operationalisierung einfließen lassen.

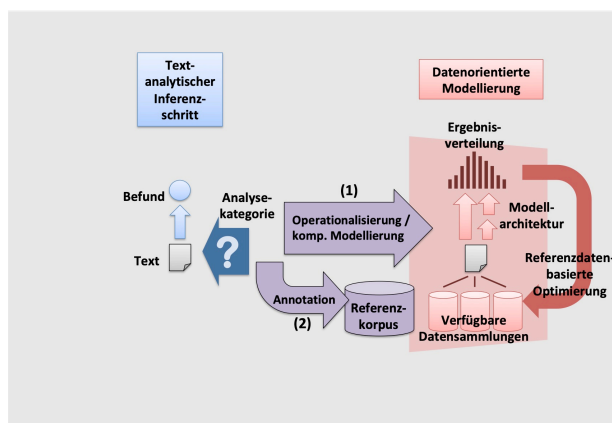


Abbildung 1: Typisches Vorgehen bei der komputationellen Modellierung nicht-trivialer Kategorien in der DH-Textanalyse

Als Ausgangspunkt skizziert Abbildung 1 ein DH-Vorgehen, das sich bei nicht-trivialen

Modellierungsaufgaben etabliert hat – in Anlehnung an ausgeprägte methodische Konventionen in der Korpus- und Computerlinguistik (vgl. u.a. Hovy/Lavid 2010, Kuhn/Reiter 2015, Stefanowitsch 2018, Kuhn 2019): Die Analysekategorie, die im Rahmen einer geisteswissenschaftlichen Gesamtfragestellung angewendet werden soll, wird konzeptuell operationalisiert – gängiger Weise in Form von präzisen Annotationsrichtlinien. Der erste zentrale Aspekt für eine effektive Praxis der Methodenkombination liegt in der **(I) Anwendung der Annotationsrichtlinien auf ein geeignetes Korpus von Referenzdaten**. Die Referenzannotation kann anschließend durch die datenbasierte Entwicklungsmethodik der Computerwissenschaften für die Modelloptimierung herangezogen werden:⁴ sie fixiert das Ziel für eine Optimierung der Vorhersagekraft von möglichen Modellarchitekturen und deren Parametrisierungen.⁵

Das bislang geschilderte Vorgehen fokussiert ausschließlich auf die technische Optimierung der Vorhersagemodelle für fixierte Referenzdaten. Ein effektiver konzeptioneller Rahmen für die Methodenkombination muss daneben Prozessen Raum geben, die eine sukzessive Verfeinerung der Analysekategorien vornehmen, um einem geisteswissenschaftlichen Fragenkomplex gerecht zu werden. Hier sind mehrere Aspekte zu unterscheiden: **(II) eine Inspektion der Vorhersageergebnisse** der (technisch optimierten) Modelle kann empirische Indikatoren zu Tage fördern, die Anlass zu einer **Revision der Operationalisierung** geben. Neben dem Entwicklungszirkel auf der rechten Seite muss es also einen Zirkel auf der linken Seite geben. Ein solches Revisionsmodell ist im Rahmen eines *Prototyping*-Ansatzes bzw. in der agilen Softwareentwicklung verbreitet – unabhängig davon, ob sich die Zielkategorisierung selbst in einem technischen Rahmen bewegt oder ob sie einem von der computerwissenschaftlichen Methodik abweichenden Rahmen entstammt.⁶

Letzteres ist allerdings bei einer komputationell/geisteswissenschaftlichen Methodenkombination der Fall. Die Gütekriterien, die zur Revision einer geisteswissenschaftlichen Analysekategorie führen, können grundsätzlich anderen methodischen Prinzipien und Vorannahmen folgen. Eine probenhalber vorgenommene Operationalisierung eines vielschichtigen Konzepts in der Textanalyse (z.B. Protagonisten in Dramen) mag sich zum Beispiel als unergiebig erweisen, obgleich die komputationelle Umsetzung entsprechend den Referenzdaten eine hohe Vorhersagequalität ermöglicht. **(III) der textanalytische Inferenzschritt** (für den eine Computermodellierung vorgenommen wird) und die **datenbasierte Modellierung** sind jeweils in **eigene methodische bzw. arbeitspraktische Rahmen** eingebettet; zentrale Aufgabe für die „Mixed methods“-

Forschung ist die Spezifikation von adäquaten Übersetzungsschritten.

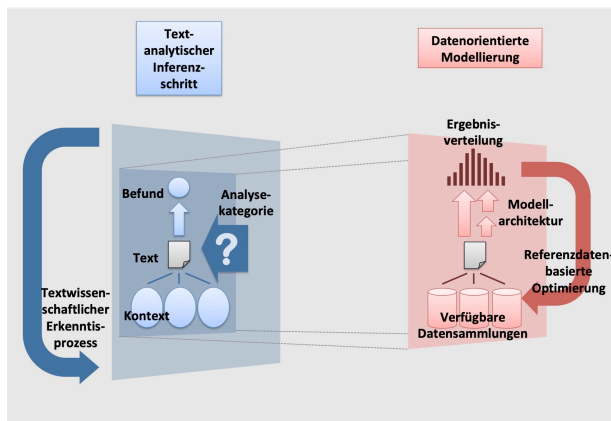


Abbildung 2: Konzeptionelle Trennung der arbeitspraktischen/methodischen Rahmen für einen textanalytischen Inferenzschritt und seine datenorientierte Modellierung; beide folgen eigenen Zyklen der Verfeinerung/Weiterentwicklung

Abbildung 2 zeigt entsprechend eine stärker ausdifferenzierte Skizze des konzeptionellen Rahmens. Der vierte Aspekt ist hier bereits angedeutet: (IV) Ein bestimmter **Analysebefund zu einem Text** hinsichtlich einer Analyse-kategorie ist **stets in Bezug auf einen angenommen Interpretationskontext** (mit potenziell vielfältigen relevanten Dimensionen) zu sehen.⁷ Ein textanalytischer Inferenzschritt (die innere blaue Box) wird also dargestellt als Ableitung eines Befundes aus einem Text, gegeben eine bestimmte Instantiierung der Kontextfaktoren, für die teilweise eigene Analyse-kategorien angesetzt werden müssen. (Zum Beispiel könnte die Operationalisierung der Kategorie *Protagonist* verwoben sein mit einer Analyse der aktiven und der passiven Präsenz der Figur, welche jeweils als eigene Kategorien zu operationalisieren sind.)

Wie Abbildung 2 suggeriert findet die Übersetzung aus der geisteswissenschaftlichen in die Sphäre der datenbasierten Computermodellierung sinnvollerweise für einzelne Inferenzschritte separat statt (obgleich wie in Fn. 5 angedeutet die Computerarchitektur für einen Schritt selbst technisch komplex sein kann). Es wird deutlich, dass bei der Bearbeitung von nicht-trivialen Fragestellungen rasch ein vielschichtiges Geflecht von Komponenten mit unterschiedlichem Status entsteht. Ein Hauptziel der hier vorgeschlagenen Konzeption liegt darin, das Augenmerk auf genau jene Statusunterschiede zu lenken, die für ein methodisch reflektiertes Vorgehen zu relevanten Vorannahmen relevant sind.

Strukturell sind die Elemente unserer Konzeption trotz der darstellbaren Komplexität vergleichsweise einfach – sie konzentrieren sich auf die Aufgabe der methodenübergreifenden Übersetzung mittels

der referenzdatengestützten Operationalisierung und komputationellen Modellierung. Für die konkrete arbeitspraktischen Projektroutine sollte also eine verhältnismäßig übersichtliche Sicht auf die relevanten Komponenten möglich sein. Die abschließende Abbildung 3 demonstriert jedoch, dass der konzeptionelle Rahmen bei Bedarf eine Schnittstelle zu einer sehr differenziert ausgearbeiteten wissenschaftstheoretischen Konzeption wie der von Danneberg/Albrecht 2017 bietet. (Aus Platzgründen können wir in diesem Abstract nicht auf Details eingehen.) Ein reflektiertes Vorgehen kann also auch auf fundamentalere Fragen etwa zur Problematisierung von divergierenden Wissensansprüchen über Disziplingrenzen hinweg eingehen.

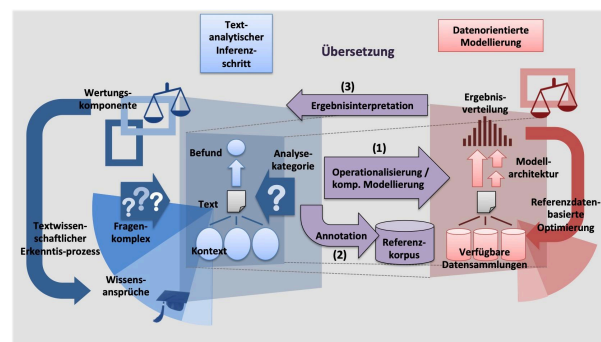


Abbildung 3: Bei Bedarf können die Wissensbereiche detailliert ausspezifiziert werden, die in die Methodenkombination gehen – aufbauend auf der wissenschaftstheoretischen Konzeption disziplinärer Praxis als Bearbeitung eines wissenschaftstheoretischen Problems nach Danneberg/Albrecht 2017 (mit den Komponenten (i) problematisierte Wissensansprüche, (ii) Wertungskomponenten, (iii) Fragenkomplex)

Fußnoten

1. Zentral für unsere methodologische Argumentation ist der Anspruch, Computermodelle für die Ableitung von einzelnen Befunden aus der Datenlage so in kombinierte Arbeitspraktiken zu integrieren, dass folgende zwei Aspekte unabhängig voneinander reflektiert werden können: (a) Wie hoch liegt die komputationelle Vorhersagequalität des Modells bei der Anwendung auf den Untersuchungsgegenstand (Texte einer bestimmten Epoche, Gattung, Inhaltsdomäne etc.) bezogen auf eine definierte Analyseaufgabe (etwa die Extraktion aller textuellen Erwähnungen der Akteure in einem Text)? (b) Wie aussagekräftig sind Befunde aus einer gegebenen, definierten Analyseaufgabe bezogen auf eine übergeordnete geisteswissenschaftliche Forschungsfrage? Eine modular aufgebaute Architektur von Analyseschritten kann entlang

dieser beiden Aspekte sukzessive für die Bearbeitung eines geisteswissenschaftlichen Forschungsgegenstands angepasst werden.

Fasst man die Idee der komputationell/geisteswissenschaftlichen Methodenkombination sehr weit, fallen natürlich alle Arbeiten aus dem Spannungsfeld der DH darunter: Bevor Standard-Werkzeuge (auch ohne eine Möglichkeit der Adaptation) etwa für explorative oder heuristische Zwecke eingesetzt werden, werden sich die Forschenden ein Bild zu deren Vorhersagequalität auf dem studienspezifischen Material machen. Und bevor in einem Machine-Learning-Projekt mit großem Zeitaufwand Computermodelle für eine Analyseaufgabe und ein bestimmtes Korpus optimiert werden, vergewissern sich die Beteiligten in der Regel, dass die ableitbaren Befunde einen relevanten Beitrag zu Fachfragen leisten. Insofern schlägt sich das Nebeneinander der beiden Aspekte mittelbar auf jede DH-Methodendiskussion nieder (vgl. etwa die Diskussion von computergestützten Analyseverfahren in unterschiedlichen DH-Fachkontexten in Reiche et al. 2014 oder die Diskussion der Rolle der Informatik in den DH bei Heyer/Niekler/Wiedemann 2014 oder Deck 2018). Die praktischen und theoretischen Implikationen für das weitere Vorgehen unterscheiden sich jedoch abhängig davon, ob eine parallel komputationell/geisteswissenschaftlich reflektierte Modularisierung der Analysekonzepte im Kern des arbeitspraktischen Vorgehensmodells für ein DH-Projekt steht oder ob beispielsweise zunächst unüberwachte Verfahren explorativ eingesetzt werden (vgl. das in Allison et al. 2011 dokumentierte Experiment). Betroffen sind also zentrale methodische Fragen, etwa zur Rolle der Visualisierung im Erkenntnisprozess (hierzu fanden und finden vielfältige Diskurse statt, vgl. z.B. Ihde 1998, Schaal/Kath/Dumm 2016, Romele et al. 2018); auch die Frage nach dem Verhältnis zwischen "building" (als Erstellen von Computerprogrammen) und "studying" in den DH (vgl. die 2011 von Stephen Ramsay ausgelöste Kontroverse, Ramsay/Rockwell 2012) erhält in einem eng verzahnten Vorgehensmodell einen differenzierten Charakter.

Gerade angesichts vielschichtiger möglicher Implikationen zum methodischen Selbstverständnis erscheint es uns für den vorliegenden Beitrag sinnvoll, die Kernideen des verzahnten Vorgehensmodells zunächst systematisch zu diskutieren. Auf eine Vertiefung spezifischer methodologischer Querbezüge müssen wir aus Platzgründen verzichten.

2. CRETA (<https://www.creta.uni-stuttgart.de>) wurde 2016 im Rahmen der BMBF-Förderung für Digital Humanities-Zentren etabliert.

3. Aus Platzgründen können wir die Umsetzung des konzeptionellen Rahmens in einzelnen Projekten im vorliegenden Vortragsexposé nicht diskutieren.

4. Falls hinreichend viele Daten annotiert wurden, kann überwachtes Lernen auf Basis einer Aufteilung in Trainings- und Testdaten zur Anwendung kommen.

Zumeist ist bei speziellen Kategorisierungsaufgaben jedoch die Menge der Referenzdaten so klein, dass sie lediglich als Testmenge dienen kann – etwa für die Adaptation von Modellen, die für vergleichbare Kategorien trainiert werden können, oder für unüberwachte Verfahren.

5. Die interne Modellstruktur wird im Rahmen eines solchen Vorgehens getrennt vom Modellierungsziel gesehen; sie kann zwar von systematischen Annahmen über die zu modellierende Aufgabe inspiriert sein, ausschlaggebend ist jedoch die Optimierung der fixierten Referenzannotationen. Häufig erzwingt die mangelnde Verfügbarkeit von unmittelbar verwendbaren Daten bzw. Annotationen eine Modellarchitektur, die von den kausalen Zusammenhängen der modellierten Domäne abweicht.

6. Einen Prototyping-Ansatz innerhalb der DH diskutieren El Khatib et al. 2019. Ein verzahntes zyklisches Vorgehensmodell ist auch in der Visualisierung/den *Visual Analytics* verbreitet (vgl. etwa Sacha et al. 2014, El-Assady et al. 2019 und die Beiträge in Butt et al. 2020). Iterative Annotationszyklen zur Optimierung der Operationalisierung werden beschrieben bei Pustejovsky/Stubbs, 2012, Bögel et. al. 2015 Gius/Jacke 2017 und Pagel et al. 2018. Eine ausführliche Diskussion des Bezugs zu *close/distant reading* in DH findet sich in Jänicke 2016, Kap. 2.

7. Wir schließen uns hier der Konzeptualisierung an, die Danneberg/Albrecht 2017 im Rahmen einer möglichst allgemeinen wissenschaftstheoretischen Charakterisierung der Inferenzschritte vornehmen, welche einer literaturwissenschaftlichen Textinterpretation zugrunde liegen. Eine Verallgemeinerung auf andere textanalytische Problemstellungen ist ohne Weiteres möglich.

Bibliographie

Allison, Sarah / Heuser, Ryan / Jockers, Matthew / Moretti, Franco / Witmore, Michael (2011): "Quantitative Formalism: An Experiment", in: *Pamphlets of the Stanford Literary Lab* 1.. <http://litlab.stanford.edu/LiteraryLabPamphlet1.pdf> [Letzter Zugriff: 05.01.2020]

Bögel, Thomas / Gertz, Michael / Gius, Evelyn / Jacke, Janina / Meister, Jan Christoph / Petris, Marco / Strötgen, Jannik (2015): "Collaborative Text Annotation Meets Machine Learning: heureCLÉA, a Digital Heuristic of Narrative", in: *DHCommons Journal* 1 10.5281/zenodo.3240591

Butt, Miriam / Hautli-Janisz, Annette / Lyding, Verena (2020): *LingVis: Visual Analytics for Linguistics*. CSLI lecture notes. Stanford: CSLI Publications.

Danneberg, Lutz / Albrecht, Andrea (2017): "Beobachtungen zu den Voraussetzungen des hypothetisch-deduktiven und des hypothetisch-induktiven Argumentierens im Rahmen einer hermeneutischen

Konzeption der Textinterpretation", in: *Journal of Literary Theory* 10(1): 1–37.

Deck, Klaus-Georg (2018): "Digital Humanities – Eine Herausforderung an die Informatik und an die Geisteswissenschaften", in: Huber, Martin / Krämer, Sybille (eds.): *Wie Digitalität die Geisteswissenschaften verändert: Neue Forschungsgegenstände und Methoden*. (= Sonderband der Zeitschrift für digitale Geisteswissenschaften, 3). PDF Format ohne Paginierung. 10.17175/sb003_002.

El-Assady, Mennatallah / Jentner, Wolfgang / Sperrle, Fabian / Sevastjanova, Rita / Hautli-Janisz, Annette / Butt, Miriam / Keim, Daniel A. (2019): "lingvis.io - A Linguistic Visual Analytics Framework", in: *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics: System Demonstrations* 13–18.

El Khatib, Randa / Wrisley, David J. / Elbassuoni, Shady / Jaber, Mohamad / El Zini, Julia (2019): "Prototyping Across the Disciplines", in: *Digital Studies/le Champ Numérique*, 8(1), 10 10.16995/dscn.282

Gius, Evelyn / Jacke, Janina (2017): „The Hermeneutic Profit of Annotation. On preventing and fostering disagreement in literary text analysis“, in: *International Journal of Humanities and Arts Computing* 11(2): 233–254.

Idhe, Don (1998): *Expanding Hermeneutics: Visualism in Science*, Evanston, Ill.: Northwestern University Press.

Jänicke, Stefan (2016): *Close and Distant Reading Visualizations for the Comparative Analysis of Digital Humanities Data*. Dissertation, Universität Leipzig. <http://www.informatik.uni-leipzig.de/~stjaenicke/dissertation.pdf> [Letzter Zugriff: 05.01.2020]

Heyer, Gerhard / Niekler, Andreas / Wiedemann, Gregor (2014): "Brauchen die Digital Humanities eine eigene Methodologie? Überlegungen zur systematischen Nutzung von Text Mining Verfahren in einem politikwissenschaftlichen Projekt", in: *1. Jahrestagung der Digital Humanities im deutschsprachigen Raum (DHd 2014)* [http://asv.informatik.uni-leipzig.de/publication/file/255/Heyer-](http://asv.informatik.uni-leipzig.de/publication/file/255/Heyer-Brauchen_die_Digital_Humanities_eine_eigene_Methodologie)

[Brauchen_die_Digital_Humanities_eine_eigene_Methodologie](http://asv.informatik.uni-leipzig.de/publication/file/255/Heyer-Brauchen_die_Digital_Humanities_eine_eigene_Methodologie) [Letzter Zugriff: 05.01.2020]

Hovy, Eduard / Lavid, Julia (2010): "Towards a 'science' of corpus annotation: a new methodological challenge for corpus linguistics", in: *International Journal of Translation*, 22(1): 13–36.

Kuhn, Jonas (2019): "Computational text analysis within the humanities: How to combine working practices from the contributing fields?", in: *Language Resources & Evaluation* 53: 565–602 <https://doi.org/10.1007/s10579-019-09459-3>.

Kuhn, Jonas / Reiter, Nils (2015): "A plea for a method-driven agenda in the Digital Humanities", in: *Proceedings of Digital Humanities 2015, Sydney, Australia, (June 2015)*.

Kuhn, Jonas / Pichler, Axel / Reiter, Nils (eds.) (erscheint): *Reflektierte algorithmische Textanalyse:*

Interdisziplinäre(s) Arbeiten in der CRETA-Werkstatt. Berlin: de Gruyter.

Krautter, Benjamin / Pagel, Janis / Reiter, Nils / Willand, Marcus (2018): "Titelhelden und Protagonisten – Interpretierbare Figurenklassifikation in deutschsprachigen Dramen", in: LitLab Pamphlets 7. https://www.digitalhumanitiescooperation.de/wp-content/uploads/2018/12/p07_krautter_et_al.pdf [Letzter Zugriff: 05.01.2020].

Krautter, Benjamin / Pagel, Janis (2019): "Klassifikation von Titelfiguren in deutschsprachigen Dramen und Evaluation am Beispiel von Lessings 'Emilia Galotti'", in: *Konferenzabstracts DHd 2019 Digital Humanities: multimedial & multimodal, Frankfurt am Main, März 2019*.

Pagel, Janis / Reiter, Nils / Rösiger, Ina / Schulz, Sarah (2018): "A Unified Annotation Workflow for Diverse Goals", in: Kübler, Sandra / Zinsmeister, Heike (eds.): *Proceedings of the Workshop on Annotation in Digital Humanities, co-located with ESSLLI 2018*.

Pustejovsky, James / Stubbs, Amber (2012). *Natural language annotation for machine learning*. Sebastopol: O'Reilly Media.

Ramsay, Stephen / Rockwell, Geoffrey (2012): "Developing Things: Notes toward an Epistemology of Building in the Digital Humanities", in: Gold, Matthew K. (ed.): *Debates in the digital humanities*. Minneapolis and London: University of Minnesota Press, 2012: 75–84 10.5749/minnesota/9780816677948.003.0010.

Reiche, Ruth / Becker, Rainer / Bender, Michael / Munson, Matt / Schmunk, Stefan / Schöch, Christof (2014): *Verfahren der Digital Humanities in den Geistes- und Kulturwissenschaften*. DARIAH-DE Working Papers. Göttingen: DARIAH-DE, 2014.

Romele, Alberto / Severo, Marta / Furia, Paolo (2018): "Digital Hermeneutics: From Interpreting with Machines to Interpretational Machines", in: *AI & Society: Knowledge, Culture and Communication*, Springer, in press. <https://hal.archives-ouvertes.fr/hal-01824173> [Letzter Zugriff: 05.01.2020]

Sachdev, Dominik / Stoffel, Andreas / Stoffel, Florian / Kwon, Bum Chul / Ellis, Geoffrey / Keim, Daniel A. (2014): "Knowledge Generation Model for Visual Analytics", in: *IEEE Transactions on Visualization and Computer Graphics* 20 (12) 1604–1613. 10.1109/TVCG.2014.2346481.

Schaal, Gary S. / Kath, Roxana / Dumm, Sebastian (2016): "New Visual Hermeneutics", in: *Cybernetics & Human Knowing* 23: 51–76.

Stefanowitsch, Anatol (2018): *Corpus Linguistics: A Guide to the Methodology* (Textbooks in Language Sciences 8). Open Review Version. Berlin: Language Science Press. www.langsci-press.org [Letzter Zugriff: 05.01.2020]