



**Constanze Baum,
Thomas Stäcker (Hg.)**

GRENZEN UND MÖGLICHKEITEN DER DIGITAL HUMANITIES

1

ZfdG Sonderband

1

Die Digital Humanities im deutschsprachigen Raum. Methoden - Theorien - Projekte

Thomas Stöcker
Constanz Baum

2

Vom Bild zum Bild – Digital Humanities jenseits des Textes

Katja Gwech

3

Die Informatik als Geisteswissenschaft

John Horbonne

4

Digital Humanities? Gibt's doch gar nicht!

Patrick Sahle

5

Vom Urheber zur Crowd, vom Werk zur Version, vom Schutz zur Öffnung? Kollaboratives Schreiben und Bewerten in den Digital Humanities

Thomas Ernst

6

Informatik und Hermeneutik. Zum Mehrwert interdisziplinärer Textanalyse

Evelyn Gus
Janina Jacke

7

Warum werden mittelalterliche und frühneuzeitliche Rechnungsbücher eigentlich nicht digital ediert?

Georg Vogler

8

Vom Zeichen zur Schrift: Mit Mustererkennung zur automatisierten Schreiberhanderkennung in mittelalterlichen und frühneuzeitlichen Handschriften

Daniel Fackel
Volker Königler
Torsten Schaffan

9

Historische Bibliometrie? Methodische Vorschläge für eine Annäherung textanalytischer Zugänge an die mediävistische Textualitätsdebatte.

Cathrin Volkhausen

10

Kontroverse Kommunikation im Umkreis der ersten Frauenbewegung. Wie können digitale Ressourcen die sprachliche Untersuchung und die Ergebnisdokumentation verbessern?

Kerstin Wolff
Alexander Geyken
Thomas Gloning

11

Exilnetz33. Ein Forschungsportal als Such- und Visualisierungsinstrument

Anna Lorenz
Theresa Buhl
Dirk Ockewill

12

What People Said: The Theoretical Foundations of a Minimal Doxographical Ontology and Its Use in the History of Philosophy

Stefan Heßbrüggen-Walter

13

Computerlinguistische Werkzeuge zur Erschließung und Exploration großer Textsammlungen aus der Perspektive fachspezifischer Theorie

Andri Binsing
Tilly Kische
Ulrich Heid
Cathleen Kattner
Jonas Kuhn

14

Neue Perspektiven für die Erziehungswissenschaften – Das Potenzial von Mixed Reality für Lehr- und Lernprozesse

Katharina Schuster
Anja Richter
Sabine Joschke

15

Semantisch angereicherte 3D-Messdaten von Kirchenräumen als Quellen für die geschichtswissenschaftliche Forschung

Felix Lange
Martin Unold

16

Bot rechnen

Walter von Pippich

17

MUSICI und MusMig. Kontinuitäten und Diskontinuitäten

Berthold Over
Torsten Rooder

18

«Losing My Religion» – Einsatz der Videomastingsdatenbank Pan.do/ra in der kunsthistorischen Analyse von Musikvideos

Thorsten Wilmanns
Eric Decker
Matthias Arnold

19

AAC-Fackel. Das Beispiel einer digitalen Mastertextedition

Hanna Biber

20

Heterogene Daten in den Digital Humanities. Eine Architektur zur forschungsorientierten Förderung von Kollektionen

Tobias Groll
Andreas Herrlich
Christoph Platte

21

Bibliographie

Zeitschrift für digitale Geisteswissenschaften

Artikel aus:

Sonderband 1 der ZfdG: Grenzen und Möglichkeiten der Digital Humanities. Hg. von Constanze Baum und Thomas Stäcker. 2015. DOI: [10.17175/sb01](https://doi.org/10.17175/sb01)

Titel:

Methoden – Theorien – Projekte

Autor/in:

Constanze Baum

Kontakt: Constanze.Baum@mww-forschung.de

Institution: Forschungsverbund MWW / Herzog August Bibliothek Wolfenbüttel

GND: [132852896](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-32852896) ORCID: [0000-0002-7214-9504](https://orcid.org/0000-0002-7214-9504)

Autor/in:

Thomas Stäcker

Kontakt: staecker@hab.de

Institution: Herzog August Bibliothek Wolfenbüttel

GND: [141905573](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-141905573) ORCID: [0000-0002-1509-6960](https://orcid.org/0000-0002-1509-6960)

DOI des Artikels:

[10.17175/sb001_023](https://doi.org/10.17175/sb001_023)

Nachweis im OPAC der Herzog August Bibliothek:

[85748091X](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-85748091X)

Erstveröffentlichung:

22.03.2016

Lizenz:

Sofern nicht anders angegeben 

Medienlizenzen:

Medienrechte liegen bei den Autoren

Letzte Überprüfung aller Verweise:

24.05.2016

GND-Verschlagwortung:

[Digital Humanities](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-32852896) |

Zitierweise:

Constanze Baum, Thomas Stäcker: Methoden – Theorien – Projekte. In: Grenzen und Möglichkeiten der Digital Humanities. Hg. von Constanze Baum / Thomas Stäcker. 2015 (= Sonderband der Zeitschrift für digitale Geisteswissenschaften, 1). PDF Format ohne Paginierung. Als text/html abrufbar unter DOI: [10.17175/sb001_023](https://doi.org/10.17175/sb001_023).

Constanze Baum, Thomas Stäcker

Methoden – Theorien – Projekte

Abstracts

Zum Geleit des Sonderbands

»[...] the moral seems clear enough: that computing belongs within the humanities (McCarty)«¹

Diskursgeschichtlich ist der *digital turn* nicht nur proklamiert, sondern auch in zahlreichen Publikationen niedergelegt und besetzt damit einen festen Platz in der Wissenschaftsgeschichtsschreibung des späten 20. und frühen 21. Jahrhunderts. Die Hinwendung zu digitalen Fragestellungen in den letzten Dezennien steht außer Frage. Längst ist klar, dass wir uns in einem Zeitalter bewegen, in dem digitale Technologien gesamtgesellschaftlich und damit auch für die geisteswissenschaftlichen Fächer an Bedeutung gewonnen haben, bestimmte Bereiche sogar dominieren und entsprechend in der Wissenschaftslandschaft auch staatliche Zuwendung und Förderung erfahren.² Kaum eine Tagungsankündigung oder ein Sammelband, der sich den Digital Humanities widmet, kommt im Geleitwort ohne einen Appell aus, der die Relevanz des Digitalen unterstreicht und dadurch an ihrer diskursgeschichtlichen Verankerung mitarbeitet.

Wir möchten betonen, dass die digitale Wende aus unserer Perspektive mehr verspricht und auch einlösen wird, als die inflationären Turns innerhalb der Geisteswissenschaften in den letzten Jahrzehnten, da sie nicht nur die Fachdisziplinen methodisch neu konturiert, sondern eine Disziplinen übergreifende Schnittstelle bietet. Über Fachdiskurse hinaus berührt dies auch die Nutzung von Infrastruktureinrichtungen wie Archive und Bibliotheken.

Von einer ›digitalen Revolution‹ für die Geisteswissenschaften zu sprechen, halten wir insofern für ungeeignet, als die Entwicklungen und Transformationsprozesse unserer Einschätzung nach weder die Episteme unserer Fachdisziplinen explosiv neu konstituiert, noch einen im Kleist'schen Sinne ›Umsturz aller Verhältnisse‹ bewirkt haben. Vielmehr handelt es sich um eine dominante technische Innovation und ihre sukzessive soziale Sedimentierung und somit in unseren Augen um eine digitale Leitkultur. Auch wenn die neu formierten Digital Humanities an inhaltliche oder technische Grenzen ihrer Möglichkeiten stoßen, sind sie schon jetzt in vielen Punkten Teil einer neuen Wissenschafts- und Wissenskultur, in der Quellen und

¹ McCarty 2003, S. 1224.

² Vgl. die Förderprogramme und -linien des Bundesministerium für Bildung und Forschung, z.B. [Forschungsinfrastrukturen für die Geistes- und Sozialwissenschaften](#), 2013, besonders das Kapitel [Forschungsinfrastrukturen in den Geisteswissenschaften](#), S. 21–33. Vgl. auch Schaal / Kath 2014, S. 334, die kritisch betonen, dass die Förderung sich zunehmend auf technische und technologische Innovationen versteife und weniger neue fachspezifische Fragestellungen unterstütze.

Dokumente folgerichtig zu Forschungsdaten werden. Die Kartierung dieser digitalen Leitkultur erstreckt sich von der Grundlagenforschung, der Datenerhebung und -erschließung über die Datenmodellierung bis hin zur Datenarchivierung. Im Sinne eines *data life cycle* können die Daten von dort aus wieder in den Forschungsprozess eingespeist werden.

Umso mehr solche Felder erschlossen werden, umso mehr müssen sich die einzelnen Disziplinen und insbesondere die Fächer, die als geisteswissenschaftliche firmieren, fragen, in welchem Maße sie auf diese Tendenzen strukturell wie methodisch reagieren wollen.³ Eine Kernfrage für das Selbstverständnis der Digital Humanities ist nach wie vor, ob sie als Instrumentarium oder als autonome Forschungsrichtung zu verstehen ist, die distinkte Alleinstellungsmerkmale vorweisen kann und von eigenständigen Erkenntnisansprüchen geleitet wird.

Klar ist freilich, dass neue digitale Verfahren und Werkzeuge sowohl Potenziale bergen als auch Probleme aufwerfen, die es für geisteswissenschaftliche Forschungen zu beachten gilt. So bietet der virtuelle Raum zwar vergleichsweise unbeschränkten Platz für die Erhebung und Visualisierung großer und übergroßer Datenmengen, aber längst nicht alles, was technisch in diesem Rahmen machbar erscheint, ist für den wissenschaftlichen Erkenntnisgewinn auch sinnvoll. Dass mit dem Erstarken solcher Ansätze auch ein ›digitales Unbehagen‹ einhergehen kann, dem es wissenschaftstheoretisch nachzuspüren gilt, halten wir geradezu für obligatorisch. Angesichts des mitunter enormen technologischen Aufwands, der für ›Versuchsaufbauten‹ und Datenerhebungen im Vergleich zu herkömmlichen Recherchen und Analysen von vielen Geisteswissenschaftlern nötig ist, erscheint es durchaus berechtigt, die Frage nach dem Erkenntnisgewinn und ebenso den moralischen Implikationen gewissermaßen prophylaktisch zu stellen. Was wird zum Beispiel in einem Datendiagramm, in einer Cloud sinnfällig, zu welchem Nutzen und mit welchen Folgen? Die Antworten werden nicht von den Daten allein kommen, reine Empirie bedarf auch in naturwissenschaftlichen Forschungen immer der Auswertung und Interpretation und obliegt angreifbaren Setzungen in der Modell- und Theoriebildung.

Digital Humanities sollten sich deshalb auch mit den Grenzen der automatisierten Verarbeitung von Text und Bild auseinandersetzen. So werden die Digital Humanities zunehmend reflektieren müssen, welche politischen, ethischen und sozialen Dimensionen die digitale, maschinenbasierte ›Erfassung‹ von Quellen und Zeugnissen ganz unterschiedlicher Provenienz tangiert. Künstlerische Arbeiten haben sich hier in jüngster Zeit kritisch zu Wort gemeldet und die vermeintliche Unschuld bzw. die Souveränität von Daten und die Anwendung und Entwicklung von Tools in einem gesamtgesellschaftlichen Rahmen zur Disposition gestellt.⁴ Ein solcher ›DH-criticism‹ müsste sich beispielsweise der Frage stellen, ob die Erschließung und Sichtung von quantitativ extrem umfangreichen, gleichsam grenzenlosen Texten oder Medien mittels *distant-reading* zu Verfahren (ver-)führt, die erhebliche Konsequenzen für unsere

³ Vgl. Schaal / Kath, S. 347, die davon ausgehen, dass »die Digitalisierung der Geisteswissenschaften ein unaufhaltsamer Prozess ist, der durch die Fachwissenschaften nur noch moderiert werden kann.«

⁴ Vgl. die Ausstellung »Nervöse Systeme. Quantifiziertes Leben und die soziale Frage«, 11.3.-9.5.2016 im Haus der Kulturen der Welt, kuratiert von Anselm Franke, Stephanie Hankey und Marek Tuszynski (Katalog im Druck).

Bewertungskulturen haben, da sie letztlich mit ähnlichen Algorithmen wie die Rasterfahndung oder Überwachungstechnologien operieren oder auf diesen basieren. Denn Quantifizierungen betreffen soziale Fragen, auch wenn sie historisch orientiert sind. Die kritische Infragestellung von algorithmischen Selbstläufern und Datenpositivismus erscheint uns als künftig stärker zu berücksichtigende Aufgabe der *scientific community*, denn jedes DH-Projekt trägt durch Annotation, Kartierung und Netzwerkbindung Mitverantwortung am datenkulturellen Wandel und dessen Folgen.

Bedingt die Hinwendung zu digitalen Gehalten und Methoden tatsächlich einen geistesgeschichtlichen Paradigmenwechsel? Ressentiments und Pioniergeist halten sich diesbezüglich in der öffentlichen Debatte seit einigen Jahren die Waage.⁵ Wir stehen also mitten in einem Prozess von Aneignung und Erkundung neuer Gebiete auf der einen Seite und fachinternen Widerständen auf der anderen. Es ist dies vielleicht keine besonders überraschende Beobachtung: Das Neue ist zugleich das Fremde und Andere, es reizt zum Experiment und zur Kritik gleichermaßen und ist eingespannt in einen Generationenkonflikt, in dem Kontinuität und Wechsel, Aufgeschlossenheit gegenüber neuen Technologien und Traditionsbeharren in einem spannungsvollen Miteinander stehen. Ob die immer wieder beschworene Krise der Geisteswissenschaften durch die Entwicklungen der Digital Humanities aufgehoben werden kann oder gar neu entfacht wird, sei dahingestellt.⁶ Vielleicht ist erkenntnistheoretischen Fächern das Krisenhafte als Ausgangspunkt einer immerwährenden Neuorientierung und kritischen Selbstüberprüfung auch immanent? In jedem Fall haben die DH die herkömmliche Fächerlandschaft durchaus produktiv irritiert.⁷ Es bleibt abzuwarten, welche Bereiche der Geisteswissenschaften die Digital Humanities besonders bereichern werden. Nach einer ersten Phase, in der besonders die Sprachwissenschaften tonangebend bei der Implementierung von DH-Anwendungen waren, positionieren sich nun zunehmend auch andere, kleinere Fächer wie die Musikwissenschaft oder Archäologie im Feld der Digital Humanities. Dies sind aber nur vorsichtige Beobachtungen, die der Bandbreite der vorhandenen Einzelprojekte sicherlich nicht gerecht werden.⁸ Es bleibt festzuhalten, dass die DH sich sukzessive konsolidieren: Lehrstühle werden im universitären Raum begründet und besetzt und erste Generationen von Studierenden durchlaufen die neu geschaffenen Curricula. Infrastrukturprojekte verlassen ihren explorativen Status und werden konstitutiver Teil der Forschungslandschaft.

Die Breite des Forschungsfeldes bestimmt dabei seinen Reichtum, macht es aber gleichzeitig schwierig, ihre Konturen genau zu erkennen.⁹ Der Begriff Digital Humanities, der erstmals in dem von Schreibman et al. 2004 herausgegebenen *Companion* benutzt wurde,¹⁰ mag leicht darüber hinwegtäuschen, dass die junge Disziplin nicht voraussetzungslos ist

⁵ Vgl. Richards 2013, der von einem »Unbehagen gegenüber den digitalen Geisteswissenschaften« spricht, das es aufzulösen gelte.

⁶ Vgl. Warwick et al. 2012.

⁷ Vgl. jüngst ausführlich mit einer Ausrichtung auf die Berliner DH-Landschaft Baillot: 2016, [halshs-01251071](#).

⁸ Eine Auswertung der Verteilung von Aktivitäten im Bereich der Digital Humanities im deutschsprachigen Raum, auf die sich hier berufen ließe, liegt unseres Erachtens nicht vor. Der [Geodatenbrowser von DARIAH-DE](#) versammelt immerhin über 120 verschiedene Projekte an mehr als 250 Standorten weltweit in zeitlicher Relation; eine Auswahl nach Fachdisziplinen ist in der Filterung aber nicht vorgesehen.

⁹ Vgl. [DHd 2020](#).

¹⁰ Schreibman et al. 2004.

und in ihren Fragestellungen und ihrer methodischen Herangehensweise eine Reihe von Mitbewerbern hat. Schon der Titel der diesem Sonderband zugrundeliegenden Tagung deutet darauf hin: Sind die Digital Humanities eher Geisteswissenschaften oder eher Informatik? Die Frage der Abgrenzung stellt sich auch mit Blick auf andere Disziplinen, etwa die Linguistik, vor allem die Computerlinguistik, aber auch die Bibliotheks- und Informationswissenschaft, die alle Arbeitsfelder mit den *Digital Humanities* teilen. Auch die historische Entwicklung der Digital Humanities, die unter Überschriften wie Computerphilologie¹¹, Historisch-Kulturwissenschaftliche Fachinformatik¹² oder Humanities Computing¹³ eine längere Tradition hat und deren Beginn oft mit den Arbeiten von Pater Busa bzw. dem *Index Thomisticus* verbunden wird,¹⁴ zeigt ein komplexes Bild unterschiedlicher Strömungen und Ansprüche. Letztlich deutet auch die jüngere Debatte, ob die Digital Humanities ›fachwürdig‹ seien, auf ein Grundproblem, dass sich darin zeigt, dass digitale Methoden und Zugriffe in unterschiedlichster Ausprägung in den geisteswissenschaftlichen Fächern zum Einsatz kommen. Gleichzeitig erfreuen sich diese Ansätze zunehmenden Interesses, was Neugründungen wie die AGs der digitalen Kunstgeschichte oder Geschichte belegen.¹⁵

2015 befinden wir uns im deutschsprachigen Raum nunmehr an einem Punkt, an dem die Erarbeitung von Positionen und Fragestellungen im Zeichen des digitalen Wandels für die geisteswissenschaftlichen Fächer im vollen Gange ist,¹⁶ was nicht zuletzt an der steigenden Zahl von Konferenzen, Workshops und Stellenausschreibungen und der großen Nachfrage einer wachsenden *scientific community* ablesbar ist. Das Ringen um die Definition dessen, was Digital Humanities eigentlich bedeutet, zeigt, wie sehr alle Beteiligten noch in Legitimierungszwängen gegenüber etablierten Fachkulturen befangen sind.¹⁷ Die Beschränkungen letzterer geraten übrigens im Zuge kulturwissenschaftlicher Weitung sowie inter- und transdisziplinärer Spreizung ebenso ins Wanken, sie entsprechen kaum mehr den herkömmlichen Zuständigkeitsfeldern. Unabhängig von solchen Diffusionsbewegungen wird gewiss auch in den Folgejahren um Eingrenzung sowie Charakterisierung des Begriffes Digital Humanities gestritten, ebenso wie weiterhin die Möglichkeiten und Schwierigkeiten der Kontextualisierung und Kategorisierung auszuloten sind, die sich durch die Digital Humanities ergeben.¹⁸ Eine definitorische Offenheit entspricht aber genau dem Impetus, mit dem die Digital Humanities angetreten sind, und sollte als Bestandteil ihrer Diskursgeschichte betrachtet werden: Offener Zugang zu Quellen, ein erweiterter Autor- oder Urheberbegriff sowie transparente Verfahren sollen Beschränkungen der geisteswissenschaftlichen Disziplinengeschichte überwinden oder aufweiten helfen und zu einer Liberalisierung des Wissensbegriffes beitragen, ohne dass dieser freilich beliebig wird (vgl. auch den [Beitrag von Thomas Ernst](#)). Diese Politik der Öffnung von Ressourcen, Methoden und Erkenntnissen im globalen virtuellen Raum sollte mit einer Offenheit in der Begriffsbestimmung von *Digital Humanities* einhergehen. Wie weit digitale Verfahren in einzelne Fachdisziplinen diffundieren

¹¹ Jannidis 1999, S. 39-60.

¹² Thaller 2001.

¹³ Vgl. insbesondere McCarty 2005.

¹⁴ Vgl. Nyhan et al. 2013.

¹⁵ [Digitale Kunstgeschichte](#); AG "Digitale Geisteswissenschaft".

¹⁶ Vgl. über den Definitionsprozess und die Aufgaben der Digital Humanities auch Rapp 2013, S. 345.

¹⁷ Vgl. Schnapp et al. 2009.

¹⁸ Zuletzt König 2016.

werden oder können, wird letztlich auch von der Offenheit letzterer gegenüber den Digital Humanities abhängen.

Dabei erscheint es insgesamt wichtig, all diesen Bestrebungen adäquate Publikationsformate beiseite zu stellen, die zur wissenschaftlichen Diskussion, Überprüfung und Konsolidierung der Forschungsansätze beitragen: konsequenterweise in digitaler Form.

Ausgehend von den hier skizzierten Entwicklungen gründete sich in Deutschland anlässlich einer Digital Humanities-Tagung in Hamburg 2012 der Verein **Digital Humanities im deutschsprachigen Raum (DHd)** als Interessensvertretung der neuen Bewegung. 2014 fand in Passau die erste Tagung des DHd mit mehr als 300 Teilnehmern statt, ein überwältigender Erfolg für die junge Gemeinschaft. Sie stand in dem hier verhandelten Sinne unter der gleichermaßen programmatischen wie polemisch zugespitzten Fragestellung *Digital Humanities – methodischer Brückenschlag oder "feindliche Übernahme"? Chancen und Risiken der Begegnung zwischen Geisteswissenschaften und Informatik*.

Mit den hier – *only online* – publizierten Beiträgen versammelt der Band eine Auswahl des Tagungsprogramms, die das breite Spektrum von Themen und Herangehensweisen verdeutlichen soll, das von den unterschiedlichsten Fachdisziplinen in Bezug auf digitale Fragestellungen entwickelt wird. Das Konferenzthema enthält jene Provokation, die John Nerbonne in einem 10-Punkte-Plan in seiner pointierten Keynote »**Die Informatik als Geisteswissenschaft**« zum Anlass genommen hat, um neue Gedanken im Hinblick auf die Digital Humanities zu formulieren und durch gezielte ›Unterstellungen‹ zum Nachdenken herauszufordern. Der Resonanz der Bewegung geht Patrick Sahle nach, er spricht in »**Digital Humanities - Gibt's doch gar nicht!**« sogar von einem ›DH-Hype‹. Bild- und Objektwissenschaften kommen in den zwanzig Beiträgen ebenso zu Wort wie die Textwissenschaften. Mit Themen und Schlagworten wie digitale Annotation, Audio-, Video-, Text-Mining, digitale Edition, Forschungsdaten, Netzwerkanalyse, Mustererkennung, Normdaten, Ontologie, Open Access, Semantic Web, Stilometrie oder 3D-Rekonstruktion lässt sich das Feld abstecken, das – um es mediatorisch zu formulieren – sowohl den Fachwissenschaften als auch den Digital Humanities ihre Forschungsgebiete zuweist, und zwar in bewusster Überschreitung bzw. Durchlässigkeit der disziplinären Grenzen. Das betrifft nicht nur die geistes- und sozialwissenschaftlichen Fächer, die Philologien, die Erziehungswissenschaften, die Geschichtswissenschaft, die Kunstgeschichte, Mediävistik oder Philosophie, sondern auch die Informatik, die Informationswissenschaften und sogar die Nachrichtentechnik. Dass gerade in der Grenzüberschreitung die maßgeblichen Möglichkeiten der *Digital Humanities* liegen, den von C.P. Snow¹⁹ beschriebenen Graben zwischen Geistes- und Naturwissenschaften zu überwinden, ist ein greifbares Resultat dieses Bandes. Die Digital Humanities bewähren sich damit als eine Praktik, die als »umbrella«²⁰ für eine konstruktive inter- und transdisziplinäre Diskussion wirksam werden kann, wo es – unbeschadet weiterer

¹⁹ Snow 1959.

²⁰ Vgl. Kirschenbaum 2012.

Definitionsbemühungen²¹ – letztlich unerheblich bleibt, mit welchen abgrenzenden bzw. definitorischen Charakteristika die Digital Humanities bedacht werden.

Was sind also angesichts der Bandbreite von Themen und Fächern greifbare Methoden, Theorien und Projekte der Digital Humanities? Der etwas wagemutig formulierte Untertitel des Geleitwortes suggeriert eine feststehende Kategorisierung, vielleicht schon Kanonisierung von Methoden und Theorien. Wir können hier jedoch lediglich umreißen, was in den Disziplinen unserer Einschätzung nach verhandelt wird, und nur eine Skizze über die vorhandene DH-Landschaft geben: Allen voran steht in den philologisch wie historisch orientierten Geisteswissenschaften die Arbeit mit digitalen oder digitalisierten Textkorpora als Ressourcen sowie die damit einhergehende Entwicklung von digitalen Werkzeugen, die interoperabel und nachnutzbar sein sollen. Die Bemühungen zielen darauf, standardisierte Annotationsverfahren wie z.B. nach den Richtlinien des STTS Tag Sets oder den Guidelines der Text Encoding Initiative (TEI) im Bereich der Textaufbereitung einzuführen. Auch sollen Standards für Metadaten etabliert werden, um Insellösungen entgegenzuwirken. Im Bereich der Texterfassung bzw. -erschließung steht nach wie vor die Qualität von OCR-Ergebnissen und anderen Verfahren zur Texterkennung zur Debatte, sei es auf der Grundlage von Drucken der Frühen Neuzeit, sei es in anderen schwierigen Textständen von Fraktur oder Handschrift (vgl. den [Beitrag von Fecker et al.](#)).

Durch den Einsatz digitaler Tools können auch im Bereich der Wortforschung und linguistischen Erschließung historischer Korpora wertvolle Erkenntnisse gewonnen werden. Denn auf der Grundlage der Erhebung und Auswertung digitaler Daten dokumentieren Ontologien oder themenbezogen ermittelte Texttypen Sprachbewegungen und Wortschatzentwicklungen.

Digitale Erschließung bedeutet auch Bestandssicherung. Sie führt uns Möglichkeiten vor, die sich für die lexikalische Erfassung, für Beobachtungen über den Wandel von Sprachgebrauch oder die Instrumentalisierung von Sprache ergeben. Sie zeigt aber zugleich die Grenzen auf, die mit einer solchen Erschließung verbunden sind: die Fehlerquotienten, die in automatisierten Verfahren bedacht werden müssen, die händische Nachbereitung, die bei großen Textkorpora mitunter noch kaum zu bewältigen erscheint.

Text verändert sich – gleiches gilt für Bild oder Musik –, wenn die Quelle mittels Daten erschlossen wird, die direkt mit ihr verbunden werden. Aus dem Scan wird durch Anreicherung von Tags der annotierte Text, der schließlich Bestandteil einer Edition ist. Jeder Eingriff stellt eine Bezeichnung, semantische Anreicherung und diesem Sinne auch Kommentierung und Deutung des Ausgangsmaterials dar; Angleichungs- und Standardisierungsprozesse sind dabei nicht nur eine Frage des Schriftsatzes, sondern auch eine Frage von Datenqualität. Gefordert ist hier neben der sorgfältigen Aufbereitung in der Einzelstudie auch ein hohes Methodenbewusstsein, um die Interoperabilität von Programmen, Tools und Contents zu gewährleisten bzw. zuzulassen.

²¹ Terras et al. 2013.

Die Visualisierung von Ergebnissen hat darüberhinaus in den letzten Jahren in den digitalen Geisteswissenschaften an Bedeutung gewonnen. Digitale Referenzierungen auf Karten erlauben beispielsweise nicht nur geodatenorientierte Visualisierungen von historischen Fakten in einem Raumkoordinatensystem. Layer bieten darüber hinaus die Möglichkeit, Sukzessionen und Sequentielles abzubilden. Genealogien und Dynastien lassen sich so mittels temporaler Verläufe in topographische Ebenen ebenso einpflegen, wie beispielsweise die Rekonstruktion von ›Soundscapes‹ durch die Visualisierung des täglichen Glockenläutens in einer Stadt. Text Mining und Topic Modelling sind zwei weitere Bereiche, die vermehrt zur Anwendung kommen, was auch der zunehmenden Verfügbarkeit von digitalisierten Textsammlungen und durchsuchbaren Volltexten geschuldet ist, die weiterhin sukzessive auf- und ausgebaut werden. Das automatisierte Auslesen von Netzwerken aus großen Datenbeständen und die computergesteuerte Kollationierung von Texten bilden Forschungsmöglichkeiten wie das *distant reading* aus. Die Bewertung solcher und anderer Verfahren im digitalen Raum können dazu beitragen, dass das Bewusstsein für quellenkritische Fragen insgesamt geschärft wird und dies auch positive Rückwirkungen auf herkömmliche, nicht computergestützte Methoden der Text- und Quellenanalyse hat.

Die in diesem Band gebündelten Beiträge widmen sich Teilaspekten der skizzierten Forschungs Herausforderungen und setzen zugleich ganz unterschiedliche Akzente in Bezug auf Grenzen und Möglichkeiten der Digital Humanities: manche gehen das Thema forciert und direkt an, indem sie die Digital Humanities selbst in Frage stellen oder subversiv zu untergraben suchen, um sie im selben Atemzug zu rehabilitieren; manche argumentieren indirekter, indem sie vorführen, wie digitale Verfahren in einem noch ungesicherten fachdisziplinären Terrain zum Einsatz kommen (z.B. der [Beitrag von Schuster et al.](#) über Mixed Reality in den Erziehungswissenschaften) oder die Belastbarkeit einer fachwissenschaftlichen Fragestellung in Verbindung mit digitalen Methoden erprobt wird (z.B. [Warum werden mittelalterliche und frühneuzeitliche Rechnungsbücher eigentlich nicht digital ediert?](#) von Georg Vogeler oder [Rot rechnen](#) von Waltraud von Pippich). Die Beiträge und die ihnen zugrunde liegenden Projekte haben mitunter einen Experimentcharakter, der naturwissenschaftlichen Versuchsaufbauten kaum mehr fern steht. Gleichzeitig gilt zu bedenken, sich den epistemologischen Konsequenzen gegenüber wachsam und kritikfähig zu verhalten: Patrick Sahle, Stefan Heßbrüggen-Walter, Evelin Gius / Janina Jacke und John Nerbonne greifen dies in ihren Beiträgen aus unterschiedlichen Perspektiven auf.

Weitere Artikel beschäftigen sich mit der Grundlegung einer digitaler Bildwissenschaft, dem Zusammenhang von Hermeneutik und Informatik, dem Urheberrecht und Konsequenzen, die aus einem kollaborativen Arbeiten erwachsen, mit der automatisierten Schreiberhanderkennung, der historischen Stilometrie und der Frage nach digitalen Ressourcen im Umfeld historischer Phänomene wie der Frauenbewegung oder deren Auswertung in Netzwerkanalysen, wie es Theresia Biehl, Anne Lorenz und Dirk Osierenski überzeugend für die Exilforschung vorführen. Welchen Nutzen Ontologien für die Philosophiegeschichte haben, beschäftigt Stefan Heßbrüggen-Walter, während Hanno Biber in einem kürzeren Beitrag die Online-Edition der *Fackel* von Karl Kraus vorstellt. Den Einsatz von 3D-Messdaten in Kirchenräumen bespricht der Beitrag von Lange und Unold, aus der Kunstgeschichte stammt auch der Beitrag von Waltraud von Pippich über

die Bestimmung und die Auswertung von Rottönen in Bildern und Zeitdokumenten. Aus der Musikwissenschaft sind ebenfalls zwei Beiträge vertreten, die sich zum einen mit dem Einsatz von Videoannotationsdatenbanken beschäftigen, zum anderen mit der Erfassung von Migrationsbewegungen von Musikern. Aus der Computerlinguistik als einem affinen Feld der Digital Humanities schließlich stammt der Beitrag von Blessing et al., während die Frage nach dem projektübergreifenden Umgang mit heterogenen Daten in dem Beitrag von Gradl et al. skizziert wird.

Den Grenzen und Möglichkeiten der Digital Humanities ist der erste Sonderband der neu gegründeten *Zeitschrift für digitale Geisteswissenschaften* (ZfdG) gewidmet. Der Auftaktband spiegelt *pars pro toto* wider, was die Zeitschrift künftig zu adressieren sucht. Es ist den Herausgebern daran gelegen, durch das so eröffnete Forum weitere Beiträge und vielleicht sogar Bände anzustoßen, denen es um die Darstellung und Bewertung relevanter digitaler Werkzeuge und Fragestellungen für die geisteswissenschaftliche Forschung geht. Zugleich soll die *Zeitschrift für digitale Geisteswissenschaften* nicht nur eine Präsentationsplattform für Projekthinhalte, sondern auch ein Diskussionsraum sein, in dem die Grenzen und Möglichkeiten der digitalen Potentiale kritisch beleuchtet und eingeschätzt werden sollen. In diesem Sinne ist der vorliegende Sonderband keine Standortbestimmung, mit dem Grenzen klar absteckt sind, sondern eine Bestandsaufnahme, die das angesprochene Ausloten und Changieren auf unterschiedlichen Ebenen vorführen will. Er fügt sich ein in den Kanon der Stimmen, die die Frage nach Chancen und Nutzen, aber auch Schwierigkeiten und Risiken digitaler Speicherung, Darstellung und Nutzung in den geisteswissenschaftlichen Disziplinen stellen.²²

Den Herausgebern des ersten Sonderbandes stellen sich – am Ende war dies vielleicht sogar vorhersehbar – aus dieser Gesamtschau heraus mehr Fragen, als dass sie schon mit Antworten aufwarten könnten. Wir verstehen dies als Herausforderung, die der Tatsache geschuldet ist, dass die einzelne Fallstudie zum Nachdenken über eine übergeordnete Problematisierung anregt und gerade durch die interdisziplinäre Qualität der Beiträge deutlich wird, mit welchem großen Spektrum an möglichen Forschungsfeldern wir es jetzt bereits zu tun haben. In der Vielfalt und Einmaligkeit zugleich liegen die Potentiale der Fächer begründet, die ihre je eigenen Anforderungen an digitale Fragestellungen entwickeln.

Wir danken an dieser Stelle ausdrücklich allen Autorinnen und Autoren, die mit ihrer fachlichen Doppel-Expertise zum Gelingen dieses ersten Sonderbands beigetragen haben, der uns Möglichkeiten aber auch Grenzen im Bereich des *digital publishing* aufgezeigt und damit formal seinen prototypischen Stellenwert für alle weiteren Veröffentlichungen der *Zeitschrift für digitale Geisteswissenschaften* erwiesen hat. Dank gilt auch dem Bundesministerium für Bildung und Forschung (BMBF), das über die Förderung des Forschungsverbundes Marbach Weimar Wolfenbüttel die Diskussion auch für die Selbstverständigung dieser Gedächtniseinrichtungen zu solchen elementaren Fragen ermöglicht, sowie dem Verband Digital Humanities im deutschsprachigen Raum (DHD), der die Gründung der *Zeitschrift für digitale Geisteswissenschaften* unterstützt.

²² Vgl. z.B. den Workshop »Grenzen überschreiten – Digitale Geisteswissenschaft heute und morgen«, veranstaltet vom Einstein-Zirkel Digital Humanities am 28. Februar 2014 in Berlin oder die Tagung »Digitale Kunstgeschichte: Herausforderungen und Perspektiven« vom 26.–27. Juni 2014 in Zürich.

Bibliographische Angaben

- A companion to digital humanities. Hg. von Susan Schreibman / Ray Siemens / John Unsworth. Malden 2004. [\[online\]](#)
[\[Nachweis im GBV\]](#)
- Anne Baillot: Einleitung zu »Berliner Beiträge zu den Digital Humanities«. In: HAL. Archive ouverte en Sciences de l'Homme et de la Société. 8. Januar 2016 (Version 1). PDF format. [\[online\]](#)
- Ausstellung »Nervöse Systeme. Quantifiziertes Leben und die soziale Frage«, 11.3.–9.5.2016 im Haus der Kulturen der Welt, kuratiert von Anselm Franke, Stephanie Hankey und Marek Tuszynski (Katalog im Druck). [\[Nachweis im GBV\]](#)
- Defining Digital Humanities. A Reader. Hg. von Melissa Terras / Julianne Nyhan / Edward Vanhoutte. Farnham 2013. [\[Nachweis im GBV\]](#)
- Digital Humanities in Practice. Hg. von Claire Warwick / Melissa Terras / Julianne Nyhan. London 2012. [\[Nachweis im GBV\]](#)
- Fotis Jannidis: Was ist Computerphilologie? In: Jahrbuch für Computerphilologie 1 (1999), S. 39–60. [\[Nachweis im GBV\]](#)
- Matthew Kirschenbaum: Digital Humanities As/Is a Tactical Term. Debates in the Digital Humanities. 2012. [\[online\]](#)
- Mareike König: Was sind Digital Humanities? Definitionsfragen und Praxisbeispiele aus der Geschichtswissenschaft. In: Digital Humanities am DHP, 17. Februar 2016. [\[online\]](#) [\[Nachweis im GBV\]](#)
- Willard McCarty: Humanities Computing. London 2005. [\[Nachweis im GBV\]](#)
- Willard McCarty: Humanities Computing. In: Encyclopedia of Library and Information Science 2003, S. 1224–1235. DOI: 10.1081/E-ELIS 120008491 [\[online\]](#)
- Julianne Nyhan / Andrew Flinn / Anne Welsh: Oral History and the Hidden Histories Project: towards a history of computing in the humanities. In: Digital Scholarship in the Humanities 30 (2013) 1, S. 71–85. [\[online\]](#) [\[Nachweis im GBV\]](#)
- Andrea Rapp: Aus Sicht der Geisteswissenschaften: Die wissenschaftliche Bibliothek als Schnittstelle zwischen digitaler Infrastruktur und geisteswissenschaftlicher Forschung. In: Evolution der Informationsinfrastruktur. Kooperation zwischen Bibliothek und Wissenschaft. Hg. von Heike Neuroth / Nobert Lossau / Andrea Rapp. Glückstadt 2013, S. 345–353. [\[Nachweis im GBV\]](#)
- Earl Jeffrey Richards: Digitale Literaturwissenschaft: Perspektiven, Probleme und Potentiale der Philologien im »digital turn«. In: Textpraxis 7 (2013). [\[online\]](#)
- Gary S. Schaal / Roxana Kath: Zeit für einen Paradigmenwechsel in der Politischen Theorie? Der Ansatz der *Neuen Visuellen Hermeneutik*. In: Die Verfassung des Politischen. Festschrift für Hans Vorländer. Hg. von André Brodocz / Dietrich Herrmann / Rainer Schmidt / Daniel Schulz / Julia Schulze Wessel. Wiesbaden 2014, S. 331–350 [\[Nachweis im GBV\]](#)
- Jeffrey Schnapp / Todd Presner / Peter Lunenfeld: A Digital Humanities Manifesto / The Digital Humanities Manifesto 2.0. 2009. [\[online\]](#)
- C. P. Snow: The two cultures and the scientific revolution. Cambridge 1959. [\[Nachweis im GBV\]](#)
- Manfred Thaller: Brücken schlagen oder zerschlagen? Das Fach Historisch-Kulturwissenschaftliche Fachinformatik an der Universität zu Köln. In: Computers – Literature – Philology (CLiP). 6.–9. Dezember 2001. Abstracts. [\[online\]](#)

Zeitschrift für digitale Geisteswissenschaften

Artikel aus:

Sonderband 1 der ZfdG: Grenzen und Möglichkeiten der Digital Humanities. Hg. von Constanze Baum und Thomas Stäcker. 2015. DOI: [10.17175/sb01](https://doi.org/10.17175/sb01)

Titel:

Vom Bild zum Bild – Digital Humanities jenseits des Textes

Autor/in:

Katja Kwastek

Kontakt:

k.kwastek@vu.nl

Institution:

VU University Amsterdam

GND:

[128627387](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-28627387)

ORCID:

DOI des Artikels:

[10.17175/sb001_002](https://doi.org/10.17175/sb001_002)

Nachweis im OPAC der Herzog August Bibliothek:

[830123016](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-830123016)

Erstveröffentlichung:

19.02.2015

Lizenz:

Sofern nicht anders angegeben 

Medienlizenzen:

Medienrechte liegen bei den Autoren

Letzte Überprüfung aller Verweise:

24.05.2016

GND-Verschlagwortung:

[Bildwissenschaft](#) | [Kunstgeschichte \(Fach\)](#) | [Digital Humanities](#) | [Bildverarbeitung](#) |

Zitierweise:

Katja Kwastek: Vom Bild zum Bild – Digital Humanities jenseits des Textes. In: Grenzen und Möglichkeiten der Digital Humanities. Hg. von Constanze Baum / Thomas Stäcker. 2015 (= Sonderband der Zeitschrift für digitale Geisteswissenschaften, 1). PDF Format ohne Paginierung. Als text/html abrufbar unter DOI: [10.17175/sb001_002](https://doi.org/10.17175/sb001_002).

Katja Kwastek

Vom Bild zum Bild – Digital Humanities jenseits des Textes

Abstracts

Dieser Beitrag widmet sich der historischen Entwicklung und den methodischen Herausforderungen der digitalen Bildforschung. Obwohl bereits früh argumentiert wurde, der Computer sei für formale Analysen besser geeignet als für ikonographische, war die digitale Bildwissenschaft lange Zeit von Projekten der ikonographischen Codierung bildlicher Darstellungen und der strukturierten Erfassung von Metadaten bestimmt. Gleichzeitig wurden schon früh Rufe nach einer ›Ideologiekritik‹ laut. Heute umfasst die digitale Bildwissenschaft ein breites Spektrum nicht nur formalanalytischer und ikonographischer, sondern auch rezeptionsästhetischer Forschungen. Ihr ideologiekritischer Ansatz kann zudem dazu beitragen, die kritische Reflexion methodischer Implikationen als Kernaufgabe der Digital Humanities zu etablieren.

This paper outlines the historical development and methodical challenges of digital visual studies. Although it was argued early on that the strength of the computer lay in formal rather than iconographic analysis, in practice the iconographic encoding of images and the collection of metadata dominated for a long time. However, scholars and practitioners of digital visual studies soon demanded a thorough ›Ideologiekritik‹ (›critique of ideology‹) of its methods. Today, the field encompasses a broad spectrum of formal and iconographical, as well as reception-based studies. Furthermore, it can help to establish the critical reflection on methodical implications as a core issue for the digital humanities.

1. Disclaimer

Ich möchte diesen Beitrag¹ mit einer Art Disclaimer beginnen: Wenn ich hier und heute als Kunsthistorikerin und als ›Advokatin des Bildes‹ spreche, dann bedeutet das keineswegs, dass eine ›digitale Kunstgeschichte‹ ausschließlich im Bereich der Sammlung, Analyse und Interpretation von Bilddaten aktiv sein sollte. Eine rein bildbasierte Wissenschaft ist wohl weder möglich noch erstrebenswert. Auch Kunsthistorikerinnen und Kunsthistoriker wenden sich mich mit Worten an ihr Publikum, publizieren ihre Forschungen in Form von Texten. Auch für die Kunstgeschichte liegt ein großes Potential sowohl in den neuen Formen der Wissenschaftskommunikation und -organisation, die digitale Medien bieten, als auch in der Text- und (Meta-)Datenanalyse.

Jedoch nehmen unsere Forschungen ihren Ausgangspunkt in historischen und zeitgenössischen Artefakten, die (meist) visuell und (noch) zum überwiegenden Teil materiell sind. Daher sind wir natürlich daran interessiert, wie sich diese visuelle und materielle Überlieferung mittels digitaler Technologien analysieren lässt und welche Rolle wiederum

¹ Da dieser Beitrag als einführende Keynote im Rahmen der DHd-Tagung in Passau konzipiert war, wird für die schriftliche Version der mündliche Charakter des Vortrags weitestgehend beibehalten.

visuelle Medien und bildbasierte Technologien – als Medium dieser Analyse selbst – spielen können.

Denn neben der meist auf Giorgio Vasari zurückgeführten, textlich fixierten Kunstgeschichtsschreibung gab es immer auch Versuche, die Geschichte der Kunst anhand von Objektsammlungen und Zusammenstellungen visueller Reproduktionen zu (re)konstruieren. Während erstere zur Einrichtung von Museen und Ausstellungen führten, reichen letztere von frühen Galeriebildern über Stichsammlungen und Gipsabgusssammlungen bis hin zu Aby Warburgs Mnemosyne-Atlas. Letztlich basieren sie alle auf der Idee des Erkenntnisgewinns durch vergleichendes Sehen.

Fragt man aber nach den Ursprüngen der digitalen Kunstgeschichte, so schließt diese weder direkt an die textuelle Analyse noch an die Tradition des vergleichenden Sehens an. Ihre Anfänge liegen vielmehr in Versuchen der Bildcodierung, die sich zunächst dem Bildinhalt (der Ikonographie) und darauf der formalen Komposition zuwandten. Damit komme ich zum zweiten Abschnitt meines Beitrags, den ich biographisch einleiten möchte.

2. Von ICONCLASS zur Ideologiekritik

Der erste internationale Kunsthistorikerkongress, den ich besucht habe, fand 1992 in Berlin statt. Auf diesem Kongress gab es einen Stand des **Deutschen Dokumentationszentrums Kunstgeschichte**, das damals noch unter dem Namen Foto Marburg firmierte. Beim Besuch an diesem Stand lernte ich vier Akronyme kennen: HiDA, MIDAS, DISKUS und ICONCLASS. **HiDA**, der ›Hierarchische Dokument-Administrator‹, ist eine Software, mittels derer **MIDAS**, das ›Marburger Informations-, Dokumentations- und Administrations-System², bereitgestellt wird. **DISKUS** steht schließlich für ›Digitales Informationssystem für Kunst und Sozialgeschichte‹, ein von der Volkswagenstiftung geförderter Datenverbund, dem sich zunächst große deutsche Museen, später auch Archive und Denkmalämter anschlossen, um ihre mittels HiDA-MIDAS erfassten Daten kontinuierlich auszutauschen.³

Ich war begeistert von der vorgestellten Vision, mit diesen Systemen eine große, kollaborative Bilddatenbank zur Kunst und Architektur zu erstellen. Da ich in Köln studierte und das dortige Rheinische Bildarchiv Projektpartner von Foto Marburg war, sprach ich sofort dort vor und wurde prompt als studentische Hilfskraft angestellt. Hier lernte ich dann **ICONCLASS** kennen, ein System zur Dezimalklassifikation von Ikonographie, das – wie ich lernen sollte – seit den 1950er Jahren durch den niederländischen Kunsthistoriker Henry van de Waal entwickelt worden war. Mithilfe dieses Systems kann man eine Abendmahlsdarstellung, die den Streit unter den Aposteln darüber, wer unter ihnen der Größte ist, zeigt, in den Dezimalcode **73D22** übersetzen, der aus der hierarchischen Folge ›7 Bibel / 73 Neues Testament / 73D Passion Christi / 73D2 Abendmahl / 73D22 Streit unter den Aposteln, wer unter ihnen der Größte ist‹ resultiert. Solche Verschlagwortungen erscheinen uns im Zeitalter von Web 2.0 vielleicht recht

² Vgl. Bove et al. 2001.

³ Vgl. Laupichler 1998, *passim*.

normativ, aber sie ermöglichen es heute, im [Online-Archiv](#) von Foto Marburg faszinierende ikonographische Suchanfragen durchzuführen.

Prägend für meine Sozialisation innerhalb dessen, was wir heute ›digitale Geisteswissenschaften‹ nennen, waren allerdings weniger DISKUS und ICONCLASS, als die Tatsache, dass die Museumsinformatik an den Kölner Museen schon damals unter der Ägide von Tobias Nagel stand. Er ist nicht nur Theologe und Kunsthistoriker, sondern auch begeisterter Fachinformatiker. Dennoch trat er vehement für eine Ideologiekritik der ›EDV im Museum‹ (so hieß es damals noch unpräzise) ein: »Nicht das Prinzip des Machbaren hat zu gelten«, so Nagel, »sondern jeder einzelne Schritt muss auf seine Konsequenzen hin untersucht werden.«⁴ Ich lernte also, bei der Anwendung digitaler Technologien immer kritisch zu hinterfragen, in welcher Relation sie zur kunsthistorischen Methodik stehen und ob sie dem zu bearbeitenden Forschungsgegenstand gerecht werden.

Die Frage nach den methodischen Konsequenzen des Einsatzes digitaler Technologien stellte zur gleichen Zeit auch der britische Kunsthistoriker William Vaughan, Spezialist für die Kunst des 19. Jahrhunderts und langjähriger Vorsitzender der bereits 1986 gegründeten, britischen [CHArt](#) (Computers and the History of Art) Vereinigung. Vaughan war davon überzeugt, dass digitale Technologien der kunsthistorischen Forschung die Möglichkeit eröffneten, sich von der Orientierung an der Linguistik als »Meisterdisziplin«⁵ wieder zu lösen.

Dies bringt uns zurück zu dem Vergleich von textbasiertem und bildbasiertem Ansatz. Während der *linguistic turn* dazu geführt habe, formanalytische Methoden zugunsten ikonographischer Analysen zu vernachlässigen, liege, so Vaughan »[i]m Hinblick auf die Computerisierung der Kunstgeschichte [...] die Ironie dieser Situation darin begründet, daß sich der formanalytische Zugriff erheblich besser automatisieren lässt als der ikonographische und der erzähltheoretische.«⁶ Seine These basiert auf der Überzeugung, dass Maschinen formale Charakteristika besser identifizieren können als ikonographische. Auch wenn diese Argumentation einleuchtet, hat sie sich auf die breite Praxis der digitalen Kunstgeschichte lange Zeit kaum ausgewirkt. Dies liegt daran, dass Vaughans These auf der Annahme einer Automatisierung der kunstwissenschaftlichen Bildanalysepraktiken beruhte. Auch wenn sich – darauf komme ich im Folgenden zurück – formanalytische Ansätze besser automatisieren lassen, so basieren unsere großen digitalen Bildrepositorien wie [prometheus](#) oder der [Marburger Bildindex](#) jedoch bis heute primär auf von Menschen vorgenommenen, textuellen Verschlagwortungen von Bildinhalten und externen Informationen. Die automatische Bildanalyse ist hingegen – noch? – auf einzelne Forschungsprojekte beschränkt.

Für mich persönlich wurde die Notwendigkeit einer ideologiekritischen Reflexion jedoch weniger hinsichtlich einer Abwägung zwischen formanalytischen und ikonographischen Analyseansätzen relevant, als vielmehr in dem Moment, als mir 1994 die Aufgabe übertragen wurde, die Neuerwerbungen des Museum Ludwig mittels HiDA-MIDAS zu inventarisieren. Eines

⁴ Nagel 1997, S. 95.

⁵ Vaughan 1997, S. 98.

⁶ Vaughan 1997, S. 98.

der ersten Werke, das ›auf meinen Tisch kam‹, war eine Arbeit des amerikanischen Künstlers James Lee Byars. Er hatte gerade den Wolfgang-Hahn-Preis gewonnen und führte eine seiner in diesem Zusammenhang vom Museum erworbenen Arbeiten anlässlich der Preisverleihung auf. Die Arbeit mit dem Titel *The Perfect Smile* besteht in einer Performance, in der der Künstler dem Publikum für einen kurzen Moment sein Lächeln schenkt. Selbstverständlich wollte Byars' Lächeln als Neuerwerbung des Museum Ludwig ordnungsgemäß inventarisiert werden. Informationen wie Inventarnummer und Ankaufsdatum ließen sich denn auch problemlos erfassen. Selbst der ›lächelnde Mann‹ ist noch mittels ICONCLASS codierbar. Doch eigentlich wird er ja nicht dargestellt, sondern er verkörpert das Werk. Und welcher Gattung ist die Arbeit zuzuordnen? Was ist ihr Material? Besteht die Arbeit wirklich im Lächeln selbst – oder doch eher in der Idee des Verkaufs eines Lächelns?

Sicher hatten die Schwierigkeiten, die ich mit der datenbanktechnischen Erfassung dieser Arbeit hatte, zum Teil damit zu tun, dass Byars' Lächeln ein Werk der Performancekunst ist, während HiDA-MIDAS seinen Ausgangspunkt in der Erfassung von Objekten nimmt. Das eigentliche Problem liegt aber darin begründet, dass Werke der Kunst sich einer eindeutigen ontologischen Festlegung oft entziehen. Nicht ohne Grund verfügt die Kunstgeschichte ja neben der ikonographischen und formalanalytischen Methode noch über viele andere, kultur- und sozialhistorische sowie rezeptionsästhetische Analysemethoden – die Frage ist nur, wie diese sinnvoll in eine digitale Kunstgeschichte zu integrieren sind. Darauf wird zurückzukommen sein.

Mir wurde spätestens bei der Erfassung von Byars' Lächeln klar, wer eigentlich von der computergestützten Erfassung am meisten profitiert: der Erfasser selbst. Datenbanklogik und individuelle Charakteristika eines Kunstwerks zusammenzubringen erfordert von der Wissenschaftlerin oder dem Wissenschaftler eine äußerst intensive Auseinandersetzung mit der jeweiligen Arbeit; sie oder er wird gezwungen, deren Eigenheiten auf den Grund zu gehen, um diese entsprechend abstrahiert in einer Datenbank repräsentieren zu können. Denn die Datenbank stellt mittels vorgegebener Felder auch genau die Fragen, die wir in der traditionellen Werkanalyse allzu gerne immer dann ignorieren, wenn sie unserem aktuellen Forschungsinteresse entgegenstehen, schwer beantwortbar oder vermeintlich ›irrelevant‹ sind.

Das hat beispielsweise auch die Künstlerin Nana Petzet realisiert und die Inventarisierung ihrer aus Alltagsgegenständen selbst gefertigten Objekte mittels HiDA-MIDAS als eigenständiges, institutionskritisches Kunstprojekt präsentiert. In ihrem *System SBF* (seit 1995) – schon der Titel ironischer Kommentar zum ›Sammeln-Bewahren-Forschen‹-Paradigma der musealen Kunstgeschichte – lud sie zur Inventarisierung ihrer Objekte mittels HiDA-MIDAS ein. Damit stellte sie die Archivare beispielsweise vor die Herausforderung, die verschiedenen Plastiktüten, aus denen ein von der Künstlerin genähter Duschvorhang bestand, minutiös nach jeweiligem Hersteller, Herstellungsort und Ikonographie zu dokumentieren oder die ›Provenienz‹ mehrfach wiederverwendeter und umfunktionierter Gegenstände – von Schenkungen der Großmutter über Umarbeitungen durch den Onkel, über partielle Zerstörung

durch einen Gast, die Reparatur durch den Besitzer bis zur Aneignung als Kunstwerk durch die Künstlerin – en détail zu erfassen (Abbildung 1).⁷



Abb. 1: © Cuxhavener Kunstverein, Nana Petzet: System SBF. Inventarisierung der Sammlung. Metainventur, 2003, Nana Petzet.

3. Maschinenlesbarkeit: Vom Digitalisat zur Formanalyse

Ich habe im ersten Teil meines Beitrags drei wichtige Fragen der digitalen Geisteswissenschaften angesprochen: Wie repräsentiere ich materielle Artefakte in Form maschinenlesbarer Information? Wie verhalten sich digitale Technologien zu tradierten Methoden des Faches? Und wie gehe ich mit Kunstprojekten um, die weder Text noch Bild, sondern ›Gebilde‹ oder sogar Prozesse sind?

Keihen wir dennoch zunächst zum zweidimensionalen Bild zurück – und zwar in Bezug auf die Frage nach seiner Maschinenlesbarkeit. Gerade in der Frühzeit der Computergraphik lag der Einwand nahe, dass das Bild in seiner traditionellen Form als individuell gefertigtes Artefakt eben nicht aus diskreten Informationseinheiten besteht, was zur Folge hat, dass bei seiner digitalen Reproduktion unweigerlich Informationen verloren gehen.

Heute stellt diese Tatsache für den Großteil der kunsthistorischen Forschung allerdings eher ein ontologisches als ein praktisches Hindernis dar. Wir sind es gewohnt, mit Reproduktionen zu arbeiten, kennen aber auch ihre Grenzen und wissen, dass bei spezifischen stilistischen oder rezeptionsästhetischen Forschungsfragen letztlich ohnehin die Konsultation des Originals notwendig ist. Dass die Reproduktionen, mit denen wir arbeiten, lange Zeit auf analoger Fotografie basierten, heute aber auf digitalen Technologien, ist für die praktischen Fragestellungen der Kunstgeschichte letztlich auch weniger bedeutend. Auch die analoge Fotografie löst ihr Motiv durch die Körnung des Films in winzige Einzelteile auf, die die Materialität des Originals wenn überhaupt nur noch abbilden, nicht verkörpern.

Der durch Digitalisierung verursachte Informationsverlust ist in der Praxis meist zu vernachlässigen, übersteigt doch die mögliche Auflösung unser Wahrnehmungsvermögen

⁷ Kat. Ausst. System SBF. Inventarisierung der Sammlung. Metainventur 2003.

um ein Vielfaches: Sollten Sie einmal ins österreichische Linz kommen, besuchen sie das sogenannte **Deep Space** des Ars Electronica Center und lassen Sie sich Leonardos *Abendmahl* auf der dortigen 16 x 9m großen Projektionsfläche anzeigen (Abbildung 2). Diese Reproduktion wurde aus über 1500 Abbildungen vom **Haltadefinizione Lab** zusammengestellt und hat eine Gesamtgröße von über 16 Milliarden Pixeln. Bei einer Größe des Originals von 4,60 x 8,80m sind das 20 Pixel pro Millimeter, in der Projektion immer noch 10. Zwar lässt sich diese Auflösung für Detailaufnahmen auch schon in der analogen Fotografie erzielen, aber eine 2:1 Gesamtproduktion in dieser Größe – zoombar und scrollbar – ist mit analogen Mitteln unmöglich.



Abb. 2: Deep Space. © Archiv Ars Electronica Linz GmbH, Fotograf Rudolf Brandstätter.

Solcherart hochauflösende Abbildungen faszinieren nicht nur – wie in Linz – ein breites Publikum, sondern sind beispielsweise für materialtechnische Untersuchungen von Einzelwerken äußerst wertvoll. Die vergleichende Analyse hingegen arbeitet interessanterweise oft gerade mit der Reduzierung der Bilddaten oder deren Verfremdung mittels Filtern. Entsprechende Projekte sind in der einschlägigen Literatur zur digitalen Kunstgeschichte gut dokumentiert,⁸ ich möchte hier daher nur kurz einige prägnante Beispiele in Erinnerung rufen:

Ich hatte anfangs bereits William Vaughans These zitiert, dass sich formale Analysen besser automatisieren ließen als ikonographische. Vaughan hat es keineswegs bei dieser theoretischen Überlegung belassen, sondern als Vorreiter auf diesem Gebiet ein eigenes Forschungsprojekt durchgeführt: Das von ihm konzipierte, nach dem frühen italienischen Kunstkritiker Giovanni Morelli benannte Programm ermöglichte bereits Ende der 1980er Jahre die Suche nach formal bzw. kompositorisch ähnlichen Bildern mithilfe eines in seiner Datenmenge stark reduzierten ›Visual Identifiers‹ (Abbildung 3).⁹

⁸ Vgl. Kohle 1997; Kohle / Kwastek 2003; Kohle 2013.

⁹ Vaughan 1997, *passim*.

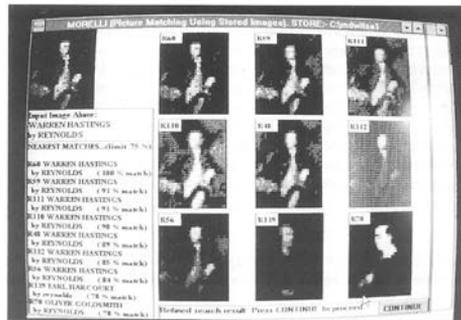


Abb. 3: »Screenshot mit Abfrageresultat aus »Morelli: Die Suche wurde ausgehend von einem Portrait Warren Hastings von Joshua Reynolds durchgeführt. Das System identifiziert andere Bilder und Reproduktionen, die mit dem Reynolds-Portrait in Beziehung stehen«. Quelle: William Vaughan: Computergeschützte Bildrecherche und Bildanalyse. In: Kunstgeschichte digital. Eine Einführung für Praktiker und Studierende. Hg. von Hubertus Kohle. Berlin 1997.

Hinsichtlich computergestützter Verfahren zur stilistischen Analyse war für viele von uns ein 2003 von Stefan Heidenreich durchgeführtes Experiment prägend: Er versuchte mittels recht einfacher Verfahren, nämlich durch Subtraktion eines weichgezeichneten Bildes vom (digitalisierten) Original, die von Heinrich Wölfflin 1915 vorgenommene Unterscheidung des linearen Stils der Renaissance vom malerischen Stil des Barock rechnerisch nachzubilden. Das verblüffend deutliche Ergebnis (Abbildung 4)¹⁰ gab Grund zu der Erwartung, dass stilistische Analysen in Zukunft vermehrt mit digitalen Technologien unterstützt werden könnten.



Abb. 4: Gegenüberstellung von zwei Frauenaktent. Aus: Heinrich Wölfflin: Kunstgeschichtliche Grundbegriffe. München 1915, S. 36/37 (oben); Subtraktion einer weichgezeichneten Version der obigen Abbildungen von eben diesen Abbildungen. In: Stefan Heidenreich: Form und Filter - Algorithmen der Bilderverarbeitung und Stilanalyse. Zeitenblicke 2 (2003), Nr. 1 (unten). Quelle: © [Zeitenblicke](#), 2003.

Tatsächlich sind in den letzten zehn Jahren die Methoden, die zur automatisierten Bildanalyse verwendet werden, maßgeblich weiterentwickelt worden und können beileibe nicht mehr auf die einfache Formel von »Stil und/oder Ikonographie« gebracht werden.

¹⁰ Heidenreich 2003.

Ein Beispiel für eine höchst komplexe, auf der Kombination ikonographischer und stilistische Fragestellungen beruhende Analyse ist das Heidelberger Forschungsprojekt zur **Gestenerkennung in mittelalterlichen Handschriften**. Ein anderes ist ein laufendes Projekt an der University of California, Riverside, das die mögliche Identifizierung von historischen Porträts mithilfe von Gesichtserkennungssoftware untersucht. Verständlicherweise betonen die dortigen Kollegen, ein solches Unterfangen sei ungleich schwieriger als die Identifizierung lebender Menschen mittels hierfür ausgerichteter Kameras, da in der Analyse künstlerischer Porträts zunächst vom Individualstil der Künstler abstrahiert werden müsse.¹¹

4. Rezeptionsforschung

Es ist zu hoffen, dass die Künstler der von den kalifornischen Kollegen untersuchten Porträts nicht ähnlich kritisch mit der Frage nach einer möglichen Identifizierbarkeit umgehen, wie der amerikanische Künstler Adam Harvey in seinem *CV Dazzle* (2012) (Abbildung 5). In einem zwischen künstlerischer Forschung, Medienaktivismus und Konzeptkunst zu lokalisierendem Ansatz entwickelt er Makeup und Frisuren, die sich gerade dadurch auszeichnen, dass sie ihren Träger vor Identifikation durch automatisierte Gesichtserkennungssysteme schützen.

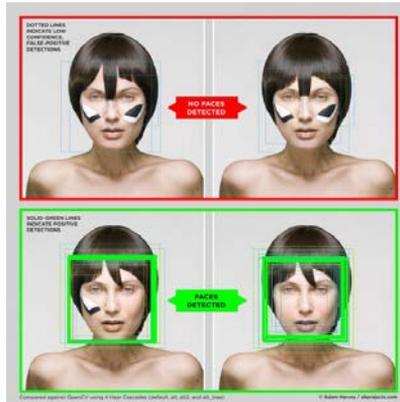


Abb. 5: *CV Dazzle*. © Adam R. Harvey, 2010.

Keinesfalls geht es mir bei dieser Gegenüberstellung darum, den interessanten Forschungsansatz der amerikanischen Kollegen zu kritisieren – ich möchte etwas anderes deutlich machen: Kunst bedarf der Wahrnehmung, und Wahrnehmung ist ein extrem komplexer neurologischer und kognitiver Prozess. Künstler instrumentalisieren ihn, Kunstwissenschaftler analysieren ihn, und Informatiker suchen ihn zu simulieren. Wie die bisher erwähnten Projekte zeigen, bieten sich digitale Technologien offensichtlich zunächst zur Behandlung von traditionellen Themen der Kunstgeschichte an: Ikonographie, Stilgeschichte,

¹¹ Miller 2013.

Zuschreibungsfragen, Motiv-Identifikation. Wie können aber Fragen des Umgangs mit Bildern, ihrer Wahrnehmung und ihres Kontexts in den Digital Humanities einen Platz finden?

Raphael Rosenberg, Kunsthistoriker in Wien, hat ausgehend von seinem Interesse an literarischen Beschreibungen von Kunstrezeptionen begonnen, tatsächliche Prozesse der Kunstbetrachtung zu untersuchen. Gemeinsam mit Kollegen aus der Psychologie nutzt er Eye-Tracking-Technologien, um in einem ›Labor für empirische Bildbetrachtung‹ Blickbewegungen von Kunstrezipienten zu analysieren (Abbildung 6).¹²



Abb. 6: Visualisierung einer Blickanalyse. Quelle: Zwischen Kunst und Alltag (Interview mit Helmut Leder und Raphael Rosenberg). © Labor für empirische Bildwissenschaft, Universität Wien, 2012.

An die Frage, wie etwas individuell wahrgenommen wird, schließt sich dann aber die Frage an, wie es individuell und kontextabhängig interpretiert wird. Möglicherweise gelangen wir hier an die Grenzen einer bildbasierten Analyse bzw. an den Punkt, wo gerade die Kombination bildbasierter und textbasierter Verfahren erfolgversprechend wird. Raphael Rosenberg stellt seine Ergebnisse historischen Bildbeschreibungen gegenüber. Hubertus Kohle lässt in seinem ARTigo-Spiel möglichst viele Internetnutzer Tags für Bilder vergeben, auch um durch deren Analyse Hinweise auf individuelle Interpretationen und Wahrnehmungsmuster zu erhalten.

Mein Eindruck ist, dass die Kunstgeschichte hier – in der Entwicklung von Methoden, die digitale Analyseverfahren und menschliche Interpretationsleistung intelligent kombinieren – noch relativ am Anfang steht, dass aber digitale Technologien dazu beitragen können, die traditionelle, auf den impliziten Betrachter gerichtete Rezeptionsästhetik stärker um eine empirisch gestützte Rezeptionsforschung zu ergänzen. Und ich möchte auch hier noch ein bisschen weiter gehen und noch einmal auf eine bereits zu Beginn gemachte Beobachtung zurückkommen, dass Kunst sich nicht auf das visuelle Artefakt beschränkt. – Was bedeutet also Rezeptionsforschung in diesem weiteren Sinne?

Kunstrezeption geschieht nicht nur durch Blickbewegungen, sondern auch durch Körperbewegungen. Während die Analyse von Bewegungen im Raum für die

¹² Betz et al. 2009, passim.

architekturwissenschaftliche Forschung, aber auch etwa im Rahmen von *curatorial studies* meines Wissens noch wenig genutzt wird, haben Künstler sich für die Möglichkeiten, die neue Technologien diesbezüglich bieten, schon früh interessiert. Gleich mit der kommerziellen Einführung von GPS-Geräten entstanden Projekte wie *Amsterdam Realtime* (2002–2003): Esther Polak hat in diesem Projekt mit Unterstützung der Waag-Society einen Stadtplan von Amsterdam erstellt, der nur auf der GPS-gestützten Aufzeichnung von realen Bewegungen von Bewohnern basiert. Auch hier interessiert uns dann aber natürlich, wie oben bereits erwähnt, nicht nur die Bewegung oder Aktion der Rezipienten selbst sondern auch deren Interpretation und Kontext. Zu diesem Zweck sind dann wiederum kombinierte Methoden notwendig.

An dieser Stelle möchte ich ein Projekt erwähnen, welches ich selbst im Rahmen meiner Forschungen zur interaktiven Medienkunst¹³ durchgeführt habe, denn gerade die partizipative und interaktive Kunst stellt die individuelle Erfahrung des einzelnen Rezipienten ins Zentrum ihrer ästhetischen Strategien. – Wie kann man diese also untersuchen, ohne– wie sonst oft in der Kunstgeschichte üblich– nur von seiner eigenen Erfahrung als einer Art idealem Betrachter bzw. idealer Betrachterin auszugehen? Gemeinsam mit zwei Kolleginnen aus Kanada und Australien habe ich hierfür eine von ihnen als Video Cued Recall bezeichnete Methode angewandt.¹⁴ Sobald die Rezeption sich nicht auf das reine Anschauen beschränkt, sondern Gesten, Bewegungen, oder sogar Sprache einschließt, lassen sich diese Aktionen per Video dokumentieren. Dies haben wir getan, um die entstandene Dokumentation anschließend den jeweiligen Rezipienten noch einmal vorzuspielen, und sie zu bitten, ihre eigene Aktion zu kommentieren. Dieser Kommentar wurde als eine Art Voice-Over mit der Dokumentation gespeichert. Auf diese Weise erstellte Dokumentationen verschiedener Werkrezeptionen ergeben zusammen mit Künstlerinterviews eine sehr interessante, multiperspektivische Dokumentation. Eine solche Dokumentation haben wir für eine Arbeit des amerikanischen Künstlerduos Tmema mit dem Titel *Manual Input Workstation* (2004–2006) erstellt. Sie ist über das Online-Archiv der Daniel Langlois Foundation zugänglich. In der Interaktion mit dieser Arbeit kann man mit einem handelsüblichen Tageslichtprojektor Schattenformationen erzeugen, die dann digital analysiert und mit animierten Formen und Sounds überlagert werden. Es geht bei dieser Arbeit weniger um das audiovisuelle Ergebnis selbst, als um die Erfahrung der Interaktion mit Formen und Tönen. Daher kann man sie letztlich nur dann angemessen analysieren, wenn man verschiedene Besucherinteraktionen beobachtet hat, idealerweise ergänzt durch Reaktionen oder Interpretationen verschiedener Besucher. Dies wird durch die oben genannte multiperspektivische Dokumentation möglich. Allerdings – und dies möchte ich ausdrücklich betonen – ist diese Form Dokumentation letztlich gar nicht auf avancierte digitale Technologie angewiesen. Während die Arbeit digital gesteuert wird, basiert die erwähnte Dokumentation nicht zwangsläufig auf digitalen Technologien sondern auf einer gewöhnlichen Videoaufnahme. Zwar machen digitale Technologien das Zusammenschneiden verschiedener Spuren um ein vielfaches einfacher, möglich war dies aber auch mit analogen Mitteln. Und damit kommen wir zurück zur Frage der Ideologiekritik: Primär muss es darum gehen, Analysemethoden zu finden, die dem Untersuchungsgegenstand und der Fragestellung am besten dienen. Das kann eine crowdbasierte Sammlung von Metadaten

¹³ Kwastek 2013, passim.

¹⁴ Müller 2008.

sein, eine graphische Wiedergabe von Bewegungsmustern, oder eben eine videogestützte Analyse.

Im letzten Teil meines Beitrags möchte ich jetzt wieder zum Bild zurückkommen, diesmal aber zum Bild als Medium der Forschung. Bezüglich des Nutzens von Visualisierungen für die Kunstgeschichte denkt man meist zunächst an die Architekturgeschichte, die digitale Modelle zur Rekonstruktion zerstörter Gebäude, zur Visualisierung von Bauphasen oder zur Untersuchung von Hypothesen bezüglich historischer Eingriffe verwendet. Dies ist ein sehr wichtiger und vielversprechender Aspekt der digitalen Kunstgeschichte, auf den ich hier aber nicht weiter eingehen möchte. Stattdessen sollen zwei andere Formen der Visualisierung angesprochen werden: die interpretierende Visualisierung am Bild und die Visualisierung sogenannter ›Big Data‹.

5. Visualisierung



Abb. 7: Jan van Eyck: *Die Madonna in der Kirche*, Ident. Nr. 525C. © Foto: Gemäldegalerie, Staatliche Museen zu Berlin, Fotograf/in: GG-Fotowerkstatt. CC BY-NC-SA 3.0.

Die Berliner Museen haben in den 1990er Jahren eine Reihe von werkmonographischen CDs herausgebracht, die interaktive Bildanalysen ermöglichen. Eine davon widmet sich der *Madonna in der Kirche* (ca. 1440) (Abbildung 7)¹⁵, dem berühmten ›Gemälde Jan van Eycks, das die Madonna nicht nur in der Kirche, sondern vielmehr *als Kirche* darstellt. Die monumentale Größe der Figur im Verhältnis zu der sie beherbergenden Architektur wurde in wegweisenden Texten interpretiert, wie etwa hier durch Otto Pächt:

»Nur zu Jan paßt es, daß die Madonna wie eine riesige Statue in einem allerdings luftigen Schrein steht, denn für ihn ist räumliche Einpassung in erster Linie nicht eine Frage der Größenverhältnisse – alle Janschen Architekturszenarien sind zu niedrig –, sondern der meßbaren Beziehung von Körpervolumen und Raumschale. [...] Von der abstrakten Idee der

¹⁵ Abbildung zu finden in der [Bildergalerie der Berliner Gemäldegalerie](#).

Personifikation der Kirche durch Maria hat sich die künstlerische Phantasie eines unentwegten Empirikers zu dem Bilde des realen Kirchengebäudes inspirieren lassen, das nun als Umwelt der Himmelskönigin beigelegt wird. So wird aus Maria als die Kirche Maria in der Kirche.«¹⁶

Gerade vor dem Hintergrund solcher interpretierender Beschreibungen ist die von den Berliner Museen angebotene Möglichkeit, die Größe der Madonna im Verhältnis zum Innenraum in einer digitalen Reproduktion zu manipulieren, interessant (Abbildung 8).¹⁷ Es handelt sich hier um eine Form des »vergleichenden Sehens«, allerdings nicht im Sinne eines Vergleichs zweier Kunstwerke, sondern eines Vergleichs der Auswirkungen von am Bild vorgenommenen Manipulationen, von imaginären Bildalternativen, die jedoch helfen, die im Original gewählte Lösung zu verstehen. Das Experimentieren mit dieser interaktiven Anwendung macht schnell deutlich, wie unterschiedlich Argumentations- und Erkenntnisprozesse in der textuellen Interpretation einerseits, der explorativen Manipulation andererseits verlaufen. Letztere kann erstere keineswegs ersetzen, sondern ist ihr komplementär.



Abb. 8: Eyck. CD: Manipulierte Reproduktion *Die Madonna in der Kirche* © Gemäldegalerie, Staatliche Museen zu Berlin auf der Grundlage von Abb. 7. CC BY-NC-SA 3.0.

Solche Formen der interaktiven Bildanalyse bleiben in der Kunstgeschichte dennoch bis heute die Ausnahme. Visualisierungen dienen in erster Linie nicht zur Analyse einzelner Kunstwerke sondern zur Darstellung der Ergebnisse von Analysen großer Datenmengen. Solche Analysen werden durch computergestützte Methoden extrem vereinfacht. Da ihre Ergebnisse jedoch oft nur schwer textuell vermittelbar sind, bietet es sich geradezu an, visuelle Darstellungen derselben anzubieten. Dies kann in Form von Zeitleisten geschehen, aber auch als Visualisierung von Netzwerkanalysen. In verschiedenen Projekten, die wir am Ludwig Boltzmann Institut *Medien.Kunst.Forschung* in Linz durchgeführt haben, ging es um die Erstellung von Themenlandschaften, die versuchen, thematische und konzeptuelle Zusammenhänge räumlich darzustellen, um die Visualisierung quantifizierbarer Entwicklungen, die etwa die Anzahl von Einreichungen von Werken in bestimmte Preiskategorien bei Wettbewerben darstellen und sogar um die Visualisierung der Verteilung von Begriffen im Text als einer Form der visuellen Diskursanalyse (Abbildung 9).¹⁸

¹⁶ Pächt 1993, S. 205.

¹⁷ Ein kurzer Film, der diese CD vorstellt, ist [online](#) abrufbar.

¹⁸ Vgl. Yavuz 2009 [[online](#)].

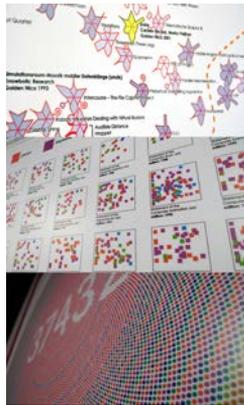


Abb. 9: Evelin Münster: »A Theme Landscape for Tagged Data«, Ludwig Boltzmann Institut *Medien.Kunst.Forschung* 2009 (oben); Jaume Nualart: »Texty«, Ludwig Boltzmann Institut *Medien.Kunst.Forschung* 2009 (Mitte); Moritz Stefaner: »X by Y«, 2009 [\[online\]](#) (unten).

All diesen Projekten gemein ist jedoch letztlich eine Visualisierung von als Text vorliegenden Daten, auch wenn diese visuelle Artefakte beschreiben oder auf deren Existenz basieren. Anders ist dies bei Projekten, die Lev Manovich, Kunsthistoriker an der City University of New York, durchführt. Seine Visualisierungen basieren auf Zusammenstellungen großer Mengen digitaler Reproduktionen. So analysiert er z. B. Reproduktionen von **Gemälden van Goghs aus verschiedenen Schaffensperioden auf deren Helligkeit und Sättigung**. Die vergleichende Anordnung in einem Koordinatensystem zeigt für die verschiedenen Schaffenssorte und damit -perioden ein verblüffend ähnliches Muster und stellt damit die gängige Auffassung in Frage, van Goghs Stil habe je nach Aufenthaltsort stark variiert. Während das van Gogh-Beispiel die kritische Überprüfung einer These darstellt, interessiert Manovich bei diesen Forschungen letztlich die Frage, inwiefern solche Bildzusammenstellungen neue Erkenntnisse liefern, ohne dass man mit einer direkten Fragestellung an sie herangegangen wäre:

»However, for media researchers, starting with well-formulated questions does not exploit what visualization is best at: exploring without preconceptions a large dataset to discover ›what is there‹ and revealing unexpected patterns, as opposed to testing already formulated ideas. We can call such process exploratory visualization.«¹⁹

Die Tragweite der Frage, wie wir große Mengen von Bildern im Zusammenhang analysieren können, wird uns in Zukunft mit Sicherheit weiter beschäftigen –, und dies schlichtweg, weil die Menge der produzierten Bilder immer noch rasant ansteigt. Ich möchte daher mit einem weiteren medienkünstlerischen Projekt schließen, das nah an rezeptionsästhetischen Fragestellungen operiert. Die französische Künstlerin Corinne Vionnet hat immer wieder beobachtet, dass Touristen Bilder berühmter Sehenswürdigkeiten sehr häufig aus fast identischen Blickwinkeln aufnehmen. Um diese amateurfotographische Praxis weiter zu analysieren, hat sie in den sozialen Medien Fotos von Sehenswürdigkeiten gesucht und in ihrer Serie *Photo Opportunities* (2005–2013) (Abbildung 10) zu großen Bildern zusammengesetzt,

¹⁹ Manovich 2013, S. 257–258.

die damit so etwas wie unser kollektives Bildgedächtnis visualisieren. Solche Projekte lassen sich durchaus als eine Form der künstlerischen Bildwissenschaft bezeichnen, und können der geisteswissenschaftlichen Forschung, gerade im Bereich der Digital Humanities, interessante Impulse geben.



Abb. 10: © Corinne Vionnet: athína, 2006, series *Photo Opportunities*.

6. Vom Bild zum Bild

Wenn ich meinem Beitrag den Titel ›Vom Bild zum Bild‹ gegeben habe, so ging es mir darum zu betonen, dass wir es in den digitalen Geisteswissenschaften mit sehr unterschiedlichen Formen von digitalen und digitalisierten Bildern zu tun haben, die durch verschiedenste digitale Methoden – und damit immer über den Umweg des Codes – in neue Bilder übersetzt werden, sei es durch die digitale Reproduktion eines Originalgemäldes, die interaktive Animation einer digitalen Reproduktion, die Visualisierung der Wahrnehmung eines Kunstwerks, die Visualisierung quantitativer Daten zum Kunstmarkt oder die Visualisierung von Bildsammlungen anhand von Bildberechnungen. Die Digital Humanities analysieren nicht nur Bilder, sondern produzieren auch neue Bilder – auch insofern betreffen die ›Digital Humanities jenseits des Texts‹ nicht nur die Kunstgeschichte sondern in gleichem Maße die Disziplinen, die sich mit der Analyse schriftlicher Überlieferung befassen, und diese durch Visualisierungen präsentieren.

Wenn ich also meinen Beitrag mit dem Disclaimer begonnen habe, dass eine digitale Kunstgeschichte selbstverständlich großen Nutzen an textbasierten Methoden der Digital Humanities hat, so möchte ich mit der Bemerkung schließen, dass andererseits das Bild ein nicht zu unterschätzender Akteur der digitalen Geisteswissenschaften im Allgemeinen ist: auch Historiker und Literaturwissenschaftler haben schon immer Bilder als Quellen herangezogen, und sie benutzen auch in zunehmendem Maße Bilder, um die Analyseergebnisse großer Datenmengen zu visualisieren und lesbar zu machen.

Bibliographische Angaben

- Kat. Ausst. Nana Petzet: System SBF. Inventarisierung der Sammlung. Metainventur. Hg. vom Cuxhavener Kunstverein; Kunstverein und Stiftung Springhornhof. Neuenkirchen, Nürnberg 2003. [\[Nachweis im GBV\]](#)
- Juliane Betz / Martina Engelbrecht / Christoph Klein / Raphael Rosenberg: Dem Auge auf der Spur: Eine historische und empirische Studie zur Blickbewegung beim Betrachten von Gemälden. In: *Image. Zeitschrift für interdisziplinäre Bildwissenschaft* 11 (2009), S. 29–41. [\[Nachweis im GBV\]](#)
- Jens Bove / Lutz Heusinger / Angela Kailus: Marburger Informations-, Dokumentations- und Administrationssystem (MIDAS). Handbuch und CD. München 2001. [\[Nachweis im GBV\]](#)
- Stefan Heidenreich: Form und Filter – Algorithmen der Bilderverarbeitung und Stilanalyse. In: *zeitenblicke* 2 (2003), Nr. 1. [\[online\]](#)
- Hubertus Kohle / Katja Kwastek: *Computer, Kunst und Kunstgeschichte*. Köln 2003. [\[Nachweis im GBV\]](#)
- Hubertus Kohle: *Digitale Bildwissenschaft*. Glückstadt 2013. [\[Nachweis im OPAC\]](#)
- Kunstgeschichte digital. Eine Einführung für Praktiker und Studierende*. Hg. von Hubertus Kohle. Berlin 1997. [\[Nachweis im GBV\]](#)
- Katja Kwastek: *Aesthetics of Interaction in Digital Art*. Cambridge, MA 2013. [\[Nachweis im GBV\]](#)
- Fritz Laupichler: MIDAS, HIDA, DISKUS – was ist das? In: *Arbeitsgemeinschaft der Kunst- und Museumsbibliotheken-News* 4 (1998), Nr. 2/3, S. 18–24. [\[online\]](#)
- Lev Manovich: Museum Without Walls, Art History without Names: Visualization Methods for Humanities and Media Studies. In: *Oxford Handbook of Sound and Image in Digital Media*. Hg. von Carol Vernallis / Amy Herzog / John Richardson. Oxford 2013, S. 253–278. [\[Nachweis im GBV\]](#)
- Bettye Miller: Research on Application of Face-recognition Software to Portrait Art Shows Promise. In: *UCR Today*, 31. Mai 2013. [\[online\]](#)
- Lizzie Muller: Towards an Oral History of Media Art. In: *Publications by the Daniel Langlois Foundation*. 2008. [\[online\]](#)
- Tobias Nagel: Zur Notwendigkeit einer Ideologiekritik der EDV im Museum. In: *Kunstgeschichte digital. Eine Einführung für Praktiker und Studierende*. Hg. von Hubertus Kohle. Berlin 1997, S. 84–96. [\[Nachweis im GBV\]](#)
- Otto Pächt: *Van Eyck. Die Begründer der altniederländischen Malerei*. 1. Auflage 1989. München 1993. [\[Nachweis im GBV\]](#)
- William Vaughan: Computergestützte Bildrecherche und Bildanalyse. In: *Kunstgeschichte digital. Eine Einführung für Praktiker und Studierende*. Hg. von Hubertus Kohle. Berlin 1997, S. 97–106. [\[Nachweis im GBV\]](#)
- Mahir Yavuz: Mapping The Archive: 30 Years of Ars Electronica Visualized in Huge Scale. *Information Aesthetics Weblog*. 16.09.2009. [\[online\]](#)

Abbildungslegenden und -nachweise

Abb. 1: © Cuxhavener Kunstverein, Nana Petzet: System SBF. Inventarisierung der Sammlung. Metainventur, 2003, Nana Petzet.

Abb. 2: Deep Space. © Archiv Ars Electronica Linz GmbH, Fotograf Rudolf Brandstätter.

Abb. 3: »Screenshot mit Abfrageresultat aus ›Morelli«: Die Suche wurde ausgehend von einem Portrait Warren Hastings von Joshua Reynolds durchgeführt. Das System identifiziert andere Bilder und Reproduktionen, die mit dem Reynolds-Porträt in Beziehung stehen.« Quelle: William Vaughan: *Computergestützte Bildrecherche und Bildanalyse*. In: *Kunstgeschichte digital. Eine Einführung für Praktiker und Studierende*. Hg. von Hubertus Kohle. Berlin 1997.

Abb. 4: Gegenüberstellung von zwei Frauenakten. Aus: Heinrich Wölfflin: *Kunstgeschichtliche Grundbegriffe*. München 1915, S. 36/37 (oben); Subtraktion einer weichgezeichneten Version der obigen Abbildungen von eben diesen Abbildungen. In: Stefan Heidenreich: *Form und Filter - Algorithmen der Bilderverarbeitung und Stilanalyse*. *Zeitenblicke* 2 (2003), Nr. 1 (unten). Quelle: © *Zeitenblicke*, 2003.

Abb. 5: *CV Dazzle*. © Adam R. Harvey, 2010.

Abb. 6: Visualisierung einer Blickanalyse. Quelle: *Zwischen Kunst und Alltag* (Interview mit Helmut Leder und Raphael Rosenberg). © Labor für empirische Bildwissenschaft, Universität Wien, 2012.

Abb. 7: Jan van Eyck: *Die Madonna in der Kirche*, Ident. Nr. 525C. © Foto: Gemäldegalerie, Staatliche Museen zu Berlin, Fotograf/in: GG-Fotowerkstatt. CC BY-NC-SA 3.0.

Abb. 8: Eyck. CD: Manipulierte Reproduktion *Die Madonna in der Kirche* © Gemäldegalerie, Staatliche Museen zu Berlin auf der Grundlage von Abb. 7. CC BY-NC-SA 3.0.

Abb. 9: Evelin Münster: »A Theme Landscape for Tagged Data«, Ludwig Boltzmann Institut *Medien.Kunst.Forschung* 2009 (oben);
Jaume Nualart: »Texty«, Ludwig Boltzmann Institut *Medien.Kunst.Forschung* 2009 (Mitte); Moritz Stefaner: »X by Y«, 2009 [[online](#)]
(unten).

Abb. 10: © Corinne Vionnet: athína, 2006, series *Photo Opportunities*.

Zeitschrift für digitale Geisteswissenschaften

Artikel aus:

Sonderband 1 der ZfdG: Grenzen und Möglichkeiten der Digital Humanities. Hg. von Constanze Baum und Thomas Stäcker. 2015. DOI: [10.17175/sb01](https://doi.org/10.17175/sb01)

Titel:

Die Informatik als Geisteswissenschaft

Autor/in:

John Nerbonne

Kontakt:

john.nerbonne@frias.uni-freiburg.de

Institution:

Universität Groningen und Freiburg Institute for Advanced Studies, Universität Freiburg

GND:

[172289548](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-172289548)

ORCID:

DOI des Artikels:

[10.17175/sb001_003](https://doi.org/10.17175/sb001_003)

Nachweis im OPAC der Herzog August Bibliothek:

[830166114](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-830166114)

Erstveröffentlichung:

19.02.2015

Lizenz:

Sofern nicht anders angegeben 

Medienlizenzen:

Medienrechte liegen bei den Autoren

Letzte Überprüfung aller Verweise:

24.05.2016

GND-Verschlagwortung:

[Informatik](#) | [Digital Humanities](#) |

Zitierweise:

John Nerbonne: Die Informatik als Geisteswissenschaft. In: Grenzen und Möglichkeiten der Digital Humanities. Hg. von Constanze Baum / Thomas Stäcker. 2015 (= Sonderband der Zeitschrift für digitale Geisteswissenschaften, 1). PDF Format ohne Paginierung. Als text/html abrufbar unter DOI: [10.17175/sb001_003](https://doi.org/10.17175/sb001_003).

John Nerbonne

Die Informatik als Geisteswissenschaft

Abstracts

Der Titel des Beitrags bezieht sich indirekt auf das Thema der 2014 Konferenz der Digital Humanities im deutschsprachigen Raum (DHD): »Digital Humanities – methodischer Brückenschlag oder ›feindliche Übernahme?‹«. Ich gehe dabei von der Unterstellung aus, dass man sich die Informatik als Urheber einer solchen gedachten Übernahme vorstellt. Im ersten Abschnitt des Beitrags wird zunächst dargelegt, dass auch die Informatik, als Hauptbastion der ›Feinde‹, die unsere Disziplinen zu übernehmen drohen, eine unsichere Reifungsperiode gehabt hat, die sie jedoch glänzend überstanden hat. Man könnte demnach Inspiration oder wenigstens Mut aus der Geschichte dieser Disziplin schöpfen. Ein weiterer Abschnitt geht – getreu dem Titel des Beitrags – der Frage nach, wieso man die Informatik als Geisteswissenschaft betrachten kann, einschließlich einer historischen Betrachtung der Entstehungsperiode der zwei Disziplinen oder Disziplinengebiete. Die historische Dimension, aber auch die Unsicherheit, die sich im Konferenzthema widerspiegelt, nehme ich zum Anlass, um in einem sich daran anschließenden Teil zehn Thesen zu präsentieren, die erläutern, worauf es ankommen wird, wenn wir mit den Digital Humanities Erfolg haben wollen.

The title of this paper indirectly refers to the theme of the 2014 Conference on Digital Humanities in German-Speaking Areas (»Digital Humanities – methodischer Brückenschlag oder ›feindliche Übernahme?‹). I assume from this title that computer science is viewed as the originator of such a deliberate conquest. The first section of this paper explores how computer science, the headquarters of the »enemy« threatening to conquer our disciplines, also experienced and overcame an uncertain period of early development. The history of this discipline could serve as a source of inspiration or, at the very least, courage. The following section explores—as indicated by the title of this paper—how computer science could be viewed as a humanities discipline, including a historical perspective on the foundation of both disciplines or disciplinary fields. The historical dimension, but also the uncertainty, that is reflected in the conference theme serves as my starting point to present ten theses in the following section that explain the necessary factors for success if we wish to succeed with the Digital Humanities.

1. Motivation

Die Unsicherheit, die im Titel der Konferenz – »Digital Humanities – methodischer Brückenschlag oder ›feindliche Übernahme?‹« – mitschwingt, ist im Zuge einer großen (informationstechnischen) Veränderung durchaus üblich: Traditionelle Fachleute verlieren an Bedeutung und neuartige Spezialisten werden wichtiger, manchmal sogar wichtiger als etablierte Autoritäten. Elizabeth Eisenstein hat die bis dato folgenreichste Änderung in der Informationstechnologie, das Entstehen der Buchdruckkunst, 1979 *en detail* dargelegt.¹ Der Einzug der Informatik in viele andere Wissenschaftsgebiete, wie z.B. die Biologie, hat ähnlich große Verschiebungen in den dazu nötigen Expertisen verursacht, wobei manch kompetente

¹ Vgl. für die weitere Reflexion zu den Parallelen zwischen den wissenschaftlichen Konsequenzen der Erfindung des Buchdrucks im 15. Jahrhundert einerseits und der Einführung der digitalen Informationsverarbeitung im 20. Jahrhundert andererseits auch meine Ausführungen in *Elektronische Incunabeln*, Nerbonne 1995.

Vertreter älterer Richtungen von den neueren Entwicklungen überrannt worden sind (so die Feldbiologie angesichts der Entstehung der Molekularbiologie und der Bioinformatik).

Zweck dieses Beitrags ist es nicht, die durch solche Entwicklungen und Umschwünge vermittelten Unsicherheiten kleinzureden oder deren Problematik gar zu leugnen, denn im aktiven Wissenschaftsdiskurs wird sich letztlich auch in diesem Bereich die Spreu vom Weizen trennen. Wenn wir – oder eben andere, jene titulierten ›Feinde‹, – wissenschaftlichen Fortschritt in den Geisteswissenschaften mittels Informationstechnologie erzielen können, dann werden sich die digitalen Geisteswissenschaften durchsetzen. Ob gerade eine feindliche Übernahme von statten geht, möchte ich bezweifeln, aber eine groß angelegte digitale Wende in den Geisteswissenschaften würde sicher einiges ändern, und nicht jeder Geisteswissenschaftler oder jede Geisteswissenschaftlerin würde gleichermaßen hiervon profitieren. Gerhard Wolf kritisiert schon jetzt, dass »Digitalprojekte Ressourcen [verschlingen], die wir dringend für unser Kerngeschäft nötig haben.«²

2. Die Klassifizierung der Wissenschaften – wozu gehört die Informatik?

1806/07 hat Hegel in seiner *Phänomenologie des Geistes* bekanntlich eine strenge Unterscheidung zwischen Natur und Geist gezogen und darüber hinaus auch abstrakteren Entitäten, wie einer Kultur oder einem Volk, einen objektiven Geist zugesprochen, den man demzufolge auch studieren kann.³ Dem nachfolgend verdanken wir Dilthey (1883) die Unterscheidung zwischen Naturwissenschaften und Geisteswissenschaften⁴: Naturwissenschaften wie Physik oder Mineralogie haben das Studium der Natur zum Ziel, d.h. alles, was ohne Einfluss des Menschen Bestand hat. Dilthey übernimmt von Hegel aber auch die Idee des kollektiven Geistes, wonach die Geisteswissenschaften nicht nur den individuellen Geist zu studieren haben, sondern auch alles, was Menschen zusammen machen, worunter beispielsweise Kulturgeschichte oder Architektur fallen. Es sei die Aufgabe der Geisteswissenschaften, so Dilthey weiter, die Erzeugnisse des menschlichen Geistes zu verstehen. Für Dilthey gehören somit auch die Sozialwissenschaften zu den Geisteswissenschaften, und er entwickelt weitere Ideen über geeignete Methoden innerhalb dieser zwei Gruppen von Disziplinen: Die Entwicklung von erklärenden Gesetzen in den Naturwissenschaften und den Gebrauch einer auf das Verstehen abzielenden Hermeneutik in den Geisteswissenschaften.⁵

Diese Unterscheidung lebt heute in der französischen Einteilung von *sciences naturelles* und *sciences humaines* fort, die wiederum die Basis für die Unterteilung in *Naturwissenschaften* und *Letteren* (›Geisteswissenschaften‹) der niederländischen Akademie der Wissenschaften (KNAW) bildet.

² Zit. nach Berg 2014, S. 17.

³ Vgl. Hegel [zuerst 1806] 1988, Kap. VI: »Der Geist ist das *sittliche* Leben eines Volks.«

⁴ Dilthey [zuerst 1883] 1959, S.14ff.

⁵ Leezenberg / de Vries 2001 referieren diese Entwicklung mit weitaus größerer Kompetenz, als ich es an dieser Stelle leisten kann.

Die Informatik beschäftigt sich mit digitaler Informationsverarbeitung, sowohl in einzelnen Prozessen (Algorithmen oder Rechnern) als auch in Organisationen. Digitale Informationsverarbeitung aber ist ein Erzeugnis des kollektiven menschlichen Geistes, genau wie der mittelalterliche Städtebau, das Blues-Schema in der Musik, der Handelsbund der Hanse oder das moderne Bankwesen. Daraus folgt, dass die Informatik als wissenschaftliches Studium der Informationsverarbeitung zu den Geisteswissenschaften gehört, und ergo eine Geisteswissenschaft ist.

Diese Zuordnung der Informatik zu den Geisteswissenschaften impliziert nicht, dass die Informatik keine exakte (mathematische) Wissenschaft ist; denn auch andere eindeutige Beispiele von nicht den Naturwissenschaften zugehörigen Fachdisziplinen machen ausführlichen Gebrauch von mathematischen Modellen. Man denke an die Ökonometrie, die Architektur oder die Phonetik. Hinzukommt, dass gewisse Subdisziplinen der Informatik wie die Softwareentwicklung⁶ oder die Software-Ergonomie, einschließlich dem Entwurf von Benutzerschnittstellen⁷, eher sozialwissenschaftliche Methoden gebrauchen.

Die Schlussfolgerung oben ist qua Struktur ein disjunktiver Syllogismus: Aus $A \vee B$, $\neg A$ folgt B . Aus A oder B und *nicht* A , folgt B . Natürlich kann man einwenden, dass die Disjunktion, die Informatik sei entweder eine Naturwissenschaft oder eine Geisteswissenschaft, gar nicht zwingend sei. Man könnte auf eine dritte Möglichkeit als fachlicher Heimat der Informatik verweisen, etwa auf die Ingenieurwissenschaft oder die Sozialwissenschaft, doch dies würde nur manchen, aber lange nicht allen, vermutlich nicht einmal den meisten Informatikern gefallen.

Der wichtigere Einwand ist jedoch, dass die Klassifizierung in Natur- versus Geisteswissenschaften obsolet geworden ist.⁸ Die methodologischen Folgen der Unterscheidung, die noch für Dilthey Relevanz hatten, gelten schon lange nicht mehr. In vielen geisteswissenschaftlichen Fachdisziplinen wurden mittlerweile Gesetzmäßigkeiten erforscht und angewendet, um Phänomene in dem jeweiligen Gebiet zu erklären. Man denke nur an das Gesetz des sinkenden Ertrags in der Ökonomie oder an die Lautgesetze der Junggrammatiker. Umgekehrt spielt Intuition vermehrt in den Naturwissenschaften eine Rolle, wenn Forscher nach den Zusammenhängen ihrer Ideen fragen: Ein Schritt, der auf das Verstehen gerichtet ist. Auch das Gedankenexperiment kann von Belang sein. Die Klassifizierung ist demnach in beiden Richtungen unzureichend. Wenn wir erklären, dass eine neue Disziplin als Naturwissenschaft und nicht als Geisteswissenschaft klassifiziert wird, dann wissen wir sehr wahrscheinlich nach eingehender Prüfung und Präzisierung des Fachgegenstands, seiner Fragestellungen und Methoden, dass die Disziplin den akzeptierten Naturwissenschaften – Physik und Chemie – ähnlicher ist als den akzeptierten Geisteswissenschaften – Geschichte und Literaturwissenschaft –, aber viel mehr als das wissen wir nicht.

⁶ Bourque / Fairley 2014.

⁷ Shneiderman / Plasant 2010.

⁸ Von dieser Grundannahme geht auch Simons in *The Sciences of the Artificial* (1969) aus, in dem er Informatik und Ökonomie zusammen in eine Klasse setzt, weil sie beide künstliche Systeme ins Auge fassen, deren Komplexität es unmöglich macht, sie nur aus ihren Bauprinzipien heraus zu verstehen.

3. Die Unsicherheit und ihre Überwindung

Auf Unterschiede zwischen den Disziplinen in Fragen von Prestige, Finanzierung und Ähnlichem möchte ich hier nicht eingehen, auch wenn klar ist, dass es manchen Kollegen vor allem darum geht.

Es ist zu bemerken, dass auch die Informatik eine Periode der Unsicherheit gekannt hat, die u.a. dadurch charakterisiert ist, dass man sich über die Klassifizierung stritt. Sollte Informatik bei der Mathematik angesiedelt werden – weil es schließlich um Rechner und das Rechnen (nicht?) geht – oder aber bei den Ingenieurwissenschaften? Als ich 1985(!) eine Stelle als Informatiker bei Hewlett-Packard annahm, einer vom Ursprung her elektrotechnischen Firma, habe ich von den Elektro-Ingenieuren gehört, dass das Fach Informatik für sie vielleicht »in« wäre, aber noch nicht substantiell. 1985 waren Computer und Informatik schon gut etabliert und wirtschaftlich bereits ein Faktor. Der Ruf, eine anerkannte Wissenschaft zu sein, musste erst noch kommen.

Wir können aber auch Inspiration aus der Art und Weise ziehen, wie andere mit der Konstitution eines neuen Fachs umgehen. John Hopcroft erzählte von seinen ersten Jahren in Princeton. Sein Abteilungsleiter hatte ihn gebeten, einen Kurs in Informatik zu geben, ohne genauer mitzuteilen, was für Material der Kurs beinhalten sollte. Hopcroft fand es merkwürdig, dass Princeton einen neuen Kurs einführen wollte, ohne dessen Inhalt zuvor zu spezifizieren. Später resümierte er jedoch: »In retrospect, I realize that people who believe in the future of a subject and who sense its importance will invest in the subject long before they can delineate its boundaries.«⁹

Man könnte es sogar radikaler formulieren, denn Hopcroft spricht von einem Lehrfach (*subject*), während viele gar nicht bereit wären, die Digitalen Geisteswissenschaften als Fach anzuerkennen. Man bemängelt die Verschiedenheit der Methoden und Themen und auch die manchmal fehlende Beziehung zu nicht-digitalen Studien.¹⁰ Auch hier sehe ich keinen Grund zur Unruhe. Viele anerkannte Fächer, z.B. Informatik, aber auch Medizin, sind von innen betrachtet sehr unterschiedlich, und es sollte klar sein, dass man mittels digitaler Studien manchmal neuartige Fragestellungen erproben will und kann – auch dort, wo die Verbindung zu traditionellen Fächern nicht auf der Hand liegt.

Von Belang ist es, dass wir wissenschaftlich zu den verschiedenen geisteswissenschaftlichen Disziplinen beitragen, und dass diese Beiträge anerkannt werden. Wir müssen auf die positiven Fälle verweisen, wenn wir den digitalen Geisteswissenschaften die gebührende Anerkennung zukommen lassen wollen. Ich habe dokumentiert, wie dies im Falle der Dialektologie einigermaßen gelungen ist,¹¹ und vorgeschlagen, dass man, um überzeugende DH-Beiträge zu finden, Projekten vier Fragen stellt, nämlich

⁹ Turing Lecture, 1987. Natürlich ist dies auch ein Grund, die Frage nach der genauen Definition von Digital Humanities recht gelassen zu stellen, vgl. Vanhoutte et al. 2013.

¹⁰ Vanhoutte et al. 2013.

¹¹ Nerbonne 2005.

- ob eine wichtige geisteswissenschaftliche Frage gestellt wird,
- ob die Informationsverarbeitung essentiell ist,
- ob man die Resultate kritisch überprüft (validiert) hat und
- ob dieses Vorgehen den Weg zu größeren Themen und Analysen freimacht.

Nicht jedes Projekt muss jeder dieser Forderungen entsprechen, auch um Raum für nicht-traditionelle Fragestellungen zu geben. Aber wir sehen vermehrt Forschungsberichte, in denen diese Fragen wenigstens teilweise bejaht werden können. Auf einige werde ich im folgenden Abschnitt kurz eingehen.

4. Zehn Thesen zur aktuellen Lage in den Digital Humanities

Erstens: Es geht bergauf mit den Digital Humanities. Die erste DHd-Konferenz in Passau (mit 320 Teilnehmern!) belegt diese These auf eindrucksvolle Weise, aber auch neue Organisationen in Italien, Japan und Australien in der *Alliance for digital Humanities Organizations* (ADHO) unterstreichen dies. Darüber hinaus werden auch Konferenzen von neuen Gruppen in Spanien (z.B. *Humanidades Digitales Hispánicas*), in den Beneluxländern (z.B. *DHBenelux*) und in der französischsprachigen Schweiz (z.B. *Humanistica*) geplant und durchgeführt.

Zweitens ist der erschwingliche Zugang zu Forschungsdaten für die verbesserte aktuelle Lage essentiell gewesen. Denn die Digital Humanities profitieren von allgemeinen Entwicklungen in der Gesellschaft, die dahin gehen, dass immer mehr in digitale Daten investiert wird. Die digitalen Geisteswissenschaften haben neue Wege eingeschlagen, und es wurde erst sukzessive der Bedarf an einer solchen Disziplin deutlich, wozu einige vorausschauende Kollegen, Institute, Bibliotheken und Firmen maßgeblich beigetragen haben. Dass so viel Aktivität erst in letzter Zeit aufgekommen ist, kurz nachdem Daten viel zugänglicher geworden sind, deutet natürlich an, dass Interesse und Potential schon vorhanden waren.

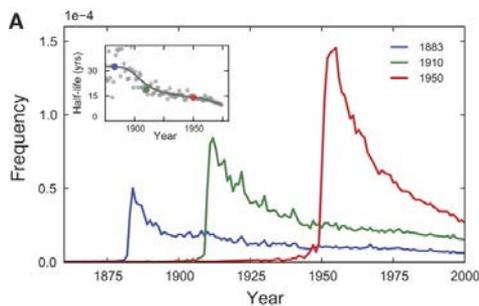
Drittens geht es um viel mehr als um den bloßen Zugang zu Daten. Echte Informationsverarbeitung ist von Anfang an unentbehrlich. Um ein Beispiel zu gebrauchen: Ohne angepasste Algorithmen kann man in alten Texten von der Zeit vor der Normalisierung der Schrift gar nicht suchen, geschweige denn Worthäufigkeitszählungen ausführen.¹² In einem rezenten Projekt in Groningen, wo Zollerhebungen am Öresund (zwischen Schweden und Dänemark) aus einer Periode von über drei Jahrhunderten digitalisiert wurden, fanden sich beispielsweise über 300 Schreibweisen für den Stadtnamen ›Kopenhagen‹. Ich werde darauf im folgenden Abschnitt über latente Strukturen noch weiter eingehen.

Viertens ist es zunehmend möglich, nicht nur relativ oberflächliche Elemente, wie z.B. Wörter (und deren Frequenz), zu analysieren, wie man dies lange Zeit bei der

¹² Gotscharek et al. 2011, S. 159-161.

Autorschaftsbestimmung gemacht hat,¹³ sondern auch die *latenten Strukturen*, wofür wir uns als Geisteswissenschaftler eher interessieren. Van Halteren et al. bestimmen Autorschaft aufgrund des Wortschatzes und aufgrund syntaktischer Merkmale, die sie mittels *Tagging* zuweisen, ein vollkommen automatisches Verfahren, das jedem Wort im Text eine syntaktische Kategorie zuordnet. Sie gebrauchen die Frequenz kurzer Sequenzen von Tags, um Autoren zu erkennen, und ziehen daraus den Schluss, dass die syntaktischen Merkmale viel distinktiver sind, als sprachwissenschaftliche Theorien dies würden erwarten lassen.¹⁴ Hirst / Feiguina verwenden partielles Parsing, wobei jeder Satz im Text (vielleicht nur partiell) zergliedert wird.¹⁵ Die Frequenz von Kategorien (und Paaren davon), die beim Parsing zugeordnet werden, wird gebraucht, um Autoren zu erkennen. Hirst / Feiguina zeigen, dass Ihre Methode imstande ist, auch sehr ähnliche Autoren zu unterscheiden, nämlich die Brontë Schwestern, Anne und Charlotte. Wiersma / Nerbonne / Louttamus entwickeln ähnliche Techniken, um Unterschiede in der Sprache von Immigranten und Muttersprachlern zu detektieren,¹⁶ und und Lučić / Blake benutzen ein anderes automatisches Verfahren, bei dem sie ihre Aufmerksamkeit insbesondere auf die Syntax in der Nähe von Nominalphrasen lenken.¹⁷

Fünftens dürfte der größte Vorteil für die digitalen Techniken dort entstehen, wo man viel Material analysieren will. Das Harvard Projekt *Culturomics*¹⁸ protzt damit, dass ihm »1,5x10⁷« Bücher zur Verfügung standen. Und auch wenn die anfängliche Begeisterung angesichts mancher Details etwas nachgelassen hat, illustriert das Projekt doch eine Art von Analyse, die den herkömmlichen Geisteswissenschaften vorenthalten ist: die Analyse auf großer Skala. Man erinnert sich gern an das Gedankenspiel, dass selbst jemand, der 80 Jahre lang 5 Bücher pro Tag läse, nur 1% des gesamten Materials kennenlernen würde. *Abbildung 1* gibt Messungen zur Frage der Resonanz einer Person wieder und sucht damit die Frage zu beantworten, wie lang auf eine Person verwiesen wird. Dass Ruhm vergänglich ist, ist sprichwörtlich und wird selbst in immer neuen Formen geäußert – so Warhols Bemerkung, dass jeder in der Zukunft 15 Minuten lang berühmt sein wird. Das Messergebnis spiegelt demnach keine neue Einsicht wider, aber es liefert zum ersten Mal eine empirisch belastbare Aussage, und dies in einer präzisen Form.



¹³ Nerbonne 2007, S. xvii.

¹⁴ Van Halteren et al. 2005, S. 71.

¹⁵ Hirst / Feiguina 2007, S. 414-415.

¹⁶ Wiersma / Nerbonne / Louttamus 2011.

¹⁷ Lučić / Blake.

¹⁸ Michel et al. 2011.

Abb. 1: »O quam cito transit gloria mundi!«, Thomas a Kempis quipped famously 500 years ago, but the *Culturomics* project not only confirmed this, but also measured how quickly fame flees, and showed that it is accelerating. The three curves show how quickly the frequency of reference to a name drops after an initial, nearly vertical climb, and the most recent curve show the steepest decline. Quelle: Michel et al. 2011.

Culturomics ist das bekannteste Projekt, in dem die Verwendung einer großen Skala die neueren Möglichkeiten zeigt, aber es gibt weitere Projekte, deren Datenbasis jenseits menschlichen Vermögens liegen, oder aber die wegen des Einsatzes digitaler Technik viel schneller ausgeführt werden können. So analysiert Jockers über 3.500 Romane aus dem 19. Jahrhundert, wobei er sich nicht auf oberflächliche Elemente konzentriert, sondern auf die Themen der Romane, die er mittels maschineller Lernalgorithmen extrahiert.¹⁹ Leinonen untersucht die SVEDIA Daten, deren Datengrundlage die Tatsache bildet, dass 1.200 Sprecher des Schwedischen jeden der 20 Vokale mehrmals ausgesprochen haben.²⁰ Sie hat darauf aufbauend den gemittelten Abstand zwischen den Vokalen für alle Sprecher bestimmt und ein Verfahren entwickelt, diese mittels multidimensionaler Skalierung (MDS) in drei Dimensionen zu repräsentieren. Die so gewonnenen MDS-Koordinaten wurden durch verschiedene Farbtintensitäten abgebildet und anschließend die älteren (ab 65 Jahren) und jüngeren Sprecher (bis 27 Jahren) getrennt voneinander auf die Karte von Schweden (inklusive der Küstenregion von Finnland, wo auch Schwedisch gesprochen wird) projiziert. Die solchermaßen erzeugten Karten (Abbildung 2) geben ein eindrucksvolles Bild davon, wie die Dialektunterschiede sich ausglätten. Weil Leinonen ein automatisiertes Verfahren zur Bestimmung der Vokalqualität (der Frequenz der Vokalformanten) gebrauchte, wurde das Projekt als Doktorarbeit überhaupt erst realisierbar. Sie erhielt für diese Arbeit 2010 den königlichen Gustav-Adolph-Preis für schwedische Volkskultur.

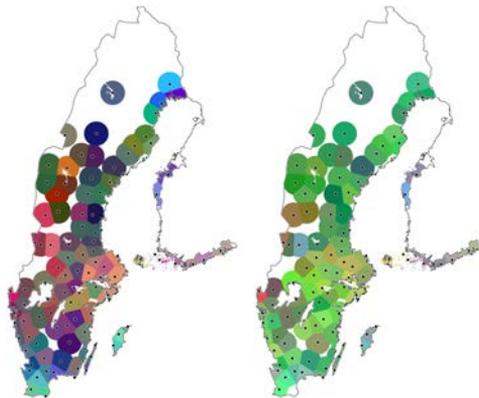


Abb. 2: Die Farben in den zwei Karten von schwedisch-sprechenden Skandinaviern spiegeln die Verschiedenheit der Aussprache wider, wobei links die Sprache von 65-jährigen steht und rechts die von 27-jährigen. Ohne automatisiertes Verfahren wäre die Fülle der Daten nicht zu analysieren gewesen. Quelle: Leinonen 2010.

¹⁹ Jockers 2013. Vgl. auch die Webseite [Macroanalysis](#), auf der Jockers sehr viel Material digital zur Verfügung stellt.

²⁰ Leinonen 2010.

Nach diesen ersten fünf Thesen will ich auch einen Blick in die Zukunft wagen. In welche Richtung werden sich die Digital Humanities bewegen und was dürfen wir in der Zukunft von ihnen erwarten?

Sechstens sage ich keinen Marsch durch (alle) (Sub-)Disziplinen voraus. Einige Disziplinen und Subdisziplinen sind sehr phänomenologisch eingestellt. Hier behält die tiefere Reflexion über die Begriffe Priorität und in solchen Fällen ist der Vorteil von digitalen Methoden minimal, im Prinzip dient er nur der Kontrolle von Konsistenz, wenn die Theorien komplexer sind. Aber selbst da kommen die Digital Humanities nur sinnvoll zum Einsatz, wenn diese Theorien formalisiert werden können. Weiterhin wird es Platz innerhalb digital angelegter Forschungsrichtungen für nicht-technisch versierte Teilnehmer geben (vgl. Punkt 9).

Siebtens haben die Geisteswissenschaften durch den Einsatz von digitalen Techniken keine »feindliche Übernahme« zu befürchten, wie der Titel der DHd-Konferenz provokant andeutet, auf die dieser Beitrag zurückgeht. Mit Forschungsfragen, etwa über die Träger von Kultur oder ob verschiedene Kulturbereiche (Alltagskultur, Hochkultur, Technologie, Mundart, Musik, Essgewohnheiten u.a.) sich entlang der gleichen Wege ausbreiten, wissen Nicht-Geisteswissenschaftler nur sehr grob etwas anzufangen. Etwas frech formuliert, behaupte ich, dass die Techniker keine Forschungsfragen innerhalb der Geisteswissenschaft formulieren können, aber auch, dass sie die Möglichkeiten der digitalen Technik innerhalb der Geisteswissenschaften unzureichend erkennen können, weil sie kein Teil dieser Forschungsgemeinschaft sind.

Achtens wird es von Vorteil sein, wenn unsere Methoden besser vergleichbar werden und wir gegenseitige Kontrolle und Kritik ausüben können. Bevor ich versuche, dies zu untermauern, will ich betonen, dass wir uns die Situation in anderen digitalen Wissenschaften vor Augen halten sollten. Der Gebrauch von digitaler Technik ist keine Panazee, nach deren Eintritt alles durchschaubar und alle Einzelstudien replizierbar werden. Die Erfahrung ist anders. Aber die »digitale Wende« bringt wohl einen Schritt in diese Richtung mit sich mit. Es ist jetzt schon so, dass digitale Geisteswissenschaftler manchmal ihre Daten austauschen und ihre Methoden am gleichen Material erproben: Van Halteren et al. stellten einen Korpus für Autorschaftserkennung zur Verfügung und andere haben ihn bereits gebraucht, um Alternativtechniken zu testen.²¹ Eder / Kestemont / Rybicki entwickelten Stylo, ein Softwarepaket für Stilometrie.²² Der wissenschaftliche Diskurs wird keinesfalls überflüssig, aber wir können viel direkter voneinander profitieren – nicht nur in Einsicht, sondern auch in Datenvorbereitung, Methoden und Analyseergebnissen.

Die Kehrseite von der Zunahme der Empirie könnte durchaus eine Minimierung der Autorität der »Weisen« bedeuten – ich spiele hierbei auf ältere, sehr respektierte Kollegen an, deren Urteilen vor allem bezüglich unserer Kultur vertraut wird. Hierüber ist schon spekuliert worden (vgl. [Abschnitt 1](#)) und obwohl die spezifischen Beispiele sich durchaus plausibel anhören, muss man gleichzeitig zugeben, dass freilich alle Wissenschaftsgebiete ihre »Weisen«

²¹ Van Halteren et al. 2005; Juola / Baayen 2005.

²² Eder / Kestemont / Rybicki 2013.

haben und auch feiern. Es wäre andererseits gut, wenn das bloße Urteil dieser ›Weisen‹ weniger tragend oder kanonisch wirken würde, weil man sich eben gewöhnen wird, mehr nach empirischen Gründen fragen zu können.

Neuntens werden die Digital Humanities mehr Zusammenarbeit mit Informatik, Statistik und anderen Nicht-Geisteswissenschaften mit sich bringen. Dies hat natürlich schon angefangen und entwickelt sich wie die Situation in vielen anderen Wissenschaften, z.B. in der Biologie und der Astronomie. Auch dort handelt es sich übrigens um keine »feindliche Übernahme«. Der Gegenstand der Wissenschaft bleibt, während die Methoden sich opportunistisch entwickeln.

Zehntens und schließlich wird die Arbeit in den Digital Humanities innovativ, aufregend und stimulierend sein. Ich hoffe, dass mehrere Kollegen in diese Richtung schauen werden!

5. Schlussbemerkung

Den Titel – Informatik als Geisteswissenschaft – habe ich in gewissem Maße missbraucht, denn ich habe mit einer gewissen Ironie argumentiert, dass die Informatik eine Geisteswissenschaft sei, und sicher nicht versucht, diese These seriös zu erhärten. Ich habe vielmehr den Einwand erhoben, dass die Klassifizierung der modernen Wissenschaften anscheinend nicht mehr im gleichen Maße sinnstiftend ist, wie sie es zu Zeiten Diltheys war.

Mein Ziel war es dabei, zu mehr Mut bei der digitalen Wende der Geisteswissenschaften aufzufordern, denn auch die Informatik – heute sicher eine blühende Wissenschaft, sowohl theoretisch als auch angewandt – hatte Geburtswehen durchzustehen, wie vielleicht alle neuen wissenschaftlichen Richtungen.

Die ersten fünf meiner 10 Thesen zielten darauf ab, zu mehr geisteswissenschaftlicher Arbeit mittels digitaler Techniken zu ermuntern. Das Interesse an DH steigt rasant, denn die Voraussetzungen hierfür haben sich in den letzten zehn Jahren enorm verbessert. Es war mir ein Anliegen, einige Missverständnisse auszuräumen, u.a. dass die Digital Humanities nur die Aufgabe haben, unsere Forschungsgegenstände digital anzubieten, und dass unsere Techniken verdammt seien, sehr oberflächlich zu bleiben. Zum Schluss habe ich ein paar erfolgreiche geisteswissenschaftliche Projekte vorgestellt, die mir wegen ihres Vermögens, viele Daten unter die Lupe zu nehmen, aufgefallen sind.

Die letzten fünf der Thesen wagten vorsichtige Vorhersagen über die Digital Humanities: Nicht jede Disziplin innerhalb der Geisteswissenschaften wird imstande und gewillt sein, von den neuen Techniken zu profitieren. Geisteswissenschaftler sind und bleiben diejenigen, die den Gegenstand (oder Gegenstände) ihres Fachs ausreichend kennen, um die Forschungsfragen zu identifizieren und die erreichten Resultate kritisch zu evaluieren. Wir werden mehr als bisher miteinander zusammenarbeiten – so die hier skizzierte Vision –, auch

mit Kollegen aus anderen, nicht-geisteswissenschaftlichen Disziplinen. Abschließend bleibt mir dem nur die wichtigste Vorhersage hinzuzufügen, nämlich: dass es sich lohnen wird.

Bibliographische Angaben

- Lilo Berg: Rotkäppchen 2.0 (Digital Humanities. Märchenhafte Chance oder Modetrend). In: Humboldt Kosmos 102 (2014), S. 12–23. [\[online\]](#)
- Wilhelm Dilthey: Einleitung in die Geisteswissenschaften. Versuch einer Grundlegung für das Studium der Gesellschaft und der Geschichte. In: Gesammelte Schriften 1. 1875–1900. Stuttgart 1959 (zuerst 1883). [\[Nachweis im GBV\]](#)
- Maciej Eder / Mike Kestemont / Jan Rybicki: Stylometry with R: A Suite of Tools. Digital Humanities 2013: Conference Abstracts. Lincoln 2013, S. 487–89. [\[online\]](#)
- Anders Eriksson: Swedia-projektet: dialektforskning i ett jämförande perspektiv. In: Folkmålsstudier 43 (2004), S. 11–32. [\[online\]](#)
- Elizabeth Eisenstein: The Printing Press as an Agent of Change. 2 Bde. Cambridge 1979. [\[Nachweis im OPAC\]](#)
- Annette Gotscharek / Ulrich Reffle / Christoph Ringlsetter / Klaus Schulz / Andreas Neumann: Towards information retrieval on historical document collections: the role of matching procedures and special lexica. In: International Journal on Document Analysis and Recognition (IJ DAR) 14.2 (2011), S. 159–171. [\[Nachweis im GBV\]](#)
- Guide to the Software Engineering Body of Knowledge, Version 3.0. Hg. von Pierre Bourque / Richard E. Fairley. IEEE Computer Society, 2014. [\[online\]](#)
- Hans Van Halteren / Harald Baayen / Fiona Tweedie / Marco Haverkort / Anneke Neijt: New machine learning methods demonstrate the existence of a human stylome. In: Journal of Quantitative Linguistics 12.1 (2005), S. 65–77. [\[Nachweis im GBV\]](#)
- Georg Wilhelm Friedrich Hegel: Phänomenologie des Geistes. Hamburg 1988 (zuerst 1806/1807). [\[Nachweis im GBV\]](#)
- Graeme Hirst / Ol'ga Feiguina: Bigrams of syntactic labels for authorship discrimination of short texts. In: Literary and Linguistic Computing 22.4 (2007), S. 405–417. [\[Nachweis im GBV\]](#)
- John E. Hopcroft: Computer science: the emergence of a discipline (Turing Award Lecture). In: Communications of the ACM 30 (1987), S. 198–202. [\[Nachweis im GBV\]](#)
- Matthew Jockers: Macroanalysis. digital methods and literary history. Champaign 2013. [\[Nachweis im OPAC\]](#)
- Patrick Juola / R. Harald Baayen: A controlled-corpus experiment in authorship identification by cross-entropy. In: Literary and Linguistic Computing 20.1 (2005), S. 59–67. [\[Nachweis im GBV\]](#)
- Michiel Leezenberg / Gerard De Vries: Wetenschapsfilosofie voor geesteswetenschappen. Amsterdam 2001. [\[online\]](#)
- Therese N. Leinonen: An acoustic analysis of vowel pronunciation in Swedish dialects. Diss. Universität Groningen 2010. [\[Nachweis im GBV\]](#)
- Ana Lužić / Catherine L. Blake: A syntactic characterization of authorship style surrounding proper names. In: LLC: the Journal of Digital Scholarship in the Humanities. Oxford 2015. [\[Nachweis im GBV\]](#)
- Jean-Baptiste Michel / Yuan Kui Shen / Aviva Presser Aiden / Adrian Veres / Matthew K. Gray / The Google Books Team / Joseph P. Pickett / Dale Hoiberg / Dan Clancy / Peter Norvig / Jon Orwant / Steven Pinker / Martin A. Nowak / Erez Lieberman Aiden: Quantitative analysis of culture using millions of digitized books. Science 331 (2011), Nr. 6014, S. 176–182. [\[online\]](#)
- John Nerbonne: Elektronische Incunabeln. Antrittsvorlesung in Groningen. 1995. [\[online\]](#)
- John Nerbonne: Computational Contributions to the Humanities. In: Literary and linguistic computing 20.1 (2005), S. 25–40. [\[Nachweis im GBV\]](#)
- John Nerbonne: The exact analysis of text. Foreword. In: Inference and Disputed Authorship: The Federalist Papers. Hg. von Frederick Mosteller / David Wallace. 3. Auflage. Stanford 2007. S. XI–XX. [\[online\]](#)
- Ben Shneiderman / Catherine Plaisant: Designing the User Interface: Strategies for Effective Human-Computer Interaction. 5. Auflage. Reading 2010 (zuerst 1987). [\[Nachweis im GBV\]](#)
- Defining Digital Humanities: A Reader. Hg. von Edward Vanhoutte / Julianne Nyhan / Melissa Terras. London 2013. [\[Nachweis im GBV\]](#)
- Wybo Wiersma / John Nerbonne / Timo Lauttamus: Automatically extracting typical syntactic differences from corpora. In: LLC: the Journal of Digital Scholarship in the Humanities 26.1 (2011), S. 107–124. DOI:10.1093/llc/fqq017. [\[Nachweis im GBV\]](#)

Abbildungslegenden und -nachweise

Abb. 1: »O quam cito transit gloria mundi!«, Thomas a Kempis quipped famously 500 years ago, but the *Culturomics* project not only confirmed this, but also measured how quickly fame flees, and showed that it is accelerating. The three curves show how quickly the frequency of reference to a name drops after an initial, nearly vertical climb, and the most recent curve show the steepest decline. Quelle: Michel et al. 2011.

Abb. 2: Die Farben in den zwei Karten von schwedisch-sprechenden Skandinaviern spiegeln die Verschiedenheit der Aussprache wider, wobei links die Sprache von 65-jährigen steht und rechts die von 27-jährigen. Ohne automatisiertes Verfahren wäre die Fülle der Daten nicht zu analysieren gewesen. Quelle: Leinonen 2010.

Zeitschrift für digitale Geisteswissenschaften

Artikel aus:

Sonderband 1 der ZfdG: Grenzen und Möglichkeiten der Digital Humanities. Hg. von Constanze Baum und Thomas Stäcker. 2015. DOI: [10.17175/sb01](https://doi.org/10.17175/sb01)

Titel:

Digital Humanities? Gibt's doch gar nicht!

Autor/in:

Patrick Sahle

Kontakt:

sahle@uni-koeln.de

Institution:

Cologne Center for eHumanities (CCeH), Universität zu Köln

GND:

[1034443909](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63864-p0033-9)

ORCID:

[0000-0002-8648-2033](https://orcid.org/0000-0002-8648-2033)

DOI des Artikels:

[10.17175/sb001_004](https://doi.org/10.17175/sb001_004)

Nachweis im OPAC der Herzog August Bibliothek:

[830166343](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63864-p0033-9)

Erstveröffentlichung:

19.02.2015

Lizenz:

Sofern nicht anders angegeben 

Medienlizenzen:

Medienrechte liegen bei den Autoren

Letzte Überprüfung aller Verweise:

24.05.2016

GND-Verschlagwortung:

[Digital Humanities](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63864-p0033-9) |

Zitierweise:

Patrick Sahle: Digital Humanities? Gibt's doch gar nicht!. In: Grenzen und Möglichkeiten der Digital Humanities. Hg. von Constanze Baum / Thomas Stäcker. 2015 (= Sonderband der Zeitschrift für digitale Geisteswissenschaften, 1). PDF Format ohne Paginierung. Als text/html abrufbar unter DOI: [10.17175/sb001_004](https://doi.org/10.17175/sb001_004).

Patrick Sahle

Digital Humanities? Gibt's doch gar nicht!

Abstracts

Die Digital Humanities sind entweder ein Forschungsfeld oder eine Disziplin, möglicherweise auch beides. Sie verfügen jedenfalls über eine gut ausgebaute Infrastruktur der Organisation, Information und Kommunikation und blicken in Bezug auf ihre vielfältigen Forschungsansätze teilweise auf lange Traditionen zurück. Als Bindeglied zwischen den Geisteswissenschaften und der Informatik scheint das Feld heute nicht nur für diese beiden Bereiche, sondern auch für die Organe der Forschungsförderung von besonders hoher Attraktivität zu sein. Neben der Wissenschaft haben selbst die Massenmedien in den letzten Jahren die Digital Humanities entdeckt. Die hohe Anziehungskraft des Feldes hat erfreulich integrative Tendenzen. Allerdings birgt dieser DH->Hype auch Gefahren. Diese reichen von der bloßen Aneignung des Etiketts über explizite Abwehrhaltungen bis hin zu Ignoranz und Verleugnung: »DH? Das gibt es doch gar nicht!«

The Digital Humanities are either a field of research or an academic discipline, possibly both. In any case, they are embedded in an extensive infrastructure in regard to organization, information, and communication and build upon long traditions in various areas of research. Furthermore, as a link between the humanities and computer science, the field seems to be highly attractive, not only to these areas, but also to neighboring disciplines as well as to the research funding agencies. Outside of the academic context, the mass media have discovered the digital humanities in recent years. While this attractiveness has several positive integrative effects, the current DH hype is not without certain risks, ranging from the mere misappropriation of the term to explicitly defensive or negative attitudes, and even ignorance and denial: »DH? But there is no such thing!«

1. Digital Humanities als Disziplin

Die Tagung des DHd 2014 »Digital Humanities – methodischer Brückenschlag oder ›feindliche Übernahme‹« in Passau vom 25.–28.3.2014 war die erste Fachtagung des noch jungen Digital Humanities-Verbandes für den deutschsprachigen Raum (DHd). Die hohe Zahl der eingereichten Vorträge und Poster sowie die große Publikumsbeteiligung dokumentierten eindrucksvoll die aktuelle Attraktivität und Relevanz der Digital Humanities. Aber was ist das eigentlich, dieses ›DH? Würde man sich der Frage ganz unvoreingenommen, ja im besten Sinne ahnungslos nähern, um z.B. eine vorurteilsfreie wissenschaftssoziologische Betrachtung anzustellen, dann würde man wohl die gleichen Strategien verwenden und die gleichen Merkmale untersuchen wie bei anderen wissenschaftlichen Feldern und Disziplinen. Zu einer Disziplin-Checkliste könnten dann gehören:

1.) Die **Einstiegspunkte unserer Zeit**. Wikipedia bietet einen eigenen Eintrag zu »Digital Humanities« in 16 Sprachen, Google liefert über eine Million Treffer für »Digital Humanities« bzw. »Humanities Computing« (und nationalsprachlichen Äquivalenten) als feststehenden Ausdrücken.

2.) Der Forschungsbereich ›Digital Humanities‹ schlägt sich in einer inzwischen unüberschaubaren Zahl von **Projekten** nieder, in denen er entweder den methodisch-technischen Teil der Fachforschung ausmacht oder selbst die Fragestellung liefert und den zentralen Gegenstand bildet. Digital Humanities als Disziplin ist mehr noch als andere Disziplinen projektgetrieben; ihre methodisch-theoretischen Leistungen nehmen in den allermeisten Fällen von den konkreten Forschungsproblemen der Geisteswissenschaften ihren Ausgangspunkt.

3.) Eine thematische Eingrenzung der Digital Humanities lässt sich über die einschlägigen **Tagungen** vornehmen, auf denen sich das Feld präsentiert und wo ein gemeinsamer Kommunikationsraum gebildet wird. Eine der einfachsten Antworten auf die Frage »Was ist DH?« lautet: »DH ist die Summe der auf den DH-Tagungen behandelten Themen«. Für einen solchen Zugang besteht mit der jährlichen, internationalen, zentralen Konferenz des Dachverbandes der »Digital Humanities« ein ideales Studienobjekt.¹ Der strenge Peer-Review-Prozess offenbart dabei, was von der Community als DH aufgefasst wird und was davon dem aktuellen Stand der Forschung entspricht. Hinzu kommt ein für die Beiträge dieser Kongressreihe eingeführtes Schlagwortsystem, das die Themen und ihre quantitative Bedeutung leicht fassbar macht.²

4.) Für die erweiterte Fachkommunikation besteht eine gut ausgebaute **Informationsinfrastruktur**. Die reiche Fachliteratur ist in Bibliografien erfasst³ und seit Jahrzehnten bestehen einschlägige Zeitschriften⁴. Für die digitale Welt werden Informationsportale gepflegt⁵, Blogs geführt⁶ und Nachrichten über Twitter verbreitet.

5.) Organisatorische Verfestigung und **Verbände**. Die beiden traditionsreichen Organisationen Association for Literary and Linguistic Computing / European Association for Digital Humanities (ALLC/EADH, seit 1973) und Association for Computers and the Humanities (ACH, seit 1978) sind mit anderen Vereinigungen im globalen Dachverband Association for Digital Humanities Organizations (ADHO) zusammengeschlossen. In den letzten Jahren haben sich weitere regionale und nationale Verbände gebildet. Dazu gehören die Canadian Society for Digital Humanities / Société canadienne des humanités numériques (CSDH/SCHN), die italienische Associazione per l'informatica umanistica e la cultura digitale (AIUCD), die spanische Organisation Humanidades Digitales Hispanicas (HDH), die mexikanische Red de Humanidades Digitales (RedHD), die Associação das Humanidades Digitais (AHDig) für die portugiesischsprachige Welt, die Australasian Association for Digital Humanities (aaDH),

¹ Zuletzt 2014 in Lausanne; die englische Wikipedia führt eine Liste der Tagungen seit 1990. Zur globalen Veranstaltung kommen nationale (z.B. Digital Humanities Congress 2014 in Sheffield) und regionale Konferenzen (z.B. »Grenzen überschreiten – digitale Geisteswissenschaften heute und morgen«, im Februar 2014 in Berlin) sowie zahlreiche auf einzelne Problembereiche fokussierte Tagungen und Workshops.

² Vgl. dazu die visualisierenden Untersuchungen von Weingart 2013f., die vor allem auf Akzeptanz- und Ablehnungsstrukturen zielen.

³ Vgl. z.B. die DARIAH-Bibliografie Doing Digital Humanities bzw. die allgemeine ZOTERO-Bibliografie Digital Humanities, die zuletzt rund 500 Einträge hatte.

⁴ Am wichtigsten für die Breite des Feldes wohl immer noch LLC – The Journal of Digital Scholarship in the Humanities; daneben vor allem Digital Humanities Quarterly (DHQ), neuerdings auch Journal of Digital Humanities (JDH). Historisch wichtig vor allem Computers and the Humanities (1966–2004).

⁵ Vgl. z.B. Digital Humanities Now.

⁶ Zu den vielen persönlichen Blogs kommen Community-Blogs wie der DHD-Blog und der Kollektiv-Blog zum jährlichen »Day of DH« (seit 2009, zuletzt 2014).

die Japanese Association for Digital Humanities (JADH) oder eben der DH-Verband für die deutschsprachigen Länder DHd. Geografisch übergreifend bestehen weitere Organisationen wie Centernet, einem Zusammenschluss von DH-Zentren.

6.) Der Forschungs- und Arbeitsbereich ›Digital Humanities‹ verfügt über eine sehr **lebendige Community** von Aktivisten, die sich sowohl in DH-Verbänden als auch in den traditionellen Fachverbänden in Arbeitsgruppen organisieren und Vereine oder losere Gruppen bilden. Für den deutschsprachigen Raum wären hier z.B. die **AG Digitale Geschichtswissenschaft** im Verband der Historiker Deutschlands, der **Arbeitskreis Digitale Kunstgeschichte**, das **Institut für Dokumentologie und Editorik e.V.** (IDE), der **Arbeitskreis Digital Humanities München** oder der **DH-Stammtisch Berlin** zu nennen. Der Verbreitung der Digital Humanities dienen häufig Summer Schools⁷ und die inzwischen als Veranstaltungsform etablierten »Unkonferenzen« unter dem Reihentitel **THATCamp** (The Humanities and Technology Camp). Typisch für den Versuch, Inhalte und Wesen der DH zu vermitteln, sind auch Initiativen wie **Global Outlook DH**, verschiedene DH Manifestos oder lokale Dissemination-Workshops.⁸

7.) Die Forschungs- und Projektorientierung der DH zeigt sich auch darin, dass zu den wichtigsten Katalysatoren in der institutionellen Verfestigung explizite **DH-Kompetenzzentren** gehören. Hier sind neben den international renommierten Zentren in London, Victoria oder Lincoln im deutschsprachigen Raum vor allem das Kompetenzzentrum in Trier, das ZIMIG in Graz, das GCDH in Göttingen und das Kölner CCeH zu nennen. Auch die TELOTA-Gruppe an der BBAW ist als etabliertes de-facto DH-Zentrum zu bezeichnen. Zu den jüngeren Gründungen gehören Einrichtungen in Basel (Digital Humanities Lab[or]), Tübingen (eScience-Center) oder Nürnberg-Erlangen (Interdisziplinäres Zentrum für Digitale Geistes- und Sozialwissenschaften), weitere Zentren sind angekündigt oder befinden sich im Prozess der Gründung und Etablierung.⁹ An anderen Standorten bestehen eher lose Organisationsformen unterhalb der Ebene einer expliziten Institutionalisierung in Form eines Zentrums.¹⁰

8.) Die **Professionalisierung** einer Disziplin lässt sich leicht an dedizierten Stellen bzw. Stellenausschreibungen ablesen. Während auf der Ebene der befristeten Projektmitarbeiter die Suche nach »Digital Humanities«-Experten schon länger üblich ist, findet die Professionalisierung ihren Niederschlag inzwischen auch in einer breiteren **Professoralisierung**. In den letzten Jahren sind zu den Professuren in Köln, Graz, Würzburg, Trier und Darmstadt weitere in Passau, Bern, Leipzig, Luxemburg, Basel, Göttingen und Stuttgart besetzt oder ausgeschrieben worden. Hier ist auch zu beobachten, dass neben den »Fach X plus Digital Humanities«-Lehrstühlen immer mehr eine reine »Digital Humanities«-Denomination tragen. Die Einrichtung von Professuren begleitet und verstärkt außerdem die

⁷ Die Anbieter von kontinuierlich stattfindenden Schools sind inzwischen in einem »DH Training Network« zusammengeschlossen.

⁸ Vgl. **DH Manifesto 2011**, **DH Manifesto 2014** bzw. den Workshop des Einstein-Zirkels, **Grenzen überschreiten – Digitale Geisteswissenschaft heute und morgen**, in Berlin am 28.02.2014.

⁹ Vgl. z.B. **DH an der Universität Bern**, ansatzweise Zentrumsbildungen sind auch für Heidelberg und München zu verzeichnen.

¹⁰ Hierzu zählen z.B. die wichtigen DH-Standorte Darmstadt, Hamburg, Leipzig oder Würzburg.

institutionelle Verfestigung der DH, wenn sie mit der Verwandlung von Kompetenzzentren in Universitätsinstitute einhergeht.

9.) Professuren sind außerdem ein wichtiges Element im Aufbau von **Ausbildungsprogrammen** in den Digital Humanities. Jenseits der erwähnten Summer Schools, einzelnen Lehrveranstaltungen oder Modulen in anderen Studiengängen gibt es inzwischen weltweit rund 100 Programme auf BA- und MA-Ebene, die dem DH-Bereich zuzuordnen sind.¹¹ Dabei ist die Zahl der explizit »Digital Humanities« benannten Studiengänge zwar noch recht überschaubar, auch sie nimmt aber stetig zu.

10.) Zur Kanonisierung eines Faches gehört die Selbstvergewisserung und die Sicherung der erreichten Wissensstände in Form von Überblicksdarstellungen, Einführungen und Handbüchern, die z.B. unmittelbar in grundlegenden Einführungs-Lehrveranstaltungen genutzt werden.¹² Zur Etablierung eines Faches trägt außerdem das Bewusstsein für die eigene Tradition und Geschichte, mithin die eigene **Geschichtsschreibung**, bei.¹³ Die Digital Humanities führen sich demnach selbst bis in die 1940er Jahre zurück und sehen weitere wichtige Entwicklungsschritte in den 1960er bis 1980er Jahren, bis die 1990er Jahre mit der Durchsetzung des Internet-Paradigmas ihr den wohl entscheidenden Schub gegeben haben.

11.) In einem nicht-soziologischen Ansatz könnte behauptet werden, dass sich eine neue Disziplin durch einen eigenen Forschungsbereich bestimmt, der zu den bestehenden Nachbardisziplinen hinreichend distinkt ist. Danach würde eine Disziplin durch die eigene **Definition** geschaffen. In den Digital Humanities ist es gewissermaßen Sport, dass jeder DH-Spezialist mindestens eine eigene Definition hat. Sie ist u.a. Voraussetzung für die Teilnahme am »Day of DH«, Gegenstand zahlreicher Artikel und eines eigenen Sammelbandes.¹⁴

12.) Die Jahre 2013 und 2014 scheinen in der Etablierung des Faches einen weiteren Höhepunkt markiert zu haben. Nie zuvor war auch in den **populären Medien**, also vor allem in Zeitungsartikeln und Radiobeiträgen so viel von Digital Humanities die Rede. Damit dürften die DH auch im Bewusstsein einer – an wissenschaftlichen Themen interessierten – breiteren Öffentlichkeit und der Forschungspolitik angekommen sein.

13.) Ungleich länger scheint der Weg in die **Fachsystematiken** hinein zu sein, die eine endgültige Anerkennung belegen würden. Fachreferate in Bibliotheken, ein bibliothekarischer Sammelschwerpunkt, die DFG-Fachsystematik und andere Fächerlisten wären ein weiterer

¹¹ Für einen analytischen Überblick vgl. z.B. Patrick Sahle 2013.

¹² Den eindeutigsten Handbuch-Charakter hat Schreibman 2004. 2015 soll das Handbuch vollständig neu bearbeitet noch einmal erscheinen. Als Handbuch für einen Teilbereich der DH ist auch Susan Schreibman / Ray Siemens 2007 einschlägig.

¹³ Bezeichnenderweise beginnt bereits das *Companion to Digital Humanities* (Schreibman / Siemens 2007) mit dem Kapitel *The History of Humanities Computing*. Der rezenteste Beitrag dürfte Willard McCarty's *Getting there from here. Remembering the future of digital humanities* von 2014 sein (McCarty 2014). An einem größeren Projekt zur Geschichte der DH arbeitet Julianne Nyhan. Vgl. z.B. Julianne Nyhan et al. 2013, die Digital Humanities Quarterly-Sonderausgabe 2012, die »Hidden Histories«-Bibliographie und die demnächst erscheinende Monographie von Julianne Nyhan mit dem Titel *Computation and the Humanities: towards a history of Digital Humanities*.

¹⁴ Terras et al. 2013. Eine Sammlung von Kurzdefinitionen aus dem Day-of-DH-Umfeld findet sich [online](#).

Indikator für die Wahrnehmung als eigenes Fach. Die Aufnahme in solche Systematiken ist aber für neue, kleine Fächer nicht leicht und braucht wohl noch einige Zeit.

2. Digital Humanities als Brücke und als Feld?

Mit dieser Checkliste ist der Entwicklungsstand der Digital Humanities im Jahre 2014 grob skizziert. Die DH erscheinen als voll ausgereiftes, eigenständiges Fach. Es kann kaum bezweifelt werden, dass wir es hier mit einer eigenständigen Disziplin zu tun haben. Allerdings scheint die Sache doch nicht ganz so einfach zu sein, weil diese Disziplin einen besonderen Charakter hat. Als »Brückenfach«, ihrem Wesen nach inter- oder transdisziplinär und zugleich viele spezialisiertere Ausprägungen umgreifend, bestimmt sie sich nämlich nicht nur in Abgrenzung von anderen Fächern, sondern gerade auch in ihrer Hinwendung zu ihnen. In vielen Beschreibungen und Definitionen wird deshalb trotz der oben beschriebenen Verfestigung und Institutionalisierung eher von einem »Forschungsfeld« oder einem »Forschungsbereich« als von einer Disziplin geredet.¹⁵

Dabei dürfte auch dies Teil eines normalen Entwicklungsprozesses sein. Neue Fächer entstehen in der Fokussierung auf bestimmte neue Forschungsfragen. Sie bewegen sich zunächst im Rahmen bestehender Disziplinen, vielleicht auch ihrer Randbereiche und Schnittmengen, bevor sie ihren Gegenstand als etwas Besonderes und Eigenständiges bestimmen, der in den bestehenden, etablierten Fächern nicht mehr hinreichend abgedeckt wird. Disziplinen entstehen aus einer fachlichen Bewegung, wachsen aus anderen Disziplinen heraus und führen letztlich zu einer Ausdifferenzierung der Fächerlandschaft. Damit lässt sich auch für die DH eine eher abstrakte Definition bilden:

Die Digital Humanities befassen sich mit Problemen, die über die benachbarten Einzelfächer in einem doppelten Sinne hinausgehen. Zum einen betreffen sie Fragen, die für viele Fächer gleichermaßen gelten; zum anderen betreffen sie Fragen, die von den benachbarten Fächern nicht behandelt werden, weil sie aus ihrer Sicht zu speziell sind oder Kompetenzen erfordern, die in den Fächern nicht enthalten sind.

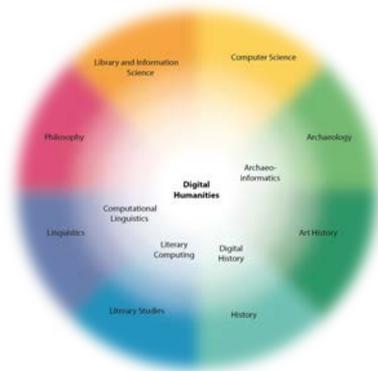
Die Digital Humanities bestimmen sich weitgehend in ihrer Positionierung und ihrem Verhältnis zu anderen Fächern. Zunächst stehen sie offensichtlich zwischen den Geisteswissenschaften und der Informatik. Dann betreffen sie grundsätzlich **alle** Geisteswissenschaften, haben aber historisch bedingt unterschiedlich starke Wurzeln und Schwerpunkte. Wenn man die Geschichte des Feldes untersucht, dann bilden traditionell z.B. die Bereiche der »digitalen Philologien« und der Digitalisierung in Library and Information Science bzw. dem Kulturerbesektor (BAM bzw. GLAM)¹⁶ einen wesentlichen Schwerpunkt. Auf der anderen Seite haben sich seit langem Bereiche wie die Computerlinguistik, aber auch die

¹⁵ Typisch sind hier die verschiedenen Wikipedia-Artikel in insgesamt 16 Sprachen, die ja versuchen einen Konsens abzubilden. Während der deutsche Beitrag »Fach« sagt, ist im Englischen von »area of research, teaching, and creation« die Rede. Andere scheinen sich - vielleicht einfach übersetzend - davon abzuleiten: »campo di studi, ricerca, insegnamento« (italienisch), »domaine de recherche, d'enseignement et d'ingénierie« (französisch), »área de investigación, enseñanza y creación« (spanisch).

¹⁶ Bibliotheken, Archive und Museen (= BAM); Galleries, Libraries, Archives, Museums (= GLAM).

Computerarchäologie oder Archäoinformatik zu weitgehend autonomen Forschungszweigen entwickelt. Insbesondere die Computerlinguistik hat seit den 1980er Jahren mit einer Reihe von Lehrstühlen und Studienprogrammen fast den Charakter einer eigenständigen Disziplin gewonnen. Zu den neueren Entwicklungen gehört, dass traditionelle Fächer in der Umstellung auf digitale Methoden und Werkzeuge Spezialbereiche ausbilden, die ebenfalls den Digital Humanities zugeordnet werden können. Beispiele für solche Schnittmengenbereiche wären ›Digital History‹ oder ›Digitale Kunstgeschichte‹.¹⁷

So wie andere Brücken- oder Schnittstellenfächer auch können die Digital Humanities in verschiedener Weise in ihrem Status bestimmt und nach ihrer Position lokalisiert werden. Während man sie auf der einen Seite – vor allem durch die dedizierten Lehrstühle und Studiengänge – als eigenständige Disziplin verstehen kann, bilden sie zugleich einen Bereich **in** anderen Fächern. Man könnte dann sagen, dass Digital Humanities-Aktivitäten, z.B. in der Abdeckung methodisch-technischer Aspekte der jeweiligen Forschung, Teil der bestehenden Fächer sind. Dabei sorgen sie dann zugleich für die Transformation traditioneller Fächer gemäß den aktuellen epistemologischen Bedingungen einer zunehmend digitalen Informationsversorgung und Informationsverarbeitung. In diesem Sinne kann DH als Skala beschrieben werden: Auf dem einen Ende kann es um den Einsatz generischer Werkzeuge in den Fächern gehen (*low end DH*), auf dem anderen Ende um die Entwicklung spezieller neuer Werkzeuge und Methoden (*high end DH*) und ihre theoretische Reflektion.¹⁸ Zugleich decken die DH damit einen weiten Bereich ab, mit dem sie sowohl **in** anderen Fächern vertreten sind, als auch eine Lücke **zwischen** den Fächern füllen und damit ihre Position als eigenständige Disziplin bestimmen. Durch die Bewegung der Geisteswissenschaften in Richtung angewandter Informatik und Informationswissenschaft zum einen und die Bewegung der Informatik in Richtung Geisteswissenschaften zum anderen lässt sich der Bereich der DH insgesamt abstecken.



¹⁷ Dabei ist nicht zu unterschlagen, dass auch der Computereinsatz in den Geschichtswissenschaften eine sehr lange Tradition hat, die eine eigene historische Aufarbeitung nahelegt. Das Etikett »Digital History« markiert insofern eher eine neue Route als den Versuch einer »Erstbesteigung«.

¹⁸ Zur low-end-high-end-Terminologie vgl. Burghardt / Wolff 2014, S. 40. Kirsch 2014 spricht von »minimalist understanding« (Werkzeugeinsatz) und »maximalist understanding« (DH als Paradigmenwechsel in der Auseinandersetzung mit unserer Kultur).

Abb. 1: 3-Sphären-Modell zur Kartierung der Digital Humanities als Schnittmenge, Brücke und eigenständigem Bereich zwischen (ausgewählten) traditionellen Disziplinen. Quelle: Eigener Entwurf. Zuerst veröffentlicht in Patrick Sahle, DH studieren! Auf dem Weg zu einem Kern- und Referenzcurriculum der Digital Humanities. DARIAH-DE Working Papers Nr. 1. Göttingen: GOEDOC 2013. [\[online\]](#)

Die Grafik (Abbildung 1) versucht deutlich zu machen, dass sich die Digital Humanities als Bereich mit unscharfen Grenzen mit anderen Bereichen überlappen. Für eine distinguierende Beschreibung lässt sich die Grafik aber auch als *Drei-Sphären-Modell* verstehen. Dabei würden die DH die Sphären der »eigenen Disziplin«, der »digital transformierten Disziplinen« (z.B. Computerlinguistik) und der »traditionellen Fächer mit digitalen Anteilen« umfassen. Diese Beschreibung könnte zum Verständnis des ganzen Feldes beitragen. Sie hilft aber auch in der Lokalisierung von Aktivitäten, Projekten oder Studienprogrammen. Die Ausdehnung und Vielfalt der Digital Humanities kann sie besonders da begreifbar machen, wo sie auf einzelne Akteure angewandt wird. Fächer werden von Menschen gebildet und getragen. So wie Disziplinen in Bewegung sind, sind es auch die Identitäten von Fachwissenschaftlern. Digital Humanities werden spätestens dann als Disziplin anzuerkennen sein, wenn eine gewisse Menge von Forschern sich selbst als »Digital Humanists« verstehen und bezeichnen. Diese Selbstbezeichnung liegt auch dann nahe, wenn sie sich aus einer entsprechenden Ausbildung oder einer Stellenaus- und -beschreibung speist. Wer einen BA und MA in »Digital Humanities« hat, wird wohl ein Digital Humanist sein; wer eine Professur in »Digital Humanities« innehat, wird wohl ein Digital Humanist sein.¹⁹ Wir sind aber auch jenseits dieser Formalia recht frei, uns zu definieren und zu etikettieren. Dabei ist es vielleicht gerade typisch für unsere Zeit, dass wir fließende, unscharfe und multiple Identitäten und Selbstbeschreibungen pflegen. Für die Digital Humanities ließe sich dies vermutlich empirisch-experimentell anlässlich der Digital-Humanities-Tagung mit einer Umfrage leicht überprüfen:

1. Auf die Frage »Was sind Sie?« würden wohl nur wenige die einfache Antwort »Digital Humanist« geben.
2. Die Frage »Sind Sie ein Digital Humanist?« würde wohl häufig mit »Hm, ja, irgendwie schon, auch, sonst wäre ich ja nicht hier« beantwortet werden.
3. »Was sind Sie?« würde wohl häufig eine solche Reaktion hervorrufen: »Ich bin [Fach X]-Fachwissenschaftler und interessiere mich für Digital Humanities / arbeite mit digitalen Methoden«.

»Früher war ich Historiker, jetzt bin ich Digital Humanist«; »Ich bin Germanist und Computerphilologe«; »Ich bin Computerlinguist und betreibe damit angewandte Informatik«; »Ich bin Informatiker und bearbeite geisteswissenschaftliche Fragestellungen«; »Ich mache digitale Kunstgeschichte« – die Selbstbeschreibungen sind so vielfältig, temporär und unscharf wie die Interessen und Aktivitäten der Akteure. Gerade weil die Digital Humanities fast immer verschiedene Dinge miteinander verbinden, liegt es hier nahe, dass Selbstbeschreibungen häufig ein oder mehrere »unds« enthalten. Digital Humanities betrifft in der Regel die

¹⁹ Für die Digital Humanities gibt es das Problem der »ersten Generation«: die ersten Vertreter des Faches müssen aus anderen Gebieten kommen, andere Ausbildungsbiografien haben und damit auch eine andere »Berufsbezeichnung«. Mit den expliziten DH-Studiengängen müsste sich dies aber umdrehen – auch wenn ein Studiengang nicht direkt zu einer Berufsbezeichnung führt. Kölner Absolventen neigen z.B. dazu zu sagen »Ich habe Informationsverarbeitung studiert«, aber auch »Ich bin Medieninformatiker/in«.

Entwicklung und Anwendung von Techniken und Methoden auf bestimmte Fachprobleme. Deshalb muss man in mindestens zwei Gebieten unterwegs sein. Darüber hinaus liegen für viele die spannendsten Fragen auf einer Metaebene: Wohin führt die Konfrontation der beiden Welten eigentlich? Wie können punktuelle Lösungen verallgemeinert werden? Was bedeuten die neuen Prozesse und Verfahren für die Geisteswissenschaften und für die Informatik insgesamt? Arbeiten wir nicht letztlich alle an der Praxis und Theorie der Erkenntnisgewinnung in der digitalen Welt?

Die inhärente Multidisziplinarität des Feldes macht es zu einem schwierigen und komplexen Gebiet. Auf der anderen Seite scheinen die Vertreter des Feldes ein besonderes Kompetenzprofil zu entwickeln, das für die Herausforderungen unserer Zeit ungewöhnlich attraktiv sein dürfte. Die Verbindung geisteswissenschaftlicher und informatischer Ansätze erscheint so spannend, dass die Digital Humanities in letzter Zeit auch außerhalb der engeren Fachzirkel eine sehr hohe Aufmerksamkeit gewonnen haben.

3. Digital Humanities als Hype

Digitale Verfahren werden in der geisteswissenschaftlichen Forschung allmählich zu einer Selbstverständlichkeit. Es gibt immer weniger neue große Projekte, die nicht digitale Komponenten enthalten, von denen dann vor allem die Methoden, die Werkzeuge, die Arbeitspraxis und die Publikationsformen betroffen sind. Die Digital Humanities gewinnen dadurch eine stetig wachsende Präsenz und Aufmerksamkeit in den Geisteswissenschaften. Diese Entwicklung ist auch von den Organisationen der Forschungsförderung (insbesondere DFG und BMBF) erkannt worden und hat in den letzten Jahren zur Ausschreibung gezielter Förderprogramme geführt. Die Verbindung spannender Fragen aus den Geisteswissenschaften mit neuen analytischen Methoden, digitalen Publikationen und eindrucksvollen Visualisierungen haben ebenso wie die wissenschaftspolitische Diskussion um die Notwendigkeit oder nur Nützlichkeit der DH auch dazu geführt, dass das Thema 2013 und 2014 in den alten Massenmedien (Zeitungen, Radio, Fernsehen) so präsent war wie noch nie. Die Digital Humanities haben dadurch eine ungeahnte Anziehungskraft entwickelt. Viele neue Akteure sind dabei, das Feld für sich zu entdecken, viele alte sind zunehmend bereit, sich unter dem gemeinsamen Label DH zu versammeln.

Aus Sicht der Community ist gerade die zuletzt genannte Tendenz sehr wichtig. Die Digital Humanities sind – zum Teil noch unter dem Vorläuferbegriff »Humanities Computing« – in den letzten Jahrzehnten manchmal als Bereich wahrgenommen worden, der vor allem von den digitalen Sprach- und Literaturwissenschaften dominiert worden ist.²⁰ In den letzten Jahren lässt sich nun eine starke Verbreiterung der Themen und eine Einbeziehung aller geisteswissenschaftlichen Teildisziplinen beobachten. Vor allem aber kommt es zu einer Annäherung von Spezialisierungsfeldern, die traditionell schon eine gewisse Autonomie entwickelt hatten. Hier sind vor allem die Computerlinguistik, aber auch die Archäoinformatik

²⁰ Diese Tendenz und Wahrnehmung scheint in den USA besonders stark zu sein. In den meisten Fällen, in denen z.B. eine »Digital Humanities Position« ausgeschrieben ist, handelt es sich tatsächlich um Stellen im Bereich »English Studies«.

zu nennen, die erst in den letzten Jahren zunehmend an den gemeinsamen Dachkonferenzen der Digital Humanities teilnehmen. Auf der anderen Seite haben ebenfalls erst in den letzten Jahren viele Akteure aus der ›normalen‹ Informatik erkannt, dass in den Geisteswissenschaften interessante Probleme zu bearbeiten sind und sich den Digital Humanities zugewandt. Die DH selbst betrachten sich als weites und offenes Feld. Unter dem Schlagwort des »Big Tent« wird ein integratives Selbstverständnis gepflegt, bei dem »alle Stämme« eingeladen sind, sich dem gemeinsamen Arbeitsbereich zuzurechnen und sich an den gemeinsamen Projekten, Konferenzen und Organisationen zu beteiligen.²¹ Der gegenwärtige Hype lässt diese Einladung für viele attraktiv erscheinen und hat damit erfreuliche **integrative Tendenzen**. Er könnte dazu führen, dass der Begriff der Digital Humanities als gemeinsamer Sammelbegriff für durchaus unterschiedlich ausgerichtete Teilfelder etabliert wird. In der Konsequenz würde dadurch auch die Wahrnehmung der DH als Disziplin und Fach gestärkt.

Nicht zu übersehen ist allerdings, dass der gegenwärtige Gipfelpunkt im Hype-Zyklus auch einige Gefahren birgt. Er zieht viele neue Akteure an, die bei näherer Betrachtung vielleicht gar nicht an der Entwicklung der DH interessiert sind, sondern ihre eigene Agenda nur unter ein attraktiv erscheinendes Label stellen wollen. Statt gegenseitiger Befruchtung bleibt es dann bei der Konzentration auf die eigenen Fragestellungen. Dies geht zuweilen einher mit einer gewissen Ignoranz gegenüber den bestehenden Traditionen, Erkenntnissen und Erfahrungen der Digital Humanities. Die Vorstellung, DH sei etwas ganz neues, begünstigt den Ansatz, das Rad immer wieder neu erfinden zu wollen. Das vage, anscheinend nicht verfestigte und nicht geschützte Etikett »Digital Humanities« kann auch zu einer gewissen Beliebigkeit führen. Es trägt dann nicht zur Substanz des Forschungsfeldes bei und hat keine nachhaltigen Auswirkungen, wenn eine Auseinandersetzung mit der Disziplin und eine Einbringung in die Community unterbleibt. Eine solche oberflächliche Hinwendung zu den DH im Hype kann dann auch erhebliche **desintegrative Effekte** haben.

4. Beobachtungen

Negative, ablehnende und desintegrative Haltungen, die häufig auf die gleichen Grundmissverständnisse zurückgehen, lassen sich aus verschiedenen Richtungen verzeichnen. Für viele Forscher aus den Geisteswissenschaften liegt der Fokus neben der Bearbeitung von konkreten Fragestellungen auf der Digitalisierung von Quellen, der Zugänglichkeit von Informationen und den digitalen Medien als Kommunikationsraum und Publikationsumgebung. Hier wird zwar auch ein theoretischer Diskurs über digitale Medien geführt, dieser ist aber weitgehend von einer Auseinandersetzung mit ihren technischen und methodischen Grundlagen abgekoppelt. Vorherrschend ist ein Interesse an unmittelbaren Lösungen für bestehende Probleme und an leicht einsetzbaren Werkzeugen. Dabei wird oft angenommen, dass die Werkzeuge nur bestehende Methoden leichter umsetzbar machen würden, selbst aber keine methodischen oder epistemologischen Implikationen hätten. Auf diesem Grundmissverständnis aufbauend werden auch die DH als reiner Dienstleister der Forschung und als Feld der ›Anwendung von Werkzeugen‹ verstanden. Weil die DH damit aber

²¹ Vgl. z.B. Pannacker 2011. Kritisch dazu Melissa Terras, in: Terras et al. 2013, S. 263–270 und Terras 2011 in ihrem Blogbeitrag »Peering Inside the Big Tent: Digital Humanities and the Crisis of Inclusion«.

auch nur **in** den traditionellen Geisteswissenschaften stattfinden würden, gäbe es gar keine eigenständige Disziplin DH. Wenn sich die Fragestellungen und Methoden in Wirklichkeit gar nicht ändern, sondern nur der Gebrauch von Werkzeugen, dann hätten wir es nur mit einer allmählichen Evolution der Geisteswissenschaften zu tun. Ein digitaler Geisteswissenschaftler wäre dann jemand, der fortgeschrittene Werkzeuge benutzt.²² Letztlich würden aber alle digital arbeiten und wenn die Aneignung der Werkzeuge allgemein durchgesetzt wäre, gäbe es auch keinen Bedarf an Digital Humanities mehr. »Die Geisteswissenschaften werden digital, die DH bilden keinen eigenen Bereich, sie sind allenfalls ein Übergangsphänomen und werden wieder verschwinden«²³, ist denn auch ein unter etablierten Geisteswissenschaftlern oft zu hörendes Credo, das zuweilen mit der plakativen Forderung nach einer Zerstörung der Digital Humanities einhergeht.²⁴

Eine ähnliche Negation kommt auch aus der anderen Richtung. Aus Sicht der angewandten Informatik geht es bei den DH meistens um technische Problemlösungen, um die Anwendung etablierter Konzepte und Ansätze aus der Informatik auf Fragestellungen der Geisteswissenschaften. Nicht zuletzt die unausrottbare Startannahme, dass es sich dabei um einfache Probleme handeln würde, die mit Standardlösungen zu behandeln seien, hat in den letzten 30 Jahren zum vielfachen Scheitern geisteswissenschaftlich-informatischer Kooperationen geführt. Zugleich ist diese Grunderfahrung eine der Ursachen für die Entwicklung der Digital Humanities. In ihrer historischen Agnostik wiederholen manche Vertreter der Informatik die Geschichte nun wieder, während andere erkannt haben, dass die Problemstellungen aus den Geisteswissenschaften häufig nicht trivial sind, sondern im Gegenteil sogar wertvolle Anstöße für Entwicklungen sowohl in der angewandten, als auch in der theoretischen Informatik selbst geben können. Mit vielen Geisteswissenschaftlern teilen sie aber die Meinung, dass ein Bereich Digital Humanities überflüssig sei, weil die Probleme aus den Fächern von der Informatik unmittelbar in Software übersetzt und damit gelöst werden könnten. Damit teilen sie aber vor allem das Grundmissverständnis, dass DH eben nicht mehr sei, als die Entwicklung und Anwendung von Werkzeugen, die keine eigenen theoretischen und methodischen Implikationen hätten. Außerdem wird hier das Problem der Kommunikation und der Übersetzung zwischen den beiden Welten übersehen, das eines der Hauptgründe für das Scheitern kooperativer Ansätze ist. Ohne das gleichzeitige Verständnis der geisteswissenschaftlichen Herangehensweisen und informatischer Bearbeitungsmöglichkeiten geht jedenfalls häufig viel Zeit verloren, bis beide Seiten über das Gleiche reden. Außerdem sind beide Seiten oft eher an konkreten Ergebnissen und exemplarischen Lösungen interessiert als an verallgemeinerbaren, nachhaltigen Entwicklungen, wie sie im Fokus eines Faches wie Digital Humanities stehen, das kontinuierlich an einem feldübergreifenden Arsenal von konzeptionellen Ansätzen, bereichsspezifischen Standards und nachnutzbaren Softwarekomponenten interessiert ist.

²² Ein »normaler« Geisteswissenschaftler würde hingegen nur triviale Werkzeuge benutzen. Diese Sicht z.B. bei Haigh 2014, S. 28: »The digital humanities movement is a push to apply the tools and methods of computing to the subject matter of the humanities« - wer heute E-Mail benutze, sei kein digitaler Geisteswissenschaftler, wer ein eigenes Web-CMS installieren und betreiben würde, aber schon.

²³ Haigh 2014, S. 28: »the allure of ›digital humanist‹ as an identity will fade over time«.

²⁴ So explizit Nordmeyer 2013 in einem Blogbeitrag: »Destroy the Digital Humanities! #dighum1213«.

Interessant ist, dass ein Teilfeld, bei dem es sich offensichtlich um eine angewandte Informatik der Geisteswissenschaften handelt, in ganz ähnlicher Weise mit den Digital Humanities und den benachbarten Disziplinen »fremdelt«: Die Computerlinguistik betrachtet sich jedenfalls häufig als echte Informatik und hat mit dieser zugleich kein Problem und kein großes Interesse an ihr, wenn man von offensichtlichen Schnittstellenthemen wie Information Retrieval, Text Mining und Machine Learning absieht. Auf der anderen Seite wird mitunter betont, dass man zwar Fragestellungen aus den Geisteswissenschaften aufgreife und Werkzeuge entwickle, die eigentlich für alle ihre Teildisziplinen relevant seien. Gleichzeitig insistieren einige Vertreter in Gesprächen jedoch darauf, dass die Computerlinguistik schon so weit eine generische Ingenieurssparte sei, dass man als Computerlinguist jedenfalls nicht automatisch auch ein Linguist sei.

Statt Integration, vielfältiger Ausrichtung und multipler Identitäten findet man auch an anderen Stellen häufig ablehnende Haltungen, Monopolisierungsversuche und separatistische Tendenzen. Innerhalb des Feldes begegnet man dieser Tendenz, wenn z.B. die DH primär als digitale Textwissenschaft oder als *literary and linguistic computing* verstanden werden, wobei alle anderen Aktivitäten dann nur exotische Randbereiche wären. Eine Ausgrenzungsstrategie liegt in dem Versuch, technische Mindesthöhen zu etablieren, nach denen z.B. niemand ein Digital Humanist sein könne, der nicht jeden Tag selbst Programmcode schreiben würde; DH sei eigentlich Software Engineering und alles andere nur Geisteswissenschaft.

Außerhalb der DH findet man von Seiten dieser Geisteswissenschaften manchmal auch starke grundsätzliche Vorbehalte, eine prinzipielle Ablehnung und verschiedene Ängste. Gegen das scheinbare Heilsversprechen der Rettung der Geisteswissenschaften durch ihre Informatisierung steht die Angst vor einer »feindlichen Übernahme«, bei der die traditionellen Methoden der Geisteswissenschaften einem Pseudo-Objektivismus und inadäquaten Ansätzen geopfert würden.²⁵ Informatisierung wird dabei als Zumutung, als Kulturfrevel und als Anfang des Untergangs der Geisteswissenschaften empfunden. Dabei sind diese Ängste nur der übersteigerte Ausdruck der an sich nachvollziehbaren Annahme, dass sich Hermeneutik eben nicht eins zu eins durch Statistik und Stochastik ersetzen ließe. Innerhalb der DH wird dies wiederum konstruktiv aufgegriffen, wenn ernsthaft nach der Operationalisierung genuin geisteswissenschaftlicher Verfahren gefragt wird, wenn der scheinbar feststehenden Zielstellung des *high performance computing* bewusst ein *minimal computing* entgegengesetzt wird oder dem *big data*-Ansatz mit der Frage nach dem Wesen von *smart data* begegnet wird.

Ein Beispiel für die »Hinwendung« der Informatik zu neuen Themenbereichen ist der *cultural heritage*-Sektor. Jenseits der Metadaten sind hier z.B. Informationssysteme, Langzeitarchivierung, semantische Informationsmodellierung, fortgeschrittene Digitalisierungsverfahren und neue Analysemethoden Gegenstand des informatischen Interesses. Inzwischen gibt es dabei aber Projekte, Netzwerke, Publikationsorgane und Tagungsreihen, in denen die angewandte Informatik das Thema allein bearbeitet, fast unter sich bleibt und der Zusammenhang zu geisteswissenschaftlichen Fragestellungen oder

²⁵ Kirsch 2014. Für den deutschsprachigen Raum zusammenfassend Kaden 2013. Initial für den »Übernahme«-Begriff Röhnert 2013.

Betrachtungsweisen mehr und mehr aus dem Blick gerät.²⁶ Aus Sicht der Informatik erscheinen die Geisteswissenschaften hier manchmal als überflüssiges Anhängsel.

Für die Entwicklung der Digital Humanities als integratives Fach und für die Haltung der verschiedenen Akteure hierzu kann auch eine Analyse der institutionellen Entwicklung, des Arbeitsmarktes und der ›Professorialisierung‹ sehr erhellend sein. In den letzten Jahren hat es eine ganze Reihe von Zentrumsgründungen gegeben, wobei eine genauere Betrachtung, welche dieser Ansätze (a) tatsächlich mit Leben erfüllt worden sind, (b) tatsächlich die DH in ihrer Breite zum Gegenstand haben und (c) mehr als nur Umetikettierungen bestehender Strukturen sind, zu sehr ernüchternden Ergebnissen führen dürfte. Ein positiveres Signal geht dagegen von den immer zahlreicher werdenden Stellenausschreibungen aus, die meistens im Kontext von Forschungsprojekten explizit an »Digital Humanists« gerichtet sind. Hier wird in der Detailbeschreibung zwar auch nach bestimmten Fachausrichtungen und speziellen Kompetenzen gefragt - angesichts eines Überhangs an Stellen gegenüber einem Mangel an qualifizierten Bewerbern sind die Ausschreibungen in den meisten Fällen aber recht offen gehalten: Sie lassen grundsätzlich offen, ob Bewerber nun Geisteswissenschaftler mit informatischen Kompetenzen, Informatiker mit geisteswissenschaftlichem Problemverständnis oder eben beidseitig gleichermaßen qualifizierte Digital Humanities-Experten sind. Ganz anders sieht die Situation bei den Leitungsstellen und Professuren aus. Ausschreibungen, die DH-Professuren zum Gegenstand haben, bei denen »das Fach in seiner ganzen Breite« abgedeckt werden soll, sind sehr selten. Häufiger bleibt der Eindruck zurück, dass unter dem allgemeinen Etikett »Digital Humanities« etwas sehr Spezielles versteckt wird. Das kann z.B. durch eine sehr lokale Definition von »Digital Humanities« erfolgen, die keinen Bezug zu der breiten internationalen Community herstellt, sondern nur eine klar begrenzte Perspektive vor Ort realisiert. Damit kann auch die Beschreibung eines engen Aufgabenprofils einhergehen, mit dem dann kein integrativer Impetus zum Fach DH hin verbunden ist. Häufig verbirgt sich hinter der DH-Überschrift sehr klar die Orientierung an einem der traditionellen Fächer. Bewerber müssen dann von vorneherein z.B. germanistische Philologen, aus dem Fach »English Studies« oder Medienwissenschaftler sein. Im Extremfall werden Interessenten, die nicht über einen bestimmten formalen Abschluss (z.B. in der Informatik) verfügen, von vornherein von einer Bewerbung ausgeschlossen – was der Idee der DH als ein Brückenfach entgegensteht. Problematisch und erhellend zugleich ist die Ansiedlung solcher Professuren an den bestehenden Strukturen. Von der Sache her liegt eine Zuordnung unmittelbar zu einer Fakultät oder einem DH-Zentrum nahe. Tatsächlich sind solche Stellen aber meistens einem Institut oder Department zugeordnet, das ein bestimmtes einzelnes Fach vertritt. In der akademischen Strukturlogik ist es dann aber sehr unwahrscheinlich, dass innerhalb eines Fachinstituts eine gänzlich fach- oder sogar fakultätsübergreifende Ausgestaltung einer Professur gewollt wird. Häufig nur umetikettierte oder zusätzlich gelabelte Stellen bewahren so einen engen Fokus auf ein bestimmtes Fach. Aus dieser Perspektive erklärt sich dann auch die Ignoranz gegenüber den DH als etabliertem globalem Forschungsfeld und die Beschränkung auf die jeweils eigene idiosynkratische Bestimmung dessen, was man vor Ort darunter zu verstehen beliebt. Nichts spricht gegen Professuren, die in dem weiten Feld der

²⁶ Nur ein Beispiel für diese Tendenz mag das EU-geförderte Doktorandenprogramm [Initial Training Network for Digital Cultural Heritage](#) sein. Von den 16 Fellows haben drei einen Bildungsanteil in Geschichtswissenschaft oder Archäologie. Fast alle anderen haben einen eher naturwissenschaftlichen Hintergrund.

DH bestimmten Bereichen gewidmet sind, es stellt sich aber die Frage, ob diese Stellen als Teil einer breiteren Disziplin und als Brücke zu diesem Bereich oder doch nur als Fortsetzung der traditionellen Fachzuschnitte ohne Verbindung zu anderen Entwicklungen gedacht werden.

Vielleicht sind dies typische Phänomene aus der Verfestigungsphase einer neuen Disziplin. Wissenschaftssystematisch und wissenschaftshistorisch könnte hier eine Vergleichsbetrachtung zu einem sehr ähnlichen Fach wie der inzwischen weithin etablierten Wirtschaftsinformatik fruchtbar sein.²⁷ Auch diese steht ja zwischen der angewandten Informatik und einem anderen Fachbereich. Auch sie wird häufig nach einem Drei-Säulen-Modell beschrieben, das aus dem Ausgangsfach, der Informatik und einem genuinen Zwischenbereich besteht. Man hat allerdings den Eindruck, dass die Existenz und Daseinsberechtigung der Wirtschaftsinformatik nicht in dem Maße in Frage gestellt wird, wie dies bei den Digital Humanities der Fall ist. Während bei der Wirtschaftsinformatik schon die Entwicklung wertvoller Problemlösungen und die Ausbildung stark nachgefragter spezialisierter Fachleute eine hinreichende Legitimation zu sein scheint, dürfte gerade die große methodische und konzeptionelle Distanz zwischen den Geisteswissenschaften und der Informatik eine der Ursachen für den beidseitigen Argwohn und die manchmal überraschend starken Vorbehalte und Ablehnungen bis hin zum ›Beißreflex‹ sein.

5. Effekte

Die digitalen Geisteswissenschaften gehen seit etlichen Jahrzehnten ihren Weg. Können die positiven oder negativen Erscheinungen des gegenwärtigen Hypes daran etwas ändern? Haben sie jenseits der theoretischen Debatten überhaupt praktische Auswirkungen? Kann es uns nicht egal sein, ob andere DH faszinierend oder abstoßend finden?

Leider steht zu befürchten, dass eine Verwendung von »Digital Humanities« als bloßes Etikett, unter dem dann ganz andere Dinge betrieben werden, der Entwicklung und Etablierung der Disziplin mehr schadet als nutzt. Denn in diesem Fall entstehen keine echten und nachhaltigen Beiträge zum Forschungsfeld und es wird eher verunklart, was eigentlich Gegenstand und Zielstellungen der DH sind. Diese Probleme aus dem Umfeld des gegenwärtigen Hypes lassen sich vielleicht auch an einem nicht unwichtigen Beispiel zeigen: den Förderlinien des deutschen Bundesministeriums für Bildung und Forschung (BMBF), mit denen in den Jahren 2013 und 2014 erhebliche Gelder für »Forschungs- und Entwicklungsvorhaben aus dem Bereich der eHumanities«²⁸ ausgelobt worden sind. Trotz der Begriffsverwendung »eHumanities« und dem Fehlen des Ausdrucks »Digital Humanities« in den Richtlinien ist davon auszugehen, dass hier thematisch die Digital Humanities angesprochen sind, da beide Begriffe synonym verwendet werden. Jedenfalls ist keine Definition einschlägig, nach der es sich bei den eHumanities um ein anderes, von den DH distinktes Feld handeln würde. Der Call richtet sich zum Teil auf den Aufbau von nachhaltigen Strukturen in den eHumanities, er geht aber ausdrücklich davon aus, dass die Fragestellungen

²⁷ Zur Entwicklungsgeschichte der Disziplin siehe z.B. das [Projekt WIGE - Wirtschaftsinformatik Genealogie](#).

²⁸ Vgl. die [Förderlinien des BMBF](#).

aus den Geisteswissenschaften kämen, die dann mit informatischen Ansätzen zu bearbeiten seien. Auch wenn hier nicht von eHumanities oder Digital Humanities geredet wird, ist das aus DH-Sicht zunächst unproblematisch, da die DH entweder tatsächlich geisteswissenschaftliche Fragestellungen aufgreifen oder – wo es um genuine DH-Probleme geht – selbst zu den Geisteswissenschaften gezählt werden können.²⁹ Auf der anderen Seite kann auch die Lösung der Probleme mit informatischen Ansätzen unproblematisch sein, da sich die DH ja als angewandte Informatik verstehen und spätestens mit den im Call genannten »informatiknahen Fächern« nichts anderes als die Digital Humanities gemeint sein können.

Dass die Ausschreibung hier sprachlich etwas unklar und interpretierbar ist, kann bei der Diversität der beteiligten Akteure nicht überraschen.³⁰ Problematisch wird dies erst durch die Effekte der verschiedenen Deutungen. Der Call geht z.B. immer von einem Team aus Wissenschaftlern aus Geisteswissenschaften und Informatik aus, das von einer entsprechenden Doppelspitze geleitet würde. Es stellt sich dann die Frage, wo und wie Anträge aus den DH im engeren Sinne zu positionieren wären. Gilt DH dann als Geisteswissenschaft – mit dem Zwang zu einem zusätzlichen Partner aus der Informatik? Oder als »informatiknahes Fach« - mit dem Zwang eines zusätzlichen Partners aus den Geisteswissenschaften? Oder fallen beide Seiten in den DH zusammen, so dass sich der DH-Teil zur Erfüllung der »Doppelspitze« verdoppeln müsste? Eine entsprechende Anfrage beim Projektträger wurde dahingehend beantwortet, dass dieses Problem in Wirklichkeit nicht bestehen würde, weil es die eHumanities ja (noch) nicht geben würde (!), sondern nur Geisteswissenschaftler und Informatiker. Das inhaltliche Verständnis auf Seiten des Projektträgers, der ja nur für die administrative Abwicklung zuständig sein sollte, mag man für irrelevant halten.³¹ Entscheidend sind letztlich die Gutachter. Aber hier zeigt sich ein ähnliches Phänomen. Zu jedem Antrag werden zwei Stellungnahmen eingeholt, die von Experten aus drei Gruppen angefertigt sind: den Geisteswissenschaften, den Digital Humanities und der Informatik (mit Computerlinguistik und Informationswissenschaft), wobei die DH-Experten nur eine sehr kleine Gruppe bilden. Anscheinend hat es immer ein Gutachten von der geisteswissenschaftlichen Seite und eines aus der Informatik oder den DH gegeben. Für etliche Anträge ist es dabei zu sehr stark divergierenden Punktzahlen gekommen.³² Die anonymen und sehr knappen Rückmeldungen lassen bei aller gebotenen Vorsicht der Interpretation den Eindruck entstehen, dass vor allem von Gutachtern aus der Informatik und Computerlinguistik Anträge mit starken DH-Anteilen rundweg abgelehnt werden, wenn nicht die je eigenen thematischen Interessen angesprochen oder in hohem Maße klassische Informatik-Institute außerhalb der Digital Humanities eingebunden sind. Dadurch kann aber eine starke Schiefelage und Chancenungleichheit entstehen, bei der bestimmte geisteswissenschaftliche Bereiche und genuine DH-Ansätze benachteiligt werden. Im Sinne dieses Beitrages ist aber vor allem zu folgern, dass mit den vom BMBF organisierten Ausschreibungen die eHumanities bzw. Digital Humanities als

²⁹ Svensson 2012, S. 42–60.

³⁰ Der Strategiebeirat deckt im Grunde die ganze Bandbreite von geisteswissenschaftlichen Disziplinen über Digital-Humanities-Experten im engeren Sinne bis zur Informationswissenschaft und der Informatik ab.

³¹ Zum Projektträger vgl. einen Kommentar von Nettelbeck 2015.

³² Dieser Eindruck ist bei vielen Gesprächen mit Antragstellern bestätigt worden. Dass ausdrückliche Nicht-Gutachten (bei denen eine echte Begutachtung eben verweigert wird) mit Null Punkten bewertet und zur arithmetischen Mittelung herangezogen werden oder bei stark divergierenden Wertungen kein drittes Gutachten eingeholt wird, betrifft nur die Frage, ob das Verfahren den allgemein üblichen Standards entspricht. Diese Frage kann aber nicht Gegenstand des Beitrags sein.

Forschungsbereich und als gemeinsamer Kommunikationsraum nicht konsequent aufgegriffen und gestärkt werden, sondern ein nur kooperatives Nebeneinander von Geisteswissenschaften und Informatik befördert wird. Dabei ist schon in den Ausschreibungsverfahren deutlich geworden, wie alle Seiten von Anfang an aneinander vorbeireden, wenn die Digital Humanities nicht als Bindeglied ernst genommen werden. In der Negation dieser Disziplin liegt zudem die Gefahr eines deutschen Sonderweges. Erfahrungsgemäß verbleiben in rein kooperativen Projekten die Ergebnisse und Erkenntnisse der geisteswissenschaftlichen bzw. der informatischen Partner in deren jeweiligen Fachdiskursen - ohne dass es zu einem Anschluss an die internationale Digital Humanities-Entwicklung kommen würde.

Auf dieser Grundlage besteht dann aber auch die Gefahr, dass es trotz des hohen Mitteleinsatzes nicht zu einer dauerhaften Stärkung und einem Ausbau der Digital Humanities in Deutschland kommt, sondern aus unverbundenen Teilen bestehende Kooperationsprojekte jenseits ihrer lokalen Problemlösung keine nachhaltigen Effekte auf die DH insgesamt haben werden. Die Betonung der autonomen Ränder hat dann jedenfalls desintegrative Effekte auf den doch eigentlich adressierten Kernbereich und fördert eher zentrifugale als zentripetale und integrative Tendenzen.

6. Digital Humanities quo vadis?

Was ist zu tun? Um die Digital Humanities nicht der Beliebigkeit anheim fallen zu lassen, kommen wir nicht umhin, auf einen klar und fest definierten Begriff zu pochen. Man wird mit dem Spagat leben müssen, dass es die DH im engeren Sinne (als eigene Disziplin) und im weiteren Sinne (als Phänomen in den anderen Fächern) gibt. Allerdings sollte auch die Vorstellung von den DH als **einer** Community nicht aufgegeben werden. Integration in dieser Fachgemeinschaft und Trennschärfe nach außen sind zwei wichtige Ziele. Die DH sind dadurch bestimmt, dass sie einen gemeinsamen Aufgabenbereich abdecken, in dem spezifische – aber eben nach außen gut anschlussfähige – Techniken, Methoden und Theorien bearbeitet werden. Dieser Bereich muss in horizontaler und vertikaler Hinsicht abgrenzbar sein. Man muss auch sagen können, was DH **nicht** ist. Dazu kann gehören, dass z.B. Digital Literacy nicht DH ist. Der Einsatz von Standardanwendungen oder allgemeiner Software bei geisteswissenschaftlichen Fragestellungen ist ebenfalls kein DH. Auf der anderen Seite sind auch generische Fragestellungen oder Ansätze aus der Informatik nicht automatisch DH. 3D-Scanning und 3D-Modelle sind z.B. nur dann DH, wenn sie nicht nur angewandt werden, sondern wenn die spezifischen Problemlagen geisteswissenschaftlicher Datenbestände oder Fragestellungen zu einer Weiterentwicklung solcher Verfahren führt oder wenn über die Rückwirkungen auf die Forschungsprozesse reflektiert wird. Die Digital Humanities sind ihrem engeren Begriff nach ein Bereich **zwischen** Geisteswissenschaften und Informatik – alles, was ohne Verbindungen nach außen **in** den Fächern passieren kann, ist nicht DH im eigentlichen Sinne.

Zum weiteren Ausbau der DH sind neben der klaren inhaltlichen Positionierung weitere Maßnahmen nötig. Dazu gehört die Verstärkung der Verbandsarbeit mit dem Ziel einer Integration nach innen und einer besseren Sichtbarkeit nach außen. Die Ergebnisse der

vielfältigen Einzelprojekte sollten stärker durch Dokumentation, Evaluation und Reflektion gesichert und zu einem gemeinsamen Wissensbestand zusammengeführt werden. Nachhaltige institutionelle Strukturen in Form von Zentren, Abteilungen und Lehrstühlen sind aufzubauen. Ausbildungsprogramme sollten verstärkt als dem gemeinsamen Feld Digital Humanities zugehörig erkennbar gemacht werden und soweit wie möglich inhaltlich abgestimmt werden. In der Vermittlung der DH sollten ihre Aktivitäten sowohl im lokalen, regionalen, nationalen als auch im internationalen Kontext verankert werden. Bei den Förderorganisationen ist darauf zu drängen, dass DH-Fachgremien eingerichtet werden, die dann auch mit Experten **aus** den DH (durchaus im weiteren Sinne) besetzt werden. Schließlich ist sicher zu stellen, dass auch DH-Förderanträge von Gutachtern aus dem entsprechenden Gebiet bewertet werden und nicht von fachfremden Wissenschaftlern ohne genauere Kenntnis des Feldes.

Dies wären jedenfalls wichtige Voraussetzungen dafür, dass auch die Digital Humanities ihre Entwicklung zu einer leistungsfähigen wissenschaftlichen Disziplin erfolgreich fortsetzen können.

Bibliographische Angaben

Manuel Burghardt / Christian Wolff: Digital Humanities: Buzzword oder Strukturwandel der Geisteswissenschaften? In: Blick in die Wissenschaft - Forschungsmagazin der Universität Regensburg 29 (2014), S. 40. [\[online\]](#)

A Companion to Digital Humanities. Hg. von Susan Schreibman / Ray Siemens / John Unsworth. Malden 2004. [\[online\]](#)

A Companion to Digital Literary Studies. Hg. von Ray Siemens / Susan Schreibman. Malden 2007. [\[online\]](#)

Defining Digital Humanities. A Reader. Hg. von Melissa Terras / Julianne Nyhan / Edward Vanhoutte. Farnham 2013. [\[Nachweis im OPAC\]](#)

Thomas Haigh: We Have Never Been Digital. In: Communications of the ACM 57 (2014), H. 9, S. 24–28. DOI: [10.1145/2644148](#).

»Hidden Histories: Computing and the Humanities c. 1965–1985«. Digital Humanities Quarterly-Sonderausgabe 6.3 (2012). [\[online\]](#)

Ben Kaden: Wer übernimmt was? Zum Verhältnis von Digital Humanities und Geisteswissenschaften. In: LIBREAS. Debatte, 12. September 2013. [\[online\]](#)

Adam Kirsch: Technology Is Taking Over English Departments - The false promise of the digital humanities, New Republic vom 2. Mai 2014 [\[online\]](#)

Willard McCarty: Getting there from here. Remembering the future of digital humanities: Roberto Busa Award lecture 2013. In: Literary & Linguistic Computing 29 (2014), H. 3, S. 283–306. DOI:10.1093/lc/fqu022. [\[Nachweis im GBV\]](#)

Joachim Nettelbeck: Abwicklung im Blindflug. In: Frankfurter Allgemeine Zeitung vom 10. September 2015. [\[online\]](#)

Philip Nordmeyer: »Destroy the Digital Humanities! #dighum1213«. Blogbeitrag vom 9.12.2013. [\[online\]](#)

Julianne Nyhan / Andrew Flinn / Anne Welsh: Oral History and the hidden histories project: towards histories of computing in the Humanities. In: Literary and Linguistic Computing 28 (2013), S. 1–15. DOI: 10.1093/lc/fqt044. [\[Nachweis im GBV\]](#)

Julianne Nyhan: Computation and the Humanities: towards a history of Digital Humanities. Cham 2016. [\[online\]](#)

William Pannacker: »Big Tent Digital Humanities« – A View From the Edge. In: The Chronicle of Higher Education vom 31. Juli 2011. [\[online\]](#)

Jan Röhnert: Feindliche Übernahme? Die Geisteswissenschaften wehren sich gegen falsche Ansprüche der Informatik, aber setzen auf die »Digital Humanities«. In: Frankfurter Allgemeine Zeitung vom 19. Juli 2013. [\[online\]](#)

Patrick Sahle: DH Studieren! Auf dem Weg zu einem Kern- und Referenzcurriculum der Digital Humanities. DARIAH-DE Working Papers Nr. 1. Göttingen 2013. [\[online\]](#)

Patrik Svensson: The digital humanities as a humanities project. In: Arts and Humanities in Higher Education 11/1-2 (2012), S. 42–60. [\[Nachweis im GBV\]](#)

Melissa Terras: Peering Inside the Big Tent: Digital Humanities and the Crisis of Inclusion. In: Blogbeitrag vom 26. Juli 2011. [\[online\]](#)

Scott Weingart: Submissions to Digital Humanities 2015 (mehrere Teile). In: the scottbot irregular. 2013f. [\[online\]](#)

Scott Weingart: Submissions to Digital Humanities 2014. In: the scottbot irregular. 2013f. [\[online\]](#)

Scott Weingart: Acceptances to Digital Humanities 2014. In: the scottbot irregular. 2013f. [\[online\]](#)

Scott Weingart: Acceptances to Digital Humanities 2013. In: the scottbot irregular. 2013f. [\[online\]](#)

Abbildungslegenden und -nachweise

Abb. 1: 3-Sphären-Modell zur Kartierung der Digital Humanities als Schnittmenge, Brücke und eigenständigem Bereich zwischen (ausgewählten) traditionellen Disziplinen. Quelle: Eigener Entwurf. Zuerst veröffentlicht in Patrick Sahle, DH studieren! Auf dem Weg zu einem Kern- und Referenzcurriculum der Digital Humanities. DARIAH-DE Working Papers Nr. 1. Göttingen: GOEDOC 2013. [\[online\]](#)

Zeitschrift für digitale Geisteswissenschaften

Artikel aus:

Sonderband 1 der ZfdG: Grenzen und Möglichkeiten der Digital Humanities. Hg. von Constanze Baum und Thomas Stäcker. 2015. DOI: [10.17175/sb01](https://doi.org/10.17175/sb01)

Titel:

Vom Urheber zur Crowd, vom Werk zur Version, vom Schutz zur Öffnung? Kollaboratives Schreiben und Bewerten in den Digital Humanities

Autor/in:

Thomas Ernst

Kontakt:

thomas.ernst@uni-due.de

Institution:

Universität Duisburg-Essen

GND:

[130500763](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-30500763)

ORCID:

[0000-0002-2170-9241](https://orcid.org/0000-0002-2170-9241)

DOI des Artikels:

[10.17175/sb001_021](https://doi.org/10.17175/sb001_021)

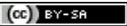
Nachweis im OPAC der Herzog August Bibliothek:

[835243028](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-835243028)

Erstveröffentlichung:

15.09.2015

Lizenz:

Sofern nicht anders angegeben 

Medienlizenzen:

Medienrechte liegen bei den Autoren

Letzte Überprüfung aller Verweise:

24.05.2016

GND-Verschlagwortung:

[Urheber](#) | [Digital Humanities](#) | [Kollaboration](#) |

Zitierweise:

Thomas Ernst: Vom Urheber zur Crowd, vom Werk zur Version, vom Schutz zur Öffnung? Kollaboratives Schreiben und Bewerten in den Digital Humanities. In: Grenzen und Möglichkeiten der Digital Humanities. Hg. von Constanze Baum / Thomas Stäcker. 2015 (= Sonderband der Zeitschrift für digitale Geisteswissenschaften, 1). PDF Format ohne Paginierung. Als text/html abrufbar unter DOI: [10.17175/sb001_021](https://doi.org/10.17175/sb001_021).

Thomas Ernst

Vom Urheber zur Crowd, vom Werk zur Version, vom Schutz zur Öffnung? Kollaboratives Schreiben und Bewerten in den Digital Humanities

Abstracts

Open Review-Plattformen und Wikis verändern die wissenschaftliche Erkenntnisproduktion, Soziale Medien und Wissenschaftsblogs die Wissenschaftskommunikation. Diese digitalen Publikationsformen stellen die Veröffentlichungsweisen und Autorschaftskonzepte der Geisteswissenschaften praktisch und kategorial vor einige Probleme. Daher ist es notwendig, den Begriff ›wissenschaftliche Autorschaft‹ angesichts neuer Formen des kollaborativen Schreibens einer Revision zu unterziehen und die Wissenschaftlichkeit der neuen Medienformate zu reflektieren. Dies kann exemplarisch an selbstkontrollierten digitalen Veröffentlichungsinfrastrukturen und Modellen eines offeneren Immaterialgüterrechts untersucht werden.

Open review projects and wikis are changing academic knowledge production, just as academic blogs and social media are influencing scholarly communication. These digital forms of publishing challenge the usual practices of publishing and the concepts of authorship within the humanities, both practically and categorically. Therefore, it is necessary to revise the concept of ›scholarly authorship‹ by differentiating new forms of academic collaborative writing and to reflect on the academic quality of the new media formats within the humanities. This can be shown when analyzing self-curated digital publishing infrastructures and open license models of intellectual property.

1. Wissenschaftliches Publizieren in digitalen Medien. Eine Einführung

»Historic man may turn out to have been literate man. An episode.«¹

Die Möglichkeiten des digitalen Publizierens verändern auch die Formen wissenschaftlicher Produktion, Publikation und Rezeption einschneidend: »Print is Dead« – so betitelt Jeff Gomez seine Reflexionen über *Books in Our Digital Age*.² Wenn man die medienwissenschaftliche Grundannahme ernst nimmt, dass unser Wissen über Menschen und Dinge auch durch seine medialen Träger präformiert wird, müsste das digitale Publizieren unser Bild vom menschlichen Subjekt und somit auch von dem des wissenschaftlichen Subjekts zumindest modifizieren. Marshall McLuhan bezeichnet bereits 1953 das Denken des prähistorischen, illiteraten und nur auf die wiederkehrenden Jahreszeiten fixierten Menschen einerseits sowie das des Menschen im elektronischen Zeitalter der Neuen Medien andererseits als eines des ›here and how‹ und stuft es somit als ähnlich geschichtsvergessen ein. Dazwischen habe es

¹ McLuhan 1953, S. 118.

² Gomez 2008.

die Gutenberg-Galaxis gegeben und mit ihr den Typus des geschichts- und selbstbewussten Menschen, der aber selbst bald Geschichte sein werde.³

Weniger medienanthropologisch als vielmehr wissenschaftspraktisch beschreibt Julia Flanders die Digital Humanities, die grundsätzlich als »critical of current tenure and reward systems« einzuschätzen seien, wobei diese bestehenden Systeme noch große Schwierigkeiten hätten, »to understand the fundamentally collaborative and interdisciplinary work of digital humanities or the new modes of scholarly communication it is engendering.«⁴ Gelassener sieht dies wiederum Valentin Groebner, der die »digitale Wissenschaftssprache« als eine »Zukunft von gestern« historisiert und sich ebenso differenziert wie skeptisch zu den enthusiastischen Prophezeiungen einer besseren, weil digitalen Wissenschaft positioniert.⁵ Die Debatte über das digitale Publizieren in der Wissenschaft ist intensiv und kontrovers, die Kritik richtet sich teils gegen das traditionelle Wissenschaftssystem und teils gegen die neuen Entwicklungen.

Innerhalb dieser Debatten interessiert sich der vorliegende Beitrag für die Potenziale des digitalen Publizierens in der Wissenschaft und beschäftigt sich mit Formen einer kollaborativen, vernetzten und interaktiven Autorschaft, der Versionier- und Wandelbarkeit von wissenschaftlichen Texten sowie dem freien Zugang zu wissenschaftlichem Wissen. Damit fokussiert er Verfahren und Themen, die bereits in der »Gutenberg-Galaxis« präsent waren, jedoch im digitalen Publizieren eine weitaus höhere Intensität und Relevanz erhalten. Zugespitzt könnte man formulieren: Dieser Beitrag untersucht, inwiefern der geisteswissenschaftliche Standard, Erkenntnisse vorrangig in *abgeschlossenen* wissenschaftlichen Werken zu produzieren, die im Regelfall *einem* wissenschaftlichen Autor oder einer Autorin zugeschrieben werden und nur in *auszugsweise zitierender* Bezugnahme von anderen Forschern weiterbearbeitet werden können, in digitalen Publikationen und Kommunikationsnetzwerken problematisch wird, wenn man ihre Potenziale zur Produktion präziserer, material- und kenntnisreicherer Texte und somit zur Produktion besserer wissenschaftlicher Erkenntnisse nutzen will.

Die Ausgangsthese ist dabei, dass die digitalen Publikationsmöglichkeiten in den Geisteswissenschaften nicht nur die bisherigen Veröffentlichungspraxen, Zitierweisen, Textsorten, Review-Verfahren und Vertragsmodelle erweitern, sondern zugleich auch unsere Vorstellung von dem ändern, was wissenschaftliche Autorschaft bzw. ein wissenschaftliches Subjekt ist und wie sich die Geisteswissenschaften selbst verstehen. Es wird dann notwendig, nicht nur die bekannten wissenschaftlichen Gattungen, wie die Monographie oder den Aufsatz, auf ihre kollaborative Autorschaft hin zu untersuchen, wie dies jüngst Nyhan / Duke-Williams unternommen haben,⁶ sondern innovative Formen digitaler wissenschaftlicher Publikationen als Gegenstände ernst zu nehmen.

³ McLuhan 1953, S. 118.

⁴ Flanders 2013, S. 213.

⁵ Groebner 2014, *passim*.

⁶ Julianne Nyhan und Oliver Duke-Williams haben in einer empirischen Analyse am Beispiel von drei Magazinen der Digital Humanities gezeigt, dass die vermutete Tendenz zur kollaborativen Autorschaft in Aufsätzen der Digital Humanities nur bedingt der Wahrheit entspricht, da tatsächlich die Einzelautorschaft dominiert, wobei allerdings die Zahl multipler Autorschaften zunimmt. Die Autoren zeigen die Notwendigkeit weiterer und differenzierter Untersuchungen auf, denn eine offizielle multiple Autorschaft eines Aufsatzes sage wenig über den kollaborativen Schreibprozess selbst aus, außerdem empfehlen sie: »other factors need

Wenn ich im Folgenden meine Argumentation aus einer medienhistorischen und diskursanalytischen Perspektive entwickle, steht methodologisch die Frage im Vordergrund, auf Basis welcher Regelsysteme und medialer Bedingungen sich der geisteswissenschaftliche Spezialdiskurs⁷ wandelt. Zunächst werde ich daher in einem medienhistorischen Rückgriff danach fragen, wie – insbesondere Ende des 18. Jahrhunderts – die für die Wissenschaft bis heute zentralen Kategorien ›Urheber‹, ›geistiges Eigentum‹ und ›Werk‹ bestimmt wurden. In einem zweiten Schritt werde ich mich mit (vor allem medien-)theoretischen Ansätzen beschäftigen, die verschiedene Kategorien der ›Gutenberg-Galaxis‹ fundamental in Frage gestellt sehen. Diese Problemlage werde ich anschließend an konkreten Beispielen der digitalen Wissenschaftskommunikation sowie digitaler Buchveröffentlichungen in den Geisteswissenschaften zu differenzieren versuchen.⁸

2. Das gedruckte Werk, der Urheber und das geistige Eigentum. Eine wissenschaftshistorische Perspektive auf die Kategorien der ›Gutenberg-Galaxis‹

Der mediale Träger ›gedrucktes Buch‹ hat spezifische materielle Eigenschaften, in fester Bindung zählen dazu der Schutzumschlag (aus Cover, Rücken, Umschlagrückseite, vorderer und hinterer Klappe), Buchdeckel, Buchrücken, Vorsatz sowie Schmutztitel, Haupttitel, Vorwort, Inhaltsverzeichnis und – im Fall wissenschaftlicher Bücher – Apparat (Register, Anmerkungen, Abbildungsverzeichnis, Literaturangaben, Glossar).⁹ Die Texte selbst sind von diesen Paratexten¹⁰ zu unterscheiden und werden in der Tradition des Kodex spezifisch angeordnet. Wissenschaftliche Erkenntnisse werden im Regelfall – da sie den Anspruch erheben, einen dauerhaften Erkenntniswert zu besitzen – als ›Werk‹ veröffentlicht, bei dem es sich – so das *Reallexikon der deutschen Literaturwissenschaft* – um ein »fertige[s] und abgeschlossene[s] Ergebnis« handelt, »das einem Autor zugehört und in fixierter, die Zeit überdauernder Form vorliegt, so daß es dem Zugriff des Produzenten ebenso entzogen ist wie dem Verbrauch durch den Rezipienten.«¹¹

Die klar strukturierte und statische Materialität des Buches und des abgeschlossenen Werks ist auch zentral für die Begründung eines geistigen Eigentums des Autors an seinen Texten. Exemplarisch sei hier an Johann Gottlieb Fichtes Argumentation von 1793 erinnert, die für den deutschsprachigen Kontext besonders folgenreich war und sich unmittelbar auf das mediale Vehikel ›Buch‹ und seine Materialität sowie eine starke Betonung des individuellen Denk- und Lektüreprozesses bezieht.¹² Fichtes entscheidende Setzung ist, dass die Menschen

to be considered in addition to publication patterns, for example, those named on the contributor pages of individual projects.« (Nyhan / Duke-Williams 2014, S. 13)

⁷Vgl. Link / Link-Heer 1990, S. 92.

⁸ Dieser Beitrag basiert auf dem Vortrag *Jenseits des wissenschaftlichen Werks und des geistigen Eigentums? Die digitale Verbreitung wissenschaftlichen Wissens*, der am 26. März 2014 auf der Konferenz *Digital Humanities – methodischer Brückenschlag oder ›feindliche Übernahme‹? Chancen und Risiken der Begegnung zwischen Geisteswissenschaften und Informatik* gehalten wurde und als [Audio-Podcast](#) abrufbar ist.

⁹Vgl. Heinold 2001, S. 328–333.

¹⁰Vgl. zu diesem Begriff grundsätzlich Genette 2001.

¹¹ Thomé 2003, S. 832.

¹²Die Ausführungen zu Fichte decken sich mit Ernst 2012, S. 45f.

notwendigerweise » *das Eigentum eines Dinges* [behalten], *dessen Zueignung durch einen Anderen physisch unmöglich ist.*« Das Buch als ›Ding‹ wäre zunächst zu differenzieren: erstens in »das *Körperliche* desselben, das bedruckte Papier«, und zweitens in »sein *Geistiges*.«¹³ Dieses Geistige wiederum wäre aufzuteilen in »das *Materielle*, den Inhalt des Buches, die Gedanken, die es vorträgt; und in die *Form* dieser Gedanken, die Art wie, die Verbindung in welcher, die Wendungen und die Worte, mit denen es sie vorträgt.« Das körperliche, gedruckte Buch kann sehr wohl seinen Besitzer wechseln und geht »durch den Verkauf [...] un widersprechlich auf den Käufer über«,¹⁴ wodurch dieser als Leser zugleich die Möglichkeit erhält, den Inhalt des Buches in seine eigenen Gedankengänge aufzunehmen. Die ursprüngliche Form dieser Gedanken allerdings, wie sie sich der Schriftsteller ausgedacht hat, kann sich »[n]iemand [...] zueignen, ohne dadurch, dass er ihre Form verändere. Die letztere bleibt auf immer sein *ausschliessendes Eigentum*.«¹⁵ Aus dieser Begründung eines ›geistigen Eigentums‹ von Autoren an ihren Texten leiten sich für sie zwei Rechte ab: Weder dürfe ihnen dieses Eigentum abgesprochen werden, noch dürfe sich jemand dieses Eigentums unrechtmäßig bemächtigen, wie dies zu Fichtes Zeiten die Nachdrucker machen, die teilweise sogar ohne Rücksprache mit den Verfassern in ihren Nachdruckversionen das Original ändern. Diese Konstruktion privilegiert die Vorstellung vom singulären, originären Autor und vom abgeschlossenen ›Werk‹, das sich zugleich als verkaufbares Objekt eignet.

Solche Begründungsfiguren eines geistigen Eigentums sind im 19. und 20. Jahrhundert auf den Feldern literarischer, journalistischer und wissenschaftlicher Textproduktion in jeweils spezifischen Regelsystemen wirkungsmächtig geworden. Sie stehen bis heute im Urheberrecht zentral, das im ersten Paragraphen u.a. »Urheber von Werken der [...] Wissenschaft«¹⁶ schützt, Urheber als »Schöpfer des Werkes« definiert und sogar im Falle einer kollektiven Autorschaft den Miturhebern weitreichende Rechte einräumt, so sind z.B. »Änderungen des Werkes [...] nur mit Einwilligung der Miturheber zulässig.«¹⁷ Dieser starke Schutz abgeschlossener wissenschaftlicher Werke und die Zentrierung auf den wissenschaftlichen Autor bzw. die Autorin korrespondieren mit der bis heute wirkungsmächtigen Vorstellung wissenschaftlicher Autorschaft, die um 1800 das »Vorläuferparadigma des ›Gelehrten‹«¹⁸ abgelöst hat, wie Felix Steiner zeigt. Während der Gelehrte vor allem überliefertes Wissen sammelte und ausbreitete, ist die Aufgabe des wissenschaftlichen Autors die Konstruktion von Wissen aus spezialisierten Forschungsarbeiten. Dieser »Paradigmenwechsel« wirkt sich laut Steiner »auf die gedankliche Architektur von wissenschaftlichen Texten entscheidend aus«: Die Wissenschaft präferiere seitdem eine »autorzentrierte[] Textordnung«, in der eine Autorinstanz – konstruiert aus Autorname, Deixis und Autorstil – als »basale[] Bedingung für die Kontextualisierung von Aussagen«¹⁹ unabdingbar ist.

Steiner zeigt zugleich, dass diese wissenschaftliche Autorinstanz keine feste unteilbare Einheit ist, indem er zwischen dem wissenschaftlichen Autor als einer sozialen Figur bzw.

¹³ Fichte 1964, S. 410.

¹⁴ Fichte 1964, S. 411.

¹⁵ Fichte 1964, S. 412.

¹⁶ UrhG § 1.

¹⁷ UrhG § 1, 7, 8 (1 & 2).

¹⁸ Steiner 2009, S. 264.

¹⁹ Steiner 2009, S. 264f.

Verantworterfigur, als innerer Figur mit mentalen Zuständen sowie als Textfigur differenziert, woraus sich sehr unterschiedliche Verknüpfungen von Aussage und Person ergeben können – so ist es beispielsweise trotz aller Autorzentrierung im deutschsprachigen Raum eine neuralgische Frage, ob man in wissenschaftlichen Texten überhaupt ›Ich‹ sagen dürfe.²⁰ Foucaults Beschreibung einer »Ego-Pluralität«²¹ von Autorschaft ließe sich allerdings am Beispiel zahlreicher geisteswissenschaftlicher Texte belegen, die zudem eine feine Balance zwischen präzisen Verweisen auf bekanntes Wissen sowie Momenten ›originärer‹ und ›neuer‹ Argumentation und Erkenntnisproduktion herstellen müssen. Als Basis dienen jeweils bestehende und ggf. modifizierte wissenschaftliche Methoden und Theorien, Ratschläge von Kollegen werden üblicherweise in Fußnoten verschoben. Diese gängigen Praktiken korrespondieren mit wissenschaftlichen Institutionen, die ›Dauerstellen‹ ad personam vergeben und sich in Bewerbungsverfahren zur Erlangung einer Professur in der Regel auf die beiden Qualifikationsschriften – Monografien mit einer individualisierten Autorschaft – und ausgewählte (Einzel-)Aufsätze konzentrieren.

3. Digitale Publikationen und Autorschaften. Der wissenschaftliche Spezialdiskurs im digitalen Wandel

In wissenschaftlichen Veröffentlichungen, die die Potenziale des digitalen Publizierens und Kommunizierens intensiv nutzen, wird jedoch diese Vorstellung eines auf ein Individuum rückführbaren Schöpfungsaktes, der sich in einem klar abgrenzbaren Werk verfestigt und verstetigt, problematisch. Die Medientheorie hat schon vor der globalen Etablierung des World Wide Web formuliert, dass – um nach McLuhan nun auch Vilém Flusser zu zitieren – ein ›Schreiben ins elektromagnetische Feld‹ den abgeschlossenen Werkcharakter von Büchern auflöse:

»Der Text ist nicht mehr, wie auf dem Papier, das Resultat eines kreativen Prozesses, sondern er ist selbst dieser Prozeß, er ist selbst ein Prozessieren von Informationen zu neuen Informationen.«²²

Auch die Buchwissenschaft weiß schon lange, dass sich digitale Texte von den statischen und geschlossenen Materialitäten des gedruckten Buchs unterscheiden, denn es werde, so Ursula Rautenberg 2001, »die über Buchblock und Einband definierte abgeschlossene Buch- und Texteinheit bei Datenbanken und in Hypertextstrukturen obsolet.«²³ Überhaupt sei auch die Rede vom »eBook« nur »metaphorisch« zu verstehen, folglich setzt Rautenberg den Begriff ›Buch‹ selbst in einfache Anführungszeichen.

Doch nicht nur die Materialität des Textes und seiner Formate ist nachhaltig erschüttert, sondern auch die Produktion der Texte sowie die wissenschaftliche Autorschaft selbst.

²⁰ Steiner 2009, S. 265f.

²¹ Foucault 1988, S. 22.

²² Flusser 1999, S. 63

²³ Rautenberg 2001, S. 10.

Schon vor einer Dekade beschrieben Medien-, Literatur- und Informationswissenschaftler wie Hans-Jürgen Bucher, Christiane Heibach, Rainer Kuhlen und Roberto Simanowski die Effekte dieser unabgeschlossenen, prozessualen Textproduktion für die literarische bzw. wissenschaftliche Autorschaft: Fortan werde es zunehmend um eine »Assoziation der Autoren«²⁴, »[k]ollaboratives Schreiben«²⁵, »die Anerkennung kollektiver Arbeit« und »Prozesssteuerung und -dynamik«²⁶ gehen, als eine Folge werde »die Trennung von Autor und Publikum« aufgehoben und es würden »die Grenzen zwischen den Kommunikationsbereichen von Medien, Politik, Wirtschaft, Wissenschaft und Privatsphäre in einem bislang nicht gekannten Ausmaß durchlässig.«²⁷ Programmatisch hat Lev Manovich 2005 Foucaults berühmte Schrift *Qu'est-ce qu'un auteur?*²⁸ unter dem Titel *Wer ist der Autor? Sampling/Remixen/Open Source* noch einen Schritt weiter gedacht und die These unterstützt, dass die »neue Medienkultur [...] eine Reihe neuer Modelle von Autorschaft mit sich [bringt], die alle unterschiedliche Formen der Zusammenarbeit beinhalten.«²⁹ In seinem Aufsatz macht Manovich allerdings eine relativ offene und noch unsortierte Differenzierung verschiedener Formen des kollaborativen Schreibens auf, dazu zählt er u.a. Kollaborationen verschiedener Autoren oder Gruppen sowie zwischen Firmen und Usern oder Autoren und Software; Lizenzierungsverfahren wie Open Source-Modelle; intertextuelle Schreibverfahren wie Remix oder Sampling.³⁰ Damit zeigt Manovich zugleich, dass die kollaborativen Potenziale digitaler Schreibverfahren sowohl eine grundsätzliche Reflexion von Kategorien wie ›Autor‹ oder ›Werk‹ notwendig machen als auch die Nutzung neuer rechtlicher (Lizenz-)Modelle.

Nun hat sich in der letzten Dekade der Wandel von gedruckten zu digitalen geisteswissenschaftlichen Veröffentlichungspraxen noch nicht umfassend vollzogen, weiterhin befinden sich die Geisteswissenschaften in einer medialen Transformation, die von Ungleichzeitigkeiten und widersprüchlichen Entwicklungen geprägt ist. Allerdings hat die Weiterentwicklung digitaler Publikationsformen sowie ihre experimentelle und breitere Nutzung und Reflexion eine differenziertere Perspektive der Medienwissenschaft auf die Veränderungen des Publizierens und der Autorschaft in der Wissenschaft ermöglicht. Im deutschsprachigen Raum wertet Mercedes Bunz diesen Wandel als eine *stille Revolution*. Bunz konstatiert insbesondere eine noch nicht gekannte kontinuierliche und schnellere Aktualisierung sowie eine Entindividualisierung der Erkenntnisprozesse in der Wissenschaft.³¹ Mit der Digitalisierung sei »eine zweite Phase der Aufklärung und Emanzipation ein[ge]treten, in der sich nach dem Individuum nun auch die Masse unter dem Diktum des kantschen ›Sapere aude!‹ formiert.«³² Kathleen Fitzpatrick beschreibt den Paradigmenwechsel, den das digitale Publizieren in der Wissenschaft mit sich bringt, mit den Schlagworten »From Product to Process«, »From Individual to Collaborative«, »From Originality to Remix«, »From Intellectual Property to the Gift Economy«, »From Text to ... Something More«.³³ Wenn man Manovich, Bunz

²⁴ Simanowski 2002, S. 46–53.

²⁵ Kuhlen 2004, S. 216.

²⁶ Heibach 2003, S. 273.

²⁷ Bucher 2004, S. 134.

²⁸ Foucault 1988.

²⁹ Manovich 2005, S. 7.

³⁰ Vgl. Manovich 2005, *passim*.

³¹ Vgl. Bunz 2012, S. 61, 114f., 125.

³² Bunz 2012, S. 133.

³³ Fitzpatrick 2011, S. 66–88.

und Fitzpatrick folgt und einen umfassenden digitalen Wandel der Geisteswissenschaften und ihrer Veröffentlichungskulturen konstatiert, müssen diese jedoch anhand konkreter Beispiele genauer auf ihre Potenziale und Probleme hin untersucht werden.

4. Open Review-Plattformen, Wikis, Wissenschaftsblogs und Soziale Medien. Eine kleine Typologie des kollaborativen Schreibens und Bewertens in der digitalen Geisteswissenschaft

Das Feld digitaler Veröffentlichungen in den Geisteswissenschaften ist vielfältig, es interessieren hier jedoch weniger wissenschaftliche Formate wie Monografien, Aufsätze oder Rezensionen, die zwar in digitalisierter Form vorgelegt werden, sich aber strukturell nicht von einer gedruckten Veröffentlichungsweise unterscheiden. In den Fokus rücken vielmehr Beispiele von digitalen Veröffentlichungspraxen, die in dieser Form nicht in eine gedruckte Variante rückübersetzbar wären und somit die Potenziale des digitalen Veröffentlichens intensiver nutzen. Hier sind vor allem zwei Bereiche interessant: erstens das kollaborative Schreiben und Bewerten, beispielsweise auf Plagiatsdokumentationsplattformen (mit teilweise anonymer Autorschaft) oder auf Open Review-Plattformen; sowie zweitens die digitale Wissenschaftskommunikation, und hier vor allem Wissenschaftsblogs und Soziale Medien. Diese Bereiche sollen im Folgenden näher untersucht werden, da sie in unterschiedlicher Weise den Wandel der Kategorien wissenschaftlicher Autorschaft, wissenschaftlicher Werke und der Geltungskraft des geistiges Eigentums repräsentieren.

4.1. Kollaboratives wissenschaftliches Schreiben und Bewerten: Open Review-Plattformen und Wikis

Exemplarisch für das kollaborative Schreiben und Bewerten in der Wissenschaft stehen einerseits Open Review-Plattformen, die transparente Formen eines akademischen Social Reading und Kommentierens zulassen, sowie andererseits Wikis, die die kollaborative Produktion und Kontrolle von Texten sowie die Klärung von Spezialproblemen ermöglichen. In diesen beiden digitalen Formaten wird, allerdings in unterschiedlicher Weise und Intensität, eine individuelle Autorschaft überstiegen und eine gemeinsame und transparente Produktion wissenschaftlicher Erkenntnisse medial unterstützt, die zudem der Grundannahme des Urheberrechts, wissenschaftliche Werke seien individuelle und schützenswerte Schöpfungen, widerspricht.

Kathleen Fitzpatrick hat ihr Buch *Planned Obsolescence. Publishing, Technology, and the Future of the Academy* in einem Open Review-Experiment kommentierbar gemacht, das bis heute einsehbar ist und auf der Kombination eines WordPress-Rahmens mit einem CommentPress-Plugin auf dem MediaCommons-Server basiert.³⁴ Ihre Erfahrungen mit dem Open Review

³⁴Vgl. Fitzpatrick: *Planned Obsolescence* [online] sowie Fitzpatrick 2011, S. 188.

bewertet die Autorin verhalten positiv: »This open review wasn't a perfect process, but it was illuminating. Forty-four unique commenters left a total of 295 comments, producing a much wider range of opinion and critique than any traditional review process could.«³⁵ Diese knapp 300 Kommentare von 44 Kommentatoren kann man zusätzlich neben die in den »Acknowledgements« genannten Unterstützer und Beiträger stellen, dazu zählen fünf Mitarbeiter des Institute for the Future of the Book, drei Mitarbeiter der New York University Press, schließlich noch ein »range of colleagues [who] made crucial contributions to this project, via Twitter discussions, conference observations, lecture invitations, and general producing, including [...]«³⁶ (hier zählt Fitzpatrick namentlich 28 Personen auf) und zahlreiche weitere Helfer. Auf diese Weise lässt sich das Ausmaß der kollaborativen Produktion im Rahmen einer Monographie transparenter beschreiben und zugleich die Frage stellen, inwiefern der individuelle Autorname »Kathleen Fitzpatrick« die alleinige Zuschreibung von *Planned Obsolescence* verdient. Diese zentrale Frage, wem überhaupt in welchem Maße die alleinige oder kollaborative Autorschaft eines Textes zugeschrieben werden kann, wird inzwischen auch im rechtswissenschaftlichen Spezialdiskurs grundsätzlicher diskutiert. Lionel Bently und Laura Biron sehen auch jenseits des digitalen Schreibens eine »disparity between legal and scientific authorship«³⁷ und fordern für die Zukunft »a more reflexive concept of authorship for attribution«³⁸ ein.



Abb. 1: Die Expanded Open Access-Edition des Buches *Debates in the Digital Humanities*. Auszug aus: Matthew K. Gold: *The Digital Humanities Moment*.

Neben dieser (in Maßen) kollaborativen Produktion des Textes ist das offene Bewertungsverfahren jedoch auch für die Rezeption des Textes relevant, ermöglicht Fitzpatrick doch eine – allerdings inzwischen nicht mehr zur Partizipation geöffnete – Form des Social Reading, indem der Leser bzw. die Leserin zwischen der Lektüre des Textes und der Hinzuziehung externer Kommentare wählen und somit eine Art »kommentiertes Lesen« vollziehen kann. Auffällig ist zudem die hybride Form: Während der finale Text seit 2011 in einer gedruckten Version von der New York University Press veröffentlicht wird, die sich alle Rechte vorbehält,³⁹ ermöglicht die Online-Version ihre Veröffentlichung »in part or in whole on condition that (1) distributed text is not sold, whether or not such sale is »for profit« and (2) distributed text bears this notice in full.«⁴⁰

³⁵ Fitzpatrick 2011, S. 189.

³⁶ Vgl. Fitzpatrick 2011, S. vii–viii.

³⁷ Bently / Biron 2014, S. 242.

³⁸ Bently / Biron 2014, S. 267.

³⁹ Fitzpatrick 2011, S. iv.

⁴⁰ Vgl. Fitzpatrick: *Planned Obsolescence* [online].

Ein anderes Beispiel für eine solche Plattform des akademischen Social Reading und Kommentierens bzw. Bewertens ist die Open Access-Edition des Sammelbandes *Debates in the Digital Humanities*,⁴¹ die sich zugleich als Diskussionsplattform der Digital Humanities begreift und verschiedene Tools zur Markierung, Bewertung und Kritik der jeweiligen Texte bereitstellt, also ein wirklich soziales Lesen und Kommentieren ermöglicht.⁴²



Abb. 2: »6 Comments on Paragraph 3«. Auszug aus der Online-Version von: Kathleen Fitzpatrick: *Planned Obsolescence. Publishing, Technology, and the Future of the Academy*.

Während es sich bei diesen Plattformen allerdings um die Lektüre, Begutachtung oder Kommentierung bekannter wissenschaftlicher Gattungen – wie Monographie, Sammelband, Aufsatz oder Miszelle – handelt, können Wikis als eine neue Form des digitalen kollaborativen Schreibens gewertet werden. Es kann hier nicht die Kontroverse um die Qualität des bekanntesten Wikis, der Wikipedia, fortgeführt werden, vielmehr soll ein anderes Beispiel wissenschaftlicher Kollaboration der letzten Jahre näher betrachtet werden, das gesellschaftlich breit und kontrovers diskutiert wurde und zugleich wissenschaftspolitische Effekte hatte, die die bestehenden wissenschaftlichen Institutionen nicht oder womöglich erst mit einer deutlichen zeitlichen Verzögerung realisiert hätten: es geht um Wikis, die als Plagiatdokumentationsplattformen genutzt werden, in diesem Fall um das **GuttenPlag Wiki**.

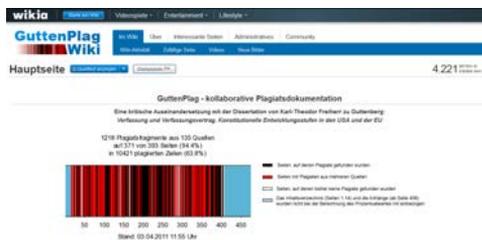


Abb. 3 : GuttenPlag Wiki: Ergebnisse der Crowd-Arbeit [online]

Auf dieser Plattform (und in anderen Projekten wie **VroniPlag**) wurde *eine Form kollaborativen textanalytischen Arbeitens in digitalen Medien* realisiert, die in der Gutenberg-Galaxis nicht möglich gewesen wäre.⁴³ Das Vorgehen lässt sich grundsätzlich als

⁴¹Vgl. Gold 2012.

⁴²Vgl. Gold 2013 [online].

⁴³Eine eingeschränkte Variante dieses Verfahrens mit einer klar zuordenbaren, allerdings eindeutig kollaborativen Autorschaft stellt das Buch *10 PRINT CHR\$(205.5+RND(1)); : GOTO 10* dar: »Ten authors collaborated to write this book. Rather than produce a collection of ten separate articles, we chose a process of communal authorship. Most of the writing was done using a wiki, although this process differed significantly from the most famous wiki-based project, Wikipedia. Our book was not written in public and

wissenschaftlich bezeichnen, da es von einer Hypothese (»das wissenschaftliche Werk x ist ein Plagiat«) ausgeht und auf Basis eines kategorialen Apparats (der z.B. Plagiatstypen wie »Komplett-«, »verschleierte«, »Übersetzungsplagiate« oder »Bauernopfer« unterscheidet) sowie in einem Verfahren intensiver und kollektiv kontrollierter Textanalysen zu einer dokumentierten und begründeten Verifikation oder Falsifikation der Ausgangsthese kommt. Für unsere Fragestellung besonders relevant ist, dass diese Plattformen sich durch Elemente auszeichnen, die der wissenschaftlichen Buchkultur tendenziell oder diametral entgegen stehen:

- Die Autorschaft ist kollaborativ und teilweise anonym, die Arbeitsgruppe versteht sich als »Schwarm« und nicht als Summe von Individuen.
- An der Arbeit können nicht nur Wissenschaftler teilnehmen, die sich zuvor durch ihre wissenschaftliche Vita oder institutionelle Zugehörigkeiten dafür qualifiziert haben.
- Der Analyseprozess wird in großen Teilen direkt öffentlich gemacht, auch die Diskussionen der Beteiligten über das Projekt sind im Regelfall transparent.
- Die Ergebnisse der Arbeit sind im Sinne des Open Access bis heute einsehbar.
- Die Arbeitsergebnisse einer freiwilligen und interdisziplinären Forschergruppe zeigen fundamentale Probleme in wissenschaftlichen Arbeiten und Erkenntnisgängen auf, die der Wissenschaftsbetrieb zuvor nicht oder nur in eingeschränkter Form thematisiert hat.

Die Aberkennungsverfahren der Dokortitel von u.a. Jorgo Chatzimarkakis, Karl-Theodor zu Guttenberg, Silvana Koch-Mehrin oder Margarita Mathiopoulos,⁴⁴ die mindestens indirekt von den Wiki-Plattformen ausgingen oder gestützt wurden und eine breite mediale Aufmerksamkeit erhalten haben, müssen hier nicht noch einmal nacherzählt werden. Es zeigt sich jedenfalls in diesen mittelbaren Effekten der Wiki-Plattformen eine Qualität der spezifischen Wissensproduktion durch Wikis und ihre Crowds: »Die *Weisheit der vielen*«, so der Medienphilosoph Stefan Münker, »triumphiert im Web 2.0 über das *Wissen der Eliten*.« Dieses kollaborative und teilweise auch anonyme Verfahren zur Wissensproduktion mache sichtbar, dass »unser[] expertenbasierte[r] Wissensbegriff« nur eine »Ideologie von Eliten« sei, die vor allem in den letzten 200 Jahren die Wissensproduktion mit einer individualisierten Autorfunktion verbunden hätten: »Wissen ist nämlich immer schon ein Resultat kollaborativer Prozesse der Informationssammlung und -verarbeitung – und war auch in der Zeit vor der digitalen Vernetzung keineswegs schlicht das Produkt der schärferen Erkenntnisfähigkeit einzelner.«⁴⁵

was not editable by the public. We benefited from comments by reviewers, and from discussions with others at conferences and in other contexts; still, the text of the book was developed by the ten of us, working together as one, and we bear the responsibility for what this book expresses.« (Montfort et al. 2013, S. 7) Das Buch wurde von der MIT Press veröffentlicht und steht zugleich frei zum [Download](#) zur Verfügung.

⁴⁴In dieser Reihung fehlt bewusst Annette Schavan: In ihrem Fall hatte eine knappe Mehrheit des VroniPlag-Schwarms entschieden, dass ihre Dissertation zwar Zitatfehler aufweise, diese jedoch nicht so gravierend seien, dass die Arbeit auch in ihrer Gesamtheit als Plagiat zu bewerten wäre. Im Gegensatz zu diesem Votum positionierte sich in einem Spiegel-Interview der Einzelbetreiber des Weblogs schavanplag, »eine Person, die sich in Faxen »Robert Schmidt« nennt, aber anders heißt. Sie ist seit längerem Mitglied im Recherchenetzwerk VroniPlag, aber nicht einmal ihre VroniPlag-Mitstreiter kennen seinen oder ihren echten Namen.« (Titz 2012; Hervorh. beseitigt, T.E.) Mit »Robert Schmidt« machte somit ein anonymen Einzelautor die Vorwürfe mit dem Ziel öffentlich, »dass Leute mit einem akademischen Betrug nach Möglichkeit nicht durchkommen sollten.« (Titz 2012) Während die kollaborative Arbeitsgruppe von VroniPlag also vorsichtige Schlüsse zog und sich öffentlich zurückhielt, löste der Einzelautor »Schmidt« das – letztlich sogar erfolgreiche – Aberkennungsverfahren aus.

⁴⁵Alle Zitate Münker 2009, S. 100.

Es ist erwartbar, dass sich prominente Vertreter der Wissenschaft gegen diese tendenzielle Aufhebung einer klaren Abgrenzbarkeit von wissenschaftlichem Spezialdiskurs und wissenschaftlicher Expertenschaft einerseits gegenüber nicht-akademischen Zusammenhängen andererseits wehren. Dies zeigt sich beispielsweise in den Empfehlungen der Hochschulrektorenkonferenz von 2013 zu einer *Guten wissenschaftlichen Praxis an deutschen Hochschulen*. Darin fordert die HRK – in einer klaren Abgrenzung vom Vorgehen der Wiki-»Schwarmintelligenz« –, dass im Falle eines Plagiatsverdachts »der Hinweisgeber (Whistle Blower) [...] höchste[] Vertraulichkeit« bewahrt und sich folglich nicht »mit seinem Verdacht an die Öffentlichkeit wendet«, was selbst ein Verstoß »gegen die Regeln der guten wissenschaftlichen Praxis«⁴⁶ sei. Die Juristen Thomas Dreier und Ansgar Ohly konstatieren zwar, dass Plagiatsdokumentationsplattformen wie das GuttenPlag Wiki »zur Dokumentation von Plagiaten erheblich beigetragen« hätten, letztlich aber nicht als wertvolles Element der Wissenschaftskultur zu akzeptieren seien: »Selbst ernannte Plagiatsjäger sind keiner Instanz verantwortlich, regelmäßig nicht einmal namentlich bekannt. Ob jede als Plagiat gekennzeichnete Stelle wirklich ein solches ist, bleibt ungeprüft.« Außerdem könne »die medial wirksame Plagiatsjagd in eine Hetzjagd« umschlagen.⁴⁷

Ganz im Gegensatz dazu sehen Bernhard Pörksen und Hanna Detel die große Leistung des GuttenPlag Wiki und seines »höchst effektive[n] Crowdsourcing« in seiner Präsentation von »Evidenzerfahrungen für das große Publikum«. Diese Plattform sei ein Paradebeispiel für eine veränderte mediale Öffentlichkeit durch digitale Medien, die nicht beim Journalismus endet, sondern auch die Wissenschaft affiziert: »Die einst klar getrennten Informations- und Wissenswelten zwischen Informant, Journalist und Gatekeeper beginnen zu changieren.«⁴⁸ Für eine solchermaßen veränderte digitale Öffentlichkeit haben PlagDoc und Martin Kotynek, die beiden Gründer des 2011 dann auch mit einem Grimme Online-Award nobilitierten GuttenPlag Wiki, im Anschluss an ihr Projekt *Empfehlungen* formuliert. Sie bezeichnen die Arbeit ihrer Community als »investigatives Crowdsourcing«,⁴⁹ das sie als ein geordnetes, methodisch begründetes und organisiertes wissenschaftliches Vorgehen darstellen. Entscheidend für den Erfolg sei erstens die technische Fundierung des Projekts gewesen (z.B. der Wechsel von GoogleDocs zu Wikia; besser sei es vielleicht, nicht-institutionell gebundene eigene Server zu nutzen) sowie zweitens eine gute und stringente Moderation der Gruppe nach klaren und transparenten Regeln. Schließlich verteidigen sie gerade die für das Wissenschaftssystem der Gutenberg-Galaxis irritierende Möglichkeit, anonym oder mit einem Pseudonym an der kollaborativen Arbeit zu partizipieren, denn »Anonymität ermöglicht eine Teilnahme ohne Hürden; Pseudonyme verbinden sich mit einer vorhandenen Identität im Internet, die anderen netzaffinen Menschen womöglich schon bekannt ist und rasch Vertrauen schafft.« Die Entscheidung von GuttenPlag, »keine Klarnamen zu verlangen, [...] war [...] im Nachhinein betrachtet essentiell.«⁵⁰ Außerdem gehe es darum, nachvollziehbar zu arbeiten und offen zu kommunizieren, weshalb alle Wiki-Seiten unter der Creative Commons Attribution-Share Alike 3.0 Unported (CC-BY-SA 3.0)-Lizenz frei verfügbar sind.

⁴⁶HRK 2013, S. 3 [online].

⁴⁷Alle Zitate Dreier / Ohly 2013, S. 166.

⁴⁸Alle Zitate Pörksen / Detel 2011, S. 61–63.

⁴⁹PlagDoc / Kotynek 2012, S. 2 [online].

⁵⁰PlagDoc / Kotynek 2012, S. 7 [online].

4.2. Digitale Wissenschaftskommunikation: Wissenschaftsblogs und Soziale Medien

Digitale Formate wie Weblogs und Soziale Medien wie Twitter und Facebook ermöglichen eine schnellere, interaktive, populäre und meinungsstarke Kommunikation in einer hohen Intensität und werden natürlich auch von Wissenschaftlern genutzt. Weblogs mit kurzen Postings und Tweets mit maximal 140 Zeichen sowie der meist vorhandenen Möglichkeit zur Kommentierung (unabhängig vom Bildungs- und Sozialstatus des Kommentierenden) stellen das Selbstverständnis der Wissenschaft – und anderer Schreibberufe wie z.B. des Journalismus – vor fundamentale Probleme. Kathrin Passig hat beispielsweise für den Journalismus eine Aufweichung der Grenzen zwischen Profis und Laien beschrieben, denn die Blogosphäre und die Sozialen Medien ersetzen »Durchsage« durch »Dialog«, weshalb die »Feedbackkanäle, die das Internet eröffnet, [...] vielfach geradezu mit Stolz ignoriert«⁵¹ würden.

Postings von Wissenschaftlern in Wissenschaftsblogs und in Sozialen Medien kann man im Regelfall nicht als *wissenschaftliche* Veröffentlichungen rubrizieren, es handelt sich eher um Kommunikation zwischen Wissenschaftlern oder aber um die popularisierte Bereitstellung wissenschaftlichen Wissens – auch für Zielgruppen jenseits der akademischen Institutionen. Dabei lässt sich allerdings die Modifikation von semi- oder populärwissenschaftlichen Textsorten, die schon in der Gutenberg-Galaxis präsent waren (wie Konferenzbericht, Essay oder Interview), beobachten, die jedoch durch die multimedialen Möglichkeiten und die viel größere Reichweite des Netzes eine wesentlich wichtigere Funktion in der Wissenschaftskommunikation erhalten.

Ein repräsentatives Beispiel für ein Portal verschiedener Wissenschaftsblogs ist de.hypotheses.org, dessen Redaktion

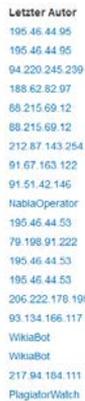
»aus der Gesamtheit der deutschsprachigen Blogs die besten Artikel für die Startseite des Portals aus[wählt]. Die Redaktion unterhält außerdem ein eigenes Blog, das Redaktionsblog, in dem sie Hinweise rund um das Bloggen gibt (z.B. Urheberrecht, Gestaltung eines Impressums etc.) sowie über Neues aus den jeweiligen Interessensbereichen der Redaktionsmitglieder berichtet«⁵²

hypotheses.org hat u.a. französische, deutsche und spanische Sektionen, in denen Wissenschaftsblogs kostenlos geführt werden können. Die für die Blogosphäre zentrale persönliche Vernetzung realisiert sich somit auf verschiedenen Levels, indem einzelne Weblogs über das Portal gehostet sowie ausgewählte Postings auf dem Portal selbst abgebildet und diskutiert werden und natürlich auch frei einsehbar und im Regelfall kommentierbar sind, auch für Leser-Autoren jenseits der wissenschaftlichen Blogosphäre.

⁵¹Alle Zitate Passig 2013, S. 72.

⁵²de.hypotheses.org. Wissenschaftliche Blogs: Redaktion [online].

Eines der deutschen hypotheses-Blogs ist das Wissenschaftsblog *Rezensieren – Kommentieren – Bloggen: Wie kommunizieren Geisteswissenschaftler in der digitalen Zukunft?*, das zur gleichnamigen Konferenz eingerichtet wurde, die Ende Januar 2013 in München stattfand. Das Blog ermöglichte den Teilnehmern bereits vor der Konferenz eigene Positionierungen zum Thema vorzunehmen und zu diskutieren sowie nach der Konferenz die Diskussion fortzuführen und Hinweise auf Medien und Beiträge zur Konferenz verfügbar zu machen. Insgesamt finden sich hier 44 Postings, die zwischen Juli 2012 und Mai 2014 freigeschaltet wurden. Exemplarisch sei hier auf das **Posting** von Fabian Steiner und Max Trecker vom 17. Februar 2014 verwiesen, das sich als eine Zusammenfassung der nachfolgenden Podiumsdiskussion *Nachwuchswissenschaftler, Verlage, Bibliotheken & Open Access. Zeitgemäßes Publizieren in den Geisteswissenschaften* vom 11. Februar 2014 versteht, wobei hier weniger der Inhalt als vielmehr die mediale Form interessiert: Das Posting verweist in 24 Hyperlinks über sich hinaus auf andere Medien, Personen, Informationen; es integriert insgesamt zehn Tweets mit teilweise starken Meinungsäußerungen (und ruft dazu Autorennamen auf wie [@hkohle](#), [@mdemanto](#), [@Erblogtges](#), [@naponaps](#), [@swimttt](#), [@SabineScherz](#)); und es lässt schließlich sechs Kommentare zu, die teilweise ebenfalls auf ergänzenden Content verweisen (einen englischsprachigen Tagungsbericht, ein Vortragsvideo, ein Tweet-Protokoll bei Storify) und in einem Fall tatsächlich einen Leser-Autor-Kommentar zur Podiumsdiskussion abbilden (Abbildung 4).



```
Letzter Autor
195.46.44.95
195.46.44.95
94.220.245.239
188.62.82.97
88.215.69.12
88.215.69.12
212.87.143.264
91.67.163.122
91.51.42.146
NabiaOperator
195.46.44.53
79.198.91.222
195.46.44.53
195.46.44.53
206.232.179.195
93.134.166.117
WikiaBot
WikiaBot
217.94.184.111
Plagiator/Welch
```

Abb. 4 : Fabian Steiner / Max Trecker: Zeitgemäß publizieren? **Blogposting** vom 17.2.2014 zur Konferenz *Rezensieren - Kommentieren – Bloggen: Wie kommunizieren Geisteswissenschaftler in der digitalen Zukunft?*

Die Texte stehen unter einer besonders offenen Creative Commons Lizenz: »All texts are published under a CC BY 3.0 license.«⁵³ Diese Lizenz stellt Textnutzern nur die Bedingung, den Namen des ursprünglichen Autors oder der Autorin zu nennen und keine neuen rechtlichen Klauseln an den Text zu binden, ansonsten darf der Nutzer bzw. die Nutzerin »das Material in jedwedem Format oder Medium vervielfältigen und weiterverbreiten« sowie »das Material remixen, verändern und darauf aufbauen und zwar für beliebige Zwecke, sogar kommerziell.«⁵⁴ Die zentrale Leistung des Urheberrechtsschutzes, gerade die Einheit eines Werkes und seinem

⁵³ Steiner / Trecker 2014 [online].

⁵⁴ Creative Commons 2014 [online].

Autor bzw. seiner Autorin die alleinige Verfügung über die Verwertung zu garantieren, wird auf diese Weise aufgehoben. Dieses Vorgehen erlaubt zugleich erst jene kollaborativen und interaktiven Formen der digitalen Wissenschaftskommunikation, die die Potenziale des digitalen Publizierens produktiv machen.

Solche Formen einer kollaborativen und interaktiven Vernetzung in der digitalen Wissenschaftskommunikation intensivieren sich in Sozialen Medien wie Twitter oder Facebook. Auf dem Microbloggingdienst Twitter hat beispielsweise die Konferenz *Digital Humanities – methodischer Brückenschlag oder ›feindliche Übernahme‹? Chancen und Risiken der Begegnung zwischen Geisteswissenschaften und Informatik* (März 2014), von der auch dieser Beitrag ausgeht, zahlreiche Tweets verschiedener Teilnehmer angeregt. Unter dem Hashtag #DHd2014 lassen sich diese Tweets retrospektiv anordnen, wobei noch unklar ist, ob sich die unter einem solchen Hashtag versammelten Tweets als ein hybrider wissenschaftlicher Text mit dissoziierter Autorschaft beschreiben ließen.

Die Forscher setzen dabei ihre Postings und Tweets direkten Kommentaren oder verschiedenen Formen des sogenannten ›User Generated Content‹ aus, die nicht zwangsläufig aus der Wissenschaft kommen müssen – wobei dies in den oben genannten Beispielen noch immer der Regelfall ist. Weblogs und Postings von Wissenschaftlern in Sozialen Medien, die auch Reaktionen aus nicht-akademischen Kontexten erhalten, stecken jedoch in einem Dilemma: Je erfolgreicher sie sind, desto stärker verlieren sie ihre spezialdiskursive Eigenheit. Für Wissenschaftsblogger oder twitternde Kollegen lässt sich dieses Dilemma entweder auffangen, indem sie die populären Potenziale der digitalen Wissenschaftskommunikation nicht nutzen und weiterhin im fachspezifischen Duktus bleiben, oder indem sie die Sozialen Medien als Verweissystem auf die eigentlichen wissenschaftlichen Werke instrumentalisieren.

5. Vom Urheber zur Crowd, vom Werk zur Version, vom Schutz zur Öffnung? Probleme und Potenziale des digitalen Publizierens in der Wissenschaft

Die geisteswissenschaftliche Arbeit lässt sich als eine gemeinsame Arbeit an Texten und an der Verbesserung von Methoden und Erkenntnissen beschreiben, dieses Vorgehen wird in digitalen Publikationen noch stärker als in der Vergangenheit auch in einzelnen Textformaten, ihrer Produktion und ihrer Rezeption abgebildet. Die Beispiele, denen sich dieser Beitrag gewidmet hat, zeigen vor allem widersprüchliche Ergebnisse: Weblogs und Soziale Medien werden auch von Geisteswissenschaftlern zur Wissenschaftskommunikation genutzt, dabei greifen die Nutzer jedoch vor allem auf kommerzielle Anbieter wie Twitter und Wikia zurück. Die intermediale und vor allem interaktive Qualität dieser Medien befördert eine wissenschaftliche Debattenkultur, an der sich die Beteiligten teilweise mit Pseudonymen beteiligen. Open Review-Plattformen treten in unseren beiden Beispielen als eine Form des hybriden Publizierens auf, die – mal vor der Buchveröffentlichung, mal nach der Buchveröffentlichung – nutzbare Social Reading- und Kommentier-Version wird auch als urheberrechtlich geschütztes Buch veröffentlicht. Die weitreichendsten Neuerungen und Effekte haben bislang die Wikis, in diesem Fall die kollaborativen

Plagiatsdokumentationsplattformen, erzielt, die allerdings von führenden Vertretern der Wissenschaftsorganisationen problematisiert werden.

Tatsächlich bringt die Transformationsphase des digitalen Wandels zahlreiche Probleme für die Veröfentlichungskultur in den Geisteswissenschaften mit sich, die noch ungeklärt sind, aber immer präziser beschrieben werden können. Dazu zählen die folgenden Punkte:

- **Das Problem der Wissenschaftlichkeit der neuen digitalen Textformate:** Ab wann kann man einen digitalen Text wissenschaftlich nennen? Ab wann ist ein Blogposting oder eine Summe von Tweets zu einem bestimmten Hashtag ein wissenschaftlicher Text? Ist es eine Fehleinschätzung oder eine zutreffende Prognose, wenn Philippe Wampfler konstatiert: »Die Vorstellung, ein Blog oder ein Wiki könnten zur Verleihung eines Dokortitels oder der Vergabe einer *venia legendi* führen, mutet absurd an, obwohl solche digitalen Publikationen sämtlichen Anforderungen an eine Dissertation oder Habilitation genügen können.«⁵⁵ Dieser Beitrag stärkt eine Differenz zwischen wissenschaftlichen Wikis und Open Review-Projekten einerseits sowie eher nicht-wissenschaftlichen Weblogs und Sozialen Medien andererseits, wobei Letztere allerdings die Wissenschaftskommunikation nachhaltig verbessern.
- **Das Problem der wissenschaftlichen Autorschaft und ihrer Leistungsbewertung:** Die geisteswissenschaftlichen Institutionen präferieren die eindeutige und im Regelfall singuläre Autorschaftszuschreibung zu einem klar abgrenzbaren Werk, die zugleich die Vergabe eines Impactfaktors an eine Veröffentlichung und somit an ihren Autor oder ihre Autorin ermöglicht. Diese Formen der Autorschaftszuschreibung und der Leistungsmessung erweisen sich in kollaborativen und versionierten Schreibprozessen jedoch als problematisch. Daher diskutiert beispielsweise momentan die **AG Digitales Publizieren** im Verband Digital Humanities im deutschsprachigen Raum sowohl die Differenzierung und Konturierung verschiedener Rollen im wissenschaftlichen Schreibprozess (von Haupt-Autorin und abgestuften Ko-Autoren über Herausgeberin und Gestalter bis zu Rechercheur und Kommentatorin etc.) als auch die Differenzierung von qualitativen und quantitativen Bewertungselementen, die einer spezifischen Version einer Einzelveröffentlichung und ggf. auch nur einer Mitarbeiterrolle zugeschrieben werden können.
- **Die Problematisierung des wissenschaftlichen Subjekts durch Crowdsourcing und den Experten-Laien-Dialog:** Beispiele wie das Open Review-Projekt *Planned Obsolescence* oder Hashtags zu Wissenschaftsthemen in Sozialen Medien wie Twitter legen offen, dass die klare Abgrenzung einer individuellen wissenschaftlichen Autorschaft in einem kollaborativ produzierten Text manchmal kaum noch zu bewerkstelligen oder sogar nicht mehr sinnvoll sein kann. Zugleich zeigt das GuttenPlag Wiki, dass die (vermutete) Teilnahme von Nicht-Akademikern an wissenschaftlichen Erkenntnisprozessen produktiv sein kann, wenn sich diese an transparenten und kontrollierten Regeln und Verfahrensweisen orientieren.
- **Das Problem der Veröffentlichungsinfrastruktur:** Twitter ist die kommerzielle Plattform eines börsennotierten Unternehmens, Wordpress und Wiki sind jeweils freie Software und die Bücher des Open Reviews sind zugleich gedruckt bei Verlagen

⁵⁵ Wampfler 2013, S. 92.

erschieden, die immerhin einen kleinen Erlös dafür erhalten. Es fehlen noch immer Vereinbarungen einer guten wissenschaftlichen Praxis, in welchem Maße und für welche Zwecke wissenschaftliche Veröffentlichungsplattformen bei eigenen, staatlichen oder kommerziellen Anbieter genutzt werden sollten.

- **Das Problem des Immaterialgüterrechts:** Die beschriebenen Beispiele haben allesamt – in unterschiedlicher Weise – den Schutz der jeweiligen Texte aufgehoben und lassen teilweise sogar die Bearbeitung und Neuveröffentlichung zu – auch innerhalb des bestehenden Urheberrechts können u.a. offene CC-Lizenzen vergeben werden. Sowohl in der Forschung als auch in der Lehre sorgen allerdings die Rechtsvorgaben des bestehenden Urheberrechts weiterhin für Probleme. Lawrence Lessig schreibt für diesen Punkt den Professoren eine eindeutige Verantwortung zu: »Die Kostspieligkeit, Unverständlichkeit und Willkür dieses Systems sprechen unserer Tradition Hohn. Juristen und Wissenschaftler sollten es als ihre Pflicht ansehen, die Art zu ändern, wie das Recht funktioniert, oder, besser noch, das Recht so zu ändern, dass es funktionieren kann.«⁵⁶

- **Das Problem der Wissenschaftsschranke und der digitalen Ungleichheit:** Da es momentan nicht nach einer umfassenden Änderung des Urheberrechts aussieht, arbeiten verschiedene Juristen, u.a. Katharina de la Durantaye, an der Frage, wie man für die Wissenschaft zusätzliche Sonderregelungen mit besonderen Zugriffsrechten auf Texte einführen könnte, die bereits unter dem Begriff der ›Wissenschaftsschranke‹ firmieren.⁵⁷ Zugleich würde sich durch weitere Sonderregelungen für die akademischen Institutionen jedoch auch die ›Digitale Ungleichheit‹, die eigentlich abgebaut werden sollte, verstärken.

Dank

Ich danke den Teilnehmern der DHD 2014-Konferenz [Digital Humanities – methodischer Brückenschlag oder ›feindliche Übernahme‹? Chancen und Risiken der Begegnung zwischen Geisteswissenschaften und Informatik](#) in Passau, die meinen Vortrag am 26. März 2014 kritisch diskutiert haben, den Mitgliedern der [AG Digitales Publizieren](#) im Verband Digital Humanities im deutschsprachigen Raum für wichtige Anregungen sowie der [Global Young Faculty III](#) für ihre besondere Unterstützung. Zudem danke ich Dr. Constanze Baum und Dr. Thomas Stäcker für kritische Lektüren des Textes und hilfreiche Hinweise.

Dieser Text sollte nicht abgeschlossen bleiben. Kritische Anmerkungen oder Ergänzungen zum Text sind sehr willkommen, damit er ggf. zu einem späteren Zeitpunkt überarbeitet oder aktualisiert werden kann: thomas.ernst@uni-due.de. Der Text kann [auf meinem Weblog](#) kritisch kommentiert werden.

⁵⁶ Lessig 2006, S. 297.

⁵⁷Vgl. de la Durantaye 2014, S. 191–264 [online].

Bibliographische Angaben

- Hans-Jürgen Bucher: Online-Interaktivität – Ein hybrider Begriff für eine hybride Kommunikationsform. In: Interaktivität. Ein transdisziplinärer Schlüsselbegriff. Hg. von Christoph Bieber / Claus Leggewie. Frankfurt/Main, New York 2004, S. 132–167. [\[Nachweis im GBV\]](#)
- Mercedes Bunz: Die stille Revolution. Wie Algorithmen Wissen, Arbeit, Öffentlichkeit und Politik verändern, ohne dabei viel Lärm zu machen. Berlin 2012. [\[Nachweis im GBV\]](#)
- Debates in the Digital Humanities. Hg. von Matthew K. Gold. Minneapolis 2012. [\[Nachweis im OPAC\]](#)
- Thomas Dreier / Ansgar Ohly: Lehren aus der Vergangenheit – Perspektiven für die Zukunft. In: Plagiate. Wissenschaftsethik und Recht. Hg. von Thomas Dreier / Ansgar Ohly. Tübingen 2013, S. 155–184. [\[Nachweis im GBV\]](#)
- Thomas Ernst: E-Books, Netzliteratur und das Urheberrecht. In: SDV – Sprache und Datenverarbeitung. International Journal for Language Data Processing 36 (2012), H. 2, S. 43–52. [\[Nachweis im GBV\]](#)
- Johann Gottlieb Fichte: Beweis der Unrechtmässigkeit des Büchernachdrucks. Ein Raisonement und eine Parabel. In: Johann Gottlieb Fichte: Gesamtausgabe der Bayerischen Akademie der Wissenschaften. Hg. von Reinhard Lauth / Hans Jacob. Band I.1.: Werke 1791–1794. Stuttgart 1964, S. 409–426. [\[Nachweis im OPAC\]](#)
- Kathleen Fitzpatrick: Planned Obsolescence. Publishing, Technology, and the Future of the Academy. New York, London 2011. [\[Nachweis im OPAC\]](#)
- Julia Flanders: The Productive Unease of 21st-century Digital Scholarship. In: Defining Digital Humanities. A Reader. Hg. von Melissa Terras / Julianne Nyhan / Edward Vanhoutte. Farnham 2013, S. 205–218. [\[Nachweis im OPAC\]](#)
- Vilém Flusser: Medienkultur. Hg. von Stefan Bollmann. 2. Auflage. Frankfurt/Main 1999. [\[Nachweis im OPAC\]](#)
- Michel Foucault: Was ist ein Autor? In: Michel Foucault: Schriften zur Literatur. Aus dem Französischen von Karin von Hofer / Anneliese Botond. Frankfurt/Main 1988, S. 7–31. [\[Nachweis im OPAC\]](#)
- Gérard Genette: Paratexte. Das Buch vom Beiwerk des Buches. Mit einem Vorwort von Harald Weinrich. Aus dem Französischen von Dieter Hornig. Frankfurt/Main 2001. [\[Nachweis im OPAC\]](#)
- Jeff Gomez: Print is Dead. Books in Our Digital Age. New York 2008. [\[Nachweis im OPAC\]](#)
- Valentin Groebner: Wissenschaftssprache digital. Die Zukunft von gestern. Konstanz 2014. [\[Nachweis im OPAC\]](#)
- Wolfgang Ehrhardt Heinold: Bücher und Büchermacher. Verlage in der Informationsgesellschaft. 5. Auflage. Heidelberg 2001. [\[Nachweis im OPAC\]](#)
- Christiane Heibach: Literatur im elektronischen Raum. Frankfurt/Main 2003. [\[Nachweis im GBV\]](#)
- Rainer Kuhlen: Kollaboratives Schreiben. In: Interaktivität. Ein transdisziplinärer Schlüsselbegriff. Hg. von Christoph Bieber / Claus Leggewie. Frankfurt/Main, New York 2004, S. 216–239. [\[online\]](#)
- Lawrence Lessig: Freie Kultur. Wesen und Zukunft der Kreativität. Aus dem Englischen von Annegret Claushues / Hartmut Pilch. München 2006. [\[Nachweis im OPAC\]](#)
- Jürgen Link / Ursula Link-Heer: Diskurs/Interdiskurs und Literaturanalyse. In: LiLi. Zeitschrift für Literaturwissenschaft und Linguistik 20 (1990), H. 77, S. 88–99. [\[Nachweis im OPAC\]](#)
- Lev Manovich: Wer ist der Autor? Sampling/Remixen/Open Source. In: Lev Manovich: Black Box – White Cube. Aus dem Amerikanischen von Ronald Voullié. Berlin 2005, S. 7–28. [\[Nachweis im GBV\]](#)
- Marshall McLuhan: Culture without Literacy. In: Explorations (1953), H. 1, S. 117–127. [\[Nachweis im GBV\]](#)
- Stefan Münker: Emergenz digitaler Öffentlichkeiten. Die Sozialen Medien im Web 2.0. Frankfurt/Main 2009. [\[Nachweis im GBV\]](#)
- Kathrin Passig: Standardsituationen der Technologiekritik. Berlin 2013. [\[Nachweis im GBV\]](#)
- Bernhard Pörksen / Hanne Detel: Evidenzerfahrungen für alle. Das kontraproduktive Krisenmanagement des Verteidigungsministers und die Logik der Skandalisierung im digitalen Zeitalter. In: Inszenierung als Beruf. Der Fall Guttenberg. Hg. von Oliver Lepsius / Reinhart Meyer-Kalkus. Berlin 2011, S. 56–69. [\[Nachweis im OPAC\]](#)
- Ursula Rautenberg: Die historisch-systematische Medienspezifik des Buches. In: Ursula Rautenberg / Dirk Wetzels: Buch. Tübingen 2001, S. 8–10. [\[Nachweis im OPAC\]](#)
- Roberto Simanowski: Interfictions. Vom Schreiben im Netz. Frankfurt/Main 2002. [\[Nachweis im GBV\]](#)
- Felix Steiner: Dargestellte Autorschaft. Autorkonzept und Autorsubjekt in wissenschaftlichen Texten. Tübingen 2009. [\[Nachweis im GBV\]](#)
- Horst Thomé: Werk. In: Reallexikon der deutschen Literaturwissenschaft. Neubearbeitung des Reallexikons der deutschen Literaturgeschichte. Hg. von Jan-Dirk Müller, gemeinsam mit Georg Braungart et al. Band III: P-Z. Berlin, New York 2003, S. 832–834. [\[Nachweis im GBV\]](#)

Philippe Wampfler: ›online first‹. Geisteswissenschaften als Social Media. In: Zürcher Jahrbuch für Wissenschaftsgeschichte. Digital Humanities. H. 9 (2013), S. 79–102. [\[Nachweis im OPAC\]](#)

Internetquellen

Lionel Bently / Laura Biron: Discontinuities between legal conceptions of authorship and social practices. What, if anything, is to be done? In: The Work of Authorship. Hg. von Mireille Van Eechoud. Amsterdam 2014, S. 237–276. [\[online\]](#)

Bundesministerium der Justiz und für Verbraucherschutz: Gesetz über Urheberrecht und verwandte Schutzrechte. §1: Allgemeines. [\[online\]](#)

Creative Commons: Namensnennung 3.0 Deutschland (CC BY 3.0 DE). [\[online\]](#)

de.hypotheses.org. Wissenschaftliche Blogs: Redaktion. [\[online\]](#)

Diverse Autoren: GuttenPlag Wiki. [\[online\]](#)

Diverse Autoren: Twitter-Hashtag #dhd2014. [\[online\]](#)

Diverse Autoren: VroniPlag Wiki. [\[online\]](#)

Katharina de la Durantaye: Allgemeine Bildungs- und Wissenschaftsschranke. Münster 2014. [\[online\]](#)

Thomas Ernst: Jenseits des wissenschaftlichen Werks und des geistigen Eigentums? Die digitale Verbreitung wissenschaftlichen Wissens. Audio-Podcast des Vortrags vom 26. März 2014 auf der Konferenz Digital Humanities – methodischer Brückenschlag oder ›feindliche Übernahme‹? Chancen und Risiken der Begegnung zwischen Geisteswissenschaften und Informatik. [\[online\]](#)

Kathleen Fitzpatrick: Planned Obsolescence. Publishing, Technology, and the Future of the Academy, auf der Plattform: Media-Commons Press. Open Scholarship in Open Formats. [\[online\]](#)

Matthew K. Gold: The Digital Humanities Moment. Die Expanded Open Access-Edition des Buches *Debates in the Digital Humanities*. [\[online\]](#)

HRK: Empfehlung der 14. Mitgliederversammlung der Hochschulrektorenkonferenz am 14. Mai 2013 in Nürnberg. Gute wissenschaftliche Praxis an deutschen Hochschulen. [\[online\]](#)

Nick Montfort / Patsy Baudoin / John Bell / Ian Bogost / Jeremy Douglass / Mark C. Marino / Michael Mateas / Casey Reas / Mark Sample / Noah Vawter: 10 PRINT CHR\$(205.5+RND(1)); : GOTO 10. Cambridge MA., London 2013. [\[online\]](#)

Julianne Nyhan / Oliver Duke-Williams: Joint and multi-authored publication patterns in the Digital Humanities. In: *Literary and Linguistic Computing* 4 (2014), H. 29, S. 1–13. [\[online\]](#)

PlagDoc / Martin Kotynek: Schwarmgedanken. Ein Versuch, Lehren aus ›GuttenPlag‹ zu ziehen. Ein Denkanstoß für künftige Recherche-Plattformen im Netz. Ein Diskussionsbeitrag zu ›investigativem Crowdsourcing‹, In: GuttenPlag Wiki, 8.6.2012. [\[online\]](#)

Fabian Steiner / Max Trecker: Zeitgemäß publizieren? Veröffentlicht von Lilian Landes am 17. Februar 2014, auf dem Weblog: Rezensieren – Kommentieren – Bloggen: Wie kommunizieren Geisteswissenschaftler in der digitalen Zukunft? [\[online\]](#)

Christoph Titz: Anonymer Schavan-Ankläger: ›Ich wollte das nicht unter den Tisch fallen lassen‹. In: Spiegel Online, 5. Mai 2012. [\[online\]](#)

Abbildungslegenden und -nachweise

Abb. 1: Die Expanded Open Access-Edition des Buches *Debates in the Digital Humanities*. Auszug aus: Matthew K. Gold: The Digital Humanities Moment.

Abb. 2: »6 Comments on Paragraph 3«. Auszug aus der Online-Version von: Kathleen Fitzpatrick: *Planned Obsolescence. Publishing, Technology, and the Future of the Academy*.

Abb. 3: GuttenPlag Wiki: Ergebnisse der Crowd-Arbeit [\[online\]](#).

Abb. 4: Fabian Steiner / Max Trecker: Zeitgemäß publizieren? [Blogposting](#) vom 17.2.2014 zur Konferenz *Rezensieren - Kommentieren - Bloggen: Wie kommunizieren Geisteswissenschaftler in der digitalen Zukunft?*

Zeitschrift für digitale Geisteswissenschaften

Artikel aus:

Sonderband 1 der ZfdG: Grenzen und Möglichkeiten der Digital Humanities. Hg. von Constanze Baum und Thomas Stäcker. 2015. DOI: [10.17175/sb01](https://doi.org/10.17175/sb01)

Titel:

Informatik und Hermeneutik. Zum Mehrwert interdisziplinärer Textanalyse

Autor/in:

Evelyn Gius

Kontakt: evelyn.gius@uni-hamburg.de

Institution: Universität Hamburg

GND: [1084241307](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63863-p0011-9) ORCID: [0000-0001-8888-8419](https://orcid.org/0000-0001-8888-8419)

Autor/in:

Janina Jacke

Kontakt: janina.jacke@uni-hamburg.de

Institution: Universität Hamburg

GND: [108423968X](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63863-p0011-9) ORCID: [0000-0001-7217-3136](https://orcid.org/0000-0001-7217-3136)

DOI des Artikels:

[10.17175/sb001_006](https://doi.org/10.17175/sb001_006)

Nachweis im OPAC der Herzog August Bibliothek:

[830166564](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63863-p0011-9)

Erstveröffentlichung:

19.02.2015

Lizenz:

Sofern nicht anders angegeben 

Medienlizenzen:

Medienrechte liegen bei den Autoren

Letzte Überprüfung aller Verweise:

31.05.2016

GND-Verschlagwortung:

[Hermeneutik](#) | [Ontologie \(Wissensverarbeitung\)](#) | [Informatik](#) |

Zitierweise:

Evelyn Gius, Janina Jacke: Informatik und Hermeneutik. Zum Mehrwert interdisziplinärer Textanalyse. In: Grenzen und Möglichkeiten der Digital Humanities. Hg. von Constanze Baum / Thomas Stäcker. 2015 (= Sonderband der Zeitschrift für digitale Geisteswissenschaften, 1). PDF Format ohne Paginierung. Als text/html abrufbar unter DOI: [10.17175/sb001_006](https://doi.org/10.17175/sb001_006).

Evelyn Gius, Janina Jacke

Informatik und Hermeneutik. Zum Mehrwert interdisziplinärer Textanalyse

Abstracts

Der Beitrag verhandelt die methodologischen Konsequenzen der interdisziplinären Zusammenarbeit zwischen Geisteswissenschaft und Informatik im Kontext des heureCLÉA-Projekts. Ziel von heureCLÉA ist es, eine ›digitale Heuristik‹ zur narratologischen Analyse literarischer Texte zu entwickeln, mit der automatisiert (1) bislang nur manuell durchführbare Annotationsaufgaben bis zu einem bestimmten Komplexitätsniveau durchgeführt und (2) statistisch auffällige Textphänomene als Kandidaten für eine anschließende Detailanalyse durch den menschlichen Nutzer identifiziert werden können.

Bei diesem Projekt müssen die disziplinären Herangehensweisen in besonderem Maß berücksichtigt werden. Im vorliegenden Beitrag werden die Ansätze dargestellt, die das bestehende Spannungsfeld zwischen (nicht-deterministischer) geisteswissenschaftlicher Hermeneutik und (deterministischer) Informatik produktiv nutzen und so über die konkrete Fragestellung des Projekts hinauswirken.

This paper discusses the methodological effects of the interdisciplinary cooperation between humanities scholars and computer scientists in the context of the project heureCLÉA. The goal of heureCLÉA is to develop a ›digital heuristic‹ that supports the narratological analysis of literary texts by (1) performing automatically annotation tasks up to a certain level of complexity that previously could only be carried out manually, and by (2) identifying statistically salient text phenomena for the subsequent detailed analysis by the human user.

In this project, it was essential to devote special attention to the specific disciplinary approaches. In this paper, we illustrate the ways in which the tensions between (non-deterministic) hermeneutics in the humanities and (deterministic) computer science can be productively employed and thus have an effect beyond the concrete research question.

1. Einleitung

Die Frage danach, was Digital Humanities sind, ist ähnlich alt wie die Digital Humanities selbst. Im Laufe der Jahrzehnte hat sich die Bezeichnung des Diskutierten mehrfach geändert: Die Digital Humanities hießen und heißen auch Computing in the Humanities, Literary Computing, Humanities Computing, eHumanities oder – im deutschen Sprachraum – ›Computerphilologie‹ und bezeichnen aus Sicht der Community mal eine Methode, mal eine Disziplin oder beliebige Kombinationen daraus.¹ Gemeinsam ist den Definitionen aber allen, dass es bei Digital Humanities um eine – wie auch immer geartete – Kombination von Aspekten der Geisteswissenschaften und der Informationswissenschaften geht. Das Zusammenspiel der beteiligten Disziplinen kann sehr unterschiedlich gestaltet sein, wobei

¹ Vgl. für eine aktuelle Übersicht über die Definition von Digital Humanities z.B. Terras et al. 2013 und dort insbesondere das Kapitel IV.

der disziplinäre Ursprung der verfolgten Fragestellung(en) eine Klassifizierungsmöglichkeit von Digital Humanities-Forschungsprojekten jenseits ihrer Definition darstellt.² Je nachdem, ob die dem betrachteten Forschungsansatz zugrundeliegenden Fragestellungen aus dem Bereich der Geisteswissenschaft oder aus der Informatik stammen, lassen sich drei Kategorien von Digital Humanities-Projekten bestimmen (vgl. Abbildung 1): Erstens Projekte, die geisteswissenschaftliche Fragestellungen verfolgen (und dafür *auch* informatische Methoden anwenden), zweitens Projekte, deren Fragestellungen aus der Informatik stammen (die *auch* mit geisteswissenschaftlichen Methoden bearbeitet werden) und drittens Projekte, die Fragestellungen sowohl aus den Geisteswissenschaften als auch aus der Informatik behandeln.

		geisteswissenschaftliche Fragestellung(en)?	
		-	+
informatische Fragestellung(en)?	-	x	Typ 1
	+	Typ 2	Typ 3

Abb. 1: Klassifikation von DH-Ansätzen (Grafik: Gius / Jacke 2014).

Dieser Beitrag beleuchtet den Mehrwert des im Forschungsprojekt **heureCLÉA** verfolgten Zugangs, der unter die dritte Kategorie von Digital Humanities-Forschungsprojekten fällt.³ Wir wollen insbesondere aufzeigen, welche Konsequenzen das Zusammenwirken von geistes- und informationswissenschaftlichen Zugängen in ein- und demselben Projekt haben kann und wie die interdisziplinäre Auseinandersetzung mit diesen Fragestellungen einen Mehrwert generiert, der durch die spezielle Situation solcher Forschungsvorhaben besonders gefördert, wenn nicht sogar erst ermöglicht wird. Dabei stehen weniger die Ergebnisse unserer Arbeit im Fokus, sondern wir wollen vielmehr die disziplinären methodischen und methodologischen Konsequenzen der Zusammenarbeit darstellen.

Im Projekt **heureCLÉA** arbeiten Informatiker und Literaturwissenschaftlerinnen an einer so genannten ›digitalen Heuristik‹, die die Analyse von literarischen Texten unterstützen soll. Die geisteswissenschaftliche Problemstellung für das Projekt **heureCLEA** ist dabei die Erschließung und Annotation von narrativen Referenzen im Feld der Grundkategorie ›Zeit‹ anhand eines Korpus von literarischen Erzählungen. Die Fragestellung seitens der Informatik bzw. der Sprachverarbeitung (auch Natural Language Processing bzw. NLP) betrifft die Neuentwicklung von Data Mining-Methoden für die bislang unbearbeitete Domäne literarischer Texte. **heureCLÉA** ist damit ein Projekt, das Fragestellungen aus der Narratologie und aus der Sprachverarbeitung vereint.

² Diese Klassifizierung wurde auch im Rahmen der DHd-Konferenz im März 2014 in Passau mehrfach diskutiert, etwa von John Nerbonne in seiner Keynote *Die Informatik als Geisteswissenschaft* und von Eva Barlösius und Axel Philipps in ihrem Konferenzbeitrag *Zur Sichtbarkeit von Street Art in Flickr. Methodische Reflexionen zur Zusammenarbeit von Soziologie und Informatik*.

³ Das Projekt **heureCLÉA** ist ein vom BMBF gefördertes eHumanities-Projekt, das von 02/2013–01/2016 an den Universitäten Hamburg und Heidelberg als Verbundprojekt durchgeführt wird (vgl. dazu auch Bögel et al.).

Ziel des Projekts ist die Entwicklung eines heuristischen Moduls, das die Funktionalität der webbasierten Textanalyse und -annotationsumgebung *CATMA*⁴ erweitert, indem es Vorschläge zu narratologischen Phänomenen in einem Text anbietet. Im Fokus stehen dabei klassische narratologische Kategorien, die das zeitliche Verhältnis zwischen Erzähltem und Erzählen thematisieren – wie etwa die Kategorien Ordnung (In welcher Reihenfolge?), Frequenz (Wie oft?) und Dauer (Wie lange?).⁵

Das Modul wird auf der Basis von drei Zugängen entwickelt: (1) Ausgangspunkt ist die Annotation der Zeitphänomene, die von geschulten Annotatoren vorgenommen wird.⁶ Dieses Markup wird anschließend mit (2) regelbasierten Verfahren⁷ sowie (3) Machine-Learning-Ansätzen kombiniert, um eine Vorhersage der Phänomene in literarischen Texten zu entwickeln.⁸

Aufgrund des Zusammenspiels von literaturwissenschaftlichen – und speziell: hermeneutischen – Verfahren und informatischen Verfahren des Data Mining stehen sich non-deterministische Zugänge zu Texten und entscheidbare bzw. deterministische Verfahren gegenüber, die nicht problemlos auf den jeweils anderen Ansatz übertragen werden können. Deshalb ist die Reproduzierbarkeit von narratologischen Analysen für die Vereinbarkeit des literaturwissenschaftlichen und des informatischen Zugangs und damit für den Erfolg des heuristischen Moduls ausschlaggebend. Andererseits ist gerade die Reproduzierbarkeit von Analyseergebnissen im Bereich der literaturwissenschaftlichen Textanalyse ein problematisches Konzept, da sie nicht ohne Weiteres mit den Prinzipien von hermeneutischen Verfahren vereinbar ist.

Das sich in diesem Gegensatz abzeichnende Problemfeld soll im verbleibenden Artikel differenzierter dargestellt werden. Dafür werden wir zuerst in **Abschnitt 2** auf die generelle Rolle der Hermeneutik in der Narratologie und ihre Bedeutung für die Wahl eines geeigneten Annotationzugangs für das *heureCLÉA*-Projekt eingehen. Anschließend beschreiben wir die in der Zusammenarbeit zwischen (hermeneutisch geprägter) Narratologie und (nicht hermeneutisch ausgerichteter) Computerlinguistik auftretenden Wechselwirkungen in Abschnitt 3. Dabei stehen insbesondere die Fragen nach Umfang und Vielfalt der analysierten bzw. genutzten Daten im Fokus, deren Problematik sich an den beiden gegensätzlichen Konzeptpaaren *Exemplarität vs. sparsity* und *Interpretationsvielfalt vs. noise* festmachen lässt. Für beide Fälle beschreiben wir die fachdisziplinären Konsequenzen sowie die entwickelten Lösungsansätze. Ein besonderes Augenmerk richten wir im letzten Teil des Abschnitts auf die Folgen für die Narratologie, die bereits jetzt, nach etwa der Hälfte der Projektlaufzeit, von erheblicher Bedeutung sind. Schließlich fassen wir die methodologischen Vorteile des interdisziplinären Ansatzes in Abschnitt 4 zusammen.

⁴ Das Tool ist [online](#) frei nutzbar und der unter der GNU general public license v3 lizenzierte Sourcecode ist [online](#) verfügbar.

⁵ Vgl. Genette 1972; Lahn / Meister 2013, S. 133–155.

⁶ Vgl. zu den genutzten Analysekatégorien die Beschreibung des *heureCLÉA*-Tagsets in den Annotationsguidelines Gius / Jacke 2014.

⁷ Zu den regelbasierten Verfahren vgl. Strötgen / Gertz 2010.

⁸ Damit ist *heureCLÉA* ein Beitrag zur *computational narratology* im Sinne von Mani, da es zu »exploration and testing of literary hypotheses through mining of narrative structure from corpora« (Mani 2013, unpag.) beiträgt.

2. Die Rolle der Hermeneutik

2.1 Hermeneutik und Narratologie

Die Hermeneutik – die »Methodenlehre des Interpretierens von Rede und Text«⁹ – dient bereits in der Antike der Auslegung religiöser Schriften und spielt heute, nach ihrer Neuprägung u.a. durch Schleiermacher und Dilthey,¹⁰ vor allem in der Literaturwissenschaft eine wichtige Rolle. Dort ist mit »hermeneutischem Textverstehen« meist eine ganz spezielle Operation gemeint: Das Herausarbeiten dessen, »was man für die Bedeutung (oder den Sinn bzw. den Inhalt) eines Werkes oder seiner Bestandteile hält.«¹¹ Der vorherrschenden Meinung zufolge zeichnet sich derartige Textverstehen unweigerlich durch eine ganz bestimmte Struktur aus, die dadurch charakterisiert ist, dass Vorannahmen von Textinterpreten, Annahmen über Textteile und Annahmen über das Textganze sich immer wieder gegenseitig beeinflussen, was zu Modifikationen bzw. zur Bestätigung dieser einzelnen Annahmen führt.¹²

Im Projekt heureCLÉA geht es weniger um die Methodenlehre des Interpretierens, sondern vielmehr um hermeneutische Textauslegung selbst im Rahmen narratologischer Textanalyse: Die Narratologie ist eine literaturwissenschaftliche Disziplin, die eine Reihe theoretischer Konzepte und Modelle für die Analyse und Interpretation erzählender Texte zur Verfügung stellt. Diese narratologischen Kategorien dienen der Bezeichnung und Verortung textueller Eigenschaften, die (a) als typisch für narrative Texte angesehen werden und (b) in unterschiedlicher Hinsicht als Basis der Textinterpretation dienen.¹³ Viele der Kategorien fassen strukturelle Phänomene, die hauptsächlich an der Textoberfläche zugänglich sind.

Die Narratologie gilt aufgrund dieser Ausrichtung auf primär textuelle Eigenschaften als Methode zur kontextunabhängigen Analyse von Erzähltexten, die ohne Bezüge zu extratextuellen Kontexten auskommt.¹⁴ Daraus wird häufig geschlossen, dass die Narratologie eine interpretationsunabhängige Methode zur Verfügung stellt.¹⁵ Kontextunabhängigkeit ist allerdings nicht gleichbedeutend mit Interpretationsunabhängigkeit – jedenfalls nicht im weiteren Sinne von »Interpretation«: Die meisten narratologisch beschriebenen Phänomene sind zwar an der Textoberfläche zugänglich, die für ihre Anwendung notwendige Rekonstruktion der sprachlichen Textbedeutung kann aber auch interpretationsabhängige Aspekte wie Disambiguierung oder Erschließung von Implikationen beinhalten und ist damit in einem weiteren Sinne des Wortes interpretativ, das heißt: Interpreten können mitunter legitimerweise zu unterschiedlichen Auffassungen darüber gelangen, was im Text

⁹ Bühler 2003, S. 3.

¹⁰ Vgl. u.a. und grundlegend Schleiermacher 2012 sowie Dilthey 1990; für eine »Philosophische Hermeneutik« vgl. Gadamer 1993.

¹¹ Bühler 2003, S. 5.

¹² Vgl. Bühler 2003, S. 13. Oft wird diese spezielle Struktur hermeneutischen Textverstehens als *hermeneutischer Zirkel* bezeichnet.

¹³ Vgl. Kindt / Müller 2003 für einen Überblick.

¹⁴ Auszuklammern sind hierbei jedoch intersubjektiv geteilte Vorannahmen wie allgemeines Weltwissen oder das Wissen über bestimmte literarische Konventionen – ohne diese im strengen Sinne extratextuellen Informationen kommt auch die narratologische Textanalyse oft nicht aus.

¹⁵ Vgl. Kindt / Müller 2003, S. 294–295.

ausgedrückt wird, da die Bedeutung nicht immer notwendig aus dem Textmaterial folgt. So ist beispielsweise sprachliches Verstehen – und damit die Anwendung nicht-deduktiver Schlüsse – notwendig, um etwa die zeitliche Ordnung von Erzählungen zu analysieren; auch wenn zunächst vielleicht der Verdacht naheliegt, dass diese sich allein aufgrund eindeutiger Textmarker wie etwa dem Tempus feststellen lässt.¹⁶ Basiert eine Erklärung auf sprachlichem Verstehen, so ist sie nicht notwendigerweise der einzige richtige Schluss, da sie auf einer Interpretation bzw. Gewichtung von ggf. widersprüchlichen Aspekten beruht und nicht ausschließlich von Phänomenen der Textoberfläche abgeleitet werden kann. Schon bei derartigen Schlüssen kann eine Variante des oben skizzierten Verfahrens hermeneutischen Textverstehens zum Tragen kommen, wenn Bedeutungshypothesen in wechselnder Bezugnahme auf die fragliche Textstelle, den weiteren intratextuellen Kontext und eventuell auf intersubjektiv geteilte Vorannahmen (wie allgemeines Weltwissen) generiert, modifiziert und bestätigt werden.

Indem wir Texte mithilfe von narratologischen Kategorien analysieren, bewegen wir uns also zwischen rein formaler Textbeschreibung und Hermeneutik: Zwar kann man generell davon ausgehen, dass narratologische Kategorien der Klassifizierung bzw. Beschreibung von Texten dienen¹⁷ und damit keinen primären Fokus auf hermeneutische Verfahren haben. Gleichzeitig ist auch die narratologische Analyse ein hermeneutisches und damit ein interpretatives Verfahren. Interpretation ist für die Narratologie genau genommen in dreierlei Hinsicht relevant:

(1) Narratologische Textbeschreibungen können potentiell als Grundlage, Prüfstein oder Heuristik umfassender Werkdeutungen dienen.¹⁸ (2) Die narratologische Analyse geht meist über eine rein formale bzw. quantitative Textbeschreibung hinaus, da – wie oben bereits dargelegt – sprachliches Verstehen involviert ist. Dieses führt zwar meist, aber nicht immer zu intersubjektiv gültigen Ergebnissen.¹⁹ (3) Es gibt Fälle, in denen der narratologischen Klassifikation eine Textdeutung vorausgehen muss, da das zu bestimmende Phänomen nur aufgrund einer umfassenden Textinterpretation erschlossen werden kann.²⁰

Während der erste Punkt nur dann zum Tragen kommt, wenn umfassende Werkinterpretationen erstellt werden, sind die beiden anderen Aspekte potentiell in jeder narratologischen Textanalyse von Belang. Damit ist auch die narratologische Textanalyse trotz ihres vergleichsweise formalen Zugangs ein eindeutig hermeneutisches Verfahren.

¹⁶ Dass dem nicht so ist, illustriert die folgende Textstelle: »Dann ist sie, ohne ein Wort zu reden, bis zu meinem Schreibtisch gekommen und hat die Blumen vor mich hingelegt. Und in der nächsten Sekunde greift sie nach den verwelkten im grünen Glas. Mir war, als griffe man mir ins Herz« (Schnitzler: Blumen [1894], S. 106). Dass der Tempuswechsel vom Präteritum ins Präsens und zurück ins Präteritum hier keinen Zeitsprung darstellt, geht nicht zwingend aus dem Textmaterial hervor und ist deshalb kein deduktiver Schluss. Allerdings scheint es in diesem Fall vernünftig, inhaltliche Aspekte, die eine chronologische Abfolge anstatt eines doppelten Zeitsprungs indizieren, stärker zu gewichten als den Tempuswechsel. Der Zeitausdruck »in der nächsten Sekunde«, der den ersten und den zweiten Satz in zeitlich enge Relation setzt, sowie die Einschätzung der Reaktion des Erzählers als unmittelbar darauf folgend wären dann ausschlaggebend für die Bestimmung der Textstelle als ›chronologisch erzählt‹. Dieser Schluss stellt offensichtlich die beste Erklärung dar und erscheint unstrittig.

¹⁷ Vgl. Kindt / Müller.

¹⁸ Vgl. bspw. Kindt / Müller 2003, S. 299–301.

¹⁹ Vgl. Jacke 2014, S. 129–130.

²⁰ Vgl. Kindt / Müller.

Entsprechend ist die Arbeit im heureCLÉA-Projekt stark von dem typischen Gegensatz zwischen Geisteswissenschaft und Informatik bzw. non-deterministischen und deterministischen Arbeitsweisen geprägt.

2.2 Hermeneutisches Markup

Bereits im Vorfeld der tatsächlichen Zusammenarbeit zwischen Narratologie und Computerlinguistik musste eine wichtige Frage geklärt werden: Wie können hermeneutische Verfahren der Textanalyse in ein adäquates maschinenlesbares Format übersetzt werden? Dabei ging es darum, einen Annotationszugang zu identifizieren, der eine nicht-deterministische Textanalyse, wie sie im Projekt heureCLÉA vorgenommen wird, unterstützt. Hierfür sind nicht alle Markup-Typen, die für die Annotation von Texten genutzt werden, geeignet.

Grundsätzlich kann man zum einen die Verbindung zwischen Markup und annotiertem Objekt, zum anderen die Funktion des Markups unterscheiden. Die Unterteilung in eingebettetes oder *inline*-Markup, das direkt in den annotierten Text geschrieben wird, und *stand-off*-Markup, das separat abgespeichert wird und durch abgespeicherte Verweise auf die jeweiligen Textabschnitte referiert, ist nicht nur technisch von Belang. Mit der Wahl der Verbindung zwischen Markup und annotiertem Objekt sind auch verschiedene methodologische Zugänge zur Textanalyse verbunden: Während mit einer fest mit dem Text verzahnten *inline*-Annotation – zumindest implizit – ein normativer Zugang zur Textanalyse und damit ein gewisser Anspruch auf objektive und zeitlose ontologische Gültigkeit verbunden ist, ist die *stand-off*-Annotation flexibler und ermöglicht vielfältige und vielschichtige Zuschreibungen an einen Text. Damit ist sie näher am oben diskutierten Gedanken der hermeneutischen Textanalyse, der seit Dilthey und Gadamer in den Geisteswissenschaften vorherrscht. Um sinnvoll literaturwissenschaftlich arbeiten zu können, braucht man insbesondere ein Markup, das verschiedene Interpretationen derselben Textstelle ermöglicht, indem es verschiedene Auszeichnungen dieser Textstelle gleichzeitig zulässt. Dabei heißt »verschieden« nicht nur »unterschiedlich«, sondern dies muss auch die Möglichkeit widersprüchlicher Annotationen mit einschließen. Wie wir in [Abschnitt 3.3](#) sehen werden, kann es nämlich durchaus vorkommen, dass zwei Analysen einer Textstelle zu Ergebnissen kommen, die sich gegenseitig ausschließen, und die Stelle entsprechend mit widersprüchlichen Annotationen versehen wird. Auch wenn solche direkten Widersprüche zumeist durch die jeweilig damit verbundenen unterschiedlichen Vorannahmen bedingt sind und es entsprechend möglich ist, die Unstimmigkeiten aufzuklären, so führt eine solche Klärung in vielen Fällen nicht zu einer übereinstimmenden Bewertung der Textstelle, sondern eher zu einer Art Einigkeit über die Uneinigkeit in der Bewertung. Entsprechend muss es für eine literaturwissenschaftliche Textanalyse möglich sein, sich gegenseitig ausschließende Annotationen für dieselbe Zeichenkette vorzunehmen.²¹

²¹ Die Tatsache, dass XML inzwischen als de facto-Standard etabliert ist, widerspricht diesen Erkenntnissen. Bei XML handelt es sich um eingebettetes deklaratives Markup, das zwar aus technischer Sicht durchaus funktional ist, aber geisteswissenschaftlichen Methoden wie der hermeneutischen Textanalyse nicht gerecht werden kann (vgl. dazu auch Schmidt 2010).

Der zweite Unterschied zwischen Annotationszugängen betrifft die Funktion des zur Verfügung gestellten Markups. Hier geht es gewissermaßen um die Frage, für wen das Markup bestimmt ist: Soll eine Maschine mit so genanntem prozeduralen Markup angewiesen werden, eine bestimmte Operation auszuführen (z.B. den Text auf eine bestimmte Weise – etwa als Überschrift – darzustellen) oder soll ein deskriptives Markup einen menschlichen Adressaten über die Kategorisierung eines bestimmten Textabschnittes (im Kontext einer Grammatik, eines Lexikons oder auch eines bestimmten pragmatischen Zusammenhangs) informieren? Vor fast zwei Jahrzehnten haben Coombs et al. bereits festgestellt, dass nur deskriptives Markup den intellektuellen Zielen wissenschaftlicher Textanalysen gerecht werden kann.²² Piez hat diesen Ansatz weiterentwickelt und den Begriff des hermeneutischen Markups geprägt:

»By ›hermeneutic‹ markup I mean markup that is deliberately interpretive. It is not limited to describing aspects or features of a text that can be formally defined and objectively verified. Instead, it is devoted to recording a scholar's or analyst's observations and conjectures in an open-ended way. As markup, it is capable of automated and semiautomated processing, so that it can be processed at scale and transformed into different representations. By means of a markup regimen perhaps peculiar to itself, a text would be exposed to further processing such as text analysis, visualization or rendition. Texts subjected to consistent interpretive methodologies, or different interpretive methodologies applied to the same text, can be compared. Rather than being devoted primarily to supporting data interchange and reuse – although these benefits would not be excluded – hermeneutic markup is focused on the presentation and explication of the interpretation it expresses.«²³

Für die Arbeit in heureCLÉA wird ein solches ›hermeneutisches Markup‹ in Form von deskriptivem *stand-off*-Markup genutzt, das in CATMA implementiert ist.²⁴ Dieses ermöglicht nicht nur mehrfache, überlappende und sogar widersprüchliche Annotationen derselben Textstelle, sondern erlaubt auch eine flexible Gestaltung der genutzten Analysekatgorien. Diese können bei Bedarf im Laufe der Annotationsdurchgänge modifiziert und angepasst werden, wobei CATMA zusätzlich durch die implementierten Suchfunktionen eine halbautomatische Veränderung der Annotationen zulässt. Diese Flexibilität von CATMA unterstützt die durch die Praxis der hermeneutischen Textanalyse bedingte rekursive Überarbeitung der Annotationen und damit den gesamten literaturwissenschaftlichen Textanalyseprozess erheblich. Deshalb wird das von CATMA zur Verfügung gestellte Annotationsformat auch im Projekt heureCLÉA genutzt.

3. Wechselwirkungen literaturwissenschaftlicher und informatischer Textanalyse

²² Vgl. Coombs et al. 1987. Diese Feststellung wird u.a. von Buzzetti und im bereits erwähnten Beitrag von Schmidt aufgegriffen (vgl. Buzzetti 2002, S. 63–64; Schmidt 2010, S. 338).

²³ Piez 2010, S. 202.

²⁴ Die flexible Gestaltung der Tagsets wird in CATMA durch die Nutzung des Feature Declaration System und der Feature Structures von TEI ermöglicht, wodurch das Markup gleichzeitig *stand-off*-Markup ist, das Überlappungen zulässt, und TEI-XML-kompatibel ist. Textbereiche (<seg>) sind mit Taginstanzen durch den Gebrauch von @ana-Attributen verlinkt, der Text selbst wird durch <ptr>-Elemente repräsentiert, die das originale Quelldokument referenzieren, welches unverändert belassen wird (vgl. weitere Informationen zur Implementierung von CATMA sowie Referenzen zu den genutzten Standards [online](#)).

Wenn zwei so unterschiedliche Disziplinen wie Literaturwissenschaft und Informatik im Rahmen eines Projekts wie heureCLÉA kooperieren, dann rücken unweigerlich einige theoretische und praktische Differenzen zwischen den Fachwissenschaften in den Fokus, die im Laufe des Projekts analysiert und konstruktiv bearbeitet werden müssen. So kommt es vor, dass bestimmte Fachtermini zwar in beiden beteiligten Disziplinen genutzt werden, diese Termini jedoch jeweils unterschiedlich definiert sind – so zum Beispiel die Ausdrücke ›Ereignis‹ und ›narrativ‹. Diesen rein terminologischen Missverhältnissen ist leicht durch das Offenlegen der fachspezifischen Definitionen Abhilfe zu schaffen: Während in der Sprachverarbeitung (Natural Language Processing bzw. NLP) die Definition von ›Ereignis‹ bzw. ›event‹ grundsätzlich möglichst allgemein gehalten wird als etwas, das auftritt oder passiert und eine zeitliche Dimension hat, ist der Ereignisbegriff in der Narratologie wesentlich detaillierter.²⁵ Als Ereignis gilt dort eine erzählte Zustandsveränderung (diese Bestimmung ist auch bekannt unter der Bezeichnung ›Ereignis I‹) bzw. eine erzählte Zustandsveränderung, der zusätzliche Eigenschaften wie bspw. Faktizität und Resultativität – also tatsächliches Eintreten und Abgeschlossenheit der Zustandsveränderung in der fiktiven Welt einer Erzählung – zukommen (auch ›Ereignis II‹).²⁶

Eine zweite Differenz erfordert hingegen komplexere Maßnahmen, um eine Zusammenarbeit zu ermöglichen. Das Problem besteht darin, dass die Qualität der vorgenommenen bzw. genutzten Analysen in den beiden Fachdisziplinen traditionell unterschiedlich bewertet wird: Das, was in der NLP als nicht geeignete Datengrundlage betrachtet wird, kann aus Sicht der Literaturwissenschaft eine methodisch einwandfreie Analyse sein.

Im Folgenden möchten wir die zwei wesentlichen Aspekte dieser Problematik vorstellen. Zum einen geht es in [Abschnitt 3.1](#) um die unterschiedliche Bewertung der geeigneten Menge an Analysedaten – also um das Spannungsverhältnis zwischen *sparsity* (Datenarmut) und exemplarischen Analysen –, und zum anderen um das ideale Maß an übereinstimmenden Analysen – also um den Widerspruch zwischen *noise* (nicht verwertbaren, widersprüchlichen Daten) und Interpretationspluralismus ([Abschnitt 3.2](#)). Dass derartige fachwissenschaftliche Unterschiede nicht nur pragmatisch aufgelöst werden können, sondern letztlich auch noch einen Mehrwert für die einzelnen Fachdisziplinen bedeuten können, soll in den [Abschnitten 3.3](#) und [3.4](#) gezeigt werden.

3.1 Zur Häufigkeit der betrachteten Phänomene

²⁵ Der Grund für den flachen Ereignisbegriff in der NLP ist hauptsächlich darin zu suchen, dass er für das Testen von NLP-Systemen genutzt wird und dann für den jeweiligen Anwendungsfall konkretisiert wird. So geschieht es etwa alljährlich im Rahmen des SemEval-Workshops im Wettbewerb »Event extraction and classification« [vgl. [2015 online](#)], bei dem sich verschiedene Systeme im automatischen Erkennen von Ereignissen messen. Da aber beide Disziplinen ihr Narrativitätskonzept maßgeblich auf Ereignissen aufbauen (vgl. bspw. Schmid 2008, S. 1–26, für den narratologischen Zugang sowie Chambers / Jurafsky 2008 für den NLP-Zugang), wäre es interessant zu untersuchen, inwiefern umfangreiche Analysen in der Narratologie ebenso eine gewisse Verflachung des Ereignisbegriffs nach sich ziehen würden. Diese Vermutung legt auch die Untersuchung in Gius 2013 nahe, in der die umfangreiche Anwendung narratologischer Kategorien nicht nur bislang unbeobachtete Interdependenzen zwischen den angewendeten Kategorien offenbart, sondern auch herausstellt, dass die Komplexität des narratologischen Theoriegebäudes zugunsten seiner Konsistenz verringert werden sollte (vgl. Gius 2015, S. 292–300 und S. 308–310).

²⁶Vgl. Schmid 2008, S. 20–27. Zu Schmid's Kriterien für ›Ereignis II‹ zählen neben Faktizität und Resultativität auch die Kriterien Relevanz, Unvorhersehbarkeit, Effekt, Irreversibilität und Nicht-Wiederholbarkeit.

In der NLP wird für die Automatisierung von Aufgaben der Texterkennung auf eine möglichst große Datenbasis zurückgegriffen, um daraus statistisch signifikante Regelmäßigkeiten abzuleiten, die etwa die Verteilung von Phänomenen in einem bestimmten Text vorhersagen können. Im Rahmen eines solchen statistischen Zugangs zur Verarbeitung natürlicher Sprache sind selten vorkommende Phänomene entsprechend problematisch, da sie nicht in ausreichender Zahl vorhanden sind, um eine verlässliche Basis für Inferenzen zu bilden. Dieses je nach Kontext als *sparse data* oder *rare events* bezeichnete Problem tritt notwendigerweise bei jeder Analyse auf: Während die Zahl der häufig vorkommenden Sprachphänomene limitiert ist und sich entsprechend mit einer geeigneten Korpuserstellung fassen lässt, ist es oft nicht möglich, ausreichend Daten für eine erschöpfende Analyse seltener Phänomene zusammenzustellen.²⁷ Je nach Problemstellung ist es deshalb in der statistischen NLP gängig, selten vorkommende Phänomene aus dem Korpus zu entfernen. Dies gilt insbesondere für jene Fälle, in denen die Parameter zur Modellierung der Phänomene ausschließlich aus einem Trainingskorpus erschlossen werden.²⁸

Während also möglichst zahlreiche Vorkommnisse der untersuchten Phänomene in der automatisierten Sprachverarbeitung das vorherrschende Desiderat sind, ist umgekehrt die exemplarische Analyse von Einzelfällen die in der Literaturwissenschaft gängige Praxis. Dafür, dass Textanalysen bzw. Interpretationen sich meist auf die Bearbeitung eines oder weniger Werke beschränken, gibt es zwei Gründe. Zum einen entspricht, gerade im Zusammenhang mit umfassenden Textinterpretationen, das Herausstellen der besonderen Eigenschaften eines Einzelwerks oder einzelnen Autors dem Selbstverständnis der Literaturwissenschaft: »[D]ie Anstrengung der Literaturwissenschaft [gilt] gerade dem Besonderen des Textes. [...] Von literarischer Bedeutung zu sprechen, heißt von der Ausnahme der Regel zu sprechen.«²⁹ Zum anderen gibt es für die exemplarische Arbeitsweise auch praktische Gründe: Selbst wenn es Literaturwissenschaftlern darum geht, durch Textanalyse und -interpretation Erkenntnisse zu erlangen, die auf bestimmte andere Texte übertragbar sein sollen, ist es oft aus praktischen Gründen nicht möglich, ein großes Textkorpus zu bearbeiten. Das liegt daran, dass literaturwissenschaftliche Textarbeit – und insbesondere stark textorientierte Zugänge wie etwa die narratologische Analyse – sehr genaue und mehrmalige Rezeption sowie textnahe Detailanalysen erforderlich machen.

3.2 Zur Vielfalt der Analysen

Für die automatische Sprachverarbeitung sollte das genutzte Datenmaterial nicht nur möglichst zahlreiche Vorkommnisse der untersuchten Phänomene aufweisen, diese sollten außerdem möglichst eindeutig sein, um den statistischen Fehler von Vorhersagen zu minimieren. Das heißt unter anderem, dass etwa bei von mehreren Personen annotierten Daten die Annotationen übereinstimmen sollten. Als Datengrundlage für Training und Evaluation von NLP-Systemen wird häufig mit einem so genannten ›Goldstandard‹ gearbeitet, der eine hohe Konsistenz des Datenmaterials garantieren soll, indem Annotationen von

²⁷ Vgl. Manning / Schütze 2001, S. 198–199.

²⁸ Vgl. Manning / Schütze 2001, S. 491.

²⁹ Jannidis et. al 2003, S. 3.

mehreren Personen nach genau festgelegten Annotationsrichtlinien manuell erstellt und verglichen werden.³⁰

Daten, die vielfältige oder gar widersprüchliche Informationen enthalten, werden hingegen als ›noisy‹ bezeichnet – und das Rauschen, das sie enthalten, wird in den meisten Zugängen entfernt, da es sich dabei um nicht-deterministisches und damit aus Sicht der NLP um nicht geeignetes Material handelt.³¹

Die Tatsache, dass literarische Texte von Interpreten unterschiedlich verstanden werden (können), wird in der Literaturwissenschaft hingegen als durch den Gegenstand und die Methode bedingt und damit als normal, wenn nicht sogar notwendig betrachtet. Für den in der Literaturwissenschaft geltenden Interpretationspluralismus gibt es mindestens drei Gründe: (1) Der wohl wichtigste Grund liegt in der Annahme der grundsätzlichen Polyvalenz literarischer Texte.³² Dieser These zufolge ist es ein charakteristisches Merkmal literarischer Texte, dass sie sich unterschiedlich deuten lassen. Die verschiedenen Deutungen sind oft gleichwertig und können einander widersprechen. Unter dieser Voraussetzung muss die generelle Vereinheitlichung unterschiedlicher Interpretationen als Missachtung der Mehrdeutigkeit literarischer Texte gewertet werden. (2) Ein weiterer Grund für den in der Literaturwissenschaft akzeptierten Interpretationspluralismus, der möglicherweise mit der Polyvalenz literarischer Texte zusammenhängt, ist die Mehrdeutigkeit des Verstehens- bzw. Bedeutungsbegriffs. Textinterpretation ist auf das Verstehen von Texten bzw. auf das Erfassen ihrer Bedeutung ausgerichtet. Dabei muss jedoch unter ›Bedeutung‹ nicht immer das Gleiche verstanden werden. Nach Danneberg liegt jeder Textinterpretation implizit oder explizit eine Bedeutungskonzeption zugrunde, die festlegt, was im Rahmen dieser Interpretation als Textbedeutung verstanden wird bzw. wie sich diese konstituiert.³³ Die Bedeutungskonzeption kann hierbei autor-, text-, rezipienten- oder kontextorientiert sein.³⁴ Liegen also zwei konkurrierende Interpretationen eines Textes oder eines Textteils vor, so kann der Grund dafür sein, dass sich die Interpretationen auf unterschiedliche Bedeutungskonzeptionen stützen. (3) Zusätzlich zu den beiden bisher genannten Punkten muss beachtet werden, dass Interpretationen immer Operationen sind, bei denen nicht-deduktive Schlüsse wie beispielsweise Abduktion (Schluss auf die beste Erklärung) vollzogen werden müssen.³⁵ Bei solchen Schlüssen geht es darum herauszufinden, was durch einen Text impliziert ist, wobei es möglich ist, dass zwei Interpreten auf Basis des gleichen Textes zu unterschiedlichen Ergebnissen kommen, die jedoch beide zulässig sein können. Eine derartige Situation liegt im oben diskutierten Beispiel zum Tempuswechsel³⁶ vor, in dem die anscheinend beste Erklärung der Textdaten diejenige ist, dass der Tempuswechsel keine Abweichung von

³⁰ Die menschliche Übereinstimmung wird also als obere Schranke für die automatischen Systeme verwendet, d.h., wenn Menschen nicht konsistent annotieren können, kann man das von einem automatischen System ebenfalls nicht erwarten.

³¹ Non-deterministische Daten würden keine Vergleichbarkeit der Systeme ermöglichen. Deshalb strebt die NLP im Normalfall eindeutige und ›richtige‹ Lösungen an, gegen die dann etwa die im Rahmen der bereits erwähnten SemEval-Wettbewerbe antretenden Systeme evaluiert werden können.

³² Vgl. für eine Analyse der Gründe der Polyvalenzthese Jannidis 2003.

³³ Vgl. Danneberg 1999, S. 101.

³⁴ Vgl. Kindt / Köppe 2008, S. 19; Fricke 1992, S. 216.

³⁵ Vgl. Jannidis 2006, S. 138–141; Jacke 2014, S. 126 und S. 130–131. Zur Methode des Schlusses auf die beste Erklärung vgl. die ausführliche Analyse von Lipton 1991.

³⁶ Vgl. das Beispiel zu Schnitzlers *Blumen* im Kommentarteil von [Abschnitt 2.1](#).

der chronologischen Darstellung bedeutet. Ob diese Folgerung allerdings tatsächlich durch den Text impliziert wird, ist in ähnlichen Fällen nicht immer vollends klar, weswegen dann auch andere Folgerungen sinnvoll sein können.

3.3 Konsequenzen

Die oben beschriebenen inkompatiblen fachspezifischen Sichtweisen auf die Frage von Umfang und Vielfalt der erzeugten bzw. genutzten Daten erforderten Maßnahmen auf beiden Seiten, um überhaupt gemeinsam am heureCLÉA-Projekt arbeiten zu können: Da die Annotation der Daten der erste Schritt bei der Entwicklung der digitalen Heuristik ist, musste als erstes auf narratologischer Seite überlegt werden, inwiefern die Problematiken von *sparsity* und *noise* bei der Auswahl und Analyse der Texte verringert werden können. Für die darauf folgenden Schritte der regelbasierten Verfahren und Machine Learning-Ansätze musste anschließend auf Seite der Computerlinguistik ein Weg gefunden werden, der auch den literaturwissenschaftlichen Prinzipien der Exemplarität und der Interpretationsvielfalt gerecht wird. Wie wir im Folgenden zeigen werden, haben diese Maßnahmen nicht nur zur Annäherung von jeweils gegensätzlichen Positionen geführt, sondern wirken sich auch – z.T. weit über das Projektziel einer digitalen Heuristik hinaus – positiv auf Fragestellungen der beiden Fachdisziplinen aus.

Das durch die Exemplarität literaturwissenschaftlicher Textanalyse bedingte *sparsity*-Problem wird im Rahmen des Projekts heureCLÉA nur durch den zweiten der in [Abschnitt 3.1](#) genannten Gründe für exemplarisches Arbeiten bedingt: In heureCLÉA spielt die durch umfassende Interpretation herauszustellende Werkbedeutung und damit die Besonderheit des einzelnen Textes keine Rolle. Die narratologische Analyse erfordert aber sehr wohl genaue und damit zeitintensive Textarbeit, weshalb keine große Menge an Texten annotiert werden kann. Gleichzeitig ist unser Anspruch an die digitale Heuristik, die auf Basis unserer Analysen erstellt wird, dass sie auf möglichst viele – unterschiedliche – narrative Texte anwendbar sein soll. Deshalb ist die Vorgehensweise in heureCLÉA trotz der praktischen Einschränkung durch den hohen Arbeitsaufwand literaturwissenschaftlicher Analysen in zweierlei Hinsicht darauf ausgerichtet, verallgemeinerbare Ergebnisse zu erzielen:

(1) Textgrundlage: Damit unser Modul letztlich nicht nur an die Eigenheiten eines Einzelwerks angepasst ist, arbeiten wir mit 21 Erzählungen unterschiedlicher Autoren. Auf diese Weise wird für eine zumindest teilweise heterogene Datengrundlage gesorgt, so dass die Chance der Übertragbarkeit der Ergebnisse auf weitere Texte gefördert wird. Um in einem ersten Schritt verwertbare Ergebnisse erzielen zu können, sind die Texte unseres Korpus nicht in jeder Hinsicht heterogen, sondern größtenteils in der gleichen Zeitspanne (um 1900) entstanden. Die weitere Verallgemeinerbarkeit der Ergebnisse kann im Anschluss – etwa mit Erzähltexten aus anderen Zeiträumen oder mit anderem Umfang – erprobt werden.

(2) Interpretieren: Um außerdem dafür zu sorgen, dass die Analysedaten in heureCLÉA nicht auf subjektiven Einzelmeinungen von Interpretinnen beruhen, wird jeder Text von mindestens zwei Annotatoren in wechselnden Teams bearbeitet. So wird gewährleistet, dass zu jedem Text

zwar nicht eine große Menge an Metadaten, doch aber immer mehr als eine Analyse vorliegt, was wiederum der Repräsentativität der Ergebnisse zuspielt.

Beide genannten Vorgehensweisen ermöglichen eine gewissermaßen ›semi-exemplarische‹ Arbeitsweise und damit die Annäherung des literaturwissenschaftlichen Umgangs an die informatische Sichtweise: Die mögliche Nicht-Repräsentativität von Daten wird zumindest teilweise als Problem betrachtet und es werden Maßnahmen getroffen, um die daraus eventuell resultierenden Schwierigkeiten zu mildern.

Die Maßnahmen, die auf literaturwissenschaftlicher Seite getroffen wurden, mildern das Problem der Datenarmut, lösen es aber nicht. Im Kontext des heureCLÉA-Projekts ist nämlich nicht nur der reine Umfang der Texte im Korpus gering, auch das Vorkommen von expliziten Zeitausdrücken in den Texten beträgt – mit etwa einem expliziten Zeitausdruck auf 15 Sätzen – nur einen Bruchteil der Zeitausdrücke in nicht-fiktionalen narrativen Texten, auf denen die Analyse von Zeitinformationen im NLP-Bereich im Normalfall aufbaut. Erschwerend kommt hinzu, dass ausgerechnet explizite Zeitinformationen, die – bei adäquater Datengrundlage – von NLP-Zugängen besonders gut erkannt werden können, bei der Analyse von Zeitphänomenen im Bereich der Narratologie nur wenig interessant sind, da es dort um wesentlich komplexere Phänomene geht, die nicht zwingend mit Zeitausdrücken in Verbindung stehen.³⁷ Für die Entwicklung der digitalen Heuristik musste also auch seitens der NLP frühzeitig ein eigener Zugang entwickelt werden.

Dafür wurde in einem ersten Schritt das Tempus anstelle der Zeitausdrücke als Grundlage für die Bestimmung von Zeitinformationen gewählt, wodurch ein mehrfacher Vorteil entstand: Durch den Fokus auf das Tempus enthält die zur Verfügung stehende Datengrundlage trotz ihrer kleinen Größe ausreichend nutzbare Daten. Außerdem entspricht diese Strategie dem Ansatz von Mani und Schiffman zur Bestimmung der zeitlichen Ordnung narrativer Texte und passt damit sehr gut zur Fragestellung des Projekts.³⁸ Die Bestimmung des Tempus erfolgt über die Kombination einer Infrastruktur in Form einer **UIMA-Pipeline**, die automatisch verschiedene komplexe Annotationen vornimmt, Teilsätze bestimmt sowie eine morphologische Analyse durchführt, und einer Heuristik, die anschließend das Tempus anhand von Textmarkern bestimmt.

Dieser kombinierte Zugang hat sich bisher als robust herausgestellt und kann außerdem im Gegensatz zu anderen Zugängen domänenunabhängig eingesetzt werden.³⁹ Im weiteren Projektverlauf werden die so generierten Annotationen von menschlichen Annotatorinnen überprüft und so in einem überwachten System eingesetzt, das die kontextabhängige Gültigkeit der Heuristik lernt. Damit ist also anscheinend eine für das Spannungsfeld Exemplarität und *sparsity* angemessene Lösung gefunden, die den beteiligten fachdisziplinären Methoden und Annahmen gerecht wird.

³⁷ Vgl. Bögel et al. 2014, S. 951.

³⁸ Vgl. Mani / Schiffmann 2005.

³⁹ Für eine ausführlichere Beschreibung des Zugangs vgl. Bögel et al. 2014.

Um hingegen das durch die Interpretationsvielfalt bedingte *noise*-Problem für die Automatisierung besser handhabbar zu machen, wurde bei der narratologischen Analyse die Menge zugelassener Interpretationen in heureCLÉA regelhaft eingeschränkt. Denn trotz des Interpretationspluralismus in der Literaturwissenschaft – und hier besteht wieder ein Anknüpfungspunkt an die Perspektive der Computerlinguistik – wird bei weitem nicht jede Interpretation als zulässig erachtet. Neben allgemeinen Kriterien wie Konsistenz und Kohärenz sowie spezifischen Regeln des jeweiligen Interpretationsansatzes gehen viele Literaturwissenschaftler davon aus, dass Interpretationen anhand des Textmaterials begründbar sein müssen – oder zumindest mit diesem vereinbar.⁴⁰ Beispiele für prominente Ansätze im immer noch vergleichsweise wenig reflektierten Bereich der Interpretationstheorie und -praxis, die die notwendige Kompatibilität von Interpretationen mit dem Textmaterial in den Fokus rücken, sind solche Ansätze, die die Methode des Schlusses auf die beste Erklärung⁴¹ oder die hypothetisch-deduktive Methode auf die Literatur übertragen. So sieht die hypothetisch-deduktive Methode Føllesdal et al. zufolge folgendes Vorgehen bei der Literaturinterpretation vor: »Zuerst wird eine Hypothese aufgestellt [...] und dann leitet man daraus eine Reihe von Konsequenzen ab, von denen sich zeigen lässt, dass sie mit dem Text übereinstimmen.«⁴²

Diesen theoretischen Grundannahmen, die die Anzahl der zulässigen Interpretationen einschränken, wird in heureCLÉA vor allem durch die Praxis Genüge getan, dass nach abgeschlossener individueller Annotation inkompatible Annotationen im Rahmen eines Vergleichsdurchgangs unter den Annotatoren diskutiert werden. Dabei müssen die individuellen Entscheidungen für widersprüchlich annotierte Textstellen begründet werden, was letztlich dazu führt, dass schlecht begründete Entscheidungen revidiert werden müssen. Bei diesem Abgleich der Annotationen zeigt sich aber auch häufig, dass der Grund für inkompatible Annotationen weniger in einer mangelhaften Begründung als vielmehr in der unzureichenden Definition der zugrundeliegenden narratologischen Kategorie besteht. Tritt dieser Fall auf, wird er zum Anlass genommen, diese theoretischen Unzulänglichkeiten möglichst zu beheben.⁴³ Wenn sich diskrepante Annotationen weder auf mangelhafte Begründungen noch auf unzureichende Definitionen des annotierten Phänomens zurückführen lassen, dann sollte es immerhin in den meisten Fällen möglich sein, den Grund für die unterschiedlichen Deutungen anzugeben, so dass jede Interpretation beispielsweise als relativ zu bestimmten Grundannahmen gesehen werden kann.⁴⁴ In solchen Fällen werden die inkompatiblen Annotationen durch eine Markierung in den Metadaten als widersprüchlich autorisiert und der individuelle Grund für die jeweilige Interpretation vermerkt. Auf diese Weise können scheinbare Inkonsistenzen in den Metadaten ohne eine erzwungene Vereinheitlichung so aufgelöst werden, dass die Daten aus computerlinguistischer Sicht verwertbar sind.

Der Umgang mit der so entstandenen ›qualifizierten‹, nicht-deterministischen Datengrundlage seitens der Computerlinguistik ist zum aktuellen Zeitpunkt noch nicht

⁴⁰ Für eine Darstellung dieser Position vgl. Kindt / Müller 2003, S. 299–300.

⁴¹ Vgl. die Erläuterung zum Schluss auf die beste Erklärung in [Abschnitt 3.2](#).

⁴² Føllesdal et al. 2008, S. 72.

⁴³ Vgl. dazu die Erläuterungen zur Definition der Prolepse in der Anmerkung zu Hebbels *Matteo*.

⁴⁴ Vgl. dazu die Erläuterungen zum Beispiel der Erzählgeschwindigkeit in [Abschnitt 3.4](#).

definiert. Allerdings steht schon jetzt fest, dass auch in diesem Fall die Vorarbeit auf narratologischer Seite nicht ausreicht, um eine vorgefertigte Lösung anzuwenden. Im Bereich des Machine Learning werden nämlich widersprüchliche Daten als Problem betrachtet, das durch geeignete Filtermethoden gelöst wird – womit die widersprüchlichen Daten entfernt und nicht weiter berücksichtigt werden. Hier muss also ein neuer Umgang mit *noisy data* entwickelt werden. Geplant ist eine Parametrisierung der modellierten Konzepte, die auch die narratologischen Erkenntnisse zu impliziten Vorannahmen mit einschließt. Anschließend soll ein feedbackgesteuertes Machine Learning stattfinden, bei dem die Annotatoren die identifizierten Parameter einstellen. So können diese ins erarbeitete Modell integriert werden.

Dieses Vorgehen wird sich auch in der zu entwickelnden digitalen Heuristik für die CATMA-Umgebung niederschlagen: Dort sollen die automatischen Annotationsvorschläge von einer vorherigen Wahl bestimmter Grundannahmen durch den individuellen Nutzer und die individuelle Nutzerin abhängig sein. Dieses *instant feedback*-Konzept legt die methodischen Annahmen, die im Modul implementiert werden, zu einem großen Teil offen. Das Modul ist dadurch keine *black box* mehr, deren Funktionsweise und Methodik verborgen bleibt, und kann darüber hinaus von den Nutzern zur Überprüfung der eigenen narratologischen Vorannahmen – und der Konsistenz bei ihrer Anwendung – herangezogen werden.

Die beschriebenen Maßnahmen zum Umgang mit den sich aus interdisziplinärer Arbeit ergebenden Problemen führen also zu erstaunlich positiven Konsequenzen für die Einzelwissenschaften: Fachwissenschaftliche Dogmen werden durch den interdisziplinären Zugang fast schon notwendigerweise überdacht und theoretische bzw. praktische Unzulänglichkeiten ebenso behoben. Außerdem ergeben sich für die beiden Fachdisziplinen auch intradisziplinär positive Effekte. Das diesbezügliche Potential für den NLP-Zugang haben wir bereits skizziert. Im folgenden Abschnitt möchten wir nun beispielhaft einige Aspekte narratologischer Begriffsbildung ausführlicher vorstellen, die durch die interdisziplinäre Arbeit angeregt wurde.

3.4 Der Mehrwert für die Narratologie

Obwohl narratologische Kategorien gemeinhin als theoretisch durchdacht und leicht operationalisierbar gelten, zeigten sich bei ihrer formalisierten Anwendung im Rahmen manueller kollaborativer Annotation in heureCLÉA einige theoretische Unzulänglichkeiten. Typischerweise werden solche Unzulänglichkeiten, wie oben bereits angemerkt, entdeckt, wenn sich die Annotatoren hinsichtlich der korrekten narratologischen Bestimmung konkreter Textstellen nicht einig sind. In den Diskussionen über die Gründe für individuelle Annotationsentscheidungen stellte sich oft die uneindeutige oder unvollständige Konzeption der jeweiligen narratologischen Kategorie als Ursache uneinheitlichen Markups heraus. Die festgestellten theoretischen Versäumnisse lassen sich in zwei Gruppen einteilen, die je unterschiedliche Problemlösungsstrategien erfordern:

(1) konzeptionelle Unvollständigheiten, die leicht durch eine Vervollständigung der Kategorie mittels funktionaler Entscheidungen behoben werden können. Stellt sich bei

der versuchten Anwendung einer Kategorie heraus, dass ihre Definition zu vage ist, um die Bestimmung einer fraglichen Textstelle vorzunehmen, müssen pragmatische Entscheidungen im Hinblick auf die Inklusion oder Exklusion bisher nicht bedachter textueller Oberflächenmerkmale getroffen werden.⁴⁵ Derartige Entscheidungen haben nur für die Anwendung der jeweiligen Kategorie Konsequenzen, nicht aber für weitere Konzepte;

(2) theoretische Unvollständigkeiten, die ihrerseits auf die unzureichende Bestimmung fundamentaler narratologischer Konzepte zurückzuführen sind. Probleme dieses Typs können nicht einfach durch pragmatische Entscheidungen behoben werden, weil die für eine Problemlösung notwendigen Setzungen auf der Ebene grundlegender narratologischer Konzepte weitreichende Konsequenzen für viele erzähltheoretische Einzelkategorien nach sich ziehen. Im Folgenden soll diese zweite Problemklasse anhand eines Beispiels erläutert werden:

In der Erzählung *Der Tod* von Thomas Mann ist bei dem Vergleich der Annotationsergebnisse im Hinblick auf die Geschwindigkeit der Erzählung⁴⁶ folgende Passage in den Fokus der Aufmerksamkeit gerückt (vgl. Abbildung 2):

»Ich habe die ganze Nacht hinausgeblickt, und mich dünkte, so müsse der Tod sein oder das Nach dem Tode: dort drüben und draußen ein unendliches, dumpf brausendes Dunkel. Wird dort ein Gedanke, eine Ahnung von mir fortleben und -weben und ewig auf das unbegreifliche Brausen horchen?«⁴⁷

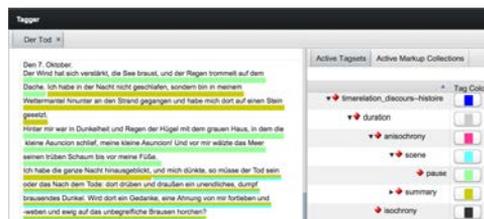


Abb. 2: Annotation von *Der Tod* in CATMA (Grafik: Gius / Jacke 2014).

Die Passage nach dem ersten Komma (»und mich dünkte, so müsse...«) wurde von einigen Annotatoren als »zeitraffend erzählt« eingeordnet, von anderen dagegen als »zeitdehnendes Erzählen«. Die Diskussion über die Gründe für die individuellen Entscheidungen hat gezeigt, dass die Annotatoren unterschiedliche Auffassungen darüber vertreten, was ein Ereignis ist. Betrachtet man mentale Vorgänge als Ereignisse, so muss man die zitierte Passage als zeitraffend klassifizieren, da die Gedanken des Erzählers in der fiktiven Welt vermutlich

⁴⁵ Ein derartiger Problemfall stellte sich an folgender Textstelle in Friedrich Hebbels Erzählung *Matteo* im Hinblick auf die Frage, ob es sich hier um eine Prolepse – bisher konzeptionalisiert als Vorgriff in der Zeit – handelt: »Sieh, morgen feire ich meine Hochzeit; zum Zeichen, daß du mir nicht mehr böse bist, kommst du auch, meine Mutter wird dich gern sehen.« (Hebbel: *Matteo* [1841], S. 466). Die Schwierigkeit ist hier dadurch gegeben, dass die angesprochene Figur am folgenden Tag nicht auf der Hochzeit erscheint. Um entscheiden zu können, ob hier eine Prolepse vorliegt, muss entschieden werden, ob dieses Konzept auch antizipierte Ereignisse fassen soll, die im Verlauf der Erzählung nicht eintreten.

⁴⁶ Unter »Erzählgeschwindigkeit« versteht man in der Narratologie das Verhältnis zwischen der Menge an Ereignissen, von denen berichtet wird, und der Zeit, die für diesen Bericht notwendig ist.

⁴⁷ Mann: *Der Tod* [1897], S. 76.

längere Zeit anhielten als die wenigen Sekunden, die in der Erzählung für ihre Wiedergabe eingeräumt werden. Ist man jedoch der Ansicht, dass es sich bei mentalen Prozessen nicht um Ereignisse handelt, so liegt in obiger Textstelle zeitdehnendes Erzählen vor: Der Bericht von Ereignissen wird verlangsamt durch die Darstellung nicht-ereignishafter Gegebenheiten. Die Frage danach, welche Konzeption von Ereignis korrekt oder sinnvoll ist, ist Gegenstand der Debatte um Narrativität: die für erzählende Texte konstitutive Eigenschaft, von Ereignissen zu berichten. Die unterschiedlichen Intuitionen der Annotatoren in Bezug auf die Definition von ›Ereignis‹ korreliert hier mit Schmidts Konzeptionen von ›Ereignis I‹, das jegliche Form von Zustandsveränderung inkludiert, und ›Ereignis II‹, das zusätzliche Kriterien anführt, die Zustandsveränderungen aufweisen müssen, um als Ereignis zu gelten.⁴⁸ Wertet man Faktizität – also die Eigenschaft einer Zustandsveränderung, in der fiktiven Außenwelt einer Erzählung tatsächlich stattzufinden – als notwendige Eigenschaft von Ereignissen, muss die zitierte Passage aus *Der Tod* als Erzählpause klassifiziert werden.

Eine Entscheidung im Hinblick auf die ›richtige‹ Narrativitätsdefinition, die für die Lösung von Annotationsproblemen im Bereich der Erzählgeschwindigkeit notwendig wäre, hätte nun nicht nur für die fraglichen Kategorien Konsequenzen, sondern beispielsweise auch für die Bestimmung des Gegenstandsbereichs der Narratologie und potenziell für eine Reihe weiterer Kategorien.⁴⁹

Obwohl die theoretische Arbeit an narratologischen Grundkonzepten zu Beginn von *heureCLÉA* nicht geplant war, hat sich im Verlauf des Projekts, wie in [Abschnitt 3.2](#) und [3.3](#) dargelegt, gezeigt, dass auch für diese komplexeren Theorieprobleme eine Lösung gefunden werden muss. Nur wenn bei widersprüchlichen Annotationen die jeweils zugrunde liegenden Basiskonzepte oder -annahmen offengelegt werden, sind die generierten Metadaten aus der NLP-Perspektive verwertbar. Aus diesem Grund arbeiten wir aktuell unter anderem an einer Analyse der literaturwissenschaftlichen Ereigniskonzeption(en), um eine Zuordnung diskrepanter Erzähldauer-Annotationen zu den verschiedenen Konzeptionen zu ermöglichen. Dieses Vorgehen ebnet den Weg für eine Parametrisierung der durch die digitale Heuristik generierten automatischen Annotationsvorschläge, wie sie in [Abschnitt 3.3](#) bereits skizziert wurde.

4. Fazit

Das Zusammenspiel von Informatik und Hermeneutik, das das Projekt *heureCLÉA* auszeichnet, scheint also sowohl für das Projektziel der digitalen Heuristik als auch für die beteiligten Disziplinen selbst äußerst produktiv zu sein. Das Faszinierende daran ist das Wechselspiel zwischen interdisziplinärer und intradisziplinärer Arbeit: Beide Disziplinen

⁴⁸ Vgl. die Erläuterungen zu ›Ereignis I‹ und ›Ereignis II‹ in [Abschnitt 3](#).

⁴⁹ Eine ähnliche Verknüpfung von Annotationsproblemen und ungeklärten narratologischen Basiskonzepten findet sich bei der Annotation von Phänomenen der zeitlichen Ordnung einer Erzählung einerseits und dem grundlegenden narratologischen Konzept der Erzählebenen. Es kann nur sinnvoll das zeitliche Verhältnis von solchen Ereignissen bestimmt werden, die sich auf derselben Erzählebene befinden. Anhand welcher Faktoren ein Ebenenwechsel festzumachen ist, wird in der narratologischen Forschung noch diskutiert (vgl. Ryan 1991, S. 175–177; Pier 2013, unpag.).

bewegen sich aufeinander zu, um etwas Gemeinsames, eine Heuristik, schaffen zu können. Dies wird aber nicht dadurch erreicht, dass die angewendeten disziplinären Zugänge soweit aneinander angenähert werden, dass ein gemeinsamer interdisziplinärer Zugang entsteht. Vielmehr macht die Konfrontation mit den Methoden, Annahmen und Begriffen der jeweils anderen Disziplin eine Reflexion der eigenen disziplinären Methodologie nötig. Die in diesem Artikel dargestellten Problematiken sind nämlich durchweg spezifisch literaturwissenschaftlich oder spezifisch computerlinguistisch, und das, obwohl – oder gerade weil? – sie durch Rahmenbedingungen erzeugt werden, die wiederum exemplarisch für das Zusammenspiel von Informatik und Geisteswissenschaften sind. So wird etwa die Reproduzierbarkeit von Analyseergebnissen, die in *heureCLÉA* durch die *noise*-Problematik in der NLP relevant wurde, in den Geisteswissenschaften traditionell nicht thematisiert, da die Geisteswissenschaften meist mit dem wesentlich schwächeren Konzept der intersubjektiven Übereinstimmung operieren, ohne dieses weiter zu bestimmen. Die durch die Zusammenarbeit gewissermaßen erzwungene Auseinandersetzung mit der Reproduzierbarkeit von Analyseergebnissen kann aber, wie wir gezeigt haben, eine relevante Theorieentwicklung nach sich ziehen. Darüber hinaus hat sich die Reproduzierbarkeit von Analyseergebnissen als eine bislang wenig erforschte Gelingensbedingung für interdisziplinäre Projekte im Bereich der Digital Humanities herausgestellt. Für die Informatik bedeutet die Zusammenarbeit u.a., dass sie einen neuen Zugang zu bekannten Problemen entwickelt, der für hermeneutische – und nicht nur für mathematisch-formale Probleme – relevant ist.

Die bisherige und weiter geplante Arbeit in *heureCLÉA* ist also sowohl von entscheidender Bedeutung für das Gelingen des Projekts als auch ein Beitrag zu disziplinären Erkenntnissen oder methodologischen Weiterentwicklungen. Die Grundlage dafür liegt in den Rahmenbedingungen des Projekts bzw. in seinem Erkenntnisinteresse: Dadurch, dass die mit der digitalen Heuristik verbundenen Fragestellungen aus beiden beteiligten Feldern kommen, wird ein methodischer Mehrwert erzielt – der wiederum die Qualität der Ergebnisse steigert, da der Arbeit an der Fragestellung eine gemeinsame Auseinandersetzung mit den Gelingensbedingungen vorausgeht.

Die methodischen Konsequenzen echter interdisziplinärer Zusammenarbeit im Bereich der Digital Humanities bedeuten zwar zuerst einen Mehraufwand, der insbesondere durch die notwendig werdenden Übersetzungs- und Anpassungsleistungen zwischen den beteiligten Disziplinen bedingt ist. Wird diese Notwendigkeit aber erkannt und mit angemessener Sorgfältigkeit und unter Berücksichtigung der jeweiligen disziplinären Annahmen und Erkenntnisse in die Forschungsarbeit integriert, so wird offensichtlich, welche Schätze mit einem Digital Humanities-Zugang – sowohl disziplinär als auch interdisziplinär – zu bergen sind.

Bibliographische Angaben

Thomas Bögel / Michael Gertz / Evelyn Gius / Janina Jacke / Jan Christoph Meister / Marco Petris / Jannik Strötgen: Collaborative Text Annotation Meets Machine Learning. heureCLÉA, a Digital Heuristics of Narrative. In: DHCommons Journal 1 (2015) [\[online\]](#).

Thomas Bögel / Jannik Strötgen / Michael Gertz: Computational Narratology. Extracting Tense Clusters from Narrative Texts, präsentiert auf der 9. Language Resources and Evaluation Conference (LREC '14). Reykjavík 2014. [\[online\]](#)

Axel Bühler: Grundprobleme der Hermeneutik. In: Hermeneutik. Basistexte zur Einführung in die wissenschaftstheoretischen Grundlagen von Verstehen und Interpretation. Hg. von Axel Bühler. Heidelberg 2003, S. 3–19. [\[Nachweis im GBV\]](#)

Dino Buzzetti: Digital Representation and the Text Model. In: New Literary History 33 (2002), S. 61–88. DOI [10.1353/nlh.2002.0003](#).

Nathanael Chambers / Dan Jurafsky: Unsupervised Learning of Narrative Schemas and Their Participants, präsentiert auf der Association for Computational Linguistics Konferenz (ACL 2009). Singapur 2009. [\[online\]](#)

James H. Coombs / Allen H. Renear / Steven J. DeRose: Markup Systems and the Future of Scholarly Text Processing. In: Communications of the ACM 30 (1987), S. 933–947. DOI [10.1145/32206.32209](#).

Lutz Danneberg: Zum Autorkonstrukt und zu einer methodologischen Konzeption der Autorintention. In: Rückkehr des Autors. Zur Erneuerung eines umstrittenen Begriffs. Hg. von Fotis Jannidis / Gerhard Lauer / Matías Martínez / Simone Winko. Tübingen 1999, S. 77–105. [\[Nachweis im OPAC\]](#)

Defining Digital Humanities. A Reader. Hg. von Melissa Terras / Julianne Nyhan / Edward Vanhoutte. Farnham 2013. [\[Nachweis im OPAC\]](#)

Wilhelm Dilthey: Die Entstehung der Hermeneutik [1900]. In: Wilhelm Dilthey: Gesammelte Schriften. Bd. 5. 8. Auflage. Stuttgart, Göttingen 1990, S. 317–338. [\[Nachweis im GBV\]](#)

Dagfinn Føllesdal / Lars Walløe / Jan Elster: Die hypothetisch-deduktive Methode in der Literaturinterpretation. In: Moderne Interpretationstheorien. Ein Reader. Hg. von Tom Kindt / Tilmann Köppe. Göttingen 2008, S. 67–78. [\[Nachweis im GBV\]](#)

Harald Fricke: Methoden? Prämissen? Argumentationsweisen. In: Vom Umgang mit Literatur und Literaturgeschichte. Positionen und Perspektiven nach der Theoriedebatte. Hg. von Lutz Danneberg / Friedrich Vollhardt. Stuttgart 1992, S. 211–227. [\[Nachweis im GBV\]](#)

Hans-Georg Gadamer: Hermeneutik. Wahrheit und Methode [1960]. In: Hans-Georg Gadamer: Gesammelte Werke, Bd. 2. 2., durchgesehene Auflage. Tübingen 1993. [\[Nachweis im GBV\]](#)

Gérard Genette: Discours du Récit. In: Gérard Genette: Figures III. Paris 1972, S. 67–282. [\[Nachweis im OPAC\]](#)

Evelyn Gius: Erzählen über Konflikte: ein Beitrag zur digitalen Narratologie. Berlin, Boston 2015. [\[Nachweis im GBV\]](#)

Evelyn Gius / Janina Jacke: Zur Annotation narratologischer Kategorien der Zeit. Guidelines zur Nutzung des CATMA-Tagsets. Hamburg 2014. [\[online\]](#)

Friedrich Hebbel: Matteo [1841]. In: Werke. Hg. von Gerhard Fricke / Werner Keller / Karl Pörnbacher. Bd. 3. München 1963, S. 465–476. [\[online\]](#)

Janina Jacke: Is There a Context-Free Way of Understanding Texts? The Case of Structuralist Narratology. In: Journal of Literary Theory 8 (2014), S. 118–139. [\[Nachweis im GBV\]](#)

Fotis Jannidis: Polyvalenz – Konvention – Autonomie. In: Regeln der Bedeutung. Zur Theorie der Bedeutung literarischer Texte. Hg. von Fotis Jannidis / Gerhard Lauer / Matías Martínez / Simone Winko. Berlin 2003, S. 305–328. [\[Nachweis im GBV\]](#)

Fotis Jannidis / Gerhard Lauer / Matías Martínez / Simone Winko: Der Bedeutungsbegriff in der Literaturwissenschaft. Eine historische und systematische Skizze. In: Regeln der Bedeutung. Zur Theorie der Bedeutung literarischer Texte. Hg. von Fotis Jannidis / Gerhard Lauer / Matías Martínez / Simone Winko. Berlin 2003, S. 3–30. [\[Nachweis im GBV\]](#)

Fotis Jannidis: Analytische Hermeneutik. Eine vorläufige Skizze. In: Heuristiken der Literaturwissenschaft. Hg. von Uta Klein / Katja Mellmann / Stefanie Metzger. Paderborn 2006, S. 131–144. [\[Nachweis im GBV\]](#)

Tom Kindt / Tilmann Köppe: Moderne Interpretationstheorien. Eine Einleitung. In: Moderne Interpretationstheorien. Ein Reader. Hg. von Tom Kindt / Tilmann Köppe. Göttingen 2008, S. 7–26. [\[Nachweis im GBV\]](#)

Tom Kindt / Hans-Harald Müller: Wieviel Interpretation enthalten Beschreibungen? Überlegungen zu einer umstrittenen Unterscheidung am Beispiel der Narratologie. In: Regeln der Bedeutung. Zur Theorie der Bedeutung literarischer Texte. Hg. von Fotis Jannidis / Gerhard Lauer / Matías Martínez / Simone Winko. Berlin, New York 2003, S. 286–304. [\[Nachweis im GBV\]](#)

Tom Kindt / Hans-Harald Müller: Zum Verhältnis von Deskription und Interpretation. Ein Bestimmungsvorschlag und ein Beispiel. In: Literatur interpretieren: interdisziplinäre Beiträge zur Theorie und Praxis. Hg. von Jan Borkowski / Stefan Descher / Felicitas Ferder / Philipp David Heine. Münster 2015. [\[online\]](#)

Silke Lahn / Jan Christoph Meister: Einführung in die Erzähltextanalyse. Stuttgart, Weimar 2013. [\[Nachweis im OPAC\]](#)

Peter Lipton: *Inference to the Best Explanation*. London, New York 1991. [[Nachweis im GBV](#)]

Inderjeet Mani: *Computational Narratology*, in: *the living handbook of narratology*. Hg. von Peter Hühn / John Pier / Wolf Schmid / Jörg Schönert. Hamburg 2013. [[online](#)]

Inderjeet Mani / Barry Schiffman: *Temporally Anchoring and Ordering Events in News*. In: *Time and Event Recognition in Natural Language*. Hg. von James Pustejovsky / Robert Gaizauskas. Amsterdam, Philadelphia 2005. [[online](#)]

Thomas Mann: *Der Tod*. In: *Thomas Mann: Große kommentierte Frankfurter Ausgabe. Werke – Briefe – Tagebücher*. Hg. von Heinrich Detering. Bd. 2.1. Frankfurt/Main 2004, S. 71–78. [[Nachweis im OPAC](#)]

Christopher D. Manning / Hinrich Schütze: *Foundations of Statistical Natural Language Processing*. 4. Auflage. Cambridge, Mass. 2001. [[Nachweis im GBV](#)]

John Pier: *Narrative Levels*. In: *the living handbook of narratology*. Hg. von Peter Hühn / John Pier / Wolf Schmid / Jörg Schönert. Hamburg 2013. [[online](#)]

Wendell Piez: *Towards Hermeneutic Markup. An Architectural Outline*. In: *Digital Humanities 2010. Conference Abstracts*, präsentiert auf der Digital Humanities Conference 2010 (DH 2010). London 2010, S. 202–205. [[online](#)]

Marie-Laure Ryan: *Possible Worlds, Artificial Intelligence, and Narrative Theory*. Bloomington 1991. [[Nachweis im GBV](#)]

Friedrich Schleiermacher: *Vorlesungen zur Hermeneutik und Kritik [1805–1833]*. In: *Friedrich Schleiermacher. Kritische Gesamtausgabe*. Bd. 4. Berlin [u.a.] 2012. [[Nachweis im OPAC](#)]

Wolf Schmid: *Elemente der Narratologie*. Berlin [u.a.] 2008. [[Nachweis im GBV](#)]

Desmond Schmidt: *The Inadequacy of Embedded Markup for Cultural Heritage Texts*. In: *Literary and Linguistic Computing* 25 (2010), S. 337–356. DOI 10.1093/llic/fqq007.

Arthur Schnitzler: *Blumen [1894]*. In: *Arthur Schnitzler: Ausgewählte Werke in acht Bänden*. Bd. 1: *Leutnant Gustl. Erzählungen 1892–1907*. Frankfurt/Main 2004, S. 98–107. [[Nachweis im GBV](#)]

Jannik Strötgen / Michael Gertz: *HeidelTime. High Quality Rule-Based Extraction and Normalization of Temporal Expressions*. In: *Proceedings of the 5th International Workshop on Semantic Evaluation*, präsentiert auf der Association for Computational Linguistics Conference. Uppsala 2010, S. 321–324. [[online](#)]

Abbildungslegenden und -nachweise

Abb. 1: Klassifikation von DH-Ansätzen (Grafik: Gius / Jacke 2014).

Abb. 2: Annotation von *Der Tod* in CATMA (Grafik: Gius / Jacke 2014).

Zeitschrift für digitale Geisteswissenschaften

Artikel aus:

Sonderband 1 der ZfdG: Grenzen und Möglichkeiten der Digital Humanities. Hg. von Constanze Baum und Thomas Stäcker. 2015. DOI: [10.17175/sb01](https://doi.org/10.17175/sb01)

Titel:

Warum werden mittelalterliche und frühneuzeitliche Rechnungsbücher eigentlich nicht digital ediert?

Autor/in:

Georg Vogeler

Kontakt:

georg.vogeler@uni-graz.at

Institution:

Universität Graz

GND:

[1035297884](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63862-p0011-9)

ORCID:

[0000-0002-1726-1712](https://orcid.org/0000-0002-1726-1712)

DOI des Artikels:

[10.17175/sb001_007](https://doi.org/10.17175/sb001_007)

Nachweis im OPAC der Herzog August Bibliothek:

[83016684X](#)

Erstveröffentlichung:

19.02.2015

Lizenz:

Sofern nicht anders angegeben 

Medienlizenzen:

Medienrechte liegen bei den Autoren

Letzte Überprüfung aller Verweise:

24.05.2016

GND-Verschlagwortung:

[Rechnungsbuch](#) | [Edition](#) |

Zitierweise:

Georg Vogeler: Warum werden mittelalterliche und frühneuzeitliche Rechnungsbücher eigentlich nicht digital ediert?. In: Grenzen und Möglichkeiten der Digital Humanities. Hg. von Constanze Baum / Thomas Stäcker. 2015 (= Sonderband der Zeitschrift für digitale Geisteswissenschaften, 1). PDF Format ohne Paginierung. Als text/html abrufbar unter DOI: [10.17175/sb001_007](https://doi.org/10.17175/sb001_007).

Georg Vogeler

Warum werden mittelalterliche und frühneuzeitliche Rechnungsbücher eigentlich nicht digital ediert?

Abstracts

Mittelalterliches und frühneuzeitliches Rechnungsschriftgut ist für die Auswertung mit dem Computer gut geeignet, da es sich leicht in Datenformate umwandeln lässt, die sich an Tabellen orientieren. Forschung und Editionspraxis haben gleichzeitig gezeigt, dass das wissenschaftliche Interesse an dieser Schriftgutart so vielfältig ist, dass sie bevorzugt digital ediert werden sollte. Dennoch kann der Beitrag nur eine einzige digitale Edition anführen, die auch eine numerische Auswertung von Rechnungsbüchern vorsieht. Während die bisherige Forschung zur Digitalen Edition sich auf philologische und paläographische Editionsprobleme konzentriert hat, entwickelt der Beitrag ein Editionsmodell, das auch den Inhalt der Rechnungen einbezieht. Das auf dem Resource Description Framework (RDF) beruhende Modell von Referenzen zwischen den verschiedenen Repräsentationsformen kann mit den Kodierungsrichtlinien der Text Encoding Initiative (TEI) serialisiert werden. Es wird schließlich demonstriert, dass so auch umfangreiche Bestände effizient digital ediert werden können.

Medieval and early modern account books are well suited for computer analysis, since they can be easily converted to tabular data formats. Scholarly research and the creation of critical editions have simultaneously demonstrated that research interest on this type of textual material is so diverse that they should ideally be digitally edited. However, this paper can only point to one single digital edition that also allows the numeric analysis of account books. While previous research on digital editions has concentrated on philologic and paleographic problems, this paper proposes an edition model that also includes the contents of the accounts themselves. Based on the Resource Description Framework (RDF), this model of references between the different forms of representation can be serialized with the encoding guidelines of the Text Encoding Initiative (TEI). Finally, this paper will demonstrate that extensive collections can also be efficiently be digitally edited in this way.

1. Mittelalterliche und frühneuzeitliche Rechnungen in der Forschung und der bisherigen Editionspraxis

Historische Rechnungsdokumente scheinen auf den ersten Blick hervorragend geeignet für die computergestützte Reproduktion: Erstens sind sie hoch strukturiert. Zweitens enthalten sie große Mengen an Einzelinformationen, die individuell nicht immer von hohem Quellenwert sind, als Aggregation jedoch Bedeutung gewinnen. Drittens enthalten sie Zahlen, mit denen gerechnet werden kann. In der Zeit zwischen 1960 und 1990, in denen die Kontaktzone zwischen Informatik und Geisteswissenschaften stark von sozialhistorischen Interessen geprägt war, war die Verbindung zwischen historischen Rechnungsdokumenten und digitalen Methoden entsprechend eng. Gerade die Quantifizierung als Methode historischer Forschung nutzte die Möglichkeiten, die Computertechnologien boten.¹ Aus den Rechnungen wurden

¹ Bergier / Solari 1963; Cerwenka / Roth 1972; Wrigley 1973; Granasztói 1971; Gundlach / Lueckerath 1976; Kropf 1969; Shorter 1971; Tilly 1973; Herlihy / Klapisch-Zuber 1978; Fossier et al. 1977; Arnold 1974.

individuelle Datenbanken und tabellarische Darstellungen erstellt. Als gemeinsamer ›Standard‹ haben sich für solche tabellarischen Darstellungen von wirtschaftshistorischen Daten einfache Tabellenformate (z.B. tabulatorgetrennte Textdateien oder kommaseparierte Werte) und das Format der Tabellenkalkulationssoftware MS Excel durchgesetzt.²

Dieser Ansatz übergeht jedoch wichtige Informationen des originalen Dokuments, die insbesondere bei mittelalterlichen und frühneuzeitlichen Rechnungen von Bedeutung sein können.³ Die jüngst erschienene kritische Edition der Stadtrechnungen von Luxemburg⁴ zeigt, dass Buchführung auch eine wichtige sprachhistorische Quelle ist, die Auskunft über Orthographie, Vokabular und Fachsprachen geben kann.⁵ Eine digitale Ressource, die nur buchhalterische Informationen wie Beträge, Konten und stichwortartige Beschreibungen der Buchungen enthält, verringert ihren sprachhistorischen Quellenwert signifikant.

Forschungen zur Verwaltungsgeschichte, zur pragmatischen Schriftlichkeit und zur Amtsbücherkunde haben gezeigt, dass Rechnungslegung und ihre Verschriftlichung im Mittelalter ein Prozess war, der mehrere Schritte umfasste: Vorbereitung der Rechnung aus informellen Dokumenten (Belege, vorläufige Rechnungsregister u.ä.), Zusammenstellung in einer Reinschrift, mündliche Rechnungslegung vor einer Rechnungskontrollinstanz oder zumindest vor einer Person, die mit Hilfe des Abakus die Rechenprozesse nachvollzog, Aktualisierung von Schuldposten im Zuge der Erstattung durch den Debitor sowie Umwandlung von Soll-Buchungen in Ist-Zahlungen insbesondere in der Steuerverwaltung und in Grundherrschaften. Mittelalterliche Rechnungen waren also nicht nur Finanzdokumente, sondern Protokolle von finanztechnischen Operationen.⁶

Auch das Layout von mittelalterlichen und frühneuzeitlichen Rechnungen war nicht stabil. Rechnungen der Verwaltung entwickelten sich im *Regnum Teutonicum* zum Beispiel von Protokollen in längeren Textblöcken zu komplexen tabellarischen Formen.⁷ Frühe Formen der doppelten Buchhaltung notierten die Debit- und Credit-Buchungen untereinander, andere benutzten dafür zwei verschiedene Teile des Rechnungsbuches, und erst in jüngerer Zeit entstand die parallele Notation in einer Zeile.⁸ Die Entwicklung der visuellen Form ist also ebenfalls Teil der Erforschung der Geschichte des Rechnungswesens.

² Vgl. z.B. Early Forward Exchange Markets, 1876–1914 (auf der Basis von Flandreau / Komlos 2006) oder die Zeitreihen der Historischen Statistik im Archiv der GESIS (Für alle angegebenen Links in diesem Beitrag wurde, wenn möglich, eine Kopie mit dem [Capturing-Dienst des Internet Archive](#) erstellt. Die Texte sind auch unter [https://web.archive.org/web/<*/\[URL\]>](https://web.archive.org/web/<*/[URL]>) abrufbar).

³ Ausführlicher zu den Eigenschaften von Rechnungen und ihrer Umsetzung in gedruckten Editionen Vogeler 2015.

⁴ Moulin / Pauly 2007–2012.

⁵ Vgl. zur sprachhistorischen Auswertung von Rechnungen auch Moulin 2012; Herborn / Mattheier 1978; Kolbeck / Rössler 2015; Ravida 2012.

⁶ Vgl. Heß 1977 und Heß 1996; Mersiowsky 2000, S. 337–340.

⁷ Die Entwicklung in spätmittelalterlichen Territorialrechnungen beschreibt Mersiowsky 2000, im Schriftgut der Steuerverwaltung Vogeler 2003–2004.

⁸ Melis 1950; Arlinghaus 2004; Yamey 1949.

Rechnungsbücher sind schließlich nicht nur eine ökonomische Quelle, sondern auch eine gute Quelle für Alltagsleben, Sozialgeschichte und Sachkultur.⁹ Diese Quelleninhalte erfordern die Verbindung der Buchungen mit Taxonomien, die nicht nach buchhalterischen Kriterien aufgebaut sind: Berufsgruppen, Warengruppen, Bevölkerungsschichten etc. Die Erschließung der Rechnungstexte allein nach finanztechnischen Gesichtspunkten muss dafür um diese Arten inhaltlicher Erschließung ergänzt werden.

Als gängige Lösung für das Problem unterschiedlicher Forschungsinteressen an Rechnungsbüchern konzentriert sich eine gedruckte Edition auf eine einzelne Perspektive: Während z.B. die erwähnten Editionen der Luxemburger Stadtrechnungen aus sprachhistorischem Forschungsinteresse heraus auf eine diplomatische Transkription Wert legen, edierte Richard Knipping mit wirtschaftshistorischem Interesse die Kölner Stadtrechnungen in tabellarischer Form.¹⁰ Im Medienkontext des Buchdrucks ist es wirtschaftlich nicht vertretbar, beiden Forschungsinteressen gerecht zu werden und eine Rechnung sowohl als Volltexttranskription als auch tabellarisch zu publizieren.

2. Digitale Editionen von Rechnungsschriftgut

Auf dem aktuellen Stand der theoretischen Diskussion über die Digitale Edition scheint das Problem gelöst: Eine kritische Edition kann als digitale Ressource mehrere Interpretationsschichten einschließen und dem Benutzer die Auswahl der für ihn nützlichen Präsentationsform überlassen.¹¹ Eine überblicksartige Recherche nach digitalen Ressourcen von mittelalterlichen und frühneuzeitlichen Rechnungsbüchern zeigt jedoch, dass diese Editionen das theoretisch formulierte Potential jeweils nur eingeschränkt realisieren. Als Beispiele für ambitionierte digitale Editionen von Rechnungsmaterial können folgende drei Projekte dienen: Mitarbeiter der Universität Lyon stellen sich der Aufgabe, eine umfangreiche Serie von Rechnungen digital zu edieren, indem sie von den 157 in zwei Archiven aufbewahrten Rechnungen der Grafen von Savoyen vom Ende des 13. Jahrhunderts Abbildungen und TEI-Transkriptionen online verfügbar machen.¹² Matthias Perstling hat das über mehrere Jahre hinweg benutzte Steirische Marchfutterurbar vom Anfang des 15. Jahrhunderts so digital ediert, dass die verschiedenen Bearbeitungsschichten in Zahlungsvermerken und Änderung der Namen der Abgabepflichtigen nachvollziehbar werden.¹³ Das Team um Jürgen Sarnowsky hat die Schuld- und Rechnungsbücher des Deutschen Ordens um 1400 unter Einschluss einer Taxonomie der Geschäftsvorfälle und Warengruppen digital ediert.¹⁴ Diese und weitere Beispiele¹⁵ bieten Faksimiles zu den Transkriptionen, erlauben eine normalisierte und auf Taxonomien aufgebaute Suche oder visualisieren die klassifizierten Informationen.

⁹ Z.B. Boshof 2008; Mersiowsky 1995; Huys 1998; Jaritz 1976; Jaritz 1984, S. 39–42; Katzler 2008; Knittler 2005. Umfangreich zum Quellenwert von Rechnungsschriftgut Hohnfeld 1912; Jappe Alberts 1958; Seidel 1911; Wirth 2010.

¹⁰ Knipping 1897–1898.

¹¹ Zum Stand der Digitalen Edition als Methode vgl. Sahle 2013.

¹² Guilleré et al. 2003ff.

¹³ Perstling 2013b, zum Datenmodell und zum Inhalt Perstling 2013a.

¹⁴ Link / Sarnowsky 2007ff. Von den Rechnungen sind auch gedruckte Editionen erschienen: Heß et al. 2008; Link / Sarnowsky 2008.

¹⁵ Boshof / Brunnhofer 2006; Lodge 2006; Comédie-Française Register Project 2007ff.; Open Domesday; Henry III Fine Rolls Project 2007–2011.

Dennoch fehlt diesen Editionen eine zentrale Komponente der Quellengattung, nämlich ihre Eigenschaft als Rechenwerk. Dagegen gibt die Edition der Rechnungsbücher des Royal Irish college of Saint George the Martyr in Alcalá¹⁶ dem Benutzer und der Benutzerin nicht nur Zugriff auf Bilder von Transkriptionen der Rechnung, sondern ermöglicht es auch, mit den Buchungsposten zu rechnen. In einer Flash-Umgebung können die originalen Texte und ihre englische Übersetzung durchsucht werden. Die Treffer werden im Kontext ihrer jeweiligen Seite angezeigt. Sie können über eine Checkbox ausgewählt und in ein »Datashheet« überführt werden, das die ausgewählten Beträge summiert. Auch wenn leider ein Zugriff auf die Quelldaten oder eine Exportfunktion der Daten fehlt, ist diese digitale Edition die einzige, deren Datenmodell und Benutzerinterface für eine essentielle Eigenschaft des Ausgangsmaterials eine adäquate Form gefunden hat, nämlich den Text als Rechenwerk zu präsentieren.

3. Ein Modell für die inhaltsorientierte Edition von Rechnungsschriftgut

Man kann einige Gründe dafür anführen, dass die Editionsweise der Rechnungsbücher des Royal Irish college of Saint George the Martyr eine Seltenheit ist. Sie ergeben sich aus primären Interessenslagen von verschiedenen Forschergruppen. Eine Erklärung für die fehlende Umsetzung des Potentials digitaler Editionen für Rechnungsschriftgut liegt in der Dominanz philologischer Editionsmethoden und in der Forschungsdiskussion über digitale Editionen.¹⁷ So sind die Richtlinien der Text Encoding Initiative (TEI)¹⁸ als De-facto-Standard digitalen Editierens z.B. mehr an komplexen Überlieferungsverhältnissen, kodikologischen und paläographischen Details oder an linguistischen Phänomenen interessiert als an der Erschließung von Propositionen. Syd Baumann und Kathryn Tomasek haben deshalb erste Vorschläge erarbeitet, wie die TEI zu erweitern wäre, um Finanztransaktionen beschreiben zu können.¹⁹ Als alternativer Standard für die Kodierung von buchhalterischen Informationen steht mit XBRL (eXtended Business Reporting Language)²⁰ ein flexibler Vorschlag von wirtschaftswissenschaftlicher Seite zur Verfügung, der mit der Global Ledger-Taxonomie auch Kategorien für die Kodierung historischer Buchhaltung anbietet.²¹ Es stellt sich die Frage, ob beide Standards miteinander in Verbindung gebracht werden können.

Ausgehend von den Überlegungen, die Manfred Thaller 2012 vorgestellt hat,²² erscheint es angemessen, die verschiedenen Forschungsinteressen an Rechnungsdokumenten nicht in einem gemeinsamen XML-Modell wie der TEI, sondern in einem RDF-Modell abzubilden,²³ das den Text als Visualisierung von Informationen (»Bild«), als Spur von Schreibhandlungen sowie als sprachlicher Ausdruck (»Spur« und »Sprache«) und als Vermittler von Bedeutungen

¹⁶ Alcalá account book 2008.

¹⁷ Ausführlicher dazu Vogeler [im Erscheinen].

¹⁸ TEI P5, 2014.

¹⁹ Bauman 2013; Tomasek / Bauman 2013; Tomasek 2012.

²⁰ [eXtensible Business Reporting Language](#); [Global Ledger Taxonomy](#).

²¹ Ausführlicher dazu Vogeler [im Erscheinen].

²² Thaller 2012.

²³ RDF steht für »Resource Description Framework« und ist ein vom W3C entwickeltes Modell zur Veröffentlichung von strukturierten Daten im Internet.

repräsentiert, und damit für unterschiedliche Forschungsfragen zugänglich macht (vgl. Abbildung 1).

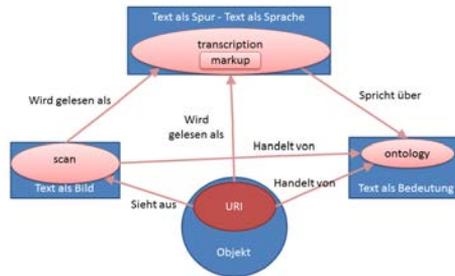


Abb. 1: Integration von Inhalt in ein Modell digitaler kritischer Editionen (Graphik: Autor).

Die Beziehung zwischen der visuellen, der textlichen und der inhaltlichen Repräsentation der Dokumente kann im Vokabular von Open Annotation Collaboration²⁴ als eine Instanz der Klasse #Annotation#[Code] verstanden werden, oder im Vokabular von CIDOC-CRM als eine #p67-Refers to#[Code]-Property.²⁵

Ein solches graphenbasiertes Modell nimmt Rücksicht darauf, dass die verschiedenen Forschungsinteressen zu Datenstrukturen führen können, die im hierarchischen Modell von Text in XML nur mit Hilfskonstrukten kodiert werden können. Das vorgeschlagene Modell schließt dabei aber die Verwendung der TEI nicht aus, denn es kann mit Hilfe von ›analytic notes‹ der TEI zumindest teilweise serialisiert werden.²⁶ Das Attribut #@ana#[Code] (für ›analysis‹) erlaubt es, beliebige TEI-Elemente mit Interpretationen zu verbinden. Das von der TEI bevorzugte linguistische oder dokumentarische Markup kann darüber auf eine Beschreibung der Buchungsfunktionalitäten, z.B. ausgedrückt als #tei:categories#[Code] oder #tei:interpretations#[Code], verweisen.

Mit einschlägigem Domänenwissen können diese Strukturen mit einfachem XSLT²⁷ in explizites RDF verwandelt werden, das auf den in der TEI vorhandenen Möglichkeiten zur Repräsentation von Taxonomien aufbaut. Dabei entstehen URIs für abstrakte Konzepte, die einerseits von den Texten repräsentiert werden und andererseits Instanzen der als Taxonomien formulierten Buchhaltungskonzepte sind. #@ana#[Code] repräsentiert dann einen Open Annotation Graph #{oa:Annotation oa:hasTarget ex:tei-frgm ; oa:hasBody ex:tei-frgm@ana-value-as-uri .}#[Code] oder direkt die CIDOC-CRM Property #p67#[Code].

Zusätzlich zu den explizit durch einen #@ana#[Code]-Verweis ausgedrückten Beziehungen kann die TEI-Kodierung weitere semantisch auswertbare Beziehungen ausdrücken. So

²⁴ Open Annotation Data Model 2013.

²⁵ CIDOC-CRM v. 5.

²⁶ Vgl. dazu den Abschnitt 17.3 *Spans and Interpretations*, in: TEI P5, 2014.

²⁷ XSLT steht für die eXtensible Stylesheet Language Transformation, die vom W3C spezifiziert wird, um XML-Daten von einer Dokumentengrammatik in eine andere zu transformieren (XSLT 2.0).

kann z.B. die Beziehung zwischen Konto und Unterkonto in der Verschachtelung von XML-Elementen `#tei:category#[Code]` kodiert werden. Ebenso können die einzelnen Buchungen Konten zugeordnet werden, indem sie als `#item#[Code]`-Elemente einer `#<tei:list ana="#gl_account #bk_konto1">#[Code]` oder als Absätze `#tei:p#[Code]` in einem Abschnitt `#<tei:div ana="#gl_account #bk_konto1">#[Code]` kodiert werden. Solche Kodierungen sind implizite Aussagen über die Bedeutung des Textes, denn Eigenschaften der Liste oder des Abschnitts können auf die einzelnen Listenelemente bzw. Absätze übertragen werden.²⁸

Weder die TEI noch die beschriebene Kodierung bieten eine Möglichkeit, die im gewählten hierarchischen XML-Dokumentmodell implizierten semantischen Referenzen formal zu explizieren. Welche Elementverschachtelungen semantische Beziehungen ausdrücken und welche nicht, ergibt sich nur aus der verbalen Beschreibungen der Bedeutung der gewählten Auszeichnung. Für die praktische Umsetzung einer Rechnungsedition ist das nicht von Belang, weil das notwendige Wissen über die Bedeutung der Kodierung bei der Entwicklung von Software zur Verarbeitung des semantischen Gehalts vorausgesetzt werden kann.

Die Frage ist relevant, wenn man TEI-Kodierung allgemein als Serialisierung des vorgeschlagenen Modells verwenden will. Ein möglicher Ansatz für eine technische Verallgemeinerung der Verwendung von XML-Strukturen als Indizien für semantische Aussagen besteht darin, die in `#@ana#[Code]` referenzierte Ontologie formal präziser z.B. als RDF-Semantics²⁹ oder als OWL³⁰ zu formulieren. Wenn dann die `#@ana#[Code]`-Referenz als eine `rdf:type`-Beziehung verstanden wird, können aus der formalen Ontologie Erwartungen an die semantische Bedeutung der XML-Struktur abgeleitet werden. So kann man z.B. in der Ontologie formulieren, dass eine RDF-Property, welche den Betrag einer Buchung kodiert, eine Buchung als RDFs-Domain erwartet. Davon ausgehend, dass solche Beziehungen in der Hierarchie der XML-Elemente ausgedrückt sind, kann angenommen werden, dass der nächste Knoten in der Vorfahrenachse im XML-Baum, der auf ein Konzept der entsprechenden Klasse verweist, als Domain zu verwenden ist.

Eine wie oben beschrieben in TEI kodierte Transkription ist klar von einer dokumentarischen Repräsentation zu trennen, welche Texte topographisch als beschriftete Bereiche einer Oberfläche beschreibt oder gar paläographisch detaillierte Beobachtungen über die graphische Gestalt der Zeichen zu kodieren versucht. Auch solche Eigenschaften von Text lassen sich kodieren und sind Teil der Kodierungsvorschläge der TEI geworden.³¹ Ihre Beziehung zur Textstruktur und den darin ausgedrückten Inhalten kann deshalb mit TEI `#facsimiles#[Code]` kodiert werden.³² Die `#@fac#[Code]`-Referenzen der TEI serialisieren aus dem obigen Modell die `#crm:p67-refers-to#[Code]`-Beziehung bzw. sind Stellvertreter für eine Annotation im Sinne von Open Annotation Collaboration. Es erscheint sogar realistisch, die Kodierungsvorschläge der TEI auch auf das auf Open Annotation

²⁸ Zum Bedeutungspotential von Listen Goody 1977.

²⁹ RDF 1.1 Semantics.

³⁰ Die formale Definition: OWL 2, 2012. Vgl. die [Übersicht](#) über alle Dokumente zu OWL.

³¹ Vgl. den Abschnitt [Representation of Primary Sources](#), in: TEI P5 2014.

³² Vgl. den Abschnitt 11.1 [Digital Facsimile](#), in: TEI P5 2014.

Collaboration aufbauende Shared Canvas-Modell abzubilden.³³ Als Referenzsystem werden dabei Koordinaten in den visuellen Repräsentationen der Handschrift verwendet. Es erscheint denkbar, aus den Koordinaten auf die inhaltliche Struktur zu schließen, denn im Laufe des späten Mittelalters haben sich Formen für die äußere Gestaltung von Rechnungsschriftgut entwickelt, welche wenigstens die einzelnen Buchungen, zunehmend auch die Beträge optisch voneinander abtrennen. Es wäre zu prüfen, ob die Funktion Spalten und Zeilen in einer tabellarischen Buchführung, wie sie sich seit dem 16. Jahrhundert zunehmend verbreitet hat, computergestützt aus der Geometrie der Textseiten ermitteln lässt.

4. Praktikabilität des Modells

Für die Implementierung solcher verallgemeinerten Transformationen fehlt noch der ›Proof of Concept‹. Denn für die Editionsarbeit ist ein theoretisch konsistentes Kodierungsmodell notwendig, um die Nachnutzung der Daten sicherzustellen. Ebenso wichtig ist jedoch, dass das Kodierungsmodell auch praktikabel ist. Gegenwärtig wird die oben beschriebene TEI-Kodierung in verschiedenen Projekten auf ihre praktische Umsetzbarkeit getestet. Es wird dabei auch nach Methoden gesucht, mit denen der Computer eine digitale Repräsentation des Dokuments »Rechnung« z.B. mit Hilfe von automatischen Umrechnungen, und die Umsetzung von Strukturen des XML-Codes in Bezüge zwischen den Sachaussagen unterstützt, sowie die zur Pflege von Taxonomien und kontrollierten Vokabularien eingesetzt werden. Die in [GAMS](#), dem Grazer auf Fedora Commons aufbauenden Digitalen Archiv für geisteswissenschaftliche Daten,³⁴ online gestellten Rechnungen, können belegen, dass der theoretische Ansatz für Rechnungseditionen realistisch umsetzbar ist.³⁵ Die fünf Beispiele aus gedruckten Editionen und neu erstellten Editionen konnten mit überschaubarem Arbeitseinsatz in das oben skizzierte TEI-Format konvertiert werden. Bei der Konversion von Textdokumenten, die als Vorlage für Drucke dienten, konnten umfangreiche Suchen mit regulären Ausdrücken eingesetzt werden, um z.B. römische Zahlen oder die Beträge am Schluss der Buchungen zu identifizieren.

Bei der Kodierung von neu erstellten Editionen ist darauf Rücksicht zu nehmen, dass sich die Technik gut in die Arbeitsabläufe der Forscherinnen und Forscher einfügen, die ihre Fachkompetenz in die Transkription und Erschließung des Textes investieren. Dafür wurde im Rahmen des Projekts der Basler Jahrrechnungen³⁶ eine TEI-Customization entwickelt, die es erlaubt, mit kurzen Bezeichnern, wie `<r:e>#[Code]` für jede Buchung, `<r:sum>#[Code]` für Summen oder `<r:b>#[Code]` für den Betrag der Buchung, die entsprechenden Textteile zu markieren. Die Mitarbeiterinnen und Mitarbeiter am Editionsprojekt haben sich darüber hinaus entschieden, die Geldbeträge mit Markup zu transkribieren, so dass `<r:lb>#[Code]`, `<r:sh>#[Code]` und `<r:d>#[Code]` für Pfund, Schilling und Pfennig stehen. In einer Transformation werden diese Angaben und die römischen Zahlen in numerische Werte und formale Kodierungen umgewandelt. Damit wird aus

³³ Shared Canvas v. 1.0.

³⁴ Vgl. Steiner / Stigler 2014, hier [Content models](#).

³⁵ Vgl. Vogeler 2013ff.

³⁶ Burghartz 2014ff.

```
»#<r:e>...<r:b><r:lb>ij<r:exp>c</r:exp> lxv</r:lb> <r:sh>v</r:sh> <r:d>j</r:d></r:b></r:e>#[Code]«
```

folgende Kodierung, die mit dem Standardschema der TEI konform ist:

```
»#<p ana="#gl_entryDetail">... <measure ana="#gl_amount" quantity="63661" unit="d" type="currency">ij<hi rend="superscript">c</hi>lxv lb v ß j d</measure></p>#[Code]«
```

Diese Buchungen werden mit Hilfe der Textstruktur in Sachgruppen, also Konten, geordnet, indem die Sequenz der Texte die einzelnen Buchungen einer Überschrift zuordnet. In der TEI werden dazu die hierarchischen Kodierungsformen von XML verwendet, welche für die Interpretation günstig sind, weil sie strikter sind als die Schriftformen des Mittelalters.³⁷

In der Testumgebung können die unterschiedlichen Editionen mit Hilfe von XSL-Stylesheets die römischen Zahlen in numerische Werte umgewandelt und diese Zahlen auf eine gemeinsame Rechnungsbasis, d.h. den Pfennig, umgerechnet werden. Die Zahlen werden als Beträge von Buchungen in einer Datenbank, d.h. konkret in einer **Sesame-Triple-Store**, gespeichert und können dort den in den TEI-Dokumenten enthaltenen Kontenhierarchien zugeordnet werden. Damit ist es möglich, effizient nach Buchungen mit schneller Berechnungen von Summen oder Darstellungen von numerischen Zeitreihen zu suchen ebenso wie eine klassische Editionsansicht oder tabellarische Präsentationen der Daten zu generieren. Die Daten selbst sind sowohl in der TEI-Kodierung wie in der Datenrepräsentation als RDF/XML zugänglich.³⁸

5. Fazit

Die angeführten Beispiele zeigen, dass das theoretisch entwickelte und in RDF formalisierbare Modell in praktischer Editionsarbeit verwendet werden kann und dabei Eigenschaften der Quelle digital repräsentiert werden können, die in der etablierten textorientierten Kodierung der TEI benachteiligt worden sind. Es entstehen also digitale Editionen, welche neben der visuellen und linguistischen Sicht auf den Text auch seine Bedeutung repräsentieren. Das an Hand von mittelalterlichen Rechnungen entwickelte Modell trägt damit dazu bei, theoretische Überlegungen aus dem Bereich der digitalen kritischen Edition, wie verschiedene Forschungsinteressen am Textobjekt modelliert werden können, in ein praktikables Modell digitaler Edition zu übersetzen. Das von der Philologie dominierte Modell kritischer digitaler Edition ist also kein Grund mehr, mittelalterliche und frühneuzeitliche Rechnungen nicht digital zu edieren und sie dabei sowohl als Informationsquelle für Sprachgeschichte als auch für ökonomische Untersuchungen zu verwenden.

³⁷ Die Vorstellung, dass Text sich grundsätzlich als eine geordnete Hierarchie von Inhaltsobjekten modellieren ließe, wird gerne mit DeRose et al. 1990 in Verbindung gebracht. Zu einer wissenschaftshistorischen Einordnung vgl. Sahle 2013, Bd. 3, S. 156–182.

³⁸ Beispiele aus der Edition mit vorläufigen Transkriptionen und einer eingeschränkten technischen Funktionalität findet man in der [Sammlung der Beispiele](#).

Bibliographische Angaben

- The Alcalá account book project, National University of Ireland. Maynooth 2008. [\[online\]](#)
- Franz-Josef Arlinghaus: Bookkeeping, Double-Entry Bookkeeping. In: Medieval Italy. An Encyclopedia. Hg. von Christopher Kleinhenz. New York 2004, S. 147–150. [\[Nachweis im OPAC\]](#)
- Klaus Arnold: Geschichtswissenschaft und elektronische Datenverarbeitung. Methoden, Ergebnisse und Möglichkeiten einer neuen Hilfswissenschaft. In: Methodenprobleme der Geschichtswissenschaft. Hg. von Theodor Schieder. München 1974 (= Historische Zeitschrift, Beiheft N.F. 3), S. 98–148. [\[Nachweis im OPAC\]](#)
- Syd Bauman: Transactionography Customized Documentation. In: Encoding Historical Financial Records Open Access Library. 20. Juli 2013. [\[online\]](#)
- Jean-Francois Bergier / Luigi Solari: Histoire et élaboration statistique. L'exemple de la population de Genève au 15e siècle. In: Mélanges d'histoire économique et social en hommage au Genève au Professeur Anthony Babel. Bd. 1. Genf 1963, S. 197–225. [\[Nachweis im GBV\]](#)
- Heidrun Elisabeth Boshof: Fest und Alltag in einem spätmittelalterlichen Damenstift. Das Reichsstift Obermünster in Regensburg im Spiegel seiner Rechnungen. In: Beiträge zur Geschichte des Bistums Regensburg 42 (2008), S. 187–258. [\[Nachweis im OPAC\]](#)
- Die Cameralia des Stadtarchivs Regensburg seit der Mitte des 14. Jahrhunderts. Bearb. von Heidrun Boshof / Regina Brunnhofer. Edition in Progress 2006ff. (= Fontes Civitatis Ratisponensis). [\[online\]](#)
- Peter Cerwenka / Paul W. Roth: Münzumlaf des 16. Jahrhunderts im Raume des östlichen Österreich. Ein Anwendungsbeispiel der elektronischen Datenverarbeitung in der historischen Forschung. Graz 1972. [\[Nachweis im GBV\]](#)
- Comédie-Française Register Project. Hg. vom Oxford Brookes und Massachusetts Institute of Technology. 2007ff. [\[online\]](#)
- Comptes de châtellenies savoyardes. Hg. von Christian Guilleré / Jean-Louis Gaulin / Marjorie Burghart. 2003ff. [\[online\]](#)
- Les comptes des consuls de Montferrand (1273–1319). Bearb. von R. Anthony Lodge 2006 (= Éditions en ligne de l'École des Chartes 16). [\[online\]](#)
- Definition of the CIDOC Conceptual Reference Model, Version 5.0.4. Hg. von Patrick Le Boeuf / Martin Doerr / Christian Emil Ore / Stephen Stead. 2011. [\[online\]](#)
- Steven J. DeRose / David G. Durand / Elli Mylonas / Allen H. Renear: What is text, really? In: Journal of Computing in Higher Education 1,2 (1990), S. 3–26. [\[Nachweis im GBV\]](#)
- Digitale Edition der Basler Jahrrechnungsbücher (1535/36 bis 1611). Hg. von Susanna Burghartz. Alpha Version 2014. [\[online\]](#)
- Early Forward Exchange Markets, 1876–1914. In: EH.net - Economic History Services [\[online\]](#)
- Edition des Steierischen Marchfutterurbars von 1414/1426. Bearb. von Matthias Perstling. Testversion [im Erscheinen]. [\[online\]](#)
- eXtensible Business Reporting Language Specifications, version 2.1. Hg. von Philipp Engel et al. 20. Februar 2013. [\[online\]](#)
- Marc Flandreau / John Komlos: Target Zones in Theory and History. Credibility, Efficiency, and Policy Autonomy. In: Journal of Monetary Economics 53 (2006), S. 1979–1995. DOI: 10.1016/j.jmoneco.2005.10.015. [\[Nachweis im GBV\]](#)
- Jack Goody: What's in a List? In: Jack Goody: The Domestication of the Savage Mind. Cambridge 1977 (= Themes in the social sciences), S. 74–111. [\[Nachweis im GBV\]](#)
- György Granasztói: Computerized Analysis of a Medieval Tax Roll. In: Acta Historica Academiae scientiarum Hungaricae 17 (1971), S. 13–24. [\[Nachweis im GBV\]](#)
- Rolf Gundlach / Carl August Lueckerath: Historische Wissenschaften und Elektronische Datenverarbeitung. Frankfurt/Main u.a. 1976. [\[Nachweis im GBV\]](#)
- Wolfgang Herborn / Klaus J. Mattheier: Sozialhistorische und sprachgeschichtliche Aspekte eines frühneuzeitlichen Rechnungsbuches der Kölner Kronenburse. In: Rheinisch-Westfälische Zeitschrift für Volkskunde 24 (1978), S. 140–182. [\[Nachweis im GBV\]](#)
- Henry III fine rolls Project. Hg. von The National Archives / King's College London. 2009. [\[online\]](#)
- David Herlihy / Christiane Klapisch-Zuber: Les Toscans et leur familles. Paris 1978. [\[Nachweis im GBV\]](#)
- Wolfgang Heß: Rechnung Legen auf Linien. Rechenbrett und Zahlisch in der Verwaltungspraxis in Spätmittelalter und Neuzeit. In: Städtisches Haushalts- und Rechnungswesen. Hg. von Erich Maschke / Jürgen Sydow. Sigmaringen 1977 (= Stadt in der Geschichte 2), S. 69–82. [\[Nachweis im OPAC\]](#)
- Wolfgang Heß: Rechnung Legen mit Rechenpfennigen. In: Numismatisches Nachrichtenblatt 45/4 (1996), S. 11–20. [\[Nachweis im GBV\]](#)

Johannes Hohfeld: Stadtrechnungen als historische Quellen. Ein Beitrag zur Quellenkunde des ausgehenden Mittelalters, dargestellt am Beispiel der Pegauer Stadtrechnungen des 14. und 15. Jahrhunderts. Leipzig 1912 (= Bibliothek der sächsischen Geschichte und Landeskunde 4,1). [\[Nachweis im GBV\]](#)

Paul Huys: Een Deense baljuwsrekening (1377-1379) als spiegeling van het dagelijks leven. In: Bijdragen tot de geschiedenis van Deinze en de Leiestreek 65 (1998), S. 185-218. [\[Nachweis im GBV\]](#)

Identifying People in the Past. Hg. von Edward A. Wrigley. London 1973. [\[Nachweis im GBV\]](#)

Informatique et histoire médiévale. Hg. von Lucie Fossier / A. Vauchez / C. Violante. Actes du Colloque de Rome, 20-22 mai 1975. Roma 1977 (= Publications de l'Ecole Française de Rome 31). [\[Nachweis im GBV\]](#)

Wybe Jappe Alberts: Mittelalterliche Stadtrechnungen als Geschichtsquellen. In: Rheinische Vierteljahrsblätter 23 (1958), S. 75-96. [\[Nachweis im GBV\]](#)

Gerhard Jaritz: Die Reimer Rechnungsbücher (1399-1477) als Quelle zur klösterlichen Sachkultur des Spätmittelalters. In: Die Funktion der schriftlichen Quellen in der Sachkulturforschung. Wien 1976 (= Veröffentlichungen des Instituts für mittelalterliche Realienkunde Österreichs 1), S. 145-249. [\[Nachweis im OPAC\]](#)

Gerhard Jaritz: Alltagsgeschichte des Mittelalters. Quellenanalysen mit Hilfe eines fachspezifischen Datenbanksystems. In: Ludwig Boltzmann Institut für Historische Sozialwissenschaft-Arbeitspapiere Nr. 1. Salzburg 1984, S. 39-42. [\[online\]](#)

Günter Katzler: Was erzählen Rechnungsbücher von der Stadt? Das Beispiel des Rechnungsbuchs Bischof Bertholds von Freising. In: Pro civitate Austriae, NF 13 (2008), S. 37-61. [\[Nachweis im OPAC\]](#)

Richard Knipping: Die Kölner Stadtrechnungen des Mittelalters mit einer Darstellung der Finanzverwaltung, 2 Bde. Bonn 1897-1898 (= Publikationen der Gesellschaft für rheinische Geschichtskunde 15). [\[Nachweis im OPAC\]](#)

Herbert Knittler: Vom Leben auf dem Lande. Die Rechnungen der Herren von Puchheim zu Horn und Göllersdorf 1444-1468. Edition und Kommentar. St. Pölten 2005 (= NÖ-Schriften 161). [\[Nachweis im GBV\]](#)

Christopher Kolbeck / Paul Rössler: Aufstieg und Fall des »Item«. Item Serialität zwischen sprachlicher Innovation und Konstanz-Item in Regensburger Rechnungsbüchern. In: Stimulus 19 (2015) [im Erscheinen].

Rudolf Kropf: Zur Auswertung von Urbaren mittels Computer. In: Mitteilungen des Instituts für Österreichische Geschichtsforschung 77 (1969), S. 141-144. [\[Nachweis im OPAC\]](#)

Federigo Melis: Storia della Ragioneria. Contributo alla conoscenza e interpretazione delle fonti più significative della storia economica. Bologna 1950. [\[online\]](#)

Mark Mersiowsky: Aspekte adligen Lebens um 1400. Frühe westfälische und rheinische Adelsrechnungen im Vergleich. In: Vestigia Monasteriensia. Westfalen - Rheinland - Niederlande. Hg. von Peter Johaneck / Mark Mersiowsky / Ellen Widder. Festschrift für Wilhelm Jansen. Bielefeld 1995 (= Studien zur Regionalgeschichte 5), S. 263-304. [\[Nachweis im OPAC\]](#)

Mark Mersiowsky: Die Anfänge territorialer Rechnungslegung im deutschen Nordwesten. Spätmittelalterliche Rechnungen, Verwaltungspraxis, Hof und Territorium (zugl. Diss. phil. Münster 1992). Sigmaringen 2000 (= Residenzenforschung 9), S. 337-340. [\[Nachweis im OPAC\]](#)

Die Rechnungsbücher der Stadt Luxemburg. Bearb. von Claudine Moulin / Michel Pauly. Luxembourg 2007-2012ff. (= Schriftenreihe des Stadtarchivs Luxemburg 1-6). [\[Nachweis im GBV\]](#)

Claudine Moulin: Zeichen und ihre Deutung. Zum handschriftennahen Edieren schriftlicher Quellen im interdisziplinären Kontext. In: Die Rechnungsbücher der Stadt Luxemburg. Hg. von Claudine Moulin / Michel Pauly. Luxembourg 2012 (= Schriftenreihe des Stadtarchivs Luxemburg 6), S. 9-17. [\[Nachweis im GBV\]](#)

Open Annotation Data Model. Community Draft. Hg. von Robert Sanderson / Paolo Ciccarese / Herbert an de Sompe. 8. Februar 2013. [\[online\]](#)

Open domesday. The first free online copy of Domesday Book. Hg. von Anna Powell-Smith / J.J.N. Palmer. University of Hull. [\[online\]](#)

OWL 2 Web Ontology Language. Structural Specification and Functional-Style Syntax. Second Edition. W3C Recommendation. Hg. von Boris Motik / Peter F. Patel-Schneider / Bijan Parsia. 11. Dezember 2012. [\[online\]](#)

Matthias Perstling: Multimediale Dokumentation und Edition mehrschichtiger Texte. Das steirisch-landesfürstliche »Marchfütterurbar« von 1414/1426. Diss. phil. Graz 2013. [\[online\]](#)

Fausto Ravidà: Graphematisch-phonologische Analyse der Luxemburger Rechnungsbücher (1388-1500). Ein Beitrag zur historischen Stadtsprachenforschung. Heidelberg 2012 (= Germanistische Bibliothek 43). [\[Nachweis im GBV\]](#)

RDF 1.1 Semantics. W3C Recommendation Hg. von Patrick J. Hayes / Peter F. Patel-Schneider. 25. Februar 2014. [\[online\]](#)

Patrick Sahle: Digitale Editionsformen. Zum Umgang mit der Überlieferung unter den Bedingungen des Medienwandels, 3 Bände. Norderstedt 2013 (= Schriften des Instituts für Dokumentologie und Editorik 7-9). [\[Nachweis im OPAC\]](#)

Schuldbücher und Rechnungen der Großschäffer und Lieger des Deutschen Ordens in Preußen, Bd. 1: Großschäfferei Königsberg I (Ordensfoliant 141). Bearb. von Cordelia Heß / Christina Link / Jürgen Sarnowsky. Köln, Weimar, Wien 2008 (= Veröffentlichungen aus den Archiven Preußischer Kulturbesitz, 62,1; zugleich Quellen und Darstellungen zur Hansischen Geschichte N.F., LX, 1). [\[Nachweis im OPAC\]](#)

Schuldbücher und Rechnungen der Großschäffer und Lieger des Deutschen Ordens in Preußen, Bd. 3: Großschäfferei Marienburg. Bearb. von Christina Link / Jürgen Sarnowsky. Köln, Weimar, Wien 2008 (= Veröffentlichungen aus den Archiven Preußischer Kulturbesitz, 62,3; zugleich Quellen und Darstellungen zur Hansischen Geschichte N.F. LIX, 3). [\[Nachweis im OPAC\]](#)

Die Schuld- und Rechnungsbücher des Deutschen Ordens. Eine synoptische Edition im Internet. Hg. von Christina Link / Jürgen Sarnowsky. Hamburg 2007ff. [\[online\]](#)

Kurt Seidel: Klosterrechnungen als Geschichtsquelle. In: Deutsche Geschichtsblätter 12 (1911), S. 291–297. [\[Nachweis im OPAC\]](#)

Shared Canvas Data Model. Hg. von Robert Sanderson / Benjamin Albritton, Version 1.0. 14. Februar 2013. [\[online\]](#)

Edward Shorter: The Historian and the Computer. A Practical Guide. Englewood Cliffs, NJ. 1971. [\[Nachweis im GBV\]](#)

Elisabeth Steiner / Johannes Stigler: GAMS and Cirilo Client: policies, documentation and tutorial. 2014. [\[online\]](#)

Early U.S. Securities Prices. Zusammengestellt von Richard E. Sylla / Jack Wilson / Robert E. Wright. In: EH.Net - Economic History Services [\[online\]](#)

TEI P5. Guidelines for Electronic Text Encoding and Interchange. Hg. von TEI Consortium. Version 2.6, 2014. [\[online\]](#)

Manfred Thaller: What is a text within the Digital Humanities, or some of them, at least? In: digital humanities 2012. Hamburg 2012. [\[online\]](#)

Charles Tilly: Computers in Historical Analysis. In: Computers and the Humanities 7 (1973), S. 323–335. [\[Nachweis im GBV\]](#)

Kathryn Tomasek / Syd Bauman: Encoding Financial Records for Historical Research. In: Journal of the Text Encoding Initiative 6 (2013). DOI: 10.4000/jtei.895. [\[online\]](#)

Kathryn Tomasek: Encoding Financial Records for Historical Research. Whitepaper 2012. NEH Ref: HD-51224-11. [\[online\]](#)

Georg Vogeler: Spätmittelalterliche Steuerbücher deutscher Territorien, Teil 1: Überlieferung und hilfswissenschaftliche Analyse. In: Archiv für Diplomatik 49 (2003), S. 165–295; Teil 2: Funktionale Analyse und Typologie. In: Archiv für Diplomatik 50 (2004), S. 57–204. [\[Nachweis im OPAC\]](#)

Rechnungen des Mittelalters im Digitalen Zeitalter. Ein Modell. Hg. von Georg Vogeler. 2013ff. [\[online\]](#)

Georg Vogeler: Digitale Edition von Wirtschafts- und Rechnungsbüchern. In: Wirtschafts- und Rechnungsbücher des Mittelalters und der Frühen Neuzeit. Hg. von Gudrun Gleba / Niels Petersen. Göttingen 2015. [\[Nachweis im GBV\]](#)

Aude Wirth: Les comptes des receveurs et leur intérêt pour l'histoire sociale. L'exemple de la Lorraine. In: Noms de métiers et catégories professionnelles. Acteurs, pratiques, discours (XVe siècle à nos jours). Hg. von Hanne Georges / Judde de Larivière Claire. Toulouse 2010, S. 97–111. [\[Nachweis im GBV\]](#)

XBRL Global Ledger Taxonomy, Specifications. 17. April 2007. [\[online\]](#)

XSL Transformations (XSLT) Version 2.0. W3C Recommendation. Hg. von Michael Kay. 23. Januar 2007. [\[online\]](#)

Basil S. Yamey: Scientific Bookkeeping and the Rise of Capitalism. In: English History Review N.S. 1 (1949), S. 99–113. [\[Nachweis im GBV\]](#)

Abbildungslegenden und -nachweise

Abb. 1: Integration von Inhalt in ein Modell digitaler kritischer Editionen (Graphik: Autor).

Zeitschrift für digitale Geisteswissenschaften

Artikel aus:

Sonderband 1 der ZfdG: Grenzen und Möglichkeiten der Digital Humanities. Hg. von Constanze Baum und Thomas Stäcker. 2015. DOI: [10.17175/sb01](https://doi.org/10.17175/sb01)

Titel:

Vom Zeichen zur Schrift: Mit Mustererkennung zur automatisierten Schreiberhanderkennung in mittelalterlichen und frühneuzeitlichen Handschriften

Autor/in:

Daniel Fecker

Kontakt: fecker@ifn.ing.tu-bs.de

Institution: Technische Universität Braunschweig

GND: [1084240270](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-64862-p0088-8) ORCID:

Autor/in:

Volker Märgner

Kontakt: maergner@ifn.ing.tu-bs.de

Institution: Technische Universität Braunschweig

GND: [1084238829](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-64862-p0088-8) ORCID:

Autor/in:

Torsten Schaßan

Kontakt: schassan@hab.de

Institution: Herzog August Bibliothek Wolfenbüttel

GND: [1027794041](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-64862-p0088-8) ORCID:

DOI des Artikels:

[10.17175/sb001_008](https://doi.org/10.17175/sb001_008)

Nachweis im OPAC der Herzog August Bibliothek:

[830167242](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-64862-p0088-8)

Erstveröffentlichung:

19.02.2015

Lizenz:

Sofern nicht anders angegeben 

Medienlizenzen:

Medienrechte liegen bei den Autoren

Letzte Überprüfung aller Verweise:

24.05.2016

GND-Verschlagwortung:

[Mustererkennung](#) | [automatische Klassifikation](#) | [Handschrift \(Grafologie\)](#) |

Zitierweise:

Daniel Fecker, Volker Märgner, Torsten Schaßan: Vom Zeichen zur Schrift: Mit Mustererkennung zur automatisierten Schreiberhanderkennung in mittelalterlichen und frühneuzeitlichen Handschriften. In: Grenzen und Möglichkeiten der Digital Humanities. Hg. von Constanze Baum / Thomas Stäcker. 2015 (= Sonderband der Zeitschrift für digitale Geisteswissenschaften, 1). PDF Format ohne Paginierung. Als text/html abrufbar unter DOI: [10.17175/sb001_008](https://doi.org/10.17175/sb001_008).

Daniel Fecker, Volker Märgner, Torsten Schaßan

Vom Zeichen zur Schrift: Mit Mustererkennung zur automatisierten Schreiberhanderkennung in mittelalterlichen und frühneuzeitlichen Handschriften

Abstracts

Für die Digital Humanities im Bereich Mediävistik und Frühnezeitforschung stellt die Digitalisierung von Handschriften ein zentrales Feld dar. Da jede Handschrift eigene Charakteristika aufweist, führt die automatische Erstellung eines maschinenlesbaren Textes durch Optical Character Recognition (OCR) anhand von Digitalisaten in den allermeisten Fällen zu fehlerhaften Ergebnissen. Andererseits können Charakteristika dieser Schrift wie Buchstabengröße und -abstand, Dichte des Schriftbildes, Neigung u.a. genutzt werden, um die Identifikation der schreibenden Hand bzw. Hände zu ermöglichen. In dem Beitrag wird gezeigt, wie die Analyse von Handschriftenabbildungen zur Identifikation der schreibenden Hand bzw. Hände genutzt werden kann. Ein Algorithmus soll sonstige paläographische oder kodikologische Befunde unterstützen und Argumente zur Veri- oder Falsifikation von unsicheren Zuschreibungen liefern.

For Digital Humanities in medieval studies and early modern studies, the digitization of manuscripts is a central field. Since each manuscript displays its own unique characteristics, the automatic generation of a machine-readable text using Optical Character Recognition (OCR) as applied to digital images leads, in most cases, to error-prone results. However, characteristics of handwriting such as the size of letters and spacing, slope, and so on can be used to identify the scribe or scribes. This paper demonstrates how the analysis of manuscript images can be used to identify the scribe or scribes. An algorithm will support additional paleographic and codicological findings and provide evidence for the verification or falsification of uncertain attributions.

1. Einleitung

»Deutschland befindet sich in einer Phase intensiv betriebener und mit einem hohen finanziellen Aufwand verbundener Digitalisierung seiner historischen Bestände. Für die Mediävistik und Frühnezeitforschung stellt hierbei die Digitalisierung der dem Mittelalter und der Renaissance entstammenden Handschriften ein zentrales Feld dar.«¹

Die Nutzung der Digitalisate allein als digitale Lesekopie durch den betrachtenden Forscher oder die Forscherin würde das Erkenntnispotential, das dem Digitalisat selbst innewohnt, schlechterdings vergeuden. Daher ist eine der ersten Fragen von ›traditionell‹ mit Handschriften arbeitenden Geisteswissenschaftlern in der Regel die, ob oder wie man aus den Bildern der Handschriften einen Text gewinnen kann.

¹ Vgl. Hays / Müller 2011, S. 416.

Die zentrale Anwendung wäre die Optical Character Recognition (OCR automatisierte Texterkennung), die der Herstellung eines maschinenlesbaren Textes aus bildhaft vorliegender Information dient. Die Güte der aus OCR-Algorithmen gewonnenen Texte ist von mehreren Faktoren abhängig, darunter auch von der Qualität der Vorlage sowie von der Qualität des Scanprozesses bzw. der daraus resultierenden Digitalisate.² Doch selbst wenn diese Faktoren theoretisch eine optimale Weiterverarbeitung möglich machen würden, böten historische und insbesondere handschriftliche Materialien weitere Herausforderungen: Schon bei Drucken mit gebrochenen, gotischen Schriften mit nur geringen Unterschieden zwischen einzelnen Buchstabenformen wie dem langen s -f- und dem normalen f produziert gängige OCR-Software in der Regel ein für eine wissenschaftliche Nutzung unbrauchbares, da zu fehlerbehaftetes Ergebnis. Vor diesem Hintergrund erscheint die Aufgabe bei Handschriften noch schwieriger. Dass OCR bei der Verarbeitung von Handschriftendigitalisaten an ihre Grenzen stoßen muss, lässt sich anhand des formalen Ablaufs einer Texterkennung veranschaulichen:

- Anfertigen eines Digitalisates in geeigneter Qualität gegebenenfalls Bildkorrekturen, wie das Geraderücken schräg aufgenommener Seiten, Glättung von Rundungen aufgrund von Materialbiegung, etc.
- Binarisierung der Farbwerte zur deutlichen Trennung von Schrift und Hintergrund
- Segmentierung der Schrift, z.B. in Linien und Einzelworte
- Mustererkennung, bei der zu erkennende Formen mit einem vorhandenen Zeichenvorrat verglichen werden
- Im Falle der Übereinstimmung Zuweisung eines Zeichencodes nach üblicher Textkodierung (=UTF-8)

Die eigentliche Herausforderung an die OCR für Handschriften liegt in Arbeitsschritt 4, da für die Zuweisung eines zu speichernden Codes in Schritt 5 hier zunächst die Zuweisung eines Codes zu einer großen Zahl von Digitalisaten manuell durchgeführt werden muss (man spricht davon, die sogenannte ›ground truth‹ für die Digitalisate anzulegen). Diese Aufgabe ist sehr zeitaufwendig aber notwendig, um das sogenannte ›Training‹ eines Klassifikators zu ermöglichen, wobei der Klassifikator an die jeweilige Schriftart angepasst wird. Der Klassifikator ist Kern des OCR-Systems und ermöglicht es, Muster (pattern) in der Vorlage Zeichen aus dem Zeichenvorrat zuzuordnen. Diese auf einer festgesetzten Wahrheit beruhende Beziehung zwischen den Bildmustern und den kodierten Zeichen ist nur für die Typen von Mustern gültig, die zum Training verwendet wurden, also z. B. für eine Schrifttype oder eine bestimmte Schriftform. Da die Handschrift jedes Schreibers³ eigene Charakteristika aufweist, welche sie zumindest von anderen Händen unterscheidbar macht, können die in den meisten Handschriftendigitalisaten aufgefundenen Muster nicht ohne weiteres eindeutig codierten Zeichen zugewiesen werden.

² Stäcker 2013, S. 123–135.

³ Den Autoren ist bewusst, dass es im frühen Mittelalter Schreiberinnen gegeben hat. Die in diesem Beitrag behandelten Beispiele stammen jedoch durchweg aus dem Kloster Weissenburg, das ein Männer-Konvent war. Dort wird es nur Schreiber gegeben haben, weshalb im Folgenden durchgehend die männliche Form gewählt wird.

Die Varianz zwischen einzelnen Handschriften, vor allem von Buchschriften wie der karolingischen Minuskel, ist als vergleichsweise klein einzuschätzen, so dass die Anwendung von OCR auf diese Schriften demnach recht gute Ergebnisse erwarten lassen dürfte, allerdings ist die Aufgabe der Textgewinnung gerade für die in dieser Schrift geschriebenen Werke relativ uninteressant. Es handelt sich dabei um zumeist wohlbekannte und gut erforschte, in der Regel auch schon kritisch edierte Texte⁴ oder um Texte, deren Varianz in den Handschriften als äußerst gering anzusetzen ist, wie z.B. Missale.

Aufgrund dieser Annahmen stand am Anfang der Kooperation zwischen dem Institut für Nachrichtentechnik an der TU Braunschweig (IFN) und der Herzog August Bibliothek Wolfenbüttel (HAB) die Überlegung, das Untersuchungsziel kurzerhand umzukehren und als Ergebnis der Analyse von Handschriftenabbildungen nicht einen durch OCR zu gewinnenden elektronischen Text anzustreben, sondern in der Schrift Merkmale zur Identifikation der schreibenden Hand zu herauszufiltern. Sollte es nämlich gelingen, anhand spezifischer Merkmale eine Schreiberhand (unter Abstraktion von nicht mehr quantifizierbaren Abweichungsquellen wie Lebensalter und Tagesform des Amanuensis oder dem Zustand der Schreibmaterialien und -utensilien) von anderen Schreiberhänden abzugrenzen, so könnte damit die wichtige Fragestellung nach dem Schreiber automatisiert werden.

Als Charakteristika dieser Schrift sollten Buchstabengröße und –abstand, Dichte des Schriftbildes, Neigung u. a. untersucht werden, aber nicht notwendig, wie in der klassischen, vom forschenden menschlichen Auge ausgehenden Paläographie, einzelne Buchstabenformen. Dabei werden in dieser Arbeit verschiedene Merkmale vorgestellt, die auf unterschiedlichen Grundprinzipien, wie der Schriftkontur, der durch die Periodizität des Schriftbildes zugrunde liegenden Textur und speziellen Schlüsselpunkten in der Schrift, beruhen.

Diese Überlegung lässt sich nun in zweierlei Richtung ausdeuten:

1. Zum einen könnte versucht werden, die Unterschiede einer Schrift in Relation zu anderen Schriften zu bringen und so eine der Grundfragen der Kodikologie und Paläographie beantworten zu helfen, die Frage nämlich, wieviele Schreiber bei der Entstehung der Handschrift mitgewirkt haben.
2. Zum anderen können mit vorhandenem Wissen über die Merkmale einer Schrift oder einer Schreiberhand andere Handschriftendigitalisate auf mögliche Identität oder Differenz hin untersucht werden. Es müsste so möglich sein, diese Schrift oder Hand in anderen Handschriften nachzuweisen bzw. die Differenz festzustellen, also einen anderen Schreiber anzunehmen. Wenn man die Schwellwerte (thresholds) entsprechend anpassen würde, sollte es auch möglich sein, Schriftfamilien voneinander zu trennen und bei einer entsprechenden Materialbasis Hinweise auf Datierung und Lokalisierung zu sammeln.

⁴ Vgl. die Textausgaben im Rahmen der [Patrologia Latina](#) (PL), eine der ersten Werke, die als elektronischer Volltext verfügbar waren, oder der [Library of Latin Texts](#) (CLCLT).

Für die Beantwortung beider Fragen ist sowohl die Festsetzung der Schwellwerte als auch die Abschätzung zentral, welche räumliche Ausdehnung der Untersuchung zugrunde gelegt wird: Die Schwellwerte müssen genutzt werden, um kleinere Abweichungen zu tolerieren. Kein Schreiber ist in der Lage, in einem handschriftlichen Text zwei völlig identische Buchstabenformen zu erzeugen. Zu genaues Betrachten des Untersuchungsgegenstandes würde somit nur noch Unterschiede erkennen lassen und keinerlei Übereinstimmung mehr zu Tage fördern.⁵ Die räumliche Ausdehnung ist insofern wichtig, da die in dieser Arbeit verwendeten Algorithmen immer ganze Seiten betrachten und daraus die Merkmalsvektoren berechnen. Unterschiede, die sich auf einer Seite befinden, beispielsweise durch Handwechsel mitten auf der Seite, können dabei allerdings nicht aufgefunden werden.

Die Erkennungsgenauigkeit muss dazu aufgrund des Trainings mit einer Handschrift in einem anderen Codex über einem zu definierenden Schwellwert (threshold) liegen, um als Indiz gewertet zu werden, dass derselbe Schreiber die Handschrift geschrieben haben könnte. Der angestrebte Algorithmus würde damit sonstige paläographische oder kodikologische Befunde unterstützende bzw. ergänzende Argumente zur Verifikation von unsicheren Zuschreibungen liefern. Im Gegenzug müsste das Unterschreiten dieses Schwellwertes Argumente für Fälschungen solcher Zuschreibungen ermöglichen.

Der typische Aufbau einer Prozesskette für die Identifikation eines Schreibers ist die einer OCR nicht unähnlich.⁶ In einem ersten Schritt werden die gescannten Bilder eines Manuskriptes mit einer Vorverarbeitung für die Extraktion von geeigneten Merkmalen aufbereitet. Anschließend werden in einem weiteren Schritt geeignete Merkmale gewonnen, die in dem finalen Schritt für eine Klassifikation verwendet werden. Abhängig von der Art der Klassifikation stehen dafür Referenzdokumente mit bekannten Schreibern zur Verfügung.

Dieser Beitrag ist wie folgt gegliedert: nach dieser Einleitung wird in **Abschnitt 2** der Schritt der Vorverarbeitung erläutert. Anschließend werden in **Abschnitt 3** unterschiedliche Methoden der Merkmalsextraktion vorgestellt. In **Abschnitt 4** werden unterschiedliche Aufgaben für die Klassifikation von Schreibern in historischen Dokumenten vorgestellt, zu denen in **Abschnitt 5** exemplarische Ergebnisse von verschiedenen Experimenten gezeigt werden.

2. Vorverarbeitung

In der Vorverarbeitungsstufe werden die gescannten Bilder der historischen Dokumente für die Schreibererkennung aufbereitet. Idealerweise sollten nach dieser Stufe die Bilder nur die reine Schrift enthalten, beispielsweise in Form von Binärbildern (Schrift schwarz, Beschreibstoff weiß). Aufgrund der Alterung der Dokumente enthalten die Seitenbilder allerdings verschiedenste Arten von Verunreinigungen, wie z.B. Wasserflecken (vgl. **Abbildung 2**) oder Texte, die von der Rückseite durchscheinen, die die Vorverarbeitung erschweren.

⁵ Vgl. Robinson 1996, S. 106, der dieses Problem im Zusammenhang mit textueller Überlieferung diskutiert und konstatiert: »In fact, the closer we look at a text, the more variation we see. It is not only that there is more variation: the text itself changes depending on how closely we are looking at it.«

⁶ Vgl. M / Idicula 2011, S. 23–33.

Des Weiteren finden sich in solchen Dokumenten oft Zusätze anderer Schreiber an den Seitenrändern. Ornamente unterschiedlichster Ausprägung können zusätzlich die automatische Merkmalsgewinnung erschweren. Um die Zuweisung von Schriftmerkmalen des Haupttextes zu verbessern, wird in einem ersten Schritt die Region des Haupttextes ermittelt. [Abbildung 1 \(links\)](#) zeigt ein solches Beispiel einer Seite eines arabischen, historischen Dokumentes mit vielen Kommentaren.⁷

Im nächsten Schritt wird die Handschrift des Haupttextes vom Hintergrund getrennt (segmentiert), um eine reine, möglichst ungestörte Version der Schrift zu erhalten. Hierfür wird eine Binarisierung angewendet, die ein farbiges Eingabebild in ein binäres Bild transformiert. [Abbildung 1 \(Mitte und rechts\)](#) zeigt ein Beispiel einer solchen Segmentierung mit einem Binarisierungsverfahren aus dem aktuellen Stand der Technik.⁸ Trotz der Wasserflecken kann die Schrift erfolgreich segmentiert werden. Einzig die rote Schrift wird nicht ausreichend gut segmentiert. Um auch in solchen Fällen die Schrift erfolgreich zu segmentieren, können spezielle Farbsegmentierungsalgorithmen verwendet werden.⁹



Abb. 1: Seite eines arabischen, historischen Dokumentes mit Haupttext und vielen Kommentaren (links); Beispiel einer Segmentierung der Handschrift mit Hilfe einer Binarisierung: Teil eines Ausgangsbildes (Mitte) und binäres Ergebnisbild (rechts) (Quelle: Autoren).

3. Merkmalsextraktion

Nach der Binarisierung erfolgt die Merkmalsextraktion. Das Ziel dieser Stufe ist es, diskriminative Merkmale aus der Handschrift für die Identifikation von Schreibern zu generieren. Allgemein wird zwischen Mikro- und Makromerkmalen unterschieden. Mikromerkmale spiegeln die feine Charakteristik der Handschrift wider. Als Beispiel hierfür gibt es z.B. Grapheme-basierte Ansätze.¹⁰ Im Gegensatz hierzu versuchen Makromerkmale die globale Charakteristik der Schrift eines Schreibers zu ermitteln. In dieser Arbeit wird der Fokus auf den zweiten Typ von Merkmalen gelegt. Dabei wird ein Merkmalsvektor für ganze Textblöcke erzeugt, der die globale Information über den Schreibstil eines Schreibers wie z.B. die Neigung oder die Krümmung der Schrift enthält. Im Folgenden werden verschiedene Arten von Merkmalen, basierend auf den unterschiedlichen Grundprinzipien Kontur, Textur und der Extraktion von Schlüsselpunkten vorgestellt.

⁷ In Fecker et al. 2014, S. 1551, wird hierfür ein Verfahren zur automatischen Detektion vorgeschlagen.

⁸ Vgl. Bar-Yosef et al. 2007, S. 89–99.

⁹ Pantke et al. 2014, S. 151–156.

¹⁰ Bulacu et al. 2007, S. 771.

3.1 Konturbasierte Merkmale

Die konturbasierten Merkmale verwenden eine Statistik über die Verteilung der Winkel, die in der Kontur einer Handschrift eines Schreibers enthalten sind.¹¹ In den binären Bildern wird diese Verteilung mit Hilfe einer Konturverfolgung ermittelt. Die Winkel werden zwischen der Verbindungsgeraden zweier Konturpunkten mit vorgegebenem Abstand und der horizontalen Achse gemessen. Sämtliche ermittelten Winkel einer Schrift werden in einem Winkelhistogramm gesammelt. Anschließend wird dieses zu einer Wahrscheinlichkeitsdichteverteilung normalisiert, welche letztendlich als Merkmalsvektor verwendet wird. Es gibt auch Beispiele für die ermittelten Wahrscheinlichkeitsdichteverteilungen zweier unterschiedlicher Schreiber. Dabei wird eine Log-Polar-Darstellung dieser Verteilungen verwendet. In Diagrammen werden jeweils die Ergebnisse zweier unterschiedlicher Texte eines Schreibers eingetragen. Trotz der unterschiedlichen Texte ändern sich die Verteilungen nur geringfügig. Ein großer Unterschied manifestiert sich dagegen zwischen den Verteilungen unterschiedlicher Schreiber.

In *Text-Independent Writer Identification and Verification on Offline Arabic Handwriting* wurde ein fester Abstand zwischen zwei Konturpunkten für die Winkelmessung verwendet.¹² Als Modifikation wurde in *Writer Identification for Historical Arabic Documents* eine variable Distanz für die Winkelmessung vorgeschlagen, welche mit Hilfe eines Fehlerkriteriums automatisch bestimmt werden kann.¹³ Diese Anpassung ermöglichte eine genauere Approximation der Kontur.

3.2 Texturbasierte Merkmale

In *Natural Image Character Recognition Using Oriented Basic Image Features* werden die Merkmale »Oriented Basic Image Features« vorgeschlagen, die auf Basis der Symmetrie und Orientierung in der lokalen Nachbarschaft von Pixeln eine Textur beschreiben.¹⁴ Hierbei handelt es sich um einen Multi-Skalen-Ansatz, basierend auf Differenzbildern von Gauß-gefilterten Bildern. Dafür werden Filter unterschiedlicher Ordnungen und Orientierungen verwendet. Basierend auf diesen Differenzbildern werden sieben Symmetrie-Merkmale konstruiert. Einigen dieser Symmetriemerkmale kann eine Orientierung zugewiesen werden. Die übrigen Merkmale sind rotationsinvariant. Sämtliche vorkommende Kombinationen von Symmetrie-Merkmal, Orientierung und Skalierungsstufe der Pixel eines Eingabebildes werden wieder zu einem Histogramm zusammengefügt. Nach einer Normalisierung zu einer Wahrscheinlichkeitsdichtefunktion ergibt sich daraus ebenfalls ein Merkmalsvektor.¹⁵

¹¹ Bulacu et al. 2007, S. 770.

¹² Bulacu et al. 2007, S. 770.

¹³ Fecker et al. 2014, S. 1551.

¹⁴ Newell / Griffin 2011, S. 191–196.

¹⁵ Vgl. Hassane / Al-Madeed 2012, S. 835–840. Dieser Ansatz erreichte in einem Wettbewerb zur arabischen Schreibererkennung in zeitgenössischen Schriften bei der Konferenz International Conference on Frontiers of Handwriting Recognition (ICFHR) im Jahr 2012 den ersten Platz.

3.3 Schlüsselpunktbasierte Merkmale

Diese Merkmale basieren auf den Scale Invariant Feature Transform (SIFT)-Deskriptoren.¹⁶ Diese basieren ebenfalls auf Differenzen von Gauß-gefilterten Bildern unterschiedlicher Skalierungen. Anders als bei den Oriented Basic Image Features werden hier über lokale Extremwerte der ermittelten Differenzbilder Schlüsselpunkte detektiert. Über die Gradienten in der Umgebung wird diesen Schlüsselpunkten eine Orientierung zugewiesen (Abbildung 2, links). Anschließend werden ebenfalls auf Basis der lokalen Orientierung und Beträgen der Gradienten in einer begrenzten Region um einen Schlüsselpunkt die Deskriptoren ermittelt (Abbildung 2, rechts).

Für Handschriften werden die Schlüsselpunkte auf Kreuzungen, Krümmungen und Spitzen detektiert. Basierend auf den Distanzen aller Deskriptoren in einem Textblock wird ein Merkmalsvektor abgeleitet.¹⁷ Die Distanzen liefern Informationen über das Verhältnis der Orientierungen und der Beträge der Gradienten um die Schlüsselpunkte, und man erhält somit Informationen über den Schreibstil eines Schreibers.

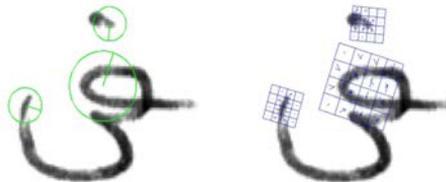


Abb. 2: Exemplare für Schlüsselpunkte in einer Handschrift. Die Mittelpunkte der Kreise deuten auf den Ort des Schlüsselpunktes, der Durchmesser auf die Skalierung, und die Orientierung ist durch die Linie gegeben (links). Deskriptoren der ermittelten Schlüsselpunkte (rechts) (Quelle: Autoren).

4. Klassifikation

Es lassen sich verschiedene Klassifikationsaufgaben im Bereich der Schreibererkennung von historischen Dokumenten identifizieren. In dieser Arbeit wird der Fokus auf die drei Klassifikationsaufgaben Intra-Manuskript-Schreiberanalyse, der direkte Vergleich zweier Manuskripte und die One-to-Many-Klassifikation eines Manuskriptes mit unbekanntem Schreiber gelegt. Diese Verfahren werden im Folgenden erläutert, wobei hier speziell die Klassifikation von mehrseitigen Dokumenten behandelt wird. Allgemein wird dabei für jede Seite eines Dokumentes ein Merkmalsvektor extrahiert. Für den Vergleich zweier Merkmalsvektoren wird in der Regel eine Distanzmetrik eingesetzt.

4.1 Intra-Manuskript-Schreiberanalyse

¹⁶ Lowe 2004, S. 91–110.

¹⁷ Fecker et al. 2014, S. 1552.

Bei der Intra-Manuskript-Schreiberanalyse wird innerhalb eines Dokumentes untersucht, ob an einem Dokument ein oder mehrere unterschiedliche Schreiber beteiligt waren. Zusätzlich können explizit die Seiten ermittelt werden, die von den einzelnen Schreibern geschrieben wurden. Ermöglicht wird dies durch die Berechnung der Distanzen sämtlicher Merkmalsvektoren eines Dokumentes zueinander. Hierdurch ergibt sich eine Distanzmatrix. In *Abbildung 5* ist schematisch eine solche Distanzmatrix eines fünfseitigen Dokumentes dargestellt. Sämtliche Elemente der Hauptdiagonale enthalten den Wert Null, da dies der Distanz von einer Seite zu sich selbst entspricht. Die Matrix ist entlang dieser Hauptdiagonalen gespiegelt und enthält somit redundante Informationen. Die Darstellung in Matrixform birgt beispielsweise bei der Visualisierung Vorteile (vgl. auch [Abschnitt 5.1](#)).

Die Distanzmatrizen lassen sich dazu verwenden, unterschiedliche Schreiber anhand der Unterschiede in den Beträgen der Differenzen zu erkennen. Um die Beträge der Differenzen einordnen zu können, können optional bereits analysierte Referenzdokumente mit bekannter Anzahl von Schreibern verwendet werden, um Schwellwerte abzuleiten.

4.2 Direkter Vergleich zweier Manuskripte

Anders als bei der Intra-Manuskript-Schreiberanalyse werden hier die Schreiber von zwei Dokumenten miteinander verglichen. Hierfür werden ebenfalls Distanzmatrizen eingesetzt. Allerdings werden hier die Distanzen der Merkmalsvektoren eines Dokumentes zu denen des anderen Dokumentes ermittelt. Um die Beträge der Distanzen einschätzen zu können, können zusätzlich Referenzdokumente für die Ermittlung von Schwellwerten verwendet werden (*Tabelle 1, Abbildung 3*).

0	D(S1,S2)	D(S1,S3)	D(S1,S4)	D(S1,S5)
D(S2,S1)	0	D(S2,S3)	D(S2,S4)	D(S2,S5)
D(S3,S1)	D(S3,S2)	0	D(S3,S4)	D(S3,S5)
D(S4,S1)	D(S4,S2)	D(S4,S3)	0	D(S4,S5)
D(S5,S1)	D(S5,S2)	D(S5,S3)	D(S5,S4)	0

Abb. 3: Tabelle 1: Schematische Darstellung einer Distanzmatrix. Zur Erläuterung: D(S1,S2) entspricht der Distanz der Merkmalsvektoren der ersten und zweiten Seite eines Manuskripts (Quelle: Autoren).

4.3 One-to-Many- Klassifikation

Bei der One-to-Many-Klassifikation wird versucht, einen unbekanntem Schreiber eines Manuskriptes mit Hilfe eines Datenbestandes von bekannten Schreibern zu identifizieren. Dazu wird in diesem Datenbestand dasjenige Manuskript ermittelt, bei dem die Handschrift die größte Ähnlichkeit mit der des unbekanntem Schreibers aufweist. Hier wird dafür die Nächster-Nachbar-Klassifikation eingesetzt. Da es sich dabei um mehrseitige Manuskripte handelt, für die ebenfalls mehrere Merkmalsvektoren extrahiert werden, existieren unterschiedliche

Strategien, diese für die Klassifikation zu kombinieren.¹⁸ Bei der Mittelung werden sämtliche Merkmalsvektoren für die untersuchten Seiten eines Manuskriptes in einen einzelnen Merkmalsvektor zusammengefasst. Dabei wird ein Schreiber dem gesamten Manuskript zugeordnet (Abbildung 4, links). Beim Voting dagegen wird jeder Seite des zu untersuchenden Manuskriptes ein Schreiber zugeordnet. Der Kandidat mit den meisten zugeordneten Seiten wird als Schreiber des Manuskriptes erkannt (Abbildung 4, rechts). Beim Voting ist zusätzlich eine Gewichtung durch die ermittelte Distanz der Nächster-Nachbar-Klassifikation möglich. Je kleiner die Distanz ist, desto höher ist die Konfidenz für die Erkennung eines Schreibers.



Abb. 4: One-to-Many-Klassifikation eines mehrseitigen Manuskriptes durch Mittelung der Merkmalsvektoren (links). One-to-Many-Klassifikation eines mehrseitigen Manuskriptes durch Voting (rechts) (Quelle: Autoren).

5. Experimente

In diesem Kapitel werden exemplarische Ergebnisse für Experimente bei der Schreibererkennung präsentiert. Die durchgeführten Experimente sind die Intra-Manuskript-Schreiberanalyse, der direkte Vergleich zweier Manuskripte und die One-to-Many-Klassifikation. Für die Experimente wurden lateinische und arabische Manuskripte verwendet.

5.1 Experimente zur Intra-Manuskript-Schreiberanalyse und direktem Vergleich zweier Manuskripte

Um beide Fragestellungen adressieren zu können, wurde folgende Arbeitsweise festgelegt und darauf basierend Testmaterial ausgewählt:

- Es soll eine Handschrift (1) für das Training ausgewählt werden, deren Schreiber bekannt ist. Die ganze Handschrift sollte erwiesenermaßen vom gleichen Schreiber geschrieben worden sein, um dem Algorithmus genügend Trainingsmaterial anbieten zu können. Diese Handschrift dient als Referenz für Hand 1.
- Der Schreiber von Handschrift 1 sollte zugleich in einer anderen Handschrift (2) nachweisbar sein, in der er allerdings nicht alleine schreibt, sondern wiederum erwiesenermaßen weitere Hände identifiziert wurden. Die anderen Schreiberhände müssten dann durch eigene Merkmalsvektoren von Hand 1 zu unterscheiden sein.

Bei der Materialauswahl wurden zwei Handschriften ausgemacht, welche die vorgenannten Bedingungen erfüllen: Cod. Guelf. 62 Weissenburg und Cod. Guelf. 63 Weissenburg. Beide Handschriften sind in der wichtigsten frühmittelalterliche Buchschrift, der karolingischen Minuskel, geschrieben, die sich durch einheitliche, relativ stark standardisierte Formen

¹⁸ Fecker et al. 2014, S. 1553.

und eine meist geringe individuelle Varianz auszeichnet. Diese allgemeinen Spezifika unterstützen die Brauchbarkeit von Digitalisaten karolingischer Handschriften ebenso wie deren meist hohes kodikologisches Niveau, das bei entsprechender Fotoaustattung nur wenig Nachbearbeitung erforderlich macht. Beide Handschriften entstammen zudem der Sammlung der im Rahmen des **Europeana Regia-Projekts** digitalisierten Codices Weissenburgenses, die weitere Vorteile aufweist: Die weitaus meisten Codices stammen aus dem Skriptorium des Klosters Weißenburg im Elsass, sind also regional und zeitlich gut einzuordnen. Mit dem den DFG-Richtlinien entsprechenden Katalog von Hans Butzmann sind diese Handschriften außerdem kodikologisch gut erschlossen. Die Beschreibungen liefern die Vorlagen der Schreiberidentifikation, die es zu verifizieren (oder falsifizieren) gilt. **Cod. Guelf. 62 Weiss.** ist dem Kolophon zufolge zwischen 819 und 826 im elsässischen Kloster Weissenburg während des Abbatats von Gerhoh von dem Mönch Waldmann geschrieben.¹⁹ In **Cod. Guelf. 63 Weiss.**, geschrieben ebenfalls in der ersten Hälfte des 9. Jahrhunderts, sind durch die klassische Paläographie drei Hände identifiziert, von denen eine die Hand Waldmanns ist. Waldmann werden die Seiten 82r bis 107r zugeschrieben.²⁰

5.1.1 Intra-Manuskript-Schreiberanalyse mit 62 Weiss.

Abbildung 5 zeigt die ermittelte Distanzmatrix in einer dreidimensionalen Darstellung. Hier zeigt sich, dass das Buch größtenteils von einem Schreiber (im Folgenden: Schreiber A) geschrieben wurde. Nur zum Ende des Manuskriptes treten variierende Schreiber (in Folgenden zusammengefasst als Schreiber B) auf, was sich in den höheren Distanzen ausdrückt. **Abbildung 6 (links)** zeigt eine Seite aus der Mitte des Manuskriptes, die von dem Hauptschreiber verfasst wurde. In **Abbildung 6 (rechts)** ist dagegen eine Seite vom Ende des Manuskriptes dargestellt, die sich im Schreibstil eindeutig von dem der Seiten in der Mitte des Buches unterscheidet.

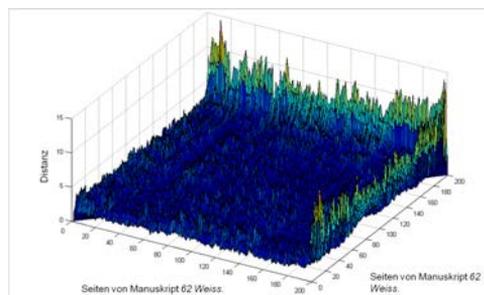


Abb. 5: Distanzmatrix für die Intra-Manuskript Schreiberanalyse von Manuskript 62 Weiss. (Quelle: Autoren).

¹⁹ Butzmann 1964, S. 202.

²⁰ Butzmann 1964, S. 203.

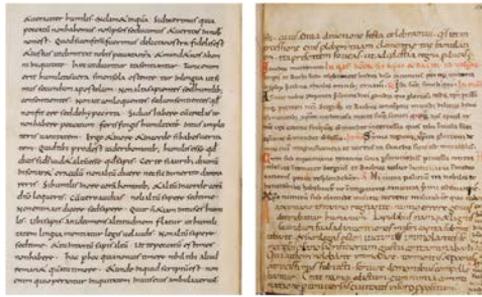


Abb. 6: Seite 20 aus dem Manuskript 62 Weiss, des Hauptschreibers (links). Seite 209 aus dem Manuskript 62 Weiss, mit variierendem Schreiber aus den letzten Seiten des Manuskripts (rechts) (Quelle: Autoren).

5.1.2 Intra-Manuskript-Schreiberanalyse mit 63 Weiss.

Ein etwas anderes Ergebnis zeigt die Distanzmatrix des Cod. Guelf. 63 Weiss. in Abbildung 7. Hier lassen sich aufgrund der Distanzunterschiede drei unterschiedliche Schreiber A, B und C identifizieren. Auffällig ist hier, dass die Unterschiede zwischen den Schreibern im Gegensatz zu den Unterschieden in Cod. Guelf. 62 Weiss. nur sehr gering sind und zu deutlich geringeren Distanzen führt. Grund hierfür dürfte die Tätigkeit im gleichen Skriptorium bzw. die Zugehörigkeit zu derselben Schule sein. Abbildung 8 (links, Mitte, rechts) zeigt von jedem Schreiber jeweils eine Seite aus dem Manuskript. Auf diesen ist die Ähnlichkeit sichtbar, im Gegenteil wird es schwierig, Differenzen zu bemerken. Die Unterschiede in der Schrift liegen jetzt nur in einzelnen Buchstabenformen wie z.B. den g-Formen, der Verwendung von Groß- und Kleinbuchstaben sowie im Schriftduktus.

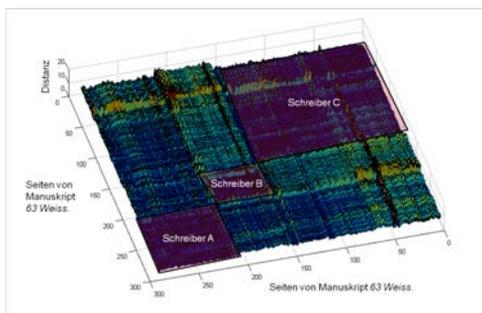


Abb. 7: Distanzmatrix für die Intra-Manuskript-Schreiberanalyse von Manuskript 63 Weiss. (Quelle: Autoren).

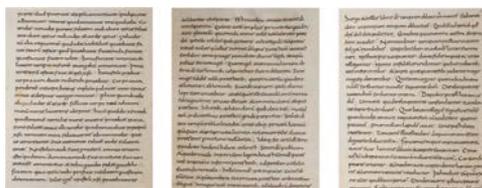


Abb. 8: Seite 94 aus Manuskript 63 Weiss, des Schreibers A (links). Seite 187 aus Manuskript 63 Weiss, des Schreibers B (Mitte). Seite 248 aus Manuskript 63 Weiss, des Schreibers C (rechts) (Quelle: Autoren).

5.1.3 Direkter Vergleich von Cod. Guelf. 62 Weiss. und 63 Weiss.

In diesem Experiment werden die Schreiber der Cod. Guelf. 62 Weiss. und 63 Weiss. miteinander verglichen. In *Die Weissenburger Handschriften* ist beschrieben, dass Teile von Cod. Guelf. 63 Weiss. von demselben Schreiber verfasst wurden, der auch 62 Weiss. verfasst hat.²¹ In der zwischen den beiden Manuskripten ermittelten Distanzmatrix, zu sehen in *Abbildung 9*, ist tatsächlich zu erkennen, dass die von dem Schreiber B geschriebenen Seiten von 63 Weiss. zu denen von 62 Weiss. eine hohe Ähnlichkeit aufweisen. Tatsächlich stimmen die mit diesem Verfahren ermittelten Seitennummern auch mit den Angaben von Butzmann überein.²²

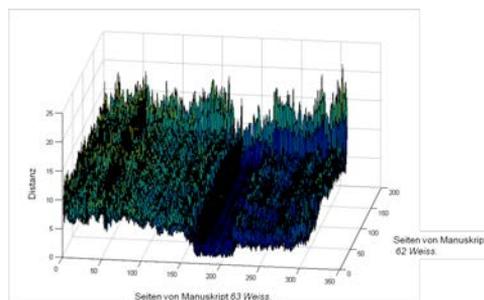


Abb. 9: Distanzmatrix für den direkten Vergleich der Manuskripte 62 Weiss. und 63 Weiss. (Quelle: Autoren).

5.2 Experimente zur One-to-Many-Klassifikation

In *Writer Identification for Historical Arabic Documents* wurden Untersuchungen zur One-to-Many-Klassifikation mit arabischen historischen Dokumenten im Rahmen des DFG-geförderten HADARA-Projektes²³ durchgeführt. Hierfür wurde eine frei-verfügbare Sammlung von arabischen, historischen Manuskripten verwendet, die Harvard Islamic Heritage Database.²⁴ Aus dieser Datenbank wurde eine Teilmenge von 60 Manuskripten ausgewählt. 43 dieser Manuskripte mit insgesamt 2313 Seiten stammen dabei von elf Schreibern, von denen jeder mindestens zwei Manuskripte geschrieben hat. Die restlichen 17 dieser Manuskripte mit insgesamt 2282 Seiten stammen von jeweils einem Schreiber. *Abbildung 10* zeigt Seiten aus drei Manuskripten dieser Datenmenge. Diese Manuskripte beinhalten die typischen Degradierungserscheinungen von historischen Dokumenten.

Für die Evaluation wird die Leave-one-out-Kreuzvalidierung eingesetzt. Bei dieser wird ein Manuskript als Testobjekt und die restlichen werden als Trainingsmenge verwendet. Hier werden nur diejenigen 43 Manuskripte, von denen ein Schreiber jeweils mindestens zwei Manuskripte verfasst hat, als Testobjekte verwendet. Die restlichen 17 Manuskripte werden

²¹ Butzmann 1964, S. 203.

²² Butzmann 1964, S. 203.

²³ Pantke et al. 2013, S. 151–156.

²⁴ Harvard University. [Islamic Heritage Project \(IHP\)](#).

in die Trainingsmenge mit aufgenommen. Für die Experimente bedeutet das, dass jedes der 43 Manuskripte einmal als Testobjekt verwendet wird. Die restlichen 59 Manuskripte dienen als Trainingsmenge mit bekannten Schreibern, von denen dem Testobjekt ein Schreiber zugeordnet werden soll.

In den Experimenten wurden jeweils die in **Abschnitt 3** und **4** vorgestellten Merkmalsextraktions- und Klassifikationsverfahren eingesetzt. Sämtliche Ergebnisse sind in Tabelle 1 (**Abbildung 3**) dargestellt. Neben der Genauigkeit der Klassifikation auf Manuskriptebene, bei der einem Manuskript genau ein Schreiber mit Hilfe der Verfahren Mittelung, Voting oder gewichtetem Voting zugeordnet wird, ist zusätzlich die Genauigkeit auf Seitenebene ermittelt worden. Dazu wird jeder Seite eines Manuskripts genau einem Schreiber zugeordnet. Am schlechtesten schneiden hier die konturbasierten Merkmale ab. Wie in *Writer Identification for Historical Arabic Documents* gezeigt, ergibt sich durch die Modifikation der Winkelmessung eine Verbesserung gegenüber dem originalen Ansatz,²⁵ jedoch reicht die Performanz nicht an die der anderen Verfahren heran.²⁶ Am besten schneiden die ebenfalls hier vorgestellten schlüsselpunktbasierten Merkmale ab. Diese sind mit der Voting-Klassifikationsstrategie besser als die texturbasierten Merkmale, sowohl auf Seitenebene als auch auf Manuskriptebene. Mit dem gewichteten Voting wird darüber hinaus auch eine perfekte Klassifikation erreicht, d.h. für jedes Manuskript wird der korrekte Schreiber erkannt. Einzig bei der Mittelung der Merkmalsvektoren schneiden die schlüsselpunktbasierten Merkmale schlecht ab. Es ist zu vermuten, dass durch die Mittelung bei diesen Merkmalen zu viel Information über einen Schreiber verloren geht.



Abb. 10: Seiten von drei Manuskripten aus der [Harvard Islamic Heritage Database](#)(Quelle: Autoren).

Merkmale	Genauigkeit Seitenebene (in %)	Genauigkeit Manuskriptebene (in %)		
		Mittelung	Voting	Gewichtetes Voting
Konturbasiert	68,6	81,0	88,1	92,9
Texturbasiert	87,6	92,9	92,9	92,9
Schlüsselpunktbasiert	92,5	59,5	97,6	100,0

Abb. 11: Tabelle 2: Ergebnisse der One-to-Many Klassifikation bei Verwendung der verschiedenen Merkmale und Klassifikationsstrategien (Quelle: Autoren).

6. Zusammenfassung

²⁵ Bulacu et al. 2007, S. 770.

²⁶ Fecker et al. 2014, S. 3054.

In diesem Beitrag wurde gezeigt, dass eine computergestützte Identifikation der schreibenden Hand bzw. Hände in historischen Dokumenten möglich ist. Hierfür wird eine Prozesskette aus Vorverarbeitung, Merkmalsextraktion und Klassifikation für die automatische Analyse gescannter Dokumente eingesetzt. Neben einer allgemeinen Vorgehensweise für die Vorverarbeitung dieser Dokumente wurden Merkmale vorgestellt, die effektiv die Charakteristika eines Schreibers beschreiben. Diese Merkmale basieren auf unterschiedlichen Charakteristika wie Kontur, Textur und Schlüsselpunkte der Handschrift. Darüber hinaus wurden diverse Einsatzgebiete für die Klassifikation von Schreibern vorgestellt, wie die Intra-Manuskript-Schreiberanalyse, der direkte Vergleich zweier Manuskripte und die One-to-Many-Klassifikation. Experimente mit diesen Verfahren zeigen vielversprechende Ergebnisse, die eine sinnvolle Unterstützung von Arbeiten im Bereich der Kodikologie und Paläographie erwarten lassen.

In der konkreten Anwendung sind vor allem zwei Szenarien denkbar:

- In einer Handschrift werden die Unterschiede der Schriften extrahiert, um die Anzahl schreibender Hände zu ermitteln;
- Im Vergleich zweier oder mehrerer Handschriften wird ermittelt, ob identifizierte Schreiber aus einer Handschrift auch in anderen Codizes tätig waren. Die Merkmale identifizierter Schriften könnten in einer Datenbank gesammelt werden und als Ausgangsmaterial für die Anwendung weiterer Algorithmen genutzt werden.

Die in den Szenarien 1 und 2 ermittelten Merkmalsvektoren könnten zur Identifikation von Schrifttypen generalisiert werden. Damit wäre es denkbar, regionale (insulare vs. kontinentale, länderspezifische) oder temporale (karolingische Minuskeln des 9. bzw. 11. Jhs.) Unterschiede zu identifizieren.²⁷

²⁷ Hiermit wird der der Deutschen Forschungsgemeinschaft (Projekt: FI 1494-3-2) für ihre finanzielle Unterstützung gedankt.

Bibliographische Angaben

Itay Bar-Yosef / Issac Beckman / Klara Kedem / Itshak Dinstein: Binarization, character extraction, and writer identification of historical hebrew calligraphy documents. In: International Journal on Document Analysis and Recognition 9 (2007), 2, S. 89–99. [\[Nachweis im GBV\]](#)

Marius Bulacu / Lambert Schomaker / Axel Brink: Text-Independent Writer Identification and Verification on Offline Arabic Handwriting. In: Proceedings of the International Conference of Document Analysis and Recognition (ICDAR). Hg. von IEEE. Parana 2007, S. 769–773. DOI 10.1109/ICDAR.2007.4377019. [\[Nachweis im GBV\]](#)

Hans Butzmann: Die Weissenburger Handschriften. Frankfurt/Main 1964 (= Kataloge der Herzog August Bibliothek Wolfenbüttel: Neue Reihe, Bd. 10). [\[Nachweis im OPAC\]](#)

Daniel Fecker / Abed Asi / Volker Märgner / Jihad El-Sana / Tim Fingscheidt: Writer Identification for Historical Arabic Documents. In: Proceedings of the International Conference on Pattern Recognition (ICPR). Stockholm 2014, S. 3050–3055. [\[Nachweis im GBV\]](#)

Abdelaali Hassane / Somaya Al-Madeed: ICFHR 2012 Competition on Writer Identification Challenge 2: Arabic Scripts. In: Proceedings of the International Conference on Frontiers in Handwriting Recognition (ICFHR). Hg. von IEEE. Bari 2012, S. 835–840. DOI 10.1109/ICFHR.2012.218. [\[Nachweis im GBV\]](#)

Thomas Haye / Stephan Müller: Mittelalter-Philologie im Internet. Digitalisierung mittelalterlicher Handschriften aus Sicht der Forschung. In: Zeitschrift für deutsches Altertum und deutsche Literatur 140 (2011), S. 416–420. [\[Nachweis im OPAC\]](#)

David G. Lowe: Distinctive Image Features from Scale-Invariant Keypoints. In: International Journal of Computer Vision 2004 (60), S. 91–110. [\[Nachweis im GBV\]](#)

Sreeraj M / Sumam Mary Idicula: A Survey on Writer Identification Schemes. In: International Journal of Computer Applications 26 (2011), Nr. 2, S. 23–33. DOI 10.5120/3075-4205. [\[online\]](#)

Andrew J. Newell / Lewis D. Griffin: Natural Image Character Recognition Using Oriented Basic Image Features. In: Proceedings of the International Conference Digital Image Computing Techniques and Applications (DICTA). Hg. von IEEE. Noosa 2011, S. 191–196. DOI 10.1109/DICTA.2011.39. [\[Nachweis im GBV\]](#)

Werner Pantke / Arne Haak / Volker Märgner: Color Segmentation for Historical Documents Using Markov Random Fields. In: Proceedings of the 6th International Conference on Soft Computing and Pattern Recognition (SoCPaR 2014). Tunis 2014, S. 151–156. DOI 10.1109/SOCPaR.2014.7007997.

Werner Pantke / Volker Märgner / Daniel Fecker / Tim Fingscheidt / Abed Asi / Ofer Biller / Jihad El-Sana / Raid Saabni / Mohammed Yehia: HADARA – A Software System for Semi-Automatic Processing of Historical Handwritten Arabic Documents. In: Proceedings of IS&T Archiving 2013, Washington D.C. 2013, S. 161–166 [\[online\]](#). [\[Nachweis im GBV\]](#)

Peter M.W. Robinson: Is There a Text in These Variants? In: The Literary Text in the Digital Age. Hg. von Richard J. Finneran. Ann Arbor (Mi) 1996, S. 99–115. [\[Nachweis im OPAC\]](#)

Thomas Stäcker: Erfahrungsbericht Helmstedter Drucke Online an der Herzog August Bibliothek Wolfenbüttel. In: Volltext via OCR. Möglichkeiten und Grenzen. Testsznarien zu den Funeralschriften der Staatsbibliothek zu Berlin – Preußischer Kulturbesitz. Hg. von Maria Federbusch / Christian Polzin. Berlin 2013 (= Beiträge aus der Staatsbibliothek zu Berlin – Preußischer Kulturbesitz Band 43), S.123–135. [\[online\]](#)

Abbildungslegenden und -nachweise

Abb. 1: Seite eines arabischen, historischen Dokumentes mit Haupttext und vielen Kommentaren (links); Beispiel einer Segmentierung der Handschrift mit Hilfe einer Binarisierung: Teil eines Ausgangsbildes (Mitte) und binäres Ergebnisbild (rechts) (Quelle: Autoren).

Abb. 2: Exemplare für Schlüsselpunkte in einer Handschrift. Die Mittelpunkte der Kreise deuten auf den Ort des Schlüsselpunktes, der Durchmesser auf die Skalierung, und die Orientierung ist durch die Linie gegeben (links). Deskriptoren der ermittelten Schlüsselpunkte (rechts) (Quelle: Autoren).

Abb. 3: Tabelle 1: Schematische Darstellung einer Distanzmatrix. Zur Erläuterung: D (S1,S2) entspricht der Distanz der Merkmalsvektoren der ersten und zweiten Seite eines Manuskripts (Quelle: Autoren).

Abb. 4: One-to-Many-Klassifikation eines mehrseitigen Manuskriptes durch Mittelung der Merkmalsvektoren (links). One-to-Many-Klassifikation eines mehrseitigen Manuskriptes durch Voting (rechts) (Quelle: Autoren).

Abb. 5: Distanzmatrix für die Intra-Manuskript Schreiberanalyse von Manuskript 62 Weiss. (Quelle: Autoren).

Abb. 6: Seite 20 aus dem Manuskript 62 Weiss. des Hauptschreibers (links). Seite 209 aus dem Manuskript 62 Weiss. mit variierendem Schreiber aus den letzten Seiten des Manuskripts (rechts) (Quelle: Autoren).

Abb. 7: Distanzmatrix für die Intra-Manuskript-Schreiberanalyse von Manuskript 63 Weiss. (Quelle: Autoren).

Abb. 8: Seite 94 aus Manuskript 63 Weiss. des Schreibers A (links). Seite 187 aus Manuskript 63 Weiss. des Schreibers B (Mitte). Seite 248 aus Manuskript 63 Weiss. des Schreibers C (rechts) (Quelle: Autoren).

Abb. 9: Distanzmatrix für den direkten Vergleich der Manuskripte 62 Weiss. und 63 Weiss. (Quelle: Autoren).

Abb. 10: Seiten von drei Manuskripten aus der [Harvard Islamic Heritage Database](#) (Quelle: Autoren).

Abb. 11: Tabelle 2: Ergebnisse der One-to-Many Klassifikation bei Verwendung der verschiedenen Merkmale und Klassifikationsstrategien (Quelle: Autoren).

Zeitschrift für digitale Geisteswissenschaften

Artikel aus:

Sonderband 1 der ZfdG: Grenzen und Möglichkeiten der Digital Humanities. Hg. von Constanze Baum und Thomas Stäcker. 2015. DOI: [10.17175/sb01](https://doi.org/10.17175/sb01)

Titel:

Historische Stilometrie? Methodische Vorschläge für eine Annäherung textanalytischer Zugänge an die mediävistische Textualitätsdebatte.

Autor/in:

Gabriel Viehhauser

Kontakt:

viehhaus@germ.unibe.ch

Institution:

Universität Bern

GND:

[139919783](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-39919783)

ORCID:

[0000-0001-6372-0337](https://orcid.org/0000-0001-6372-0337)

DOI des Artikels:

[10.17175/sb001_009](https://doi.org/10.17175/sb001_009)

Nachweis im OPAC der Herzog August Bibliothek:

[83016927X](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-83016927X)

Erstveröffentlichung:

19.02.2015

Lizenz:

Sofern nicht anders angegeben 

Medienlizenzen:

Medienrechte liegen bei den Autoren

Letzte Überprüfung aller Verweise:

24.05.2016

GND-Verschlagwortung:

[Mediävistik](#) | [literarischer Stil](#) | [Textanalyse](#) |

Zitierweise:

Gabriel Viehhauser: Historische Stilometrie? Methodische Vorschläge für eine Annäherung textanalytischer Zugänge an die mediävistische Textualitätsdebatte.. In: Grenzen und Möglichkeiten der Digital Humanities. Hg. von Constanze Baum / Thomas Stäcker. 2015 (= Sonderband der Zeitschrift für digitale Geisteswissenschaften, 1). PDF Format ohne Paginierung. Als text/html abrufbar unter DOI: [10.17175/sb001_009](https://doi.org/10.17175/sb001_009).

Gabriel Viehhauser

Historische Stilometrie? Methodische Vorschläge für eine Annäherung textanalytischer Zugänge an die mediävistische Textualitätsdebatte.

Abstracts

Da die mediävistische Textualitätsdebatte von einer deutlichen Zurückhaltung gegenüber der Instanz des Autors geprägt ist, stehen die Voraussetzungen für eine Akzeptanz des Einsatzes stilometrischer Methoden auf dem Gebiet der mittelhochdeutschen Literatur nicht günstig. Doch eröffnet gerade das Spannungsfeld von Tradition und Individualität, das für diese Texte von Bedeutung ist, der Stilometrie Anwendungsmöglichkeiten, die zu einer gegenseitigen Erhellung von Methode und Fachdisziplin führen können. Der Artikel zeigt anhand von Beispielanalysen, wie mittelhochdeutsche Texte sowohl durch individuelle als auch durch gattungsspezifische Stilfaktoren bestimmt sind und was bei der Quantifizierung dieser Kategorien zu beachten ist.

The debate on medieval textuality is shaped by an obvious reservation about the category of authorship. Therefore, the conditions for the acceptance of stylometric methods in the field of Middle High German literature might not seem too favorable. However, it is exactly this tension between tradition and individuality, which bears great significance for those texts, that offers opportunities for the application of stylometric methods, which can lead to a mutual elucidation of method and discipline. The paper provides examples of the ways in which Middle High German texts are influenced by individual as well as genre-specific style factors. Furthermore, it elaborates on the issues that must be addressed for a proper quantification of these categories.

1. Zur Problematik des Stilbegriffs in der Mediävistik

Im Rahmen der Digital Humanities stellt die Stilometrie, also die Anwendung quantitativer Methoden zur Erfassung und Klassifizierung stilistischer Merkmale von Texten, eine jener Unterdisziplinen dar, die zur Zeit eine gesteigerte Aufmerksamkeit der Forschung für sich beanspruchen kann. Das Potential und die Beliebtheit der Methode belegt nicht zuletzt eine große Zahl von Publikationen, die sich auch für den deutschsprachigen Raum nachweisen lässt.¹ Aufgrund ihrer Aktualität bietet sich die Stilometrie demnach als Untersuchungsfeld geradezu an, wollte man, wie dies im Folgenden geschehen soll, exemplarisch der Frage nachgehen, ob die Digital Humanities einen methodischen Brückenschlag zu den traditionellen Geisteswissenschaften ermöglichen können oder aber die Gefahr einer »feindlichen Übernahme« der Geisteswissenschaften durch die Computerphilologie besteht.² Was die Methode diesbezüglich vor allem interessant macht, ist der Umstand, dass sie in letzter Zeit offenkundig auch in der traditionellen Literaturwissenschaft verstärkt wahrgenommen wird. Dies scheint zumindest für die germanistische Mediävistik zuzutreffen, dem Fachgebiet,

¹ Ungeachtet ihrer derzeitigen Konjunktur handelt es sich bei der Stilometrie freilich nicht um eine gänzlich neue Methode. Einen Überblick über die ältere Forschung bietet Schöch 2014, S. 133f.

² Ich greife damit die vom Verband [Digital Humanities im deutschsprachigen Raum](#) (DHD) auf der Passauer Jahrestagung ausgegebene Grundfrage auf.

dem ich mich näher zuwenden will. Als symptomatisch für das Interesse kann hier etwa die Aufnahme eines einschlägigen Aufsatzes von Mike Kestemont in die traditionsreiche *Zeitschrift für deutsches Altertum* gewertet werden³ oder der programmatische Artikel von Manuel Braun in der letztjährigen Ausgabe der *Zeitschrift für Literaturwissenschaft und Linguistik*, die unter dem Titel »Turn, Turn, Turn? Oder: Braucht die Germanistik eine germanistische Wende?« den gegenwärtigen Stand der Fachdisziplin reflektiert.⁴

Fast sieht es also so aus, als könnte sich nach der mittlerweile recht gut etablierten digitalen Editorik die nächste Teildisziplin der Digital Humanities anschicken, ihren Platz auf dem Feld der althergebrachten Germanistik zu erobern. Jedoch stehen die Voraussetzungen für eine Akzeptanz der Stilometrie wesentlich ungünstiger, als dies bei der Editionswissenschaft der Fall war. Denn rückblickend lässt sich feststellen, dass die Etablierung der digitalen Editorik nicht zuletzt deswegen gelingen konnte, weil sich die Methode in vorzüglicher Weise zur Beantwortung von Fragestellungen eignete, die bereits im Fokus der literaturwissenschaftlichen Diskussion standen: Die Möglichkeiten einer dynamischen, von Autorpositionen freigehaltenen Betrachtungsweise von Texten trafen sich mit den Anforderungen einer stark poststrukturalistisch beeinflussten Philologie, die Konzepte von Subjekt, Autorschaft und Textfestigkeit zunehmend in Frage stellte.⁵ Selbst nach dem Abklingen der heftig geführten Debatte um die sogenannte *New Philology* und der partiellen Rücknahme allzu radikaler Extrempositionen ist nicht von der Hand zu weisen, dass die digitale Darstellung gerade für mittelalterliche Texte besondere Vorteile bietet, da diese nur handschriftlich und in den allermeisten Fällen nur in späteren Abschriften, nicht aber in einem vom Autor selbst angefertigten Original überliefert sind.⁶ Auch wenn man aufgrund dieser Umstände die Bedeutung von Autorschaft nicht gleich völlig in Abrede stellen möchte, wie dies poststrukturalistische Extrempositionen tun, so bleibt doch der Befund, dass das Autororiginal in der Überlieferung nicht mehr direkt greifbar ist. Statt nun dessen ungeachtet gezwungenermaßen nur einen, möglicherweise sogar hypothetisch rekonstruierten ›Original-Text bieten zu müssen, wie dies früher bei den nur gedruckt vorliegenden Editionen der Fall war, lässt sich mittels einer digitalen Editionsdarstellung die gesamte handschriftliche Überlieferung mit einbeziehen. Dies ermöglicht den Blick auf eine dynamische Entwicklungsgeschichte des Textes im Verlauf seiner Tradierung, die nicht allein durch den Autor bestimmt wird, wodurch aber grundlegende theoretische Fragen wie jene, was die Integrität eines Textes oder die Rolle des Autors überhaupt ausmachen, nicht nur neu angegangen werden können, sondern sich regelrecht aufdrängen.

Demgegenüber müssen die primären Anwendungsgebiete der Stilometrie aus der Sicht der traditionellen Literaturwissenschaft zumindest auf den ersten Blick deutlich rückwärtsgewandt, wenn nicht gar altmodisch anmuten. Nur allzu leicht etwa könnten stilistische Untersuchungen

³ Kestemont 2013, S. 1–33.

⁴ Braun 2013, S. 83–90. Braun beschäftigt sich in dem Artikel mit den Möglichkeiten computerunterstützter Zugangsweisen in der Literaturwissenschaft insgesamt. Sein Fazit, dass der Einfluss der Digital Humanities zu einer Wiederannäherung von Sprach- und Literaturwissenschaft führen kann, betrifft auch die Stilometrie im engeren Sinn.

⁵ Von grundlegendem Einfluss waren hier etwa Foucault 1988 sowie Cerquiglini 1989. Vgl. zu den methodisch-technologischen Voraussetzungen digitaler Editionen den Überblick bei Sahle 2013, insbesondere Teil 1, S. 191ff.

⁶ Zur Debatte um die *New Philology* vgl. Stackmann 1994 sowie Glessgen / Lebsanft 1997.

zur Identifizierung von Autorschaft, die bislang einen der ergebnisträchtigen Zweige der Stilometrie darstellen, Assoziationen an mittlerweile als überholt geltende Ansätze zur Klärung von ›Echtheitsfragen‹ und anderen wertästhetischen Problemstellungen hervorrufen.⁷ Denn die Frage nach der Beschaffenheit von Autorstilen wurde in der traditionellen Philologie durchaus bereits gestellt, gerade nämlich bei der heute verpönten Rekonstruktion von Originaltexten, zu der sich die ältere Forschung aufgrund ihrer Ausrichtung auf feste Textgestalten genötigt sah. Da es den Vertretern der klassischen Textkritik vor allem darum ging, den einen, ›richtigen‹ Autortext aus der ihrer Ansicht nach entstellten handschriftlichen Überlieferung zu rekonstruieren, erfolgte die Einrichtung der Ausgabentexte nicht selten unter der Maßgabe, ob eine bestimmte Formulierung dem jeweiligen Autor stilistisch zuzutrauen sei oder nicht. Die Frage nach dem ›richtigen‹ Text traf sich dabei mit einem überhöhten Stilideal, das davon ausging, dass die besten Autoren sich auch durch einen besonders qualitätsvollen Stil auszeichnen sollten, und daher stilistisch Minderwertiges aus dem Text auszuschneiden sei. Auf dieser Grundlage wurden Entscheidungen getroffen, die mitunter höchst subjektiv waren, was diese Methode nicht zu Unrecht in Verruf gebracht hat.

Es zeigt sich also, dass die Wissenschaftsgeschichte der Mediävistik nicht die besten Anschlussvoraussetzungen für eine autorzentrierte Stilometrie bietet. Gerade eine Fixierung auf die Instanz ›Autor‹ könnte den erkenntnistheoretischen Mehrwert in Frage stellen, der sich durch die Anwendung der Computertechnologie mit ihrer Möglichkeit zur Öffnung und Perspektivierung von Texten ergeben hat. In dieser Hinsicht erscheint die Stilometrie auf den ersten Blick der erfolgreichen digitalen Editorik fast schon diametral entgegengesetzt.

Dieser Befund überrascht umso mehr, wenn man bedenkt, dass gerade der vielleicht wirkmächtigste Ansatz zur Rehabilitierung quantitativer textanalytischer Methoden in den Geisteswissenschaften, das von Franco Moretti in die Diskussion eingebrachte Konzept des *Distant Reading*, eigentlich mit dem Anspruch angetreten ist, die Fokussierung auf Höhenkammliteratur und Autorgenies zu unterlaufen: Durch die Anwendung quantitativer Methoden, so paraphasiere ich Moretti, werde es nämlich möglich, ein realistischeres Bild der gesamten Schriftproduktion zu erhalten, das sich nicht auf wenige elitäre Spitzenprodukte beschränken muss, sondern, aufgrund der erhöhten Verarbeitungskapazitäten des Computers, die breite Masse der tatsächlich vorhandenen Texte erschließen kann.⁸ Dieser reizvolle, provokante Ansporn ginge verloren, würde sich die Stilometrie allzu sehr auf die Aufdeckung von Autorschaftszugehörigkeiten beschränken.

Doch wird gerade an den Arbeiten von Moretti deutlich, dass sich die Stilometrie durchaus auch zur Klärung nicht-autorbezogener Fragestellungen verwenden lässt. So hat sich die stilometrische Forschung etwa schon bald – und nicht erst seit Moretti – an der Klassifizierung von Texten in Hinblick auf ihre Gattung versucht, und zwar allein schon aus methodischen Gründen.⁹ Der Einbezug dieser Kategorie liegt deshalb nahe, weil sich gerade die Zugehörigkeit

⁷ Dieses forschungsgeschichtliche Erbe erschwert den Umgang der Mediävistik mit dem Stilbegriff insgesamt, vgl. hierzu Haustein 2011, S. 46. Zum Verhältnis von Stilometrie und Stilforschung jetzt grundlegend Jannidis 2014, insbesondere S. 178–183.

⁸ Vgl. Moretti 2000; Moretti 2005, S. 3f.

⁹ Vgl. etwa Biber 1992, Moretti 2005 und jüngst insbesondere Jockers 2013, Jannidis / Lauer 2014 sowie Schöch 2014.

von Texten zu unterschiedlichen Gattungen als größter Störfaktor bei der Differenzierung von Individualstilen erwiesen hat. Ganz offenbar ist der Stil eines Textes nicht nur von der Persönlichkeit des Autors abhängig, sondern auch von überindividuellen Faktoren, wie sie etwa die Gattungen darstellen.

Gerade dieses Spannungsfeld zwischen Tradition und Individualität, welches sich in der Determiniertheit der Texte zwischen Autor- und Gattungsstil zeigt, eröffnet nun aber sehr wohl Anschlussmöglichkeiten an Fragestellungen, die wieder mehr ins Zentrum aktueller Diskussionen in den traditionellen Literaturwissenschaften führen. Denn Morettis Versuch einer Entdifferenzierung zwischen Spitzenprodukten und breiter Masse an Texten erscheint bei näherer Hinsicht der spezifisch mittelalterlichen Literaturauffassung vielleicht gar nicht völlig unangemessen. So wurde bei der Beschreibung mittelalterlicher Literatur immer wieder die Notwendigkeit betont, die einzelnen Texte nicht nach dem Maßstab einer genieästhetischen, auf Originalität abzielenden Literaturproduktion zu beurteilen, sondern den Eigenwert einer speziellen »Ästhetik der Identität«¹⁰ anzuerkennen: Mittelalterliche Texte zielen, anders als wir dies vom modernen Literaturbetrieb seit der Goethezeit gewohnt sind, gar nicht darauf ab, völlig Neues oder Individuelles zu schaffen, sondern knüpfen bewusst an Traditionen an. Statt um das Neuerfinden einer Geschichte geht es vielmehr um das *Wiedererzählen* althergebrachter Stoffe.¹¹ So ist beispielsweise der *Iwein* von Hartmann von Aue, einer der wichtigsten Artusromane der mittelhochdeutschen Literatur, über große Passagen hinweg eine ziemlich genaue Übertragung seiner französischen Quelle, des *Yvain* von Chrétien de Troyes. Die Eingriffe Hartmanns in den Text sind zwar durchaus aussagekräftig, führen aber nicht zu einer völligen Lösung von seiner Vorlage. Dieser Mangel an Originalität wird dabei nicht als Manko empfunden, sondern im Gegenteil, gerade die Wiederaufnahme des Altbewährten, immer schon Gültigen, rechtfertigt erst das literarische Tun. Aufgrund dieser grundsätzlichen Ausrichtung mittelalterlicher Literatur kommt der Einordnung der Texte in Traditions- und damit Gattungszusammenhänge tendenziell eine größere Bedeutung zu als bei modernen Texten -- und dementsprechend, so ließe sich zumindest vermuten, verringert sich die Relevanz des Individualstils eines einzelnen Autors. Dass die Instanz »Autor« deswegen aber nicht gleich völlig verloren geht, hat die sich am Poststrukturalismus abarbeitende Autorschaftsdebatte in der Mediävistik ebenfalls deutlich gezeigt.¹² Auch das Bewusstsein für literarische »Spitzenprodukte« fehlt im Mittelalter nicht, so wird etwa – um wieder bei Hartmann als Beispiel zu bleiben – dieser im bekannten »Literaturkatalog« aus dem *Tristan* Gottfrieds von Straßburg neben anderen als herausragender Autor gepriesen und gerade für seinen besonders klaren Stil gelobt.¹³

2. Fallstudien

¹⁰ Der Ausdruck »Ästhetik der Identität« wurde von Lotman 1993, S. 410ff. geprägt.

¹¹ Vgl. Worstbrock 1999, S. 128–130.

¹² Vgl. exemplarisch Schnell 1998.

¹³ Vgl. die *Tristan*-Verse 4589–4852. Im »Literaturkatalog« unterbricht Gottfried seine Erzählung von Tristan, um eine Reihe seiner Meinung nach vorbildlicher Autoren zu nennen. Hartmann werden dabei insbesondere seine »kristallinen wortelin«, also seine kristall(klaren) Worte, zugutegehalten (Vers 4629).

2.1 Voraussetzungen

Vor diesem Hintergrund könnte es nun gerade besonders interessant erscheinen, auch mittelalterliche Texte in Hinblick auf ihre stilometrische Auswertbarkeit zu überprüfen.

Dem stehen jedoch gewichtige praktische Gründe entgegen: Stilometrische Verfahren beruhen im Wesentlichen auf der computerunterstützten Auszählung von Worthäufigkeiten, auf deren Grundlage Texte mit ähnlichem Wortgebrauch nach statistischen Verfahren sortiert werden. Zwar sind mittlerweile immer mehr mittel- und frühneuhochdeutsche Texte in elektronischer Form verfügbar, was die Grundlage für die Wortfrequenzfassung bietet, die Vergleichbarkeit des Wortgebrauchs in diesen Texten ist jedoch mit erheblichen Schwierigkeiten behaftet. Denn im mittelalterlichen Deutsch gibt es keine festgelegte Orthographie, die Schreibung ein- und desselben Wortes kann variieren, einerseits regional, weil die Schreiber aus unterschiedlichen Dialektgebieten stammen, andererseits auch zeitbedingt, weil sich die Schreibsprache im Verlauf des Mittelalters erheblich verändert hat. Darüber hinausgehend können die Schreiber der Handschriften auch unabhängig von Dialekten und Zeitstufen unterschiedliche Schreibgewohnheiten aufweisen, was sogar soweit geht, dass ein und dasselbe Wort in derselben Handschrift unterschiedlich geschrieben auftritt. Zwar werden diese Differenzen in den meisten der heute verfügbaren Textausgaben durch den Herausgeber ausgeglichen und die Schreibung bis zu einem gewissen Grad vereinheitlicht, doch stellt gerade diese Normalisierung einen weiteren Störfaktor dar, denn auch für diese Vereinheitlichung gibt es keine bis in alle Details festgelegten Regeln, weshalb sie je nach Herausgeber unterschiedlich ausfallen kann. Schließlich ist noch ein weiterer Punkt zu nennen, der bei der Untersuchung vieler mittelhochdeutscher Texte ins Gewicht fällt: Die wichtigsten literarischen Werke sind in Versen abgefasst, also auch epische bzw. erzählende Texte, die zumeist in Reimpaarversen, manchmal sogar in Strophen gebunden sind. Die Besonderheiten der metrischen Struktur und der Reimbindung wirken sich ebenfalls auf den Wortgebrauch aus, und auch dies kann die Vergleichbarkeit der Texte erschweren.¹⁴

2.2 Höfische Epik

Trotz dieser Schwierigkeiten, mit denen eine stilometrische Analyse mittelalterlicher Texte konfrontiert ist, habe ich in einer ersten Annäherung zu eruiieren versucht, welche Resultate sich mit den bereits vorhandenen Textkorpora erzielen lassen. Ziel der im Folgenden beschriebenen Fallstudien ist es also ausdrücklich nicht, endgültige Ergebnisse zu liefern, sondern zunächst heuristisch zu ermitteln, welche Probleme genauer in den Blick zu nehmen sind. Die Grundlage für meinen ersten Versuch bildete ein Korpus von einigen wichtigen epischen Texten der mittelhochdeutschen Literatur, das ich mithilfe des von Maciej Eder, Mike Kestemont und Jan Rybicki entwickelten *Stylo-Packages*¹⁵ für das Statistikprogramm R¹⁶ einer

¹⁴ Dazu in **Abschnitt 2.2** ein Beispiel. Allgemein zu möglichen Fehlertypen in Textkorpora und deren Auswirkungen auf die stilometrische Analyse Eder 2013, S. 604ff. Eders Versuche zeigen, dass sich stilometrische Auswertungen erstaunlich robust gegenüber *Noise* in den herangezogenen Textkorpora erweisen (vgl. insbesondere Eder 2013, S. 610ff.).

¹⁵ Eder / Kestemont / Rybicki 2013.

¹⁶ R Core Team 2013.

stilometrischen Analyse ausgesetzt habe. Konkret herangezogen habe ich die heute erhaltenen Werke der drei Klassiker der höfischen Epik, den *Tristan* Gottfrieds von Straßburg, den *Parzival* und den *Willehalm* Wolframs von Eschenbach sowie die epischen Werke Hartmanns von Aue, nämlich die beiden Artusromane *Erec* und *Iwein* sowie die Legendendichtungen *Der arme Heinrich* und *Gregorius*.¹⁷ Um eine weitere zeitliche Streuung zu erreichen, wurden zudem noch der *Eneas*-Roman Heinrichs von Veldeke, der eine Art Vorläufer der höfischen Klassik darstellt, in die Analyse mit einbezogen sowie die Werke Konrads von Würzburg und Ulrichs von Etzenbach, die zeitlich etwas später anzusiedeln sind.

Eine Clusteranalyse dieser Texte, die auf der Grundlage der 200 häufigsten Wörter erstellt wurde,¹⁸ bietet ein auf den ersten Blick erstaunlich eindeutiges Bild (Abbildung 1):

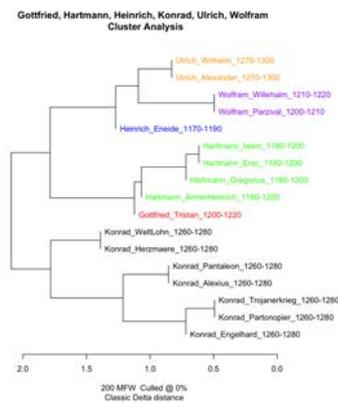


Abb. 1: Clusteranalyse der höfischen Epik (Graphik: Autor).

Die Werke der Autoren werden vom Computer ohne Ausnahme richtig sortiert, zum Teil scheinen sich sogar Texte derselben Gattung zusammenzuordnen, so bilden etwa die beiden Artusromane *Erec* und *Iwein* einen eigenen Unterzweig in der Hartmann-Gruppe.¹⁹ Grundsätzlich scheint jedoch das Autorensignal das Gattungssignal zu überwiegen, so befinden sich etwa der *Parzival* und der *Willehalm*, die unterschiedlichen Gattungen angehören, auf ein und demselben Wolfram-Zweig. Diese Ergebnisse sind freilich mit äußerster Vorsicht zu behandeln: Denn natürlich spielen die bereits erwähnten Einflussfaktoren wie Schreiber, Schreibdialekt und Herausgeber bei dieser Sortierung eine Rolle, und ich werde im Folgenden noch näher auf ein Beispiel für eine solche Verzerrung eingehen. Allerdings wäre selbst unter Berücksichtigung dieser Faktoren nicht unbedingt eine solch klare Verteilung zu

¹⁷ Zu den verwendeten Ausgaben vgl. das Literaturverzeichnis.

¹⁸ Die Analyse beruht auf einer Auszählung der prozentual am häufigsten auftretenden Wörter in den einzelnen Texten. Je ähnlicher zwei Texte sich in Bezug auf ihren Wortgebrauch verhalten, desto näher werden sie zusammensortiert. Zur Berechnung der Nähe kommen statistische Methoden zum Einsatz, im vorliegenden Fall das von Burrows 2002 entwickelte Delta-Distanzmaß, das mit standardisierten Wortfrequenzen arbeitet.

¹⁹ Versuche mit anderen Distanzmaßen (Manhattan Distance, Eder's Delta, vgl. dazu die Dokumentation bei Eder / Rybicki / Kestemont 2014, S. 14ff.) liefern jeweils ein ähnliches Bild. Eine signifikante Änderung ergibt sich aber bei der Berechnung der Darstellung auf der Grundlage von 1000 Wörtern: Hier sprengt der *Tristan* die Gruppe der Hartmann-Werke und bildet mit dem *Armen Heinrich* einen eigenen Zweig. Zu den möglichen Gründen hierfür siehe [Abschnitt 2.2](#).

erwarten, da ja die Werke mancher der hier aufgeführten Autoren in ganz unterschiedlichen Handschriften überliefert sind und auch von unterschiedlichen Editoren herausgegeben wurden. So ist beispielsweise die Ausgabe von Hartmanns *Iwein* nach Handschriften des 13. Jahrhunderts erstellt, der nebengeordnete *Erec* jedoch nach einer Handschrift, die erst aus dem 16. Jahrhundert stammt. Diese Differenz in der handschriftlichen Basis spielt also offensichtlich keine genügend große Rolle, als dass sie eine Zusammenordnung der beiden Texte verhindern könnte.²⁰

2.3 Wolfram und Hartmann

Die Ergebnisse erscheinen jedenfalls vielversprechend genug, um diesen Zusammenhängen genauer nachzugehen. In einem weiteren Versuch habe ich daher Wolframs *Parzival* und die beiden Artusromane Hartmanns, den *Erec* und den *Iwein*, mithilfe des von John Burrows und Hugh Craig entwickelten Zeta-Tests einer kontrastiven Analyse unterzogen. Dabei werden die einzelnen Texte in Abschnitte gleicher Länge zerteilt (hier in Abschnitte zu 2000 Wörtern), die jeweils von einer Textgruppe (Wolfram bzw. Hartmann) im Vergleich zur anderen konstant bevorzugten Wörter ermittelt und schließlich die einzelnen Textpartien in eine graphische Darstellung gebracht, bei der wieder Textabschnitte mit ähnlichem Wortgebrauch näher sortiert werden.²¹ Als Testgruppe sind der Analyse noch Wolframs *Willehalm*, Gottfrieds *Tristan* und Hartmanns Legendedichtungen hinzugefügt worden (Abbildung 2).

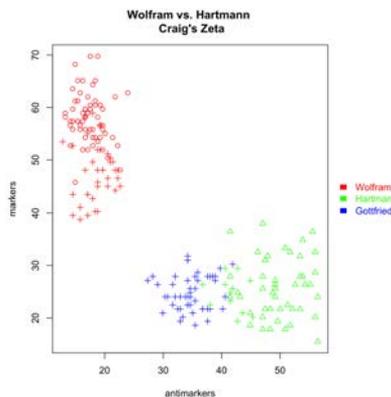


Abb. 2: Kontrastive Analyse der Werke Hartmanns, Wolframs und Gottfrieds (Graphik: Autor).

In der Darstellung repräsentiert jedes Symbol (Kreis, Kreuz und Dreieck) jeweils einen Textabschnitt. Bei den roten Kreisen handelt es sich um Textpartien aus Wolframs *Parzival* (also aus der ersten Textgruppe), bei den grünen Dreiecken um Abschnitte aus Hartmanns

²⁰ Freilich sind die Ausgaben des *Iwein* und des *Erec* nach ähnlichen Maßgaben normalisiert worden, doch auch in Bezug auf den Herausgeberfaktor gilt ähnliches wie für die Handschriften: Trotz zum Teil unterschiedlicher Ausgabenkonventionen werden die Werke gleicher Autoren sortiert.

²¹ Die Auswertung erfolgte wieder mit dem Stylo-Package von Eder / Kestemont / Rybicki 2013. Zum Zeta-Test vgl. Burrows 2007 sowie Craig / Kinney 2009. Anders als das in [Abschnitt 2.2](#) zur Anwendung gebrachte Delta-Distanzmaß zielt Craig's Zeta weniger auf die absolute Worthäufigkeit ab als auf die Konstanz des Wortgebrauchs über die einzelnen Abschnitte hinweg.

Artusromanen (der zweiten Textgruppe). Kreuze stellen Abschnitte aus der Testgruppe dar (*Willehalm*, *Tristan* und Hartmanns Legenden), wobei die Partien aus dem *Willehalm* rot, die aus dem *Tristan* blau und die aus Hartmanns Legenden grün eingefärbt sind.

Überwiegend zeigt sich wieder ein relativ klarer Unterschied zwischen den Autoren, die Textpartien sind nicht weit über das Raster verteilt, sondern ordnen sich entsprechend den Gruppierungen zusammen. Der *Willehalm* erscheint zusammen mit dem *Parzival* deutlich von Hartmann abgesetzt und die Legenden Hartmanns finden sich bei seinen Artusromanen. Einzig beim *Tristan* gibt es Unschärfen, da er sich teilweise mit den Werken Hartmanns vermischt.

Spätestens an diesem Punkt wäre es nun interessant zu wissen, welche Wörter für diese Sortierung verantwortlich sind. Eine Auflistung der für die Differenzierung der beiden Korpora aussagekräftigsten Wörter sieht folgendermaßen aus (Abbildung 3):

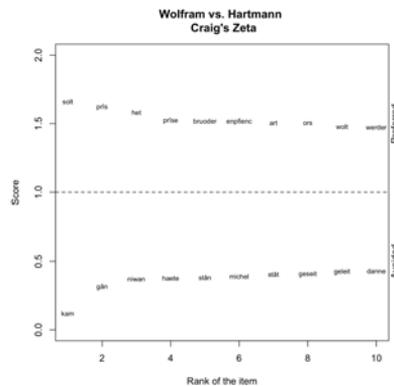


Abb. 3: Liste der von Wolfram gegenüber Hartmann bevorzugten bzw. vermiedenen Wörter (Graphik: Autor).

Ausgangspunkt der Darstellung ist das Wolfram-Korpus,²² über dem horizontalen Strich in der Mitte liegen die von Wolfram im Vergleich zu Hartmann konstant häufig gebrauchten Wörter, unter dem Strich die von Wolfram vermiedenen Wörter. Die erste Position unter den vermiedenen Wörtern nimmt das Wort ›kam‹ ein, das im Mittelhochdeutschen die 1. und 3. Pers. Sg. Präteritum des neuhochdeutschen Verbs ›kommen‹ ausdrückt. Das Wort ist deutlich von der Mittellinie abgesetzt, wird also von Wolfram gegenüber Hartmann in besonders auffälliger Weise gemieden.²³ Doch ist gerade dies kein Beleg dafür, dass Hartmann etwa öfter von ›kommen‹ sprechen würde als Wolfram, denn die 1. und 3. Pers. Sg. von ›komen‹ kann im Mittelhochdeutschen auch ›kom‹ lauten. Und genau diese Form ›kom‹ tritt nun bei Wolfram anstelle von ›kam‹ auf, wodurch sich der scheinbare Mangel an ›kam‹-Formen in den Wolfram-Texten erklärt. Die Verteilung der Formen ›kom/kam‹ ist regional bedingt, ›kam‹ ist nämlich die

²² Im Unterschied zur vorigen Darstellung (Abbildung 2) wurde bei dieser Analyse der *Willehalm* gemeinsam mit dem *Parzival* in die erste Textgruppe eingeordnet. Die zweite Textgruppe bilden wie zuvor *Erec* und *Iwein*.

²³ Der Zahlenwert auf der y-Achse gibt den Zeta-Score der aufgelisteten Wörter an. Grundsätzlich kann der Zeta-Score einen Wert von 0 bis 2 annehmen. Ein Zeta-Score von 0 würde bedeuten, dass das Wort in allen Abschnitten der Hartmann-Texte vorhanden ist und in jedem Abschnitt der Wolfram-Texte fehlt, ein Zeta-Score von 2 den genau umgekehrten Fall.

Variante, die tendenziell im südwestdeutschen (alemannischen) Raum bevorzugt wird, aus dem Hartmann stammt.²⁴ Dass die Verwendung von ›kam‹ tatsächlich auf Hartmann zurückgeht und nicht erst durch spätere Abschreiber seiner Werke eingeführt wurde, lässt sich in einigen Fällen sogar sichern, nämlich dann, wenn das Wort auch im Reim auftritt, und dort auf Wörter reimt, die nicht nur im Alemannischen, sondern auch überregional auf die Silbe #am enden.²⁵ Die Schreiber des Mittelalters hatten zwar grundsätzlich keine Bedenken, den Text, den sie kopierten, ihrem eigenen Dialekt anzupassen, im Fall eines Reimpaars hätten sie jedoch den Text regelrecht umschreiben müssen, um etwa die Form ›kom‹ in ›kam‹ zu ändern.

Nun eignet sich so gesehen das Wortpaar ›kom‹ versus ›kam‹ bei der direkten Gegenüberstellung tatsächlich hervorragend zur Differenzierung zwischen Hartmann und Wolfram,²⁶ allerdings ist damit über einen möglichen Individualstil Hartmanns noch wenig ausgesagt.²⁷ Zudem dürfte dieser letztlich regionale Einflussfaktor eine mögliche Klassifizierung nach Autorschaften bei einem Korpus mit mehreren Autoren unterlaufen. So erscheint es etwa durchaus denkbar, dass das vom Computer berechnete Näheverhältnis von Gottfried und Hartmann, das in den oben gezeigten Clusterdarstellungen ersichtlich ist, darauf zurückgeht, dass auch Gottfried – und weite Teile der *Tristan*-Überlieferung – wie Hartmann aus dem südwestdeutschen Raum stammen.

Und schließlich bleibt die Verteilung der Formen ›kom‹ versus ›kam‹ anfällig für Schreiber- und Herausgeber-Einflüsse, zumindest solange sie nicht im Reim auftreten. Gleich dieses erste Wort in der Rangliste bietet also einen Beleg für die oben beschriebenen Störfaktoren, die die Autorzuordnung behindern können. Für weitergehende stilometrische Untersuchungen könnte es sich daher als sinnvoll erweisen, mit lemmatisierten Texten zu arbeiten, um diese Verzerrungen auszuschließen. Leider gibt es dafür aber noch nicht ausreichend offen zugängliche lemmatisierte E-Texte, die zur Analyse herangezogen werden können. Zudem lässt sich am Beispiel zeigen, dass die Lemmatisierung noch nicht alle systematischen Probleme löst, die die unterschiedliche Schreibung mit sich bringt: Blickt man nämlich zurück auf die Reimbindung des Wortes und betrachtet man nach der Gruppe der Hartmann-Texte nun jene der Wolfram-Texte, dann zeigt sich, dass Wolfram ›kom‹ im Reim nie gebraucht. Dies hat seinen guten Grund, der darin liegt, dass im Mittelhochdeutschen ein Reimwort auf #om sehr viel schwerer zu finden ist als ein Reim auf #am.²⁸ Das heißt also, dass ein Autor, der die Form ›kom‹ verwendet, diese im Reim weniger zum Einsatz bringen kann. Eine hohe Frequenz von ›kom‹ wäre so gesehen noch auffälliger als eine gleich hohe Frequenz von ›kam‹ -- und dieser Unterschied ginge bei der Lemmatisierung verloren.²⁹

²⁴ Vgl. hierzu Paul 2007, S. 253 (§M 79, Anm. 4); zu Unschärfen in der regionalen Verteilung der Formen vgl. Klein 1985, S. 137f.

²⁵ Auffälligerweise hat Hartmann den Reim auf ›kam‹ jedoch in seinem zweiten Artusroman, dem *Iwein*, im Verlauf des Textes immer mehr vermieden. Vgl. hierzu Schirokauer 1923, S. 13.

²⁶ Dieser Umstand ist natürlich der älteren Stilforschung nicht entgangen, vgl. etwa Schirokauer 1923, S. 13.

²⁷ Zur grundsätzlichen Problematik, dass sich die in der stilometrischen Analyse ermittelten Stilmarker nicht zu einem Gesamtbild des Individualstils eines Autors zusammenfügen lassen vgl. Jannidis 2014, S. 191.

²⁸ Auf diesen Umstand hat mich Prof. Klaus Schmidt von der [mittelhochdeutschen Begriffsdatenbank](#) aufmerksam gemacht. Ich danke Herrn Prof. Schmidt für seine umfangreichen Anregungen zum vorliegenden Thema.

²⁹ Ein weiterer Störfaktor ergibt sich dadurch, dass ›kom‹ anders als ›kam‹ auch die 2. Pers. Sg. des Imperativs ausdrücken kann.

Neben diesen problematischen Fällen gibt es in der Wortliste nun aber auch Marker, die eindeutig nicht auf regionale oder Schreibereinflüsse zurückzuführen sind. Ein Beispiel hierfür wäre das Wort ›prîs‹, das in der Liste der von Wolfram bevorzugten Wörter etwas überraschend an zweiter Stelle aufscheint. Für ›prîs‹, was dem Neuhochdeutschen ›Preis‹ im Sinne von Lob, Ruhm und Anerkennung entspricht, gibt es keine doppelte Wortform, der ›prîs‹ wird also tatsächlich viel häufiger in den Wolfram-Texten thematisiert als bei Hartmann.

Die Frage ist nun, ob ›prîs‹ also ein Wort darstellt, das besonders typisch für den Individualstil von Wolfram ist oder ob andere Faktoren bei dieser Verteilung eine Rolle spielen. Zur Beantwortung dieser Frage ist es zunächst von Relevanz, dass ›prîs‹ ein inhaltlich bedeutendes Substantiv ist und kein lediglich füllendes Funktionswort. Die jüngere stilometrische Forschung präferiert für die Unterscheidung von Autorstilen bekanntlich eher inhaltsleere Funktionswörter wie Artikel, Pronomen oder Konjunktionen, da diese von den Autoren tendenziell unbewusst eingesetzt werden.³⁰ ›prîs‹ als semantisch aussagekräftiges Wort hingegen wäre so gesehen eher unter den topic-bezogenen Markern einzuordnen, also unter den Worten, deren Verwendung beispielsweise von der Thematik oder der Gattung beeinflusst wird.³¹ Nun ist ›prîs‹ im Sinne von Belohnung ritterlicher Bewährung zweifelsohne ein Wort, das man in höfischer Artusliteratur erwarten würde, aber es bleibt immer noch auffällig, dass das Wort in den Artusromanen Hartmanns viel seltener auftaucht. Und ebenso erklärungsbedürftig wäre, wieso Wolfram ›prîs‹ sowohl im *Parzival* als auch im *Willehalm* häufiger gebraucht, die streng genommen gar nicht zur selben Gattung gehören.³² Viel mehr als eine Gattungsvorgabe scheint sich hier also eine Art Dichterkonzept Wolframs anzudeuten, der den höfischen Leitbegriff ›prîs‹ gattungsübergreifend zum Einsatz bringt.³³

3. Gattungsstil

Diese Zwischenstellung des Wortes ›prîs‹ zwischen den einzelnen Stilebenen bringt mich nun wieder zurück zu der Frage, ob mittelalterliche Literatur vermehrt von Traditionsvorgaben beeinflusst ist und zurück zu dem vielschichtigen Begriff der Gattung. Die Verwendung von ›prîs‹ bei Wolfram hat gezeigt, dass es offensichtlich gewisse thematische oder konzeptuelle Vorgaben gibt, die quer zu diesen Gattungen liegen. *Parzival* und *Willehalm* gehören aus

³⁰ Vgl. Kestemont 2014, S. 60–61. Für einen Überblick über die bei der Autorschaftsattribuierung in der Forschung verwendeten Marker vgl. Holmes 1994.

³¹ Zur Problematik der Differenzierung vgl. jedoch Schöch 2014, S. 147ff.

³² Der *Willehalm* ist anders als der Artusroman *Parzival* ein Text aus der Chanson de Geste-Tradition. Freilich sind beide Texte von derselben höfischen Grundhaltung getragen.

³³ Dass sich Wolframs Wortschatz gegenüber anderen zeitgenössischen Dichtern durch verschiedene Besonderheiten auszeichnet, ist schon früh in der Forschung bemerkt worden (vgl. dazu zusammenfassend Hartmann 2011, S. 147), ebenso, »daß Wolframs Dichtung von terminologischen Leitmotiven durchzogen ist« (Mohr 1954, S. 175). Im Zentrum von Wolframs Wortschatzes stehen »bestimmte Leitwörter [...], die gewissermaßen das ethisch-theologische Rückgrat seines Erzählprogramms bilden und immer wieder aufgerufen werden« (Hartmann 2011, S. 150). Heiko Hartmann nennt als Beispiel für solche Leitwörter »kiusche« (Keuschheit), »riuwe« (Reue), »triuwe« (Treue) und »zwîvel« (Zweifel), deren Stellung im Werk schon genauer untersucht wurde. Insofern ist es überraschend, dass in der Liste der von Wolfram bevorzugten Worte nun gerade ›prîs‹ als erstes Substantiv auftritt, das bislang, soweit ich sehe, noch vergleichsweise wenig Beachtung gefunden hat. Hier lassen sich Anknüpfungspunkte für eine qualitative Auswertung der stilometrischen Methode finden: Natürlich können die Leitbegriffe nicht unabhängig von ihrem Kontext und ihrer Semantik betrachtet werden und die Häufigkeit eines Wortes sagt noch nicht alles über seinen Stellenwert aus, dennoch wäre es den Versuch wert, zunächst einmal rein quantitativ festzustellen, welche Wörter sich aufgrund ihrer Frequenz als Leitwörter aufdrängen.

literaturwissenschaftlicher Sicht zwar unterschiedlichen Gattungen an, sie sind aber aus sprachlicher Sicht durch ein ähnliches Vokabular geprägt, mit dem Wolfram auf seine spezielle Art ein höfisches Idealbild entwirft. Das heißt aber, dass der stilistische Befund nicht immer ausreichend zur Differenzierung von Gattungen ist, zumindest von Gattungen, wie sie in der Literaturwissenschaft eingeführt sind.

Aus literaturwissenschaftlicher Sicht sind Gattungen nämlich vielschichtige Gebilde, die sich nicht oder nicht nur durch einen ähnlichen Stil auszeichnen, sondern auch durch ähnliche Thematik, ähnliche Figurenkonstellationen oder andere Faktoren bestimmt sein können. Gattungen sind so gesehen Sammelbegriffe, die ganz unterschiedliche Aspekte vereinen, die nicht immer auf derselben kategorialen Ebene liegen müssen. So kann etwa das rein äußerlich-formal hervorgehobene Sonett (eine Gedichtform mit zweimal vier und zweimal drei Zeilen) ebenso als Untergattung gelten wie beispielsweise der sich durch seine Hauptfigur auszeichnende Alexanderroman (ein episches Werk, welches über das Leben Alexanders des Großen erzählt). Letztlich – so hat die literaturwissenschaftliche Gattungstheorie bereits seit den 1970er Jahren herausgearbeitet – sind Gattungen in erster Linie institutionell bestimmt, sie sind Ordnungsbegriffe, die in die literarische Diskussion eingeführt werden und den Erwartungshorizont von Autoren und Lesern determinieren.³⁴

Das heißt aber auch, dass der literaturwissenschaftliche Gattungsbegriff nicht unbedingt mit dem sprachwissenschaftlichen Begriff von Textsorten deckungsgleich sein muss. Dass es ratsam ist, bei der textstatistischen Analyse zwischen eher literaturwissenschaftlich bestimmten *genres* und eher linguistisch bestimmten *text types* zu differenzieren, hat bereits Douglas Biber erkannt, der sich in den 90er Jahren mit der stilometrischen Gattungsbestimmung beschäftigt hat und damit als Vorreiter auf dem Gebiet gelten kann. *Genres* sind für Biber, »text varieties that are readily recognized and ›named‹ within a culture (e.g., letters, press editorials, sermons, conversation)«, entsprechen also dem institutionell bestimmten, literaturwissenschaftlichen Modell, während der Begriff *text types* für »varieties that are defined linguistically (rather than perceptually)«³⁵ vorbehalten ist. Zwar können sich sowohl *genres* als auch *text-types* durch gewisse stilistische Eigenarten auszeichnen, aber nur *text-types* sind per definitionem dadurch determiniert:

»Both genres and text types can be characterized by reference to co-occurring linguistic features, but text types are further defined quantitatively such that the texts in a type all share frequent use of the same set of co-occurring linguistic features.«³⁶

Genres müssen dagegen nicht unbedingt sprachlich kohärent sein: »Genres have a perceptual basis in a given culture, but they are not necessarily linguistically coherent.«³⁷

³⁴ Vgl. grundlegend Voßkamp 1977 sowie Jauß 1977.

³⁵ Biber 1992, S. 332.

³⁶ Biber 1992, S. 332.

³⁷ Biber 1992, S. 339.

Das Problem scheint mir nun zu sein, dass die stilometrische Analyse eher dafür geschaffen ist, *text types* zu untersuchen, dass es aus literaturwissenschaftlicher Sicht aber eigentlich aufschlussreicher wäre, *genres* bestimmen zu können. Aus literaturwissenschaftlicher Sicht sind Gattungen hochgradig unfeste Größen, die vor allem aufgrund ihrer historischen Veränderlichkeit nur schwer zu fassen sind. Nur allzu oft scheinen sich Texte der gattungsmäßigen Einordnung zu widersetzen bzw. die einmal festgelegten Gattungsgrenzen zu überschreiten. Diese Schwierigkeiten bei der Kategorisierung hat die Gattungsforschung dazu gebracht, Gattungen als historisch offene Kategorien mit prototypensemantischen Kernen zu definieren. Und einmal mehr ist die Situation im Mittelalter besonders prekär, denn anders als wir das vom heutigen Literaturbetrieb gewöhnt sind, ist Literatur im Mittelalter kein ausdifferenziertes, institutionalisiertes Teilsystem der Gesellschaft. Es fehlen die festen Rahmenbedingungen, die wir von einem modernen Literaturbetrieb kennen, wo die Zugehörigkeit von Texten zu Gattungen viel stärker reflektiert werden, als dies im Mittelalter der Fall war. Mittelalterliche volkssprachige Textsorten hingegen sind eher durch eine sich immer wieder neu formierende, mündliche Aufführungssituation determiniert als durch eine feste Vorstellung davon, wie Gattungen aussehen müssen.³⁸

Ich möchte an meinem letzten Beispiel zeigen, was dieser Umstand für die Kategorienbildung bedeuten kann. Ich gehe dazu von der Gattung der Epik zum Bereich der Lyrik über. Die mittelhochdeutsche Lyrik wird traditionell in zwei Unterarten eingeteilt, in den Minnesang und den Sangspruch. Die Abgrenzung zwischen diesen beiden Unterarten ist äußerst unscharf und letztlich nur thematisch bedingt. Minnesang ist grundsätzlich Dichtung, in der von höfischer Liebe, also der so genannten Minne, die Rede ist. Demgegenüber kann der Sangspruch nur negativ definiert werden, er gilt als alles das im Rahmen mittelalterlicher Lieddichtung, was nicht Minnesang ist.³⁹

Wollte man sich nun mit stilometrischen Methoden an einer genaueren Differenzierung dieser beiden Formen versuchen, so wird ersichtlich, dass eine Auszählung der häufigsten Wörter für den Minnesang ein durchaus interessantes Ergebnis bringt, das schon in einer Darstellung der bekannten Minnesang-Anthologie *Minnesangs Frühling* als Wordcloud auf dem ersten Blick ins Auge fällt (*Abbildung 4*):⁴⁰

³⁸ Vgl. Grubmüller 1999, S. 195; Bleumer / Emmelius 2011, S. 1.

³⁹ Vgl. den Überblick bei Tervooren 1995, S. 2. Den aktuellsten in einer langen Reihe von Abgrenzungsversuchen zwischen Sangspruch und Minnesang bietet Schnell 2013.

⁴⁰ Bei der Wordcloud, die mithilfe des Programms *Voyant-Tools* erstellt wurde (Sinclair / Rockwell 2012) werden die häufigsten im Text auftretenden Wörter in einer Schriftgröße entsprechend ihrer Frequenz dargestellt: Je häufiger ein Wort auftritt, umso größer erscheint es in der Darstellung.

4. Schluss

Ich breche hier meine Beispielreihe ab und fasse meine Überlegungen zusammen: Ich hoffe mit meinen Explorationsen gezeigt zu haben, dass sowohl traditionelle Literaturwissenschaft als auch Stilometrie bei gegenseitiger Kenntnisnahme voneinander profitieren können. Die traditionelle Literaturwissenschaft wird wohl erst dann bereit sein, sich auf quantifizierende Verfahrensweisen einzulassen, wenn diese den Anschluss an ihre aktuellen Fragestellungen suchen. Wenn diese Verbindung aber hergestellt ist, dann – so legen die hier gezeigten vorläufigen Ergebnissen zumindest nahe – könnte die Stilometrie durchaus ihren Beitrag zu theoretischen Debatten der Mediävistik leisten, etwa wenn es um die Frage geht, ob der Individualstil und damit die Profilierung von Autorschaft vielleicht nicht doch eine größere Rolle spielt als mitunter in der mediävistischen Forschung angenommen. Umgekehrt sollte die Stilometrie darauf achten, was die traditionelle Literaturwissenschaft in Hinblick auf die verwendeten Kategorien zu sagen hat. Denn tut sie das nicht und überblendet leichtfertig quantitative Textanalyse und Hermeneutik, dann lassen sich zwar messbare Ergebnisse erzielen, die aber an den Objekten der Literaturwissenschaft vorbeigehen. Erst die Herausarbeitung der Differenzen schafft die Voraussetzung für eine gegenseitige Befruchtung der Disziplinen und ermöglicht einen ›methodischen Brückenschlag‹, der über die bloße ›feindliche Übernahme‹ hinausgeht.⁴⁴

⁴⁴ Vgl. den Titel der [DHD-Tagung](#) in Passau, auf die dies Bezug nimmt: »Digital Humanities – methodischer Brückenschlag oder ›feindliche Übernahme‹? Chancen und Risiken der Begegnung zwischen Geisteswissenschaften und Informatik«.

Bibliographische Angaben

Primärtexte

- Gottfried von Straßburg: Tristan. Nach dem Text von Friedrich Ranke neu hg., ins Neuhochdeutsche übersetzt, mit einem Stellenkommentar und einem Nachwort von Rüdiger Krohn. Stuttgart 1980. [\[Nachweis im GBV\]](#)
- Hartmann von Aue: Der arme Heinrich. Mittelhochdeutsch/Neuhochdeutsch. Hg. von Ursula Rautenberg, übersetzt von Siegfried Grosse. Stuttgart 1993. [\[Nachweis im GBV\]](#)
- Hartmann von Aue: Erec. Mittelhochdeutscher Text und Übertragung von Thomas Cramer. Frankfurt/Main 1972. [\[Nachweis im OPAC\]](#)
- Hartmann von Aue: Gregorius. Hg. von Hermann Paul. 13., neu bearbeitete Auflage, besorgt von Burghart Wachinger. Tübingen 1984. [\[Nachweis im GBV\]](#)
- Hartmann von Aue: Iwein. Eine Erzählung. Hg. von Georg F. Benecke / Karl Lachmann. Neu bearbeitet von Ludwig Wolff. 7. Ausgabe. Berlin 1968. [\[Nachweis im OPAC\]](#)
- Heinrich von Veldeke: Eneasroman. Mittelhochdeutsch/Neuhochdeutsch. Nach dem Text von Ludwig Ettmüller ins Neuhochdeutsche übersetzt, mit einem Stellenkommentar und einem Nachwort von Dieter Kartschoke. Stuttgart 1986. [\[Nachweis im GBV\]](#)
- Konrad von Würzburg: Engelhard. Hg. von Ingo Reiffenstein. 3., neu bearbeitete Auflage der Ausgabe von Paul Gereke. Tübingen 1982. [\[Nachweis im GBV\]](#)
- Konrad von Würzburg: Kleinere Dichtungen. Bd. 1: Der Welt Lohn – Das Herzmaere – Heinrich von Kempten. Hg. von Edward Schröder. 9. Auflage. Berlin 1968. [\[Nachweis im GBV\]](#)
- Konrad von Würzburg: Die Legenden. Bd. 2. Hg. von Paul Gereke. Tübingen 1926. [\[Nachweis im OPAC\]](#)
- Konrad von Würzburg: Pantaleon. Hg. von Winfried Woesler. 2. Auflage. Tübingen 1974. [\[Nachweis im OPAC\]](#)
- Konrads von Würzburg: Partonopier und Meliur. Aus dem Nachlasse von Franz Pfeiffer, hg. von Karl Bartsch. Nachdruck der Ausgabe von 1871. Berlin 1970. [\[Nachweis im OPAC\]](#)
- Konrad von Würzburg: Der Trojanische Krieg. Nach den Vorarbeiten K. Frommanns und F. Roths zum ersten Mal hg. von Adelbert von Keller. Stuttgart 1858. [\[Nachweis im OPAC\]](#)
- Jean L. C. Putmans: EDV-Text von ›Des Minnesangs Frühling‹. Lorch 1993 (= Göppinger Arbeiten zur Germanistik 600). [\[Nachweis im GBV\]](#)
- Ulrich von Etzenbach: Alexander. Hg. von Wendelin Toischer. Tübingen 1888. [\[Nachweis im OPAC\]](#)
- Ulrich von Etzenbach: Wilhelm von Wenden. Hg. von Hans-Friedrich Rosenfeld. Berlin 1957. [\[Nachweis im OPAC\]](#)
- Wolfram von Eschenbach: Willehalm. Nach der gesamten Überlieferung kritisch hg. von Werner Schröder. Berlin, New York 1978. [\[Nachweis im OPAC\]](#)
- Wolfram von Eschenbach: Werke. Hg. von Karl Lachmann. 5. Auflage. Berlin 1891. [\[Nachweis im OPAC\]](#)

Sekundärliteratur

- Alte und neue Philologie. Hg. von Martin-Dietrich Glessgen / Franz Lebsanft. Tübingen 1997. [\[Nachweis im OPAC\]](#)
- Douglas Biber: The Multi-Dimensional Approach to Linguistic Analyses of Genre Variation: An Overview of Methodology and Findings. In: Computers and the Humanities 26 (1992), H. 5/6, S. 331–345. [\[Nachweis im GBV\]](#)
- Hartmut Bleumer / Caroline Emmelius: Generische Transgressionen und Interferenzen. Theoretische Konzepte und historische Phänomene zwischen Lyrik und Narrativik. In: Lyrische Narrationen – narrative Lyrik. Gattungsinterferenzen in der mittelalterlichen Literatur. Hg. von Hartmut Bleumer / Caroline Emmelius. Berlin, New York 2011, S. 1–39. [\[Nachweis im OPAC\]](#)
- Manuel Braun: Kodieren, Annotieren, Theoretisieren. Zur Wiederannäherung von Literatur- und Sprachwissenschaft über Korpora. In: Lili. Zeitschrift für Literaturwissenschaft und Linguistik 43 (2013), H. 172, S. 83–90. [\[Nachweis im OPAC\]](#)
- John Burrows: ›Delta‹: a Measure of Stylistic Difference and a Guide to Likely Authorship. In: Literary and Linguistic Computing 17 (2002), H. 3, S. 267–287. [\[Nachweis im GBV\]](#)
- John Burrows: All the Way Through: Testing for Authorship in Different Frequency Strata. In: Literary and Linguistic Computing 22 (2007), H. 1, S. 27–47. [\[Nachweis im GBV\]](#)

- Bernard Cerquiglini: *Eloge de la variante. Histoire critique de la philologie*. Paris 1989. [[Nachweis im OPAC](#)]
- Hugh Craig / Arthur Kinney: *Shakespeare, Computers, and the Mystery of Authorship*. Cambridge 2009. [[Nachweis im GBV](#)]
- Maciej Eder / Mike Kestemont / Jan Rybicki: *Stylometry with R: a suite of tools*. In: *Digital Humanities 2013: Conference Abstracts*. University of Nebraska-Lincoln, NE, S. 487–489. [[Nachweis im GBV](#)]
- Maciej Eder: *Mind Your Corpus: systematic errors in authorship attribution*. In: *Literary and Linguistic Computing* 28 (2013), H. 4, S. 603–614. [[Nachweis im GBV](#)]
- Maciej Eder / Jan Rybicki / Mike Kestemont: *Stylo: a package for stylometric analyses*. August 17, 2014. [[online](#)]
- Michel Foucault: *Was ist ein Autor?* In: *Michel Foucault: Schriften zur Literatur*. Aus dem Französischen von Karin von Hofer / Anneliese Botond. Frankfurt/Main 1988, S. 7–31. [[Nachweis im OPAC](#)]
- Klaus Grubmüller: *Ich als Rolle. »Subjektivität« als höfische Kategorie im Minnesang?* In: *Höfische Literatur -- Hofgesellschaft – Höfische Lebensformen um 1200*. Hg. von Gert Kaiser / Jan-Dirk Müller. Düsseldorf 1986, S. 387–408. [[Nachweis im OPAC](#)]
- Klaus Grubmüller: *Gattungskonstitution im Mittelalter*. In: *Mittelalterliche Literatur und Kunst im Spannungsfeld von Hof und Kloster. Ergebnisse der Berliner Tagung, 9.–11. Oktober 1997*. Hg. von Nigel F. Palmer / Hans-Jochen Schiewer. Tübingen 1999, S. 193–210. [[Nachweis im OPAC](#)]
- Heiko Hartmann: *Darstellungsmittel und Darstellungsformen in den erzählenden Werken*. In: *Wolfram von Eschenbach. Ein Handbuch*. Hg. von Joachim Heinze. Berlin, Boston 2011, S. 145–220. [[Nachweis im OPAC](#)]
- Jens Hausteine: *Mediävistische Stilforschung und die Präsenzkultur des Mittelalters. Mit einem Ausblick auf Gottfried von Straßburg und Konrad von Würzburg*. In: *Textprofile stilistisch. Beiträge zur literarischen Evolution*. Hg. von Ulrich Breuer / Bernhard Spies. Bielefeld 2011, S. 43–60. [[Nachweis im GBV](#)]
- David I. Holmes: *Authorship Attribution*. In: *Computers and the Humanities* 28 (1994), H. 2, S. 87–106. [[Nachweis im GBV](#)]
- Fotis Jannidis / Gerhard Lauer: *Burrow's Delta and Its Use in German Literary History*. In: *Distant Readings. Topologies of German Culture in the Long Nineteenth Century*. Hg. von Matt Erlin / Lynne Tatlock. Rochester 2014, S. 29–54. [[Nachweis im GBV](#)]
- Fotis Jannidis: *Der Autor ganz nah. Autorstil in Stilistik und Stilometrie*. In: *Theorien und Praktiken der Autorschaft*. Hg. von Matthias Schaffrick / Marcus Willand. Berlin, Boston 2014, S. 169–195. [[Nachweis im GBV](#)]
- Hans Robert Jauf: *Theorie der Gattungen und Literatur des Mittelalters*. In: *Hans Robert Jauf: Alterität und Modernität der mittelalterlichen Literatur. Gesammelte Aufsätze 1956–1976*. München 1977, S. 327–358. [[Nachweis im OPAC](#)]
- Matthew Jockers: *Macroanalysis. digital methods and literary history*. Champaign 2013. Champaign 2013. [[Nachweis im OPAC](#)]
- Mike Kestemont: *Arthur's Authors. A Quantitative Study of the Rhyme Words in the Middle Dutch Arthurian Epic*. In: *Zeitschrift für deutsches Altertum* 142 (2013), H. 1, S. 1–33. [[Nachweis im OPAC](#)]
- Mike Kestemont: *Function Words in Authorship Attribution. From Black Magic to Theory?* In: *forProceedings of the 3rd Workshop on Computational Linguistics Literature (ClfL)*. Göteborg 2014, S. 59–66. [[online](#)]
- Thomas Klein: *Ermittlung, Darstellung und Deutung von Verbreitungstypen in der Handschriftenüberlieferung mittelhochdeutscher Epik*. In: *Deutsche Handschriften 1100–1400. Oxforder Kolloquium 1985*. Hg. von Volker Honemann / Nigel F. Palmer. Tübingen 1988, S. 110–167. [[Nachweis im OPAC](#)]
- Erich Kleinschmidt: *Minnesang als höfisches Zeremonialhandeln*. In: *Archiv für Kulturgeschichte* 58 (1976), S. 35–76. [[Nachweis im OPAC](#)]
- Jurij M. Lotman: *Die Struktur literarischer Texte*. 4., unveränderte Auflage. München 1993. [[Nachweis im GBV](#)]
- Wolfgang Mohr: *Hilfe und Rat in Wolframs Parzival*. In: *Festschrift für Jost Trier zu seinem 60. Geburtstag am 15. Dezember 1954*. Hg. von Benno von Wiese / Karl Heinz Borch. Meisenheim 1954, S. 173–197. [[Nachweis im GBV](#)]
- Franco Moretti: *Conjectures on World Literature*. In: *New Left Review* 1 (2000), S. 54–68. [[Nachweis im GBV](#)]
- Franco Moretti: *Graphs – Maps – Trees*. London, New York 2005. [[Nachweis im GBV](#)]
- Hermann Paul: *Mittelhochdeutsche Grammatik*. 25. Auflage. Tübingen 2007. [[Nachweis im OPAC](#)]
- R Core Team: *A language and environment for statistical computing*. R Foundation for Statistical Computing. Wien 2013. [[online](#)]
- Patrick Sahle: *Digitale Editionsformen. Zum Umgang mit der Überlieferung unter den Bedingungen des Medienwandels*, 3 Bände. Norderstedt 2013 (= *Schriften des Instituts für Dokumentologie und Editorik* 7–9). [[Nachweis im OPAC](#)]
- Arno Schirokauer: *Studien zur mittelhochdeutschen Reimgrammatik*. In: *Beiträge zur Geschichte der deutschen Sprache und Literatur (PBB)* 47 (1923), S. 1–126. [[Nachweis im OPAC](#)]
- Rüdiger Schnell: *»Autor« und »Werk« im deutschen Mittelalter. Forschungskritik und Forschungsperspektiven*. In: *Neue Wege der Mittelalter-Philologie. Landshuter Kolloquium 1996*. Hg. von Joachim Heinze / L. Peter Johnson / Gisela Vollmann-Profe. Berlin 1998 (= *Wolfram-Studien* 15), S. 12–73. [[Nachweis im OPAC](#)]

Rüdiger Schnell: Minnesang und Sangspruch im 13. Jahrhundert. Gattungsdifferenzen und Gattungsinterferenzen. In: Transformationen der Lyrik im 13. Jahrhundert. Wildbader Kolloquium 2008. In Verbindung mit Eckart Conrad Lutz und Klaus Ridder hg. von Susanne Köbele. Berlin 2013 (= Wolfram-Studien 21), S. 287–348. [\[Nachweis im OPAC\]](#)

Christoph Schöch: Corneille, Molière et les autres. Stilometrische Analysen zu Autorschaft und Gattungszugehörigkeit im französischen Theater der Klassik. In: Literaturwissenschaft im digitalen Medienwandel. Hg. von Christof Schöch / Lars Schneider. Philologie im Netz, Beiheft 7 (2014), S. 130–157. [\[online\]](#)

Stéfan Sinclair / Geoffrey Rockwell / the Voyant Tools Team: Voyant Tools (web application). 2012. [\[online\]](#)

Karl Stackmann: Neue Philologie? In: Modernes Mittelalter. Neue Bilder einer populären Epoche. Hg. von Joachim Heinze. Frankfurt/Main 1994, S. 398–427. [\[Nachweis im GBV\]](#)

Peter Strohschneider: Tanzen und Singen. Leichs von Ulrich von Winterstetten, Heinrich von Sax sowie dem Tannhäuser und die Frage nach dem rituellen Status des Minnesangs. In: Mittelalterliche Lyrik. Probleme der Poetik. Hg. von Thomas Cramer / Ingrid Kasten. Berlin 1999, S. 197–231. [\[Nachweis im OPAC\]](#)

Helmut Tervooren: Sangspruchdichtung. Stuttgart, Weimar 1995. [\[Nachweis im OPAC\]](#)

Wilhelm Voßkamp: Gattungen als literarisch-soziale Institutionen. Zu Problemen sozial- und funktionsgeschichtlich orientierter Gattungstheorie und -historie. In: Textsortenlehre – Gattungsgeschichte. Hg. von Walter Hinck. Heidelberg 1977, S. 27–44. [\[Nachweis im OPAC\]](#)

Franz Josef Worstbrock: Wiedererzählen und Übersetzen. In: Mittelalter und frühe Neuzeit. Übergänge, Umbrüche und Neuansätze. Hg. von Walter Haug. Tübingen 1999, S. 128–142. [\[Nachweis im OPAC\]](#)

Abbildungslegenden und -nachweise

Abb. 1: SClusteranalyse der höfischen Epik (Graphik: Autor).

Abb. 2: Kontrastive Analyse der Werke Hartmanns, Wolframs und Gottfrieds (Graphik: Autor).

Abb. 3: Liste der von Wolfram gegenüber Hartmann bevorzugten bzw. vermiedenen Wörter (Graphik: Autor).

Abb. 4: Wordcloud Minnesangs Frühling (Graphik: Autor).

Zeitschrift für digitale Geisteswissenschaften

Artikel aus:

Sonderband 1 der ZfdG: Grenzen und Möglichkeiten der Digital Humanities. Hg. von Constanze Baum und Thomas Stäcker. 2015. DOI: [10.17175/sb01](https://doi.org/10.17175/sb01)

Titel:

Kontroverse Kommunikation im Umkreis der ersten Frauenbewegung. Wie können digitale Ressourcen die sprachliche Untersuchung und die Ergebnisdokumentation verbessern?

Autor/in:

Kerstin Wolff

Kontakt: wolff@addf-kassel.de

Institution: Stiftung Archiv der deutschen Frauenbewegung, Kassel

GND: [172963486](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63862-p0011-9) ORCID:

Autor/in:

Alexander Geyken

Kontakt: geyken@bbaw.de

Institution: BBAW/CLARIN-D

GND: [1084139936](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63862-p0011-9) ORCID:

Autor/in:

Thomas Gloning

Kontakt: g909@uni-giessen.de

Institution: JLU Gießen/BBAW/CLARIN-D

GND: [135943043](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63862-p0011-9) ORCID:

DOI des Artikels:

[10.17175/sb001_010](https://doi.org/10.17175/sb001_010)

Nachweis im OPAC der Herzog August Bibliothek:

[830184120](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63862-p0011-9)

Erstveröffentlichung:

19.02.2015

Lizenz:

Sofern nicht anders angegeben 

Medienlizenzen:

Medienrechte liegen bei den Autoren

Letzte Überprüfung aller Verweise:

24.05.2016

GND-Verschlagwortung:

[Frauenbewegung](#) | [Digital Humanities](#) | [Informationssystem](#) | [Computerunterstützte Lexikographie](#) |

Zitierweise:

Kerstin Wolff, Alexander Geyken, Thomas Gloning: Kontroverse Kommunikation im Umkreis der ersten Frauenbewegung. Wie können digitale Ressourcen die sprachliche Untersuchung und die Ergebnisdokumentation verbessern?. In: Grenzen und Möglichkeiten der Digital Humanities. Hg. von Constanze Baum / Thomas Stäcker. 2015 (= Sonderband der Zeitschrift für digitale Geisteswissenschaften, 1). PDF Format ohne Paginierung. Als text/html abrufbar unter DOI: [10.17175/sb001_010](https://doi.org/10.17175/sb001_010).

Kerstin Wolff, Alexander Geyken, Thomas Gloning

Kontroverse Kommunikation im Umkreis der ersten Frauenbewegung. Wie können digitale Ressourcen die sprachliche Untersuchung und die Ergebnisdokumentation verbessern?

Abstracts

Die erste Frauenbewegung ist eine der großen sozialen Bewegungen des 19. und frühen 20. Jahrhunderts mit einem beträchtlichen Kommunikationsaufkommen. Wir fragen in diesem Beitrag, wie sich sprach-, themen- und kommunikationsgeschichtliche Forschungen in diesem Bereich durch Digital Humanities-Verfahren unterstützen lassen. Wir stellen die Konzeption eines integrierten Dokumentationssystems vor, das eine lexikalische Dokumentation, ein Volltext-Korpus und sprach- bzw. kommunikationsgeschichtliche Untersuchungen verbindet, und zeigen, wie sich die Forschungsarbeit an eine große Forschungsinfrastruktur (CLARIN-D) anbinden lässt.

The first women's movement of the 19th and early 20th centuries was one of the major social movements which produced an immense body of texts. In this paper, we explore the ways in which research into the history of language use, topic development, and forms of communication in the first women's movement can be supported by Digital Humanities methods. We present an integrated documentation system that combines lexical documentation, a full-text corpus and linguistic investigations (such as on argumentation, word usage, etc.), and we show how this type of research activity can be integrated into a large infrastructure project (CLARIN-D).

1. Einleitung

In diesem Beitrag stellen wir laufende Arbeiten und Planungen für ein Forschungsprojekt zum Sprachgebrauch und zu Formen der Kommunikation im Umkreis der sogenannten ersten oder alten Frauenbewegung (1865–1933) vor.¹ Zentraler Gegenstand sind Texte von Autorinnen, die der ersten Frauenbewegung zuzuordnen sind. Wir berücksichtigen für bestimmte Gesichtspunkte, z.B. intertextuelle Bezüge, aber auch Texte von Gegnerinnen und Gegnern. Im Mittelpunkt dieses Beitrags steht die Frage, inwiefern digitale Ressourcen und Verfahren die sprachlichen und kommunikationsgeschichtlichen Untersuchungen und die darauf bezogene Ergebnispräsentation unterstützen bzw. verbessern können. Es werden darüber hinaus Überlegungen angestellt, wie solche Arbeiten an ein Infrastrukturprojekt (Clarín-D) angebunden werden können.

¹ Zu den Vorarbeiten gehören insbesondere: (i) ein vom Land Hessen (HMWK) gefördertes Pilot-Projekt (6/2011–12/2011); (ii) eine Gießener Magisterarbeit zu Thematisierungspraktiken und Wortgebrauch im Diskurs um die Mädchenschulreform; (iii) ein laufendes Vorhaben zu Wortgebrauch und Thematisierungspraktiken im Diskurs um das Frauenwahlrecht 1850–1918. Gegenstand dieser Arbeit ist auch die Zusammenstellung eines thematischen Korpus und eines digital nutzbaren und facettiert erschlossenen Glossars; (iv) die Integration von digitalen Volltexten aus dem Themenbereich der Frauenfragen um 1900 in das [Deutsche Textarchiv](#) als Teil des CLARIN-D-Verbunds; (v) die Ausarbeitung von kommunikations- und sprachgeschichtlichen Fallstudien sowie von Prototypen der lexikalischen Dokumentation.

Wir gehen hierfür in fünf Schritten vor: Zunächst stellen wir die erste Frauenbewegung kurz vor (**Abschnitt 2**). Daran anschließend präzisieren wir unsere wesentlichen sprach- und kommunikationsgeschichtlichen Fragestellungen (**Abschnitt 3**). Aufbauend darauf stellen wir die Konzeption eines integrierten digitalen Dokumentationssystems für sprach- und kommunikationsgeschichtliche Befunde vor (**Abschnitt 4**) und fragen, wie digitale Ressourcen die sprachlichen und kommunikationsgeschichtlichen Untersuchungen selbst fördern können (**Abschnitt 5**). Abschließend präsentieren wir Überlegungen zur Anbindung solcher Arbeiten an die Infrastruktur, die im Clarin-D-Projekt derzeit aufgebaut wird (**Abschnitt 6**).

2. Die erste Frauenbewegung

Die erste Frauenbewegung gehörte zu den großen sozialen Strömungen des 19. und frühen 20. Jahrhunderts. Mitglieder waren im Wesentlichen Frauen, auch wenn Männer aus dem proletarischen und liberalen Spektrum – vor allem in der Mitte des 19. Jahrhunderts – diese Bewegung auch mittrugen. Den Kernziel der Partizipation von Frauen an der bürgerlichen Gesellschaft dienten u.a. die Verbesserung der Lage der Frauen im Bereich der Bildung und der Beteiligungsmöglichkeiten in den Bereichen Politik und Erwerbsarbeit. Eng verbunden mit den Diskussionen um die Rechte von Frauen waren Auseinandersetzungen um Lebensentwürfe von Frauen und um Fragen des Geschlechterverhältnisses.

Die Bewegung formte verschiedene Flügel aus und setzte sich aus einer dezidiert bürgerlichen Strömung (mit einer gemäßigten und einer radikaleren Richtung) und einer proletarischen zusammen. Hinter den verschiedenen Flügeln standen unterschiedliche Emanzipationskonzepte und sich teilweise widersprechende methodische Ansichten. Zu den prominenten Autorinnen des bürgerlichen Flügels gehörten etwa Helene Lange, Gertrud Bäumer, Lida Gustava Heymann, Anita Augspurg und Auguste Kirchhoff, die zentrale Autorin im Lager der proletarischen Frauenbewegung war zweifellos Clara Zetkin.²

Die erste Frauenbewegung des 19. und beginnenden 20. Jahrhunderts war *auch* eine kommunikative Bewegung. Denn um ihre Forderungen im öffentlichen Raum zu thematisieren und zu begründen, nutzten die Vertreterinnen (und auch ein paar wenige männliche Unterstützer) der Frauenbewegung ganz unterschiedliche Medien und kommunikative Formen. Dazu gehörten mündliche Formen, wie z.B. Reden oder Redebeiträge in Gremien, vor allem aber handelt es sich um medial verbreitete, geschriebene Texte. In eigenen Zeitschriften, in Petitionen, in umfangreichen Begleitschriften zu Petitionen, in Sitzungsprotokollen, privaten Briefen, Flugschriften und Monographien legten die Autorinnen ihre Anliegen, Motive, Hoffnungen, Forderungen und Argumente dar und versuchten so, die Notwendigkeit von Veränderungen der Geschlechterordnung zu begründen.³ Ein fünfbändiges *Handbuch der Frauenbewegung*, das 1901–1906 von Helene Lange und Gertrud Bäumer herausgegeben wurde (**Abbildung 1**), systematisiert aus einer zeitgenössischen Perspektive den erreichten Stand der Diskussion in den unterschiedlichen Themenbereichen:

² Zur Frauenbewegung vgl. die Überblicksdarstellungen von Schaser 2006 und Gerhard 1991.

³ Zu den Zeitschriften vgl. Wolff 2007.



Abb. 1: Helene Lange, Frontispiz, aus: Lange 1928, Bd. 1 (links); *Handbuch der Frauenbewegung*, Titelblatt, aus: Lange / Bäumer 1901–1906, Bd. 4 [\[online\]](#) (rechts).

Eine über 700 Seiten umfassende Bibliographie *Die Frauenfrage in Deutschland. Strömungen und Gegenströmungen 1790-1930*, die 1934 von Hans Sveistrup und Agnes von Zahn-Harnack herausgegeben wurde, dokumentiert die Texte aus dem Umkreis der ersten Frauenbewegung und das erweiterte kommunikative Umfeld der einzelnen Fragen, Streitpunkte und Themenbereiche (Abbildung 2).

**DIE FRAUENFRAGE
IN DEUTSCHLAND**

**Strömungen und Gegenströmungen
1790–1930**

Sachlich geordnete und erläuterte Quellenkunde

Herausgegeben von **1934**

Hans Sveistrup und Agnes v. Zahn-Harnack

Abb. 2: Bibliographie zur Geschichte der Frauenbewegung 1790–1930 (1934) (Quelle: Thomas Gloning, privat).

Die erste Frauenbewegung im deutschsprachigen Raum ist eingebettet in einen mehrsprachigen Quellenraum, in dem nicht nur deutsche Quellen, sondern auch Bezugstexte aus der europäischen Tradition der Aufklärung (z.B. Rousseau, Wollstonecraft), der europäischen Frauenbewegung und aus dem nordamerikanischen Diskurs des 19. Jahrhunderts für argumentative Zwecke genutzt wurden. Dieser Befund verweist auf die enge Verflechtung weltweiter Frauenbewegungen, die in der Gründung des ICW (= International Council of Women, 1888) einen organisatorischen Rahmen fanden, auf die auch private Vernetzung der Protagonistinnen und auf eine damit verbundene dichte Kommunikation.⁴

⁴ Vgl. Schöck-Quinteros 2007.

Zentrale Themenfelder, Streitpunkte und Forderungen in der Hochphase der ersten Frauenbewegung um 1900 betreffen vor allem die folgenden vier Bereiche⁵:

- (i) politische Partizipation (u.a. aktives und passives Wahlrecht),
- (ii) Bildung (u.a. Mädchenschulwesen, Hochschulzugang),
- (iii) Erwerbsarbeit (u.a. Zugang zu beruflichen Tätigkeiten, gerechter Lohn) und
- (iv) Sexualmoral (u.a. Prostitutionsdiskurs, §218, uneheliche Mutterschaft).

Einzelne Themen der ersten Frauenbewegung um 1900 sind in ihren sozial- und ideengeschichtlichen Grundzügen recht gut erforscht. Dagegen ist eine umfassende und detaillierte Untersuchung und Dokumentation von kommunikativen Verfahren, Thematisierungspraktiken sowie des Sprachgebrauchs der Autorinnen und der Texte im Umkreis der ersten Frauenbewegung bisher nach wie vor ein Desiderat.

3. Kommunikations- und Sprachgeschichte der alten Frauenbewegung. Fachliche Ziele und Fragestellungen

Unser fachliches Ziel ist es, einen Beitrag zur bislang nicht bzw. nur in Ansätzen geschriebenen Kommunikations- und Sprachgeschichte der ersten Frauenbewegung und zu einer Themen- bzw. Thematisierungsgeschichte zahlreicher Fragen zu leisten, bei denen die Ordnung der Geschlechterverhältnisse im Zentrum stehen.⁶ Im Vordergrund unseres Vorhabens steht dabei die interne und externe Kommunikation im Umkreis der ersten Frauenbewegung.

Die systematische Analyse des Sprachgebrauchs und der kommunikativen Strukturen lässt sich drei großen Untersuchungsbereichen zuordnen: (i) Diskursorganisation; (ii) Handlungsformen, Argumentationsweisen und Thematisierungspraktiken; (iii) Wortgebrauch.

Fragestellungen dieser Art sind im Schnittpunkt von historischer Diskursanalyse, historischer Argumentationsforschung und historischer Lexikologie und Semantik zu verorten. Dabei sind im Einzelnen auch zahlreiche weitere Disziplinen berührt, z.B. historische Soziolinguistik, historische Textlinguistik, Bildungsgeschichte, Geschlechtergeschichte, Sozialgeschichte.

Die genannten Untersuchungsbereiche lassen sich mit Leitfragen konkretisieren, die auch als Grundlage für die Beurteilung dienen sollen, inwiefern einzelne Digital Humanities-Werkzeuge und -Methoden für die jeweiligen Teilfragen und Teilaufgaben erheblich und nützlich sein können.

⁵ Vgl. Wischermann 2003, S. 17 und passim; Mädchenschulgeschichte(n) 2008; Handbuch der Frauenbewegung 1901–1906.

⁶ Vgl. hierzu auch Gloning 2012.

3.1 Diskursorganisation

Texte und Kommunikationsbeiträge im Rahmen der ersten Frauenbewegung sind in aller Regel keine isolierten Einzelereignisse. Vielmehr sind sie auf übergeordnete Fragestellungen und Themen bezogen und weisen vielfach Bezüge zu thematisch verwandten Beiträgen, zu gegnerischen Schriften und Textpassagen in Schriften von Gewährsleuten, die zur Stützung eigener Positionen genutzt werden, auf. Einzelne Kommunikationsbeiträge sind Teile größerer Einheiten, die sich als Diskurse modellieren und auf die sich unterschiedliche Methoden der linguistischen Diskursforschung anwenden lassen. Besonders produktiv erscheinen uns dafür die methodischen und beschreibungspraktischen Vorschläge der Düsseldorfer Schule um Georg Stötzel, Martin Wengeler und Dietrich Busse, die bislang in Breite vor allem für eine Sprachgeschichte der Nachkriegszeit genutzt wurden.⁷

Im Hinblick auf den Aspekt der Diskursorganisation lassen sich folgende Leitfragen formulieren, die zum einen auf eine kommunikationsgeschichtliche Kartographie der thematischen Stränge zielen, die darüber hinaus aber auch beschreibungspraktische Grundlagen für die Auswahl von Texten und die Entscheidung darüber liefern, in welcher Weise und mit welcher Detailgenauigkeit einzelne Texte herangezogen werden sollen. Zu diesen Leitfragen gehören insbesondere die folgenden:

- (i) Wie lassen sich einzelne thematische und diskursive Stränge abgrenzen und in ihren Verläufen und diskursiven Zusammenhängen rekonstruieren?
- (ii) Welche Einzeltexte gehören jeweils zu einem bestimmten diskursiven Strang?
- (iii) Welche Medien und textuellen Darstellungsformen werden in einem bestimmten diskursiven Strang jeweils (von wem) genutzt?
- (iv) Welche intertextuellen Bezüge sind zwischen einzelnen Beiträgen und zwischen Teilen von Beiträgen eines diskursiven Strangs erkennbar?
- (v) Mit welchen sprachlichen Verfahren etablieren die Autorinnen Bezüge zu anderen Texten, die entweder zur Stützung eigener Positionen oder aber als Beispiele für gegnerische Positionen angeführt werden?
- (vi) Welche Beteiligungsrollen sind in einzelnen Diskursen und Teildiskursen erkennbar?

3.2 Handlungs- und Argumentationsstrukturen

Die Analyse von Formen des sprachlichen Handelns ist ein zentraler Schwerpunkt der Untersuchung einzelner Texte. Dabei wird gefragt, welche sprachlichen Handlungen überhaupt genutzt werden (z.B. etwas fordern, etwas behaupten, eine Behauptung stützen, jemandem etwas vorwerfen, etwas bestreiten, etwas veranschaulichen), in welcher Abfolge sprachliche Handlungen typischerweise oder in Einzelfällen sequenziert werden und mit welchen sprachlichen Mitteln sie dabei realisiert werden können. Auch die Frage, welche thematischen Aspekte im Rahmen des sprachlichen Handelns zur Sprache kommen, ist ein

⁷ Vgl. exemplarisch Busse / Teubert 1994; Stötzel / Wengeler 1995; Böke / Jung / Wengeler 1996.

wichtiger Untersuchungsaspekt.⁸ Wenn es um strittige Aspekte geht, spielen natürlich Formen des Argumentierens und des Stützens eine zentrale Rolle. Die unterschiedlichen Werkzeuge der (historischen) Argumentationsforschung sind deshalb von besonderer Bedeutung für unser Vorhaben.

In sehr offener Form kann man diesen Untersuchungsbereich mit folgenden Leitfragen ansprechen:

- (i) Was sind zentrale sprachliche Handlungen und Handlungssequenzen? Was ist ihre kommunikative Rolle?
- (ii) Wie tragen unterschiedliche kommunikative Verfahren und Thematisierungspraktiken zur Etablierung und Stützung bestimmter Sichtweisen bei?
- (iii) Welche sprachlichen Mittel und Verfahren werden dafür im Einzelnen benutzt?
- (iv) Welche Argumentationsformen werden gebraucht, um Sichtweisen und Forderungen zu stützen?
- (v) Welche thematischen Aspekte werden im Rahmen des sprachlichen Handelns genutzt, um Forderungen zu stützen, Sichtweisen einzuführen, Gegner zu diskreditieren usw.?

3.3 Wortgebrauch

Der Wortgebrauch spielt, wie in anderen Diskursen auch, eine zentrale Rolle, nicht nur für die Bewältigung thematischer und textfunktionaler Aufgaben, sondern auch für die Erzeugung bestimmter Sichtweisen bei der Behandlung umstrittener Fragen. Ein Wort wie »Überbürdung« gehört z.B. in den thematischen Umkreis der Diskussion um die Reform der Mädchenbildung.⁹ Zu den textfunktionalen sprachlichen Mitteln gehören zum Beispiel die Konnektoren (»nicht nur«, »auch«, »aber«) oder Aufzählungs- und Textstrukturierungsmittel wie »zunächst«. Eine spezifische Sichtweise ist an Wörter wie »Eigenart«, »Geschlechtsunterschied«, »intellektuelle Grenzlinie« oder »Arbeitsteilung« gebunden, die z.B. von Helene Lange im Hinblick auf unterschiedliche Lebens- und Wirkungssphären von Männern und Frauen postuliert wird. Helene Lange bewegt sich damit im Rahmen einer verbreiteten Denkfigur, die man als Differenzhypothese bezeichnen kann. Aber auch traditionelle lexikologische Fragestellungen, z.B. Fragen nach der Rolle von Wortbildungen (»Antistimmrechtlerin«) oder von Fremdwörtern (»vicarieren«) müssen mit einbezogen werden. In Bezug auf solche und andere Aspekte des Wortgebrauchs in Texten der ersten Frauenbewegung kann man folgende Leitfragen formulieren:

- (i) Wie lässt sich der Wortgebrauch dieser Texte in seiner spezifischen Funktionalität charakterisieren und dokumentieren?
- (ii) Wie tragen unterschiedliche Formen des Wortgebrauchs dazu bei, Sichtweisen auf Geschlechterverhältnisse zu konstituieren und neue Forderungen zu stützen, durchzusetzen oder ihre Umsetzung zu verhindern?

⁸ Zum Verhältnis von Handlungsstruktur und Themenstruktur vgl. Schröder 2003; Fritz 2013.

⁹ Vgl. das Ariadne-Themenheft Mädchenschulgeschichte(n) 2008.

- (iii) Welche Funktion haben unterschiedliche Wortschatzeinheiten bei der Organisation dieser Texte?
- (iv) Wie unterscheidet sich das lexikalische Profil von Texten aus dem Diskurs um Frauenfragen im Umkreis der ersten Frauenbewegung von Texten aus anderen Domänen des Sprachgebrauchs?
- (v) Welche Rolle spielen Wortbildung, Metaphorik, lexikalische Kreativität und andere Verfahren als kommunikative Ressourcen?

Damit sind die wesentlichen Bereiche kurz charakterisiert, in denen unsere fachlichen Ziele liegen. Wir kommen nun zur Unterstützung der Arbeit durch digitale Ressourcen und betrachten diese Frage zunächst vom Ende her: Wie sollen und können die Befunde zu den gerade genannten Forschungsbereichen formuliert, publiziert und in geeigneter Form dokumentiert werden? Damit sind wir bei einer ersten Form der Unterstützung eines solchen Projekts durch eine integrierte digitale Publikations- und Dokumentationsumgebung.

4. Eine integrierte Publikations- und Dokumentationsumgebung

Eine integrierte digitale Dokumentations- und Publikationsumgebung umfasst drei wesentliche Säulen bzw. Komponenten, deren Bestandteile untereinander vernetzt sind: (i) ein strukturiertes digitales Textkorpus; (ii) eine lexikalische Komponente zur Charakterisierung und Dokumentation des Wortgebrauchs; (iii) eine Komponente mit sprach- und kommunikationsbezogenen Untersuchungen.

4.1 Digitale Quellentexte und das TeF-Korpus

Im Rahmen unserer Arbeiten haben wir erste Quellentexte als Volltexte erfasst und in standardkonformer Weise gemäß den Richtlinien der Text Encoding Initiative in der aktuell gültigen Version aufbereitet (TEI-P5). Bei der Auswahl haben wir uns bislang an einzelnen Themensträngen (Diskussion um die Mädchenschulreform im Anschluss an die sogenannte *Gelbe Broschüre*), an größeren Themenbereichen (Wahlrechtsfrage) und an wichtigen Autorinnen (z.B. Hedwig Dohm, Helene Lange) orientiert.

Die Textkodierung erfolgt im sogenannten Basisformat des Deutschen Textarchivs der BBAW (DTABf), das für die Aufbereitung historischer Texte mit dem Ziel entwickelt wurde, eine einheitliche und interoperable Annotation für eine Vielzahl verschiedener Texte zu schaffen. Das DTABf ist eine echte Teilmenge des TEI-P5-Standards und wird im Infrastrukturprojekt CLARIN-D als Best-Practice-Modell für historische Korpora genutzt.¹⁰ Alle Texte durchlaufen einen Qualitätssicherungszyklus über den web-basierten und das kollaborative Arbeiten ermöglichenden [DTAQ-Bereich des Deutschen Textarchivs](#).

¹⁰ CLARIN-D Benutzerhandbuch.

Abbildung 3 und Abbildung 4 zeigen zwei Screenshots, auf denen die wesentlichen Eigenschaften dieser Umgebung zu erkennen sind: a) die Annotation der Texte im DTABF-Format, verbunden mit der Möglichkeit, daraus unterschiedliche Ausgabeformate zu erzeugen, insbesondere das in CLARIN-D genutzte TCF-Format, welches für Annotations- und Text-Mining-Aufgaben genutzt wird, eine HTML-Ansicht für die Lektüre auf dem Desktop sowie ein epub- oder mobi-Format für die Ansicht auf mobilen Endgeräten; b) das synoptische Text/Bild-Arrangement und c) das DTAQ-Modul für die Korrekturarbeiten am Text und am Markup.



Abb. 3: Arbeitsumgebung im DTA/DTAQ (Deutsches Textarchiv, Qualitätssicherung) (Screenshot: Thomas Gloning).



Abb. 4: Arbeitsumgebung im DTA/DTAQ (Deutsches Textarchiv, Qualitätssicherung) (Screenshot: Thomas Gloning).

Ziel der Korrektur im DTAQ ist die Schaffung einer stabilen Version, die dann auch dauerhaft referenziert werden kann. Dies ist zum Beispiel standardisiert über einen CTS-Server (Canonical Text Services for big data) möglich, wie er derzeit in dem Leipziger Projekt [The Library of Billion Words](#) (Crane / Heyer) entwickelt wird.

Ein weiterer wichtiger Aspekt besteht darin, dass wir einzelne Texte nicht nur ins DTA einspeisen, sondern zusammen mit einer kommentierenden Einleitung bzw. Untersuchung als ›richtige‹ Veröffentlichung in der Elektronischen Bibliothek der Justus-Liebig-Universität Gießen publizieren (werden). Dieser Schritt hängt mit der unterschiedlichen Art der Gewichtung bzw. Zählung von elektronischen Texten und ›richtigen‹ Publikationen im akademischen

Belohnungssystem zusammen, aber auch mit der bibliothekarischen Katalogisierung und dem zusätzlichen Anreiz, der dadurch geschaffen werden kann. Im Hinblick auf die DH-Unterstützung bedeutet dies unter anderem, dass wir ein geeignetes Satzsystem benötigen, das aus einer XML-kodierten Quelle ein vernünftiges Satzprodukt erstellt. Obwohl es eine Reihe von Systemen gibt (z.B. Tustep, Indesign, XML-Print), ist diese Aufgabe in der praktischen Umsetzung nicht trivial. Sie ist nur dann vollautomatisch möglich, wenn die Texte im Vorfeld einheitlich bezüglich ihrer TEI-P5-Kodierung annotiert wurden. Ein Beispiel hierfür liefert die an der BBAW entwickelte Digitale Arbeitsumgebung für das BBAW-Editionsvorhaben [Schleiermacher in Berlin 1808–1834](#).

Die Frage der Textmenge und des Mengengerüsts steht vorläufig nicht im Vordergrund, weil wir derzeit themenorientiert arbeiten und die gewählten Themenbereiche sukzessive ausbauen. Der weitere Ausbau geht in die Richtung eines strukturierten Korpus zu Texten der ersten Frauenbewegung (TeF-Korpus), in dem alle wesentlichen thematischen Stränge abgedeckt sein sollen. Mit Hilfe zusätzlicher Metadaten (neben Autorin, Datierung etc.) sollen dann auch Markierungen zu spezifischen Diskursbereichen (z.B. Wahlrechtsdiskurs), zu einzelnen Themensträngen, zu den genutzten Text- und Medientypen erfolgen, ggf. auch um Autorinnen den bereits erwähnten ›Flügel‹ innerhalb der Frauenbewegung zuordnen zu können.

4.2 Dokumentation des Wortgebrauchs

Die zweite Komponente des integrierten Publikations- und Dokumentationssystems dient der Charakterisierung und der Dokumentation des Wortgebrauchs. Hier beschreiben wir in erster Linie die lexikalischen Mittel, die für die kommunikativen Ziele innerhalb der Diskurse im Umkreis der ersten Frauenbewegung erheblich sind. Dazu gehören, um nur drei Beispielgruppen zu nennen, Bezeichnungen für zentrale Forderungen (»Frauenstimmrecht«), Bezeichnungen, die bestimmten Denkfiguren und Grundannahmen zuzuordnen sind (z.B. »Sphäre«, »Herd«), oder auch der polemische Wortschatz (z.B. »Frauenlandsturm« als abschätzige Bezeichnung der Gegnerinnen und Gegner, als es um die Petitionen zum *Bürgerlichen Gesetzbuch* um 1900 ging).

Der Kern der lexikalischen Beschreibungen sind strukturierte Wortartikel, in denen die einzelnen Verwendungsweisen (Lesarten) von Wörtern lexikographisch beschrieben und mit geeigneten Belegen auf den textuellen Gebrauch bezogen werden. Darüber hinaus werden die einzelnen Bedeutungspositionen durch Deskriptoren intern markiert, so dass eine thematische, funktionale, gruppenspezifische usw. Erschließung des Wortgebrauchs ermöglicht wird.

Im Folgenden geben wir einen Ausschnitt aus einem solchen Artikel: Die wesentlichen Informationspositionen lassen sich entsprechend den Vorgaben der TEI-P5 in Form- und Bedeutungsbeschreibungen unterteilen. Zu den Formangaben gehören das Lemma (genauer die Lemmazcichengestaltangabe) und die grammatischen Angaben, zur Bedeutungsbeschreibung die Lesarten, Bedeutungsparaphrasen und Belege sowie die

Markierungen zu unterschiedlichen Gebrauchsdimensionen (z.B. diasystematische Merkmale, Bezüge zu argumentativ-ideologischen Systemstellen). Über die Belegstellenangabe, wenn diese eindeutig auf den Ursprungstext über einen persistenten Identifizierer referenzierbar ist, ist eine Verlinkung mit dem Volltext möglich. Die Markierungen dienen der systematischen Erschließung des Wortgebrauchs, können aber auch genutzt werden, um einzelne Verwendungsweisen von Ausdrücken mit denjenigen Passagen in den Untersuchungen zu verlinken, in denen der entsprechende Aspekt (z.B. Wortbildung oder eine bestimmte Denk-Systemstelle) behandelt wird. Die Verwendungsweise des Substantivs *Herd*, die sich auf die Systemstelle »Differenzhypothese« bezieht, würde im Druckbild etwa so aussehen:

Eine Verlinkung ist aber nicht nur vom Wortartikel aus denkbar. Umgekehrt lässt sich über das Stichwort und ggf. die Verwendungsweisen-Zählung aus den monographischen Darstellungen heraus auf Artikel und Artikelteile verweisen. Wir greifen damit eine Darstellungsform auf, die Alfred Schirmer in seinem *Wörterbuch der Kaufmannssprache* (1911) entworfen und damals mit typographischen Mitteln realisiert hatte. Wenn die Belegstellenangaben in den Wortartikeln seiten- und zeilengenau mit den Volltexten verbunden werden, erweitern wir auf diese Weise traditionelle diskursanalytische Darstellungsformen¹¹ um eine lexikographische Komponente und um eine Anbindung an die Volltexte.

Eine wichtige Zielsetzung der lexikalischen Dokumentation ist es darüber hinaus, die Resultate und Befunde systematisch »anspruchbar« zu machen für die laufenden und zukünftigen Wörterbuchprojekte, z.B. beim DWDS, beim IDS oder auch an den Akademien. Hierfür ist es nötig, entsprechende Kodierungsschemata zu entwerfen (im Falle von XML wären dies XML-Schemata), mit denen die Anbindung an diese Projekte über Schnittstellen automatisierbar möglich ist. Als Beispiel dient der bereits oben aufgeführte Wortartikel für *Herd*. Der Artikel ist hier in die Werkstattsprache des DWDS (vgl. *Abbildung 5*) überführt.¹² Diese ist, was die lexikographischen Merkmale angeht, verlustfrei in den TEI-P5-Standard überführbar. Die Metadaten hingegen, wie beispielsweise Quelle und Zeitstempel, können durch eine Anpassung (customization) von TEI-P5 abgebildet werden. Über die Elemente <Verweis>, <Ziel_url> und <Ziellesart> lässt sich ein Artikel aus dem Tef-Wörterbuch direkt auf die Lesart Nr. 1 des entsprechenden Artikels im DWDS-Wörterbuch referenzieren (*Abbildung 5*).

¹¹ Z.B. im Band *Kontroverse Begriffe* (Stötzel / Wengeler 1995).

¹² Geyken 2014.

Wir verlassen nun den Bereich der digital integrierten Ergebnisdokumentation und kommen zu der Frage, wie sich die sprachlichen und kommunikationsbezogenen Untersuchungen, die ja den wissenschaftlichen Kern des Projekts darstellen, mit digitalen Mitteln unterstützen lassen.

5. Möglichkeiten der digitalen Unterstützung der Untersuchungen

Unser Ziel besteht nicht primär darin, die Möglichkeiten der digitalen Geisteswissenschaften bzw. Digital Humanities (DH) im Bereich der »avantgardistischen« Anwendungen zu erweitern. Wir sind primär an der Frage interessiert, wie sich fachlich vorgegebene Zielsetzungen durch die Anwendung von DH-Ressourcen und -Werkzeugen unterstützen lassen. Und wir sind im Hinblick auf eine kritische Prüfung des Mehrwert-Versprechens auch interessiert an den typischen Schwierigkeiten sowie an Grenzen der Anwendbarkeit, der Nützlichkeit und der Effizienz von DH-Methoden.

Unsere Fragestellungen und Zielsetzungen sind zum einen projektspezifisch, sie haben darüber hinaus aber exemplarischen Charakter. Wir betrachten unser Thema auch als stellvertretend bzw. typisch für eine ganze Klasse von diskursorientierten, pragmatischen und lexikologischen bzw. historisch-semantischen Fragestellungen und der damit verbundenen Nutzung digitaler Textkorpora und Werkzeuge. Es ist damit ein paradigmatischer Fall für die Anforderungen, die geisteswissenschaftliche Projekte dieser Art an die Datennutzungsmethoden der DH stellen.

Wir stellen hier zunächst basale Anwendungen von DH-Verfahren zusammen, deren Nutzen in der Korpuslinguistik und in der Texttechnologie unumstritten und weithin akzeptiert ist.

5.1 Digitale Texte

Digitale Texte werden zusammen mit den darauf bezogenen Konkordanzwerkzeugen und Rechercheverfahren in erster Linie im Bereich der lexikalischen Analyse genutzt. Zu den fachlichen Zwecken in diesem Bereich gehören u.a. die Unterscheidung von Verwendungsweisen und ihre semantische Beschreibung, die Ermittlung spezifischer Diskursfunktionen von Wörtern, die Ermittlung von Wortbildungszusammenhängen und die Beurteilung von Frequenzverhältnissen innerhalb von Texten, Textgruppen, aber auch vor dem Hintergrund eines Referenzkorpus.

5.2 Formen des Markup und der Annotation

Formen des Markup und der automatischen Annotation spielen – unabhängig von ihrer Rolle für Standardisierung, Interoperabilität, Nachnutzbarkeit – für fachliche Interessen im

engeren Sinne zunächst eine eher vermittelte bzw. vorbereitende Rolle, etwa bei kombinierten Abfragen, in denen Wortarten berücksichtigt werden sollen und bei denen ein Part-of-Speech-Tagging genutzt wird. Mit automatischen Mitteln lassen sich Wortformen lemmatisieren und ggf. zerlegen, Eigennamen erkennen, mit Thesauri kann ein semantisches Tagging vorbereitet werden, auch stehen eine Reihe von Parsern zur Verfügung, mit denen sich syntaktische Annotationen erstellen lassen, die u.a. für die Analyse von Konstruktionsmustern genutzt werden können.

Höherwertige Formen der Annotation, gerade dann, wenn diese verlässlich sein sollen, müssen jedoch nach wie vor händisch erstellt werden. Hierzu gehören insbesondere:

- die Markierung von Verwendungsweisen im Rahmen der lexikalischen Dokumentation im Hinblick auf semantische, kommunikative und argumentative Aspekte des Gebrauchs der betreffenden Ausdrücke (sie erlaubt es, die beträchtliche Komplexität des Wortschatzes überschaubar und für Abfragen zugänglich zu machen; z.B. »Suche alle Personenbezeichnungen, mit denen eine abwertende Festlegung verbunden ist«);
- die Auszeichnung von Formen der Metaphorik und anderer rhetorischer bzw. rhetorisch-lexikalischer Strategien;
- die Auszeichnung unterschiedlicher Argumenttypen und ihrer Rolle in größeren argumentativen Kontexten;
- die Auszeichnung von spezifischen Diskursphänomenen wie z.B. intertextuelle Bezüge zu gegnerischen oder stützenden Texten.

Zu den Herausforderungen in diesem Bereich gehört es, dass relevante Textteile von ganz unterschiedlicher Größe sein können und dass unterschiedliche Parameter der Textorganisation auch zu überlappenden Annotationsstrukturen führen und somit inline in XML nicht kodiert werden können. Das Problem der überlappenden Annotationsstrukturen ist jedoch in einem stand-off- und Mehrebenenverfahren lösbar, z.B. in Systemen wie ANNIS oder WebAnno.

5.3 Editorische Erschließung

Auch die digitale editorische Erschließung von Diskursthemen gehört mit zu den fachlichen Aufgaben, die im Sinne einer nachhaltigen und vielfältigen Nutzbarkeit der Texte in Form von kommentierten, thematischen Spezialkorpora geschehen kann.

5.4 Formen der Wortschatz- und Wortgebraucherschließung

Zu den zentralen Unterstützungsformen gehört auch die Nutzung korpuslinguistischer und lexikographischer Werkzeuge, mit denen digitale Texte bearbeitet, ausgewertet und verglichen werden können. Im Vordergrund stehen dabei semantische Aspekte, Fragen des diskurspezifischen Gebrauchsprofils von Wörtern, morphologische Zusammenhänge in

diskursspezifischen Wortschatzsektoren, Frequenzverhältnisse und Frequenzentwicklungen sowie der Abgleich von Texten des TeF-Korpus mit zeitgenössischen Referenzkorpora.

5.5 Argumentations- bzw. Kontroversen-Visualisierung

Verfahren der Argumentations- bzw. Kontroversen-Visualisierung sind ein Bereich, den wir noch nicht erprobt haben, dessen Anwendbarkeit auf historische Auseinandersetzungen wir aber ins Auge fassen.¹³

6. Die Anbindung an das Infrastrukturprojekt CLARIN-D

Unser Projekt zeigt in exemplarischer Weise auf, wie schon bei den Vorarbeiten die Anbindung an eine Infrastrukturumgebung (in unserem Fall Clarin-D) mit bedacht werden kann. Sie betrifft (i) das TeF-Korpus und (ii) die lexikalische Dokumentation.

6.1 Das TeF-Korpus im Deutschen Textarchiv

Das Zentrum Sprache der Berlin-Brandenburgischen Akademie der Wissenschaften (BBAW) ist Partner im Clarin-D-Verbund und damit den Zielen der Standardisierung, der Nachhaltigkeit, der Interoperabilität und der freien Nutzbarkeit verpflichtet. Ebenfalls Teil des Zentrum Sprache der BBAW ist das Deutsche Textarchiv (DTA), welches die Aufnahme wichtiger historischer Volltexte in hoher Qualität mit dem Ziel verfolgt, ein Referenzkorpus für das ältere Neuhochdeutsch für die Forschung bereitzustellen. Die Zusammenarbeit des Projekts mit dem DTA liefert für beide Parteien Vorteile: Das Projekt erhält eine sprachtechnologische Expertise, die es allein niemals in dieser Form erreichen könnte; umgekehrt wird das DTA durch die Zusammenarbeit sozusagen ›gefüttert‹ mit neuen Texten aus einem ganz bestimmten Segment des Sprachgebrauchs, das bisher deutlich unterrepräsentiert war.

Eine solche Art von Zusammenarbeit könnte ein Standardmodell für neue korpusgestützte Projekte im Bereich der Geisteswissenschaften sein, bei denen oft das Problem besteht, dass jedes Projekt ›das Rad neu erfindet‹, nicht selten aber auch hinter den etablierten Standards für das Datenmanagement zurückbleibt. Kooperationen dieser Art setzen allerdings voraus, dass eine Einrichtung wie das DTA langfristig als wissenschaftliche Zentraleinrichtung für historische Korpus-texte und als Ansprechpartner für neue Projekte zur Verfügung steht (und entsprechend finanziert wird).

6.2 Schnittstellen der lexikalischen Dokumentation

¹³ Vgl. hierzu etwa die [Arbeiten von Michael Hoffmann](#). Auf der Mailingliste [ArgThry](#) wurden immer wieder Fragen und Probleme der digitalen Visualisierung von Argumentationen behandelt, so zum Beispiel die Konkurrenz in der Darstellung von sequenziellen und hierarchischen Zusammenhängen im Rahmen von Argumentationen.

Im Clarin-D-Projekt sind mehrere Zentren vertreten, zu deren Aufgaben die nachhaltige Pflege digitaler lexikalischer Ressourcen gehört. Zu nennen sind hier vor allem das Institut für Deutsche Sprache (Mannheim) und die Arbeitsstelle des Digitalen Wörterbuchs, einem Langzeitvorhaben der BBAW (Berlin).

An Projekte wie das unsere, das spezifische lexikalische Ergebnisse produzieren und dokumentieren soll, kann man folgende Anforderungen stellen:

- (i) Aufbereitung der lexikalischen Dokumentation in strukturierter Form, die den Best-Practice-Richtlinien entspricht, wie sie im Projektverbund von Clarin-D formuliert werden. Im Hinblick auf die Quellentexte bedeutet dies, dass die Daten bezüglich Metadaten und Textdaten im DTABf-Format vorliegen müssen.
- (ii) Besonders wichtig, auch für die interne Vernetzung von Dokumenten, sind eindeutige und praktikabel adressierbare Referenzierungen auf Seiten-, Paragraphen- und ggf. auch Zeilenebene, die auch für Schnittstellen von auswärtigen Systemen nutzbar sind. Auf diese Weise können die Befunde, die in unseren Spezialuntersuchungen dokumentiert sind, auch von allgemeineren Darstellungen aus digital angesprochen werden. Zum Beispiel könnten etwa Wortschatzdokumentationen zum Diskurs um das Frauenwahlrecht auch vom DWDS, von den IDS-Seiten oder von woerterbuchnetz.de aus angesteuert und verlinkt werden.

7. Zusammenfassung und Ausblick

Gegenstand des vorliegenden Beitrags war es, einen Überblick über die laufenden und geplanten Arbeiten zum Sprachgebrauch und zu Formen der Kommunikation in der ersten Frauenbewegung zu geben. Wir haben zunächst die wichtigsten fachlichen Fragestellungen im Bereich der Diskursorganisation, der Handlungs- und Argumentationsformen sowie des Wortgebrauchs erläutert und daran anschließend zu zeigen versucht, welche Rolle eHumanities- bzw. DH-Werkzeuge bei der Ergebnisdokumentation und bei den fachlich orientierten Untersuchungen spielen können.

Eine wichtige Frage war auch die, wie solche Arbeiten rechtzeitig und schon bei der Planung an Infrastruktur-Umgebungen wie Clarin-D angebunden werden können, um die Befolgung fachlicher DH-Standards sicherzustellen.

Die Ausführungen zu diesem spezifischen Projekt betrachten wir gleichzeitig als ein Diskussionsangebot über die Möglichkeiten und Erfordernisse von Projekten, deren Ziele und Gegenstände thematisch unterschiedlich, strukturell aber ähnlich gelagert sind.

Bibliographische Angaben

- Dietrich Busse / Wolfgang Teubert: Ist Diskurs ein sprachwissenschaftliches Objekt? Zur Methodenfrage der historischen Semantik. In: Begriffsgeschichte und Diskursgeschichte. Hg. von Dietrich Busse / Fritz Hermanns / Wolfgang Teubert. Opladen 1994, S. 10–28. [[Nachweis im OPAC](#)]
- Gerd Fritz: Dynamische Texttheorie. Gießen 2013. [[online](#)]
- Ute Gerhard: Unerhört. Die Geschichte der deutschen Frauenbewegung. Reinbeck 1991. [[Nachweis im OPAC](#)]
- Alexander Geyken: Methoden bei der Wörterbuchplanung in Zeiten der Internetlexikographie. In: Lexicographica 30 (2014), S. 73–111. [[Nachweis im GBV](#)]
- Thomas Gloning: Diskursive Praktiken, Textorganisation und Wortgebrauch im Umkreis der ersten Frauenbewegung. In: Historische Pragmatik. Hg. von Peter Ernst. Berlin, Boston 2012, S. 127–146 (= Jahrbuch für germanistische Sprachgeschichte 3). [[Nachweis im GBV](#)]
- Mädchenschulgeschichte(n). Die preußische Mädchenschulreform und ihre Folgen. Ariadne 53/54 (2008). [[Nachweis im GBV](#)]
- Suane Haaf / Alexander Geyken / Frank Wiegand: The DTA ›Base Format: A TEI Subset for the Compilation of a Large Reference Corpus of Printed Text from Multiple Sources. In: Journal of the Text Encoding Initiative (JTEI), 8 (2015). DOI: 10.4000/jtei.1114 [[online](#)]
- Handbuch der Frauenbewegung. Hg. von Helene Lange / Gertrud Bäumer. Band I–V. Berlin 1901–1906. [[Nachweis im GBV](#)]
- Helene Lange: Die höhere Mädchenschule und ihre Bestimmung. Begleitschrift zu einer Petition an das preußische Unterrichtsministerium und das preußische Abgeordnetenhaus. Berlin 1887. [[Nachweis im GBV](#)]
- Helene Lange: Kampfzeiten. Aufsätze und Reden aus vier Jahrzehnten. 2 Bände. Berlin 1928. [[Nachweis im GBV](#)]
- Öffentlicher Sprachgebrauch. Praktische, theoretische und historische Perspektiven. Georg Stötzel zum 60. Geburtstag gewidmet. Hg. von Karin Böke / Matthias Jung / Martin Wengeler. Opladen 1996. [[Nachweis im GBV](#)]
- Politische Netzwerkerinnen. Internationale Zusammenarbeit von Frauen 1830–1960. Hg. von Eva Schöck-Quinteros / Anja Schüler / Annika Wilmers / Kerstin Wolff. Berlin 2007. [[Nachweis im GBV](#)]
- Angelika Schaser: Frauenbewegung in Deutschland. Darmstadt 2006. [[Nachweis im GBV](#)]
- Thomas Schröder: Die Handlungsstruktur von Texten. Ein integrativer Beitrag zur Texttheorie. Tübingen 2003. [[Nachweis im GBV](#)]
- Georg Stötzel / Martin Wengeler: Kontroverse Begriffe. Geschichte des öffentlichen Sprachgebrauchs in der Bundesrepublik Deutschland. In Zusammenarbeit mit Karin Böke, Hildegard Gorny, Silke Hahn, Matthias Jung, Andreas Musolff, Cornelia Tönnesen. Berlin, New York 1995. [[Nachweis im GBV](#)]
- Hans Sveistrup / Agnes von Zahn-Harnack: Die Frauenfrage in Deutschland. Strömungen und Gegenströmungen 1790–1930. Sachlich geordnete und erläuterte Quellenkunde. Burg 1934. [[Nachweis im GBV](#)]
- Ulla Wischerman: Frauenbewegungen und Öffentlichkeiten um 1900. Netzwerke – Gegenöffentlichkeiten – Protestinszenierungen. Königstein 2003. [[Nachweis im GBV](#)]
- Kerstin Wolff: Ein ungewöhnlicher Schreib-Ort? Frauenrechtlerinnen im deutschen Kaiserreich und ihr politisches Schreiben im Frauenverein – Eine Annäherung. In: Frauen in der literarischen Öffentlichkeit 1780–1918. Hg. von Caroline Bland / Elisa Müller-Adams. Bielefeld 2007, S. 121–142. [[Nachweis im GBV](#)]

Abbildungslegenden und -nachweise

Abb. 1: Helene Lange, Frontispiz, aus: Lange 1928, Bd. 1 (links); *Handbuch der Frauenbewegung*, Titelblatt, aus: Lange / Bäumer 1901–1906, Bd. 4 [[online](#)] (rechts).

Abb. 2: Bibliographie zur Geschichte der Frauenbewegung 1790–1930 (1934) (Quelle: Thomas Gloning, privat).

Abb. 3: Arbeitsumgebung im DTA/DTAQ (Deutsches Textarchiv, Qualitätssicherung) (Screenshot: Thomas Gloning).

Abb. 4: Arbeitsumgebung im DTA/DTAQ (Deutsches Textarchiv, Qualitätssicherung) (Screenshot: Thomas Gloning).

Abb. 5: Anbindung an das DWDS-Wörterbuch (Quelle: Alexander Geyken).

Abb. 6: Verbindung von Wortgebrauchsdokumentation, Untersuchungen, Volltext. (Eigene Darstellung: Thomas Gloning).

Zeitschrift für digitale Geisteswissenschaften

Artikel aus:

Sonderband 1 der ZfdG: Grenzen und Möglichkeiten der Digital Humanities. Hg. von Constanze Baum und Thomas Stäcker. 2015. DOI: [10.17175/sb01](https://doi.org/10.17175/sb01)

Titel:

Exilnetz33. Ein Forschungsportal als Such- und Visualisierungsinstrument

Autor/in:

Theresia Biehl

Kontakt: biehl@uni-trier.de

Institution: Universität Trier

GND: [1084603756](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63862-p0011-9) ORCID:

Autor/in:

Anne Lorenz

Kontakt: lorenz@dla-marbach.de

Institution: Deutsches Literaturarchiv Marbach

GND: ORCID:

Autor/in:

Dirk Osierenski

Kontakt: osierens@informatik.uni-halle.de

Institution: Martin-Luther-Universität Halle-Wittenberg

GND: [1084240947](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63862-p0011-9) ORCID:

DOI des Artikels:

[10.17175/sb001_011](https://doi.org/10.17175/sb001_011)

Nachweis im OPAC der Herzog August Bibliothek:

[830202544](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63862-p0011-9)

Erstveröffentlichung:

19.02.2015

Lizenz:

Sofern nicht anders angegeben 

Medienlizenzen:

Medienrechte liegen bei den Autoren

Letzte Überprüfung aller Verweise:

24.05.2016

GND-Verschlagwortung:

[Exil](#) | [Brief](#) | [Semantisches Netz](#) |

Zitierweise:

Theresia Biehl, Anne Lorenz, Dirk Osierenski: Exilnetz33. Ein Forschungsportal als Such- und Visualisierungsinstrument. In: Grenzen und Möglichkeiten der Digital Humanities. Hg. von Constanze Baum / Thomas Stäcker. 2015 (= Sonderband der Zeitschrift für digitale Geisteswissenschaften, 1). PDF Format ohne Paginierung. Als text/html abrufbar unter DOI: [10.17175/sb001_011](https://doi.org/10.17175/sb001_011).

Theresia Biehl, Anne Lorenz, Dirk Osierenski

Exilnetz33. Ein Forschungsportal als Such- und Visualisierungsinstrument

Abstracts

Auf der Grundlage der computergestützten quantitativen Datenauswertung einerseits und der qualitativen philologischen Texterschließung andererseits sieht das Projekt *Vernetzte Korrespondenzen. Erforschung und Visualisierung sozialer, räumlicher, zeitlicher und thematischer Netze in Briefkorpora* die Entwicklung eines internetbasierten, generischen Forschungsportals für Briefkorpora vor, das diese insbesondere in ihren netzartigen Zusammenhängen präsentiert. Unterschiedliche Such- und Visualisierungsinstrumente schaffen damit neue Zugangsmöglichkeiten zu einem inhaltlich spezifischen Korpus – den Briefen von deutschsprachigen Schriftstellern und Kulturschaffenden, die in der Zeit der nationalsozialistischen Herrschaft ins Exil gezwungen wurden.

In combining computer-supported methods of quantitative data analysis with qualitative philological text analysis approaches such as indexing and semantic interpretation of texts, the project *Epistolary Networks: Visualising multi-dimensional information structures in correspondence corpora* aims to develop a generic, web-based research platform for the visualisation of and research on social, spatial, and personal networks of letters. Instruments like multiple visualisations and a cascading search create new means of accessing one specific corpus – the letters of German-speaking writers and intellectuals forced into exile during the Nazi regime.

1. Vom Brief zum Netzwerk

Seinen Roman *Georg* beendet Siegfried Kracauer 1934 im Pariser Exil, veröffentlicht wird er fast vierzig Jahre später posthum im Suhrkamp Verlag.¹ Zwischenzeitlich dokumentieren Kracauers Korrespondenzen mit Freunden und weniger nahestehenden Kollegen und Verlegern seine Suche nach einem geeigneten Verlag, die sich in der Zeit der nationalsozialistischen Herrschaft beinahe aussichtslos gestaltete. In einem Brief an den Kunsthistoriker Julius Meier-Graefe vom 24. August 1933 schreibt Kracauer: »Ich bin durch wochenlange Korrespondenz über Abfindung, Möbel usw. so auf dem Hund, dass ich kaum an meinem Roman arbeiten kann. Dabei ist er meine einzige Hoffnung.«²

Dass sich in Kracauers Fall letztlich ein Ausweg findet und sich doch eine Möglichkeit zur Publikation ergibt, wird bei der Lektüre seiner Briefe schnell ersichtlich. Woran jedoch seine exilierten Kollegen und Korrespondenzpartner ihre letzte Hoffnung knüpfen oder wer wo zur gleichen Zeit dieselben Verlage nennt, legt erst ein Texterschließungsprozess offen, der die Netzstruktur der Briefkorpora berücksichtigt.

¹ Erstveröffentlichung in Kracauer 1973, Band 7.

² Brief von Siegfried Kracauer an Julius Meier-Graefe vom 24.8.1933, Deutsches Literaturarchiv Marbach, Nachlass Siegfried Kracauer.

Das interdisziplinär angelegte Projekt *Vernetzte Korrespondenzen*³, das im Folgenden vorgestellt wird, nimmt sich dieser Aufgabe an. Unter Zuhilfenahme quantitativer Auswertungsmethoden aus der Informatik sieht das Projekt vor, Korrespondenzen deutschsprachiger Kulturschaffender aus der Exilzeit von 1932–1950 inhaltlich-philologisch zu erschließen, auf einem Forschungsportal bereitzustellen und den Netzwerkcharakter der Briefe, ihre soziale, zeitliche, räumliche und thematische Dimension durch unterschiedliche Formen der Visualisierung sicht- und erforschbar zu machen (Abbildung 1).

Mit dem Aufbrechen linearer Repräsentationsstrukturen durch das digitale Medium vollzieht sich auch ein Wandel in der Editionsphilologie. Während Briefeditionen traditionell einen Autor als Ordnungsgröße wählen und in den Mittelpunkt stellen, verschiebt sich nun der Fokus hin zu einer autordezentrierten Perspektive. Die digitale Edition stellt dem Ordnungssystem der Print-Ausgabe multiperspektivische und multidimensionale Darstellungsmöglichkeiten gegenüber und erweitert dadurch die Erschließung von Briefkorpora um die Repräsentation ihres genuinen Netzwerkcharakters.

Dass die Darstellung von Briefen in ihrer netzartigen Struktur durchaus vorteilhaft ist, zeigt sich bei der Betrachtung der Briefe Kracauers und seiner emigrierten Kollegen. Denn diese geben nicht nur Einblicke in die Erfahrungswelt des Exils deutschsprachiger Intellektueller, in der die identitätsstiftenden kommunikativen Strukturen plötzlich auseinanderbrechen und den Briefen in ihrer Ersatzfunktion eine existentielle Bedeutung zuweisen. Sie veranschaulichen gleichzeitig, wie sich neue, für das Exil spezifische Themen in den Korrespondenzen verbreiten und welche Motive die briefliche Kontaktaufnahme zu anderen exilierten Schriftstellern und Kulturschaffenden mehrheitlich bestimmen. Gerade in der durch Zensur und Verlust geprägten Zeit des Exils, als allein aus dem deutschsprachigen Raum etwa eine halbe Million Menschen in die verschiedensten Länder fliehen, zeugt das hohe Aufkommen und die Ausdehnung des Schriftverkehrs über nationale und kulturelle Grenzen hinweg⁴ von der per se netzwerkbildenden Funktion des Briefes.

Dennoch wurden Netzwerke und speziell Korrespondenznetzwerke von der Exilforschung bislang nur in Ansätzen bearbeitet.⁵ Wesentliche Gründe dafür scheinen die quantitative Fülle des Materials und die qualitative, thematische und gedankliche Heterogenität der Briefe zu sein, die einer Erschließung und Erforschung mit traditionellen literaturwissenschaftlichen Methoden entgegenstehen. Mit der interdisziplinären Entwicklung eines Portals, das der Erforschung und Visualisierung solcher Korrespondenznetze dient, soll der Beweis angetreten werden, dass sich die Erschließungsmethoden der Literaturwissenschaft und Editionsphilologie durch Verfahren der Informatik sinnvoll ergänzen und unterstützen lassen. Der Frage, ob und

³ Das Projekt ist ein vom Bundesministerium für Bildung und Forschung (BMBF) im Rahmen der Ausschreibung eHumanities gefördertes Verbundvorhaben. Verbundpartner sind das Trier Center for Digital Humanities an der Universität Trier (Leitung: Dr. Thomas Burch, Dr. Vera Hildenbrandt, Prof. Dr. Claudine Moulin), das Deutsche Literaturarchiv Marbach (Leitung: Dr. Roland S. Kamzelak) und das Institut für Informatik der Martin-Luther-Universität Halle-Wittenberg (Leitung: Prof. Dr. Paul Molitor, Dr. Jörg Ritter).

⁴ Zur »Briefschreibwiederbelebung« im Exil sowie zum regen Schriftverkehr zwischen den verschiedenen Exilländern bei gleichzeitiger »Abnahme der Brieffrequenz zwischen Exil und Heimat« vgl. Evelein 2011, S. 21–23 und S. 28f. Auch Wende 1996, S. 172f. und Kucher et al. 2011, S. 13 konstatieren eine Ausweitung des Briefverkehrs mit Beginn des Exils.

⁵ Vgl. Dogramaci / Wimmer 2011, S. 9f.

wie Kracauers Roman *Georg* in den Korrespondenzen verhandelt wird, lässt sich schließlich auf unterschiedliche Weise nachgehen. Eine konkrete Suche nach Verlagen oder Verlegern, die in diesem Kontext erwähnt werden, ist genauso denkbar wie das heuristische Erfassen der Häufigkeit, mit der der Roman in den Briefen erwähnt wird.

Dementsprechend kommt der Informatik eine zweifache Funktion zu: Sie hat zum einen insofern Anteil am philologischen Texterschließungsprozess als sie Werkzeuge bereitstellt, die die inhaltliche Erschließung unterstützen. Zum anderen leistet sie einen genuinen Beitrag, indem sie automatisierte Verfahren zur Aufbereitung und Auswertung der Daten und ihrer Visualisierung entwickelt.

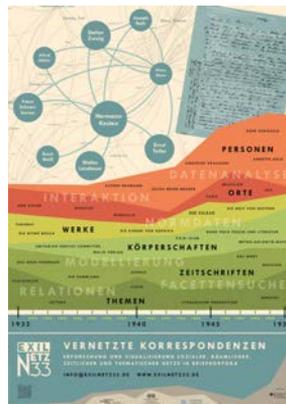


Abb. 1: Auf der DHd 2014 in Passau präsentiertes und mit dem DHd-Poster-Award prämiertes Poster des Projekts *Vernetzte Korrespondenzen*.

Anhand einer beispielhaften Auswahl verschiedener Briefe wird im Folgenden der Weg des Verbundprojektes vom Texterschließungsprozess bis zur Entwicklung der Such- und Visualisierungsmechanismen des Forschungsportals nachgezeichnet. Dabei illustriert die in den Briefen dokumentierte Verlagssuche für Kracauers Roman *Georg*, wie sich eine Forschungsfrage zu einem Briefkorpus, insbesondere zu einem Korrespondenznetzwerk, durch das Zusammenspiel von geisteswissenschaftlichen und computergestützten Erschließungsmethoden konturieren und beantworten lässt.

2. Zusammenspiel von Informatik und Geisteswissenschaften im Editionsprozess

Kern des Projektes bildet ein Korpus ausgewählter Exilbriefe (Abbildung 2) aus den Beständen des [Deutschen Literaturarchivs Marbach](#), des [Literaturarchivs der Münchner Stadtbibliothek Monacensia](#) und der [National Library of Israel](#) in Jerusalem. Durch die Verbindung dieser drei Archivbestände in einem Portal werden verstreut liegende Nachlässe zusammengeführt, auseinandergerissene Korrespondenzzusammenhänge wiederhergestellt und einzelne Briefe in einen größeren Kontext eingeordnet, der ihren Inhalt möglicherweise

in neuem Licht erscheinen lässt. Der generische Aufbau des Portals ermöglicht zudem die Integration weiterer Nachlässe.⁶

Den digitalen Methoden der Editionsphilologie entsprechend wird jeder Brief zunächst ins digitale Medium übertragen. Leitend sind dabei zum einen die projekteigenen Editionsrichtlinien, die – mit einigen Abweichungen⁷ – der diplomatischen Wiedergabe der Brieftexte verpflichtet sind. Zum anderen regelt ein projektspezifisches Schema die strukturelle und inhaltliche Auszeichnung der Briefe in XML, wobei das Auszeichnungsschema den Richtlinien der TEI folgt.

Die briefspezifischen Informationen wie Verfasser, Empfänger, Datum und Schreibort werden wie üblich als Teil des `Tags<sourceDesc>` innerhalb des TEI-Headers abgelegt. Ebenfalls Teil von `<sourceDesc>` sind beschreibende Angaben zum Dokument, also Hinweise zur Provenienz (besitzende Institution, Signatur, Copyright), zum Umfang, zum Typ (Brief, Postkarte, Telegramm oder Sonstiges) und zu dessen Eigenschaften (Manuskript, Typoskript oder Typoskript mit handschriftlichen Ergänzungen). Weist das Dokument weitere Besonderheiten auf, etwa ein auffälliges Briefpapier, werden diese im TEI-Header vermerkt. Auf der Ebene der Briefstruktur werden Schreibort, Schreibdatum, Empfänger und Verfasser sowie Anrede und Grußformel ausgezeichnet. Eine Besonderheit dabei ist die Typisierung der Anrede anhand einer sechsstufigen Skala, die von ›besonders förmlich‹ bis ›sehr persönlich‹ reicht.⁸ Über diese Kategorisierung kann bei der späteren Auswertung des Korpus z. B. ermittelt werden, wieviel Prozent der Briefe innerhalb des Gesamtnetzes – oder auch nur innerhalb eines Teilnetzes – förmlich gehalten sind und in welchen Briefen ein eher persönlicher Umgangston gepflegt wird. Zudem kann gezeigt werden, wie sich die Anrede innerhalb eines Briefwechsels im Zeitverlauf verschiebt, sodass Veränderungen im persönlichen Verhältnis zweier Korrespondenzpartner erforscht werden können.

Betrachtet man etwa die Anredeformen verschiedener An- und Von-Briefe, die Kracauers Verlagssuche für seinen Roman *Georg* thematisieren, lassen diese zweifellos auf die jeweils zugrunde liegende Beziehung der Korrespondenzpartner schließen. Den 20 Jahre älteren Kunsthistoriker Meier-Graefe, der sich innerhalb seiner Netzwerke für Kracauers Roman einsetzt, adressiert Kracauer, der in diesem Briefwechsel die Rolle des dankbaren Bittstellers innehat, durchaus förmlich mit »Sehr geehrter« oder »Sehr verehrter, lieber Herr Meier-Graefe, Sie« (Kategorie F 2). Dagegen eröffnet Meier-Graefe in der Rolle des hilfsbereiten Fürsprechers⁹ und verständnisvollen Ratgebers seine Briefe mit »Lieber Herr Kracauer, Sie« (Kategorie F 1) einen Deut persönlicher. Zieht man weitere Korrespondenzen vergleichend hinzu, zeigt sich,

⁶ Im Sinne eines möglichst vielgestaltigen und weitgreifenden Netzwerks ist diese sogar wünschenswert.

⁷ Fehlende Leerzeichen in Typoskripten werden ergänzt. Buchstabenausfälle oder -dopplungen in Typoskripten, die mechanisch bedingt sind, werden stillschweigend korrigiert.

⁸ Grundlage der Typisierung sind die gegenüber dem Adressaten gebrauchte Anredeformel sowie die im Brief verwendete pronominale Anrede (›duzen‹ oder ›siezzen‹), anhand derer jeder Brief genau einer Kategorie zugeordnet wird. Die im Projekt erstellte sechsstufige Skala reicht von F 3 ›besonders förmlich‹ bis P 1 ›sehr persönlich‹. Dabei vertreten die mit ›F‹ (für ›förmlich‹) bezeichneten Typen verschiedene Grade der Höflichkeitsform, die mit ›P‹ (für ›persönlich‹) bezeichneten Typen stehen für einen vertraulicheren Umgangston.

⁹ Das Angebot Meier-Graefes, »Wo auch immer Sie glauben, ein Wort von mir könne Ihnen helfen, stehe ich natürlich zur Verfügung« (Brief von Julius Meier-Graefe an Siegfried Kracauer vom 30.3.[1933], Deutsches Literaturarchiv Marbach, Nachlass Siegfried Kracauer), bringt das Verhältnis auf den Punkt.

dass Kracauer Thomas Mann – einen Kontakt, den Meier-Graefe im Zuge der Verlagssuche vermittelt – förmlich mit »Sehr geehrter Herr Mann, Sie« (F 2) anspricht und Thomas Mann gleichsam paritätisch mit »Sehr verehrter Herr Kracauer, Sie« (F 2) antwortet. Vertraulicher ist Kracauer in der Korrespondenz mit seinem langjährigen Lektor Max Tau, den er etwa mit »Lieber Freund Max, Sie« (P 3) adressiert. Im Zeitverlauf bleiben die Anredeformen der betrachteten Korrespondenzen dagegen konstant und liefern somit keine Hinweise auf eine Veränderung der jeweiligen Beziehung.

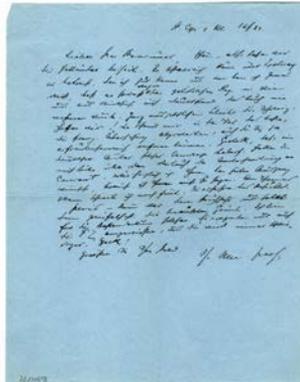


Abb. 2: Faksimile eines Briefes von Julius Meier-Graefe an Siegfried Kracauer vom 16.11.1934. © Deutsches Literaturarchiv Marbach.

Die inhaltliche Erschließung des Brieffextes erfolgt durch die Auszeichnung relevanter Informationen wie Personen, Orte, Werke, Periodika und Körperschaften, die mit entsprechenden Registern verknüpft werden. Dabei werden nicht nur explizite Nennungen berücksichtigt, sondern auch implizite. Die Passage, »Kommt für Ihren neuen Roman {Georg} nicht einer der beiden holländischen Verlage {Allert de Lange, Querido-Verlag}, die Sie natürlich dem Namen nach kennen, in Frage. Der Leiter {Kesten, Hermann} des einen [...] sitzt in Paris. Ich kenne ihn, weiß aber im Augenblick nicht seinen Namen [...]. Sie kennen den Herrn {Kesten, Hermann} sicher, er war früher in Berlin«¹⁰, erweist sich für den mit dem Korpus Unvertrauten als kryptisch, erschließt sich jedoch über die hier in geschweiften Klammern ergänzten Registerverknüpfungen.

Jeder Registereintrag wird mit der zugehörigen Nummer der **Gemeinsamen Normdatei** (GND) versehen, sofern eine Nummer bereits vorhanden ist. Bei Orten werden die Geokoordinaten aufgenommen, sodass jene bei der späteren Visualisierung auf einer Karte angezeigt werden können. Zur Unterstützung dieses Arbeitsschritts wurden im Projekt zwei Werkzeuge entwickelt: Der sogenannte Geohamster durchforstet verschiedene Datenbestände nach Geo-Koordinaten sowie GND-Nummern, bietet weitere Orientierungshilfen zur Identifikation des gesuchten Orts und liefert die gewünschten Ergebnisse auf einen Blick.¹¹

¹⁰ Brief von Julius Meier-Graefe an Siegfried Kracauer vom 17.3.[1934], Deutsches Literaturarchiv Marbach, Nachlass Siegfried Kracauer.

¹¹ Durchsucht werden die Gemeinsame Normdatei (GND) der Deutschen Nationalbibliothek, die Online-Datenbanken **Geonames** und **Getty Thesaurus of Geographic Names** sowie für weiterführende

Ganz ähnlich arbeitet der sogenannte Normhamster, der für die Suche nach GND-Nummern zu Personen und Werken optimiert ist. Er bietet die zusätzliche Möglichkeit, die Suche auf einen bestimmten Zeitraum einzugrenzen. Die Anreicherung mit Normdaten dient einerseits der globalen Identifizierung der Registereinträge und eröffnet andererseits die Möglichkeit zur Einbindung von oder zur Vernetzung mit externen Ressourcen wie etwa dem entsprechenden Datensatz der Deutschen Nationalbibliothek, der häufig weiterführende Informationen enthält, oder mit lexikalischen Nachschlagewerken wie der [Neuen Deutschen Biographie](#). Die Registereinträge selbst werden nach einem einheitlichen Muster erläutert. Die strukturiert hinterlegten Informationen, etwa die Zugehörigkeit einer Person zu einer Berufsgruppe oder die geographische Lokalisierung einer Körperschaft, ermöglichen zusätzliche Formen der visuellen Datenaufbereitung wie etwa Gruppierungen.

Das Projekt geht jedoch einen Schritt weiter und bietet, zusätzlich zu den erwähnten Registern, eine thematische Erschließung der Briefe über einen auf das Themenfeld ›Exil‹ zugeschnittenen Thesaurus. Erarbeitet wird dieser Thesaurus anhand von Briefen aus der Zeit des Exils, die einem *close reading* unterzogen und deren Inhalte anschließend auf möglichst passende Schlagworte, wie etwa ›Heimat‹, ›Geldnot‹, ›Innenleben‹, ›Erfolg‹ oder ›Solidarität‹, gebracht werden. Repräsentiert ein Schlagwort auch die Inhalte weiterer Briefe, gilt dies als Indiz, dass es sich um ein für das Korpus einschlägiges Schlagwort handelt. Mit diesem induktiven Verfahren ist sukzessive ein Thementhesaurus aufgebaut worden, der inzwischen ca. 500 Schlagworte umfasst und zugleich offen für weitere Modifikationen und Ergänzungen ist. Dieser ›Exilthesaurus‹ bildet die Basis für die thematische Verschlagwortung der Briefe, die über ein teilautomatisches Verfahren erfolgt. Dazu wird zunächst die Sprache eines jeden Briefes ermittelt. Im Korpus vorhanden sind – hier äußern sich zugleich die Eigenheiten des Exils, was teils zu erzwungenen, teils zu freiwilligen Sprachwechseln führt – Briefe auf Deutsch, auf Französisch und auf Englisch. Im Anschluss daran wird jeder Brief in seine Wortformen zerlegt und mit einem Part-of-Speech-Tagger¹² bearbeitet, der – soweit er diese erkennt – für jede Wortform die Wortart und die lemmatisierte Form ermittelt. Aus dem so bearbeiteten Text werden nun alle Substantive extrahiert und in einem weiteren Schritt alle Lemmata entfernt,¹³ die bereits im Zuge des XML-Markups als Personen, Orte usw. ausgezeichnet sind. Schließlich wird für jedes Substantiv die inverse Dokumentfrequenz ermittelt, um dessen Spezifik innerhalb eines Briefes im Vergleich zum Gesamtkorpus zu ermitteln. Unter Ausnutzung bereits vorhandener Thesauri (z. B. [OpenThesaurus](#)) und Wortlisten (z. B. [Wortschatz des Deutschen](#)) und der darin abgelegten Synonyme wird für jedes extrahierte Stichwort eine Auswahlliste generiert, die die Zuordnung eines mutmaßlich passenden Schlagworts aus dem Thementhesaurus vorschlägt: Etwa die Verknüpfung des Stichworts ›Gefühl‹ mit dem Schlagwort ›Innenleben‹ oder die Verknüpfung des Stichworts ›Vaterland‹ mit dem Schlagwort ›Heimat‹. Es ergeben sich jedoch auch falsche oder missverständliche Zuweisungsvorschläge, etwa die angebotene Verbindung des Stichworts ›Prädikat‹ mit dem Schlagwort ›Zensur‹ oder des Stichworts ›Beitrag‹ – gemeint

Informationen die [deutschsprachige Wikipedia](#), die im Zuge dezentraler Wissensgenerierung eine gute erste Orientierung bieten kann (vgl. Steinsiek 2013, S. 231). Zu den Möglichkeiten technischer Unterstützung bei der Ermittlung von Normdaten siehe grundlegend Andert et al. 2014.

¹² Verwendet wird der [TreeTagger](#) (vgl. grundlegend Schmid 1999 und Schmid 1994).

¹³ Der Ausschluss bereits getaggtter Passagen kann nicht in einem früheren Schritt erfolgen, da POS-Tagger kontextsensitiv arbeiten.

ist eine Zeitschriftenveröffentlichung – mit den Schlagworten ›Unterstützung‹, ›Honorar‹. Daher schließt sich an die thematische Vorstrukturierung ein weiterer manueller bzw. intellektueller Arbeitsschritt an, in dem unter Berücksichtigung des jeweiligen Kontextes Zuweisungsvorschläge als sinnvoll akzeptiert oder als unbrauchbar abgelehnt werden.¹⁴ Im letzten Fall können eigene Zuweisungen zu einem Schlagwort des Thementhesaurus erfolgen, auch gegebenenfalls nötige, sich aus dem Stichwortbefund ergebende Erweiterungen des Exilthesaurus sind möglich. Durch die Verknüpfung von Briefen mit den Einträgen dieses Exilthesaurus entsteht ein mächtiges inhaltliches Erschließungsinstrument. Denn erst die thematische Anreicherung ermöglicht, dass bei der Suche – etwa nach dem Stichwort ›Diskretion‹ – nicht nur die rein syntaktischen Treffer einer bloßen Volltextsuche ausgeliefert werden, sondern auch Briefstellen wie »Inzwischen schon ist es uns aber von der allergrößten Wichtigkeit, daß Sie Direktor W. sofort sagen: er möchte über den Existenzplan Lilis gegenüber jedermann das strengste Stillschweigen bewahren.«¹⁵

3. Forschungsportal

Weder das eingangs genannte Zitat Kracauers noch die bloße Zeitspanne, die zwischen der Entstehung und der Veröffentlichung seines Romans *Georg* liegt, geben Aufschluss über die immensen Anstrengungen, mit denen sich Kracauer, aber auch die anderen Exilanten immer wieder alternative Publikations- und Verbreitungswege im Exil eröffnen müssen. Zudem kann erst im Vergleich mit weiteren Korrespondenzen beurteilt werden, ob Kracauers Briefe besonders häufig um das Thema ›Verlagssuche‹ kreisen und ob diesem somit eine besondere Bedeutung zukommt. Erst durch die inhaltliche Erschließung des Korrespondenznetzes – zu dem Empfehlungsschreiben an Zeitungsredakteure und Verleger genauso gehören wie schriftliche Absprachen zur Gründung eigener Verlage – lässt sich überblicken, wie sich ein Thema zeitlich und räumlich ausbreitet und welche spezifische Bedeutung es für die im Korpus angelegte Epoche des Exils hat.

Kracauer informiert Meier-Graefe nicht nur über Thomas Manns Zusage, bei der Verlagssuche behilflich zu sein, er schickt ihm auch eine Kopie seines Schreibens an Mann. Inzwischen denkt Meier-Graefe an der französischen Riviera über die Verlage der ihm bekannten und ebenfalls emigrierten Schriftsteller im Nachbarort nach und zieht als weitere Möglichkeit Verlage in Betracht, die wiederum im Briefwechsel zwischen den Autoren Stefan Zweig und René Schickele genannt werden. Jenseits der stets mitzudenkenden Einschränkung, dass sich die Überlieferungslage insbesondere bei Nachlässen Exilierter als schwierig erweist, zeigt sich am Beispiel von Kracauers Verlagssuche, welchen Vorteil die Darstellung der zu erschließenden Korrespondenz in Form einer Netzstruktur mit sich bringt. Was sich hier als interpretative Aussage nur mit einem isolierten Blick auf die Korrespondenzen eines oder zwischen zwei Briefpartnern für das Thema ›Verlagssuche‹ formulieren lässt, erhält durch die

¹⁴ Ist für die Stichwort-Schlagwort-Zuordnung der Kontext des jeweiligen Briefes notwendig, kann dieser über eine beim Stichwort hinterlegte Verknüpfung aufgerufen werden.

¹⁵ Brief von Siegfried Kracauer an Max Tau vom 30.01.1935, Deutsches Literaturarchiv Marbach, Nachlass Siegfried Kracauer.

im Projekt vorgesehene zusätzliche Auswertung eines ganzen Korrespondenznetzes eine neue Größenordnung.

Der Ort, an dem dieses Korrespondenznetz zumindest virtuell sichtbar wird, ist das im Rahmen des Projektes entwickelte Forschungsportal. Es bietet dem Forscher verschiedene Möglichkeiten, um sich dem bereitgestellten Korpus zu nähern – vom eher entdeckenden Einstieg durch Visualisierungen, über gezielte Suchanfragen, bis hin zur Kombination aus beidem. So lässt sich die Frage, wie häufig ein bestimmter Registerbegriff – etwa der Kracauer-Titel *Georg* – vorkommt, zusätzlich zeitlich und räumlich eingeschränkt, mit den Suchinstrumenten des Forschungsportals in Zahlen und unterschiedlichen Darstellungen beantworten und präsentieren. Zugleich kann umgekehrt die informationstechnologische Texterschließung helfen, Forschungsfragen aufzuwerfen, die im spielerischen Umgang mit verschiedenen Such- und Visualisierungsinstrumenten entstehen und sich in der klassischen Texterschließung noch nicht überblicken lassen. Wesentliche Elemente des Forschungsportals sind daher die kaskadierende Suche, ein dynamischer Zeitregler, Netzwerkvisualisierungen und geographische Visualisierungen, Listen, Diagramme und Matrizen sowie die in edierter, teilweise auch faksimilierter Form vorliegenden Briefe selbst.

Zur Vermeidung unnötiger Hürden durch Installations- und Wartungsaufwand wird das Forschungsportal in Form einer Webanwendung implementiert, für deren Nutzung lediglich ein moderner Webbrowser benötigt wird. Dieser Browser (Client) kommuniziert über einen Webserver, der ihm die Daten zur Darstellung des Korrespondenznetzes übermittelt. Weiterhin übernimmt der Server rechenintensive Aufgaben, wie z. B. das Auffinden sämtlicher Beziehungen zwischen einer bestimmten Person und einem bestimmten Ort anhand der Briefmetadaten.

Der größte Teil der Anwendung wird jedoch clientseitig ausgeführt, d. h. im Browser des Nutzers. Bei einer klassischen Website erhält der Browser eine in HTML beschriebene Seite vom Server und stellt diese dem Anwender in graphischer Form dar. Eine Aktion des Nutzers durch Klick auf einen Link oder durch das Absenden eines Formulars führt zum Laden einer neuen Seite. Für das Erkunden des Datenbestands aus verschiedenen Blickwinkeln – durch den Wechsel zwischen verschiedenen Darstellungsformen, dem Färben, Gruppieren und Filtern der dargestellten Daten – ist jedoch eine sofortige visuelle Rückmeldung auf die angewandte Aktion erforderlich. Hierzu werden alle mit dem Datenbestand verknüpften Elemente umgehend aktualisiert, wie z. B. die aktuelle graphische Darstellung (Korrespondenzgraph, Karte, Briefliste etc.), die Anzeige statistischer Informationen oder die Legende. Um dies zu ermöglichen, werden sämtliche auf der Seite angezeigten Komponenten anhand der vom Server übermittelten Daten durch den Client erzeugt und bei Aktionen des Nutzers aktualisiert. Hierbei wird die Seite über das DOM, eine Schnittstelle für den Zugriff auf und die Bearbeitung von Dokumenten, verändert. Da die Bearbeitung des DOMs zeitaufwendig ist, versucht man Änderungen am DOM möglichst gering zu halten. Als für die Umsetzung des Portals geeignet hat sich das JavaScript-Framework *Mithril* erwiesen, welches mit einem virtuellen DOM arbeitet und zum Aktualisieren der Seite nur die Unterschiede zwischen dem virtuellen DOM und dem eigentlichen DOM auf letzteres überträgt. Durch diese Technik entfällt größtenteils die Trennung zwischen dem Code, welcher zur erstmaligen Darstellung der Seite dient, sowie dem

Code zur Aktualisierung der Seite, da in beiden Fällen das virtuelle DOM neu aufgebaut wird. Eine Kommunikation über das Netzwerk erfolgt somit nur in den seltenen Fällen, in denen weitere umfangreiche Daten, wie z. B. der vollständige Text zu einem Brief, benötigt werden. Solche Daten werden bei Bedarf im Hintergrund nachgeladen.¹⁶

Für die Verarbeitung und Darstellung der Briefe im Forschungsportal erweist sich die Auszeichnung in TEI-konformem XML aufgrund der Menge an verfügbaren Werkzeugen zur Weiterverarbeitung der Daten und der einfachen Konvertierung in HTML als vorteilhaft. Bestehende Werkzeuge ermöglichen u. a. das Extrahieren von Daten aus XML mit der Anfragesprache XPath, das Ändern von Tags und Attributen sowie das Entfernen sämtlicher Tags aus einem Dokument, d. h. die Umwandlung in reinen Text. Letztere Umwandlung dient dabei einer Volltextsuche aller im Portal verfügbaren Briefe.

Die Darstellung der Briefe selbst erfolgt durch eine projektspezifische Transformation von XML in HTML. Die graphische Darstellung der Textformatierungen wird in CSS angegeben, sodass Formatierungen wie Unterstreichungen, Hochstellungen, zentrierter, rechts- und linksbündiger Text im transkribierten Brieftext zu sehen sind. Weiterhin können mit Registereinträgen verknüpfte Textstellen hervorgehoben werden, sofern dies vom Nutzer gewünscht wird.

Während sich Brieftexte, Listen und Tabellen sehr gut in HTML und CSS beschreiben lassen, gibt es für andere Darstellungsformen, wie Graphen oder Karten mit SVG und Canvas zwei geeignetere Möglichkeiten: Bei Canvas handelt es sich um ein HTML-Element, welches eine Zeichenfläche für Rastergraphiken zur Verfügung stellt; bei SVG hingegen um ein Format für Vektorgraphiken, welche in HTML-Seiten eingebettet werden können. Da SVG ebenso wie die aktuellen HTML-Standards in XML spezifiziert ist, lassen sich Elemente im SVG mittels JavaScript über die Schnittstelle des DOMs genauso bearbeiten wie HTML-Elemente. Der Nachteil ist jedoch die geringere Performanz gegenüber Canvas. Änderungen am DOM sind verhältnismäßig zeitaufwendig, insbesondere bei vielen Elementen auf der Seite.

Realisiert werden die Netzwerkvisualisierungen und Diagramme mit **D3**, einer JavaScript-Bibliothek zum Binden von Daten an das DOM. Zusätzlich erleichtert diese Bibliothek die Implementierung von Funktionen wie Zeichnen und Beschriften von Achsen, Skalierung oder Zoomen, die in Graphiken oft vorkommen. Bisher sind zwar alle mit D3 erzeugten graphischen Darstellungen des Portals SVG-Graphiken, jedoch ist D3 nicht auf SVG beschränkt. Sollten sich im Laufe der Entwicklung Probleme hinsichtlich der Performanz bei bestimmten Visualisierungen ergeben,¹⁷ so bleibt die Möglichkeit offen, die betroffene Darstellung oder

¹⁶ Dies geschieht mittels AJAX. Ein Verfahren, bei dem ein sogenannter XML-HTTP-Request an den Webserver gesendet wird und erhaltene Daten (in unserem Fall im JSON-Format) mittels JavaScript in die Seite eingebaut werden.

¹⁷ Wünschenswert ist eine Bildwiederholfrquenz von 60Hz bei der Interaktion mit Visualisierungen, wie z. B. beim Zoomen oder Verschieben. Aufgrund komplexer Berechnungen sowie durch viele DOM-Elemente, wie bei großen Graphen in SVG der Fall, kann die erreichte Bildwiederholfrquenz jedoch deutlich niedriger sein, wodurch sich Probleme bei der Performanz ergeben. Zwar ist die Arbeit mit dem Portal auch bei einer Frequenz von 20Hz–30Hz noch möglich, jedoch mit wahrnehmbaren ruckartigen Aktualisierungen der Graphiken.

auch nur einzelne Teile davon auf einem Canvas-Element zu zeichnen, gegebenenfalls auch hardwarebeschleunigt.¹⁸

Um dem Anwender zunächst einen Überblick über das Netzwerk zu verschaffen, folgt die Auswahl der zu visualisierenden Daten einem subtraktiven Verfahren: Ausgehend vom Gesamtnetzwerk kann durch das Anwenden von Filtern die Auswahl nach und nach auf ein immer kleineres Teilnetz eingeschränkt werden (Abbildung 3). Hierbei ergibt sich jedoch das Problem der Darstellung großer Graphen für den Einstieg in das Portal. Da sich diese hinsichtlich ihrer Struktur nur in ungenügendem Maße darstellen lassen, wird meist das Anwenden von Filtern oder einer hierarchischen Gliederung vorgeschlagen.¹⁹ Für die Darstellung des im Projekt erarbeiteten Exilbriefnetzes kommen derzeit, in Abhängigkeit von der jeweiligen Darstellungsform, folgende Methoden zum Einsatz:

- Filterfunktionen zum Fokussieren bestimmter Daten, z. B. durch Suche
- Gruppierungen z. B. nach Berufen
- verschiedene Detailstufen, von makroskopischer zu mikroskopischer Ansicht

Während Filterfunktionen es ermöglichen, eine Auswahl des Netzes zu betrachten (z. B. nur Briefe einer Person), kann durch Hinzunahme einer Gruppierung mit den bestehenden Daten und derselben Darstellungsart ein anderes Muster erkennbar werden. Bei Darstellungsformen, die ein Heranzoomen erlauben, wie Graphen oder Karten, werden verschiedene Detailstufen genutzt. Hierbei werden, je nach Tiefe der Zoomstufe, Informationen ein- oder ausgeblendet, sodass sowohl mikro- als auch makroskopische Betrachtungen der visualisierten Daten möglich werden. Diesem Ansatz folgt auch der Zeitregler. Hier wird zunächst der gesamte Zeitraum der vorhandenen Briefe jahresgenau angezeigt, mit der Möglichkeit, Daten auf einen Bereich einzuschränken. Für eine zeitlich genauere Einschränkung besteht die Möglichkeit, in einen Bereich des Zeitreglers zu zoomen, wobei die Beschriftung des Zeitraumes je nach Zoomstufe monats- oder tagesgenau angezeigt wird. Ein neuer Zeitraum lässt sich damit exakter auswählen bzw. ein bestehender noch verfeinern.

¹⁸ In diesem Fall durch Nutzung der WebGL-Schnittstelle.

¹⁹ Vgl. Koutsofios et al. 1999, S. 457–461.

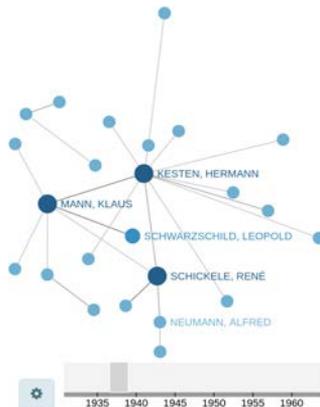


Abb. 3: Visualisierung eines Korrespondenznetzes für den Zeitraum 2.8.1936 bis 29.12.1938. © Projekt *Vernetzte Korrespondenzen*.

Einem komplexen Thema wie der Verlagssuche für Kracauers Roman *Georg* kann sich der Nutzer auf verschiedene Weise nähern. So lässt sich die Suche durch die Anwendung unterschiedlicher Filter auf die ausgezeichneten Registerbegriffe Personen, Orte, Werke, Körperschaften usw. einschränken und durch kombinierte Suchen etwa herausfinden, welche Verlage im Kontext der Veröffentlichung im Gespräch sind. In der Suchmaske kann darüber hinaus nach den verschiedenen Rollen ›Schreiber‹, ›Empfänger‹ oder ›erwähnte Person‹ differenziert werden. Auf diese Weise lässt sich anzeigen, in welchen Briefen Thomas Mann Kracauer nennt und in welchen er von ihm genannt wird. In diesem Kontext ist es auch denkbar, die Zugehörigkeit zu bestimmten Personengruppen, etwa Berufsgruppen abzufragen und beispielsweise direkt nach Verlegern zu filtern. Entsprechend können nicht nur der Schreibort eines oder mehrerer Briefe oder die erwähnten Orte ausgegeben werden, sondern es lässt sich auch nach Verlagssitzen suchen.

In Kombination mit dem dynamischen Zeitregler bedeutet das etwa für das Thema der Verlagssuche, dass sich die Verbreitung der Exilverlage und die durch die Verfolgung bedingte räumliche Verschiebung der Schreiborte graphisch darstellen und vor dem historischen Hintergrund der politischen Entwicklungen nachvollziehen lassen. Außerdem können Suchanfragen mit unterschiedlichen Parametern auch in einer Parallelansicht visualisiert werden. Im unmittelbaren Vergleich offenbart sich schnell, dass Kracauers Roman trotz seiner Versicherung, eine französische Ausgabe stehe schon in Aussicht,²⁰ bei Gallimard seinerzeit nicht mehr erscheinen kann und dass auch derselbe Verlag den dritten Band von Thomas Manns *Joseph-Tetralogie* erst 1948 publiziert. Dies deutet zugleich auf die Zensurpolitik der deutschen Besatzer hin. Das angezeigte Ergebnis lässt sich wiederum in Bezug setzen zu den Verlagen, die im Kontext weiterer Werke, z. B. Kracauers *Offenbach-Biographie*, genannt werden.

²⁰ Brief von Siegfried Kracauer an Julius Meier-Graefe vom 16.3.1934, Deutsches Literaturarchiv Marbach, Nachlass Siegfried Kracauer.

Auf Grundlage der kaskadierenden Suche ist immer die Möglichkeit gegeben, nicht nur zwischen verschiedenen Filtern hin und her zu wechseln, sondern die Anwendung der Filter in ihrer Abfolge nachzuvollziehen. Zusätzlich erleichtern in das Portal eingebundene, gängige Tools wie Breadcrumbs dem Nutzer die Navigation.

Vorstellbar sind letztlich alle Suchkombinationen, die sich aus den ausgezeichneten Registerbegriffen und den briefspezifischen Daten ergeben. Die Suchergebnisse lassen sich in Form strukturierter Listen oder in den Erkenntnisprozess stärker visuell unterstützenden Darstellungsformen wie Matrizen, Diagrammen, Netzwerkvisualisierungen oder Karten ausgeben und nach entsprechender Auswahl bis zum einzelnen Brief zurückverfolgen. Weiterhin werden nutzerspezifische Einstellungen, wie die gewählte Darstellungsart, angewandte Suchfilter, Zeitfilter, Zoomstufe etc., nach Möglichkeit in der URL gespeichert, um sowohl einen Wiedereinstieg in das Portal als auch eine Zusammenarbeit durch Weitergabe der aktuellen Ansicht zu gestatten. Der Anwender kann diese URL als Lesezeichen im Browser hinterlegen oder an eine andere Person z. B. via E-Mail weitergeben. Durch den Aufruf einer solchen URL wird die damit verbundene, zuvor betrachtete Ansicht des Netzwerks wiederhergestellt. Eine integrierte Funktion für den graphischen Export ermöglicht überdies die Dokumentation von Arbeitsergebnissen sowie die weitere Nutzung außerhalb des Portals.

4. Fazit und Ausblick

Die hier primär zu illustrativen Zwecken verfolgten Fragen nach den Ereignissen und Dynamiken, die im Zuge von Kracauers Verlagssuche virulent werden, lassen erimmen, welches Potential ein Forschungsportal birgt, das den Brief nicht als isoliertes Dokument, sondern zuallererst in seiner netzartigen Struktur präsentiert. Erst durch das Zusammenführen von Korrespondenzen und deren inhaltlicher Auszeichnung und graphischer Vernetzung werden Zusammenhänge sichtbar, die dem Nutzer herkömmlicher Briefeditionen – ob gedruckt oder digital – verborgen bleiben.

Der sich hieraus ergebende Wunsch, auch die verborgenen Netze anderer Korpora sichtbar und damit erforschbar zu machen, ist verständlich und – entsprechendes inhaltliches Markup vorausgesetzt – dank der generischen Konzeption der Portals auch möglich.

Zugleich enthält die Idee des Netzwerks ein Moment des Nie-Abgeschlossenen- der beinahe unendlichen Erweiterbarkeit, weshalb auch dem Projekt ein gewissermaßen vorläufiger, offener Charakter eigen ist. Abschließend folgt daher ein kurzer Ausblick: Naheliegender wäre zunächst die Ausdehnung des inhaltlichen Markups auf die Erläuterungen der Registereinträge und die Stellenkommentare der Briefe, denn auch diese bergen Informationen, die für weitere Gruppierungen und Vernetzungen der Briefe genutzt werden könnten. Zu erwägen wäre auch die Integration zusätzlicher Korpora, die der Zeit des Exils vorgängig oder nachfolgend sind und über gemeinsame Verbindungsknoten – ob personeller oder thematischer Art – das Korrespondenznetzwerk über Epochengrenzen hinweg aufspannen. Über weitere externe Vernetzungen könnten die Briefe schließlich in Kontexte eingeordnet werden, die ihnen selbst entspringen, jedoch weit über sie hinausweisen.

Bibliographische Angaben

Martin Andert / Jörg Ritter / Paul Molitor: Optimized platform for capturing metadata of historical correspondences. In: *Literary and Linguistic Computing* 6 (2014), S. 1–10. DOI [10.1093/lc/fqu027](https://doi.org/10.1093/lc/fqu027).

Burcu Dogramaci / Karin Wimmer: Vorwort. In: *Netzwerke des Exils*. Hg. von Burcu Dogramaci / Karin Wimmer. Berlin 2011, S. 9–11. [[Nachweis im GBV](#)]

Johannes Evelein: Briefkultur und Exil. In: *Erste Briefe aus dem Exil 1945–1950*. Hg. von Primus-Heinz Kucher / Johannes F. Evelein / Helga Schreckenberger. München 2011, S. 15–31. [[Nachweis im GBV](#)]

Eleftherios E. Koutsofios / Stephen C. North / Russell Truscott / Daniel A. Keim: Visualizing Large-Scale Telecommunication Networks and Services. In: *VIS '99 Proceedings of the Conference on Visualization '99*. Los Alamitos 1999, S. 457–461. [[Nachweis im GBV](#)]

Siegfried Kracauer: *Schriften*. Band 7. Hg. von Karsten Witte. Frankfurt/Main 1973. [[Nachweis im OPAC](#)]

Primus Kucher / Johannes F. Evelein / Helga Schreckenberger: Einleitung. In: *Erste Briefe aus dem Exil 1945–1950*. Hg. von Primus-Heinz Kucher / Johannes F. Evelein / Helga Schreckenberger. München 2011, S. 9–14. [[Nachweis im GBV](#)]

Helmut Schmid: Improvements in Part-of-Speech Tagging with an Application to German. In: *Natural Language Processing Using Very Large Corpora*. Hg. von Susan Armstrong / Kenneth Church / Pierre Isabelle / Sandra Manzi / Evelyne Tzoukermann / David Yarowsky. Dordrecht 1999, S. 13–26. [[Nachweis im GBV](#)]

Helmut Schmid: Probabilistic Part-of-Speech Tagging Using Decision Trees. In: *Proceedings of International Conference on New Methods in Language Processing*. Manchester 1994, S. 44–49. [[online](#)]

Angela Steinsiek: Alles Wikipedia? Kommentieren heute am Beispiel der Jean Paul-Brief-Edition. In: *Brief-Edition im digitalen Zeitalter*. Hg. von Anne Bohnenkamp / Elke Richter. Berlin, Boston 2013, S. 229–235. [[Nachweis im OPAC](#)]

Frank Wende: Briefe aus dem Exil. 1933–1945. In: *Der Brief. Eine Kulturgeschichte der schriftlichen Kommunikation*. Hg. von Klaus Beyrer / Hans-Christian Täubrich. Frankfurt/Main 1996, S. 172–183. [[Nachweis im OPAC](#)]

Abbildungslegenden und -nachweise

Abb. 1: Auf der DHD 2014 in Passau präsentiertes und mit dem DHD-Poster-Award prämiertes Poster des Projekts *Vernetzte Korrespondenzen*. Layout & Design: Michael Lambert, Trier Center for Digital Humanities.

Abb. 2: Faksimile eines Briefes von Julius Meier-Graefe an Siegfried Kracauer vom 16.11.1934. © Deutsches Literaturarchiv Marbach.

Abb. 3: Visualisierung eines Korrespondenznetzes für den Zeitraum 2.8.1936 bis 29.12.1938. © Projekt *Vernetzte Korrespondenzen*.

Zeitschrift für digitale Geisteswissenschaften

Artikel aus:

Sonderband 1 der ZfdG: Grenzen und Möglichkeiten der Digital Humanities. Hg. von Constanze Baum und Thomas Stäcker. 2015. DOI: [10.17175/sb01](https://doi.org/10.17175/sb01)

Titel:

What People Said: The Theoretical Foundations of a Minimal Doxographical Ontology and Its Use in the History of Philosophy

Autor/in:

Stefan Heßbrüggen-Walter

Kontakt:

early.modern.thought.online@gmail.com

Institution:

National Research University Higher School of Economics, Moskau

GND:

[141879335](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63862-p0011-9)

ORCID:

[0000-0002-9383-026X](https://orcid.org/0000-0002-9383-026X)

DOI des Artikels:

[10.17175/sb001_001](https://doi.org/10.17175/sb001_001)

Nachweis im OPAC der Herzog August Bibliothek:

[826998763](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63862-p0011-9)

Erstveröffentlichung:

19.02.2015

Lizenz:

Sofern nicht anders angegeben 

Medienlizenzen:

Medienrechte liegen bei den Autoren

Letzte Überprüfung aller Verweise:

24.05.2016

GND-Verschlagwortung:

[Wissensrepräsentation](#) | [Philosophie](#) | [Ontologie](#) |

Zitierweise:

Stefan Heßbrüggen-Walter: What People Said: The Theoretical Foundations of a Minimal Doxographical Ontology and Its Use in the History of Philosophy. In: Grenzen und Möglichkeiten der Digital Humanities. Hg. von Constanze Baum / Thomas Stäcker. 2015 (= Sonderband der Zeitschrift für digitale Geisteswissenschaften, 1). PDF Format ohne Paginierung. Als text/html abrufbar unter DOI: [10.17175/sb001_001](https://doi.org/10.17175/sb001_001).

Stefan Heißbrüggen-Walter

What People Said: The Theoretical Foundations of a Minimal Doxographical Ontology and Its Use in the History of Philosophy

Abstracts

Üblicherweise beruht das Design von Ontologien auf der Annahme, dass die Bedeutung einer Proposition sich aus der Bedeutung ihrer Elemente (Begriffe) und ihrer syntaktischen Struktur ergibt. Die Reichweite dieses ›Prinzips der Kompositionalität‹ ist jedoch innerhalb der Semantik strittig. Die Gegner des Prinzips verteidigen den Primat der Satzbedeutung und leiten die Bedeutung von Begriffen aus ihrem Beitrag zur Satzbedeutung ab. Angesichts dieses Sachstandes argumentiert der Aufsatz zugunsten eines Zugangs zum Design von Ontologien, der keine Stellungnahme in dieser Debatte voraussetzt. Die hier vorgeschlagene ›minimale doxographische Ontologie‹ dient als heuristisches Werkzeug zur Erfassung unbekannter oder komplexer Gegenstandsbereiche. In ihr werden Satzbedeutungen als unanalysierbar angesehen und auf einen Träger des propositionalen Inhalts (Personen oder Texte) bezogen. Die Stärken eines solchen Ansatzes werden zunächst anhand eines vereinfachten Beispiels erörtert, einer Analyse von juristischen Begriffsdefinitionen alkoholischer Getränke. Ein komplexerer Anwendungsfall betrifft die doxographische Analyse von Debatten in der Geschichte der frühneuzeitlichen Philosophie. Schließlich erörtert der Aufsatz kurz, wie ein solcher Ansatz erweitert werden kann, indem Ontologien als hermeneutische Werkzeuge zur Deutung von Quellen der Philosophiegeschichte verwendet werden.

Traditionally, ontology engineering is based on the presumption that the meaning of a proposition results from the combination of the meaning of its elements (concepts) and its syntactical structure. The reach of this ›principle of compositionality‹ is, however, a contested topic in semantics. Its opponents defend the primacy of propositional meaning and derive the meaning of concepts from their contribution to propositional meaning. In this situation, this paper argues for an approach to ontology design that does not presuppose a stance in this debate. The proposed ›minimal doxographical ontology‹ is intended as a heuristic tool charting unknown or complex domains. It regards propositional meaning as atomic and relates it to a bearer of propositional content (persons or texts). The strengths of such an approach are first discussed in a simplified example, the analysis of legal stipulations on alcoholic beverages. A more complex use case concerns the doxographical analysis of debates in the history of early modern philosophy. In closing, the paper sketches briefly how this approach may be extended using ontologies as hermeneutic tools in the interpretation of sources from the history of philosophy.

1. Introduction

In recent years, technologies of knowledge representation that are usually subsumed under the heading of the ›semantic web‹ have been used within the digital humanities in disciplines as diverse as literary studies (e. g. regarding [the ontology of fictional characters](#)), philosophy ([the Wittgenstein ontology](#)), or history ([LODE](#), an ontology for the description of historical events). Thinking about the ›semantic web‹ comes most naturally to digital humanists approaching the discipline from what could be called a ›cultural heritage‹ angle, e. g. librarians, archivists, or curators. In these areas, the production of meaningful metadata is part of everyday workflows; the transition from cataloguing guidelines to machine-readable metadata

standards to semantic web languages like RDF comes quite naturally and is an important step in fighting the ›siloesation‹ of digital collections by embedding them in a web of ›linked open data‹.¹

Whether ontologies are an important part of the tool set of the digital humanist is still a disputed question.² Some of the more radical defenders of ontologies are unperturbed by this criticism. They contend that ontologies are capable of not just representing or modelling knowledge, but that they capture features of mind-independent reality.³ This may even be true for artefacts.⁴

Whether semantic web technologies are a good fit for a given use case is, however, first of all a technological problem that probably should not be solved on purely philosophical grounds. Nevertheless, philosophy may be able to contribute to some foundational debates in the digital humanities, if its function is not taken to consist in the provision of foundations, but of ›maieutic impulses‹ that help to explicate hidden presuppositions and stimulate to rethink unacknowledged biases and blind spots.

In this spirit, this paper discusses one such unacknowledged presupposition of ontology design. Both the knowledge to be modelled in an ontology and the ontology itself are necessarily articulated in propositional form. Correspondingly, the formal structure of ontology languages (like the Web Ontology Language OWL) are consciously modelled on central premises of formal semantics, first and foremost the ›principle of compositionality‹. This ancestry, however, may not be as innocuous as it seems, because philosophers of language discuss controversially whether the impact of the principle of compositionality is limited by a second principle, the ›context principle‹. Thus, those interested in the capabilities of ontologies for modelling knowledge must first clarify the possible impact of these debates on ontology design.

The most appropriate strategy in such uncharted territory is ›risk avoidance‹. Accordingly, it may be possible to use the tools of the semantic web in an unassuming and modest manner, as a heuristic tool for mapping vague, complicated, or partially unknown domains. I will first discuss the *raison d'être* for such a modest approach using a somewhat contrived example and show some problems we encounter in trying to extend the well-known wine ontology. It has already been shown how the modelling of certain domains can profit from analytical restraint, namely if we desist from analysing propositional content into component terms and bind this content to the existence of concrete spatiotemporal entities as their ›bearers‹. The fruitfulness of such a minimal ontology of discourses depends, however, on use cases in ›real life‹. Hence, the proposed ›ontology of what people said‹ is applied to a ›doxographical map‹ of a spatially and temporally circumscribed discourse in the history of philosophy, the debate about the proper definition of the term ›philosophy‹ in early modern Iberian philosophy.

¹ Cf. Kemmann 2014, section »Embracing Technology«.

² Cf. Kohle 2013, p. 26f. with further references.

³ Cf. Smith 2004, *passim*.

⁴ Cf. Jansen 2013, *passim*.

The history of philosophy is, of course, more than just doxography. It should ideally be complemented by interpretations of what people said. In my conclusion, I will sketch how we may use the resources of semantic web technologies to describe the conceptual hierarchies that are implicated by what philosophers (and, possibly, others) have to say. However, it is important to keep in mind that in the light of the foundational discussions of the first part of this paper, such ontologies will always be interpretations, leaving room for controversy and dissent that is probably inevitable if we try to capture the meaning of a text. This is true regardless of the medium we use to express our findings.

2. Ontologies, Compositionality, Contextuality

In a first approximation, ontologies can be defined as »explicit formal specifications of the terms in the domain and relations among them«. ⁵ Such a specification determines »a common vocabulary for researchers who need to share information in a domain. It includes machine-interpretable definitions of basic concepts in the domain and relations among them«. ⁶ Moreover, such a specification is supposed to be indispensable for »analyzing domain knowledge«. ⁷ OWL knows two sorts of concepts: classes that »provide an abstraction mechanism for grouping resources with similar characteristics« ⁸ and that are defined in so-called »class axioms«, and properties which are defined in so-called »properties axioms«. ⁹ The third category of statements to be found in an OWL document concerns facts about individuals. ¹⁰ All statements are composed out of classes, properties, and constants as building blocks. These building blocks must have been defined beforehand: their intension must be known, before statements can be constructed.

This means that ontology engineering is firmly rooted in a theory of meaning based on Frege's »principle of compositionality«: the »[...] meaning of a complex expression is determined by its structure and the meanings of its constituents«. ¹¹ This nexus raises interesting questions. Those who believe that ontologies may be capable of modelling knowledge in a given domain without being committed to the stronger view that they capture features of mind-independent reality may be content to limit the scope of compositionality to a given language. ¹² Or they could maintain that it only applies to the formal language which is used for articulating the model of a domain, because artificial languages can be construed in such a way as to exhibit compositionality as a feature. Those who subscribe to a more realist interpretation of concepts in an ontology might probably have to accept the stronger thesis of »cross-linguistic compositionality«: »For every complex expression *e* in *L*, the meaning of *e* in *L* is functionally

⁵ Noy / McGuinness 2001, p. 1.

⁶ Noy / McGuinness 2001, p. 1.

⁷ Cf. Noy / McGuinness 2001, p. 2.

⁸ Dean / Schreiber 2004, section 3.

⁹ Cf. Dean / Schreiber 2004, section 3.2.3.

¹⁰ Cf. Dean / Schreiber 2004, section 4.

¹¹ Szabó 2013, section 1.

¹² In this case, the definition of compositionality is formulated relative to a given language. Cf. Szabó 2013, section 1.1: »For every complex expression *e* in *L*, the meaning of *e* in *L* is determined by the structure of *e* in *L* and the meanings of the constituents of *e* in *L*.«

determined through a single function for all possible human languages by the structure of e in L and the meanings of the constituents of e in L .¹³

But even if people may have reasoned disagreements about the scope of compositionality, the ›compositionalist‹ bias is apparently built into the very notion of an ontology as a ›common vocabulary for researchers‹. However, this first Fregean principle conflicts with a second also discussed in relation to his philosophy of language, the ›principle of contextuality‹ (or ›context principle‹): »The meaning of an expression is determined by the meanings of all complex expressions in which it occurs as a constituent.«¹⁴ So whereas compositionalists hold that the meaning of a proposition is the sum total of its parts with the semantic contribution of the structure of the proposition, contextualists presume that propositional meaning comes first and that the meaning of the constituents of a proposition depends on their role in all other propositions in which they are contained.¹⁵ But, again, the scope of this priority claim must be determined precisely. In the context of this paper, it is helpful to follow Stainton and to distinguish three different understandings of the priority expressed in context principles, namely methodological, metasemantical, and ›psychological‹ interpretations of this priority of propositional over conceptual meaning.¹⁶

In a *methodological* perspective, we assume that an analysis of the meaning of subsentential expressions must take into account the context of the proposition they appear in. This understanding of propositional priority may even be compatible with compositionality, because we can understand how the meaning of a subsentential expression appearing in a proposition that we understand can be isolated and transferred into new contexts, allowing us to express a proposition that we had not yet understood. In other words, we may need both compositionality and contextuality of meaning in order to explain linguistic creativity, the capability of expressing new thoughts by recombining elements which we already understand.¹⁷

But this does not mean that we are necessarily committed to the stronger *metasemantic* thesis that propositional meaning is in some substantial sense the only (or only the most relevant) source for the meaning of subsentential expressions. If this stronger thesis were applicable to the methodology of ontology design, the project as such might well be hopeless, because the recombination of terms could always lead to mutations in meaning that are unforeseeable for the designer.

The *psychological* thesis states that competent speakers of a natural language cannot grasp the meaning of subsentential expressions in isolation. Proponents of this view are probably sceptical with regard to the attempt to represent online resources by applying subsentential expressions to them: for them, tagging as such cannot be a meaningful linguistic activity.

¹³ Szabó 2013, section 1.4.

¹⁴ Szabó 2013, section 1.6.4.

¹⁵ Cf. Szabó 2013, section 1.6.4: »Compositionality is about bottom-up meaning-determination, while the context principle about top-down meaning-determination.«

¹⁶ Cf. Stainton 2006, p. 109f. Helpful discussions of the role of both principles in Frege can be found in Janssen 2001, *passim*. Lemanski 2013, *passim*, shows convincingly that the context principle is an essential element of (neo-)Aristotelian logic.

¹⁷ Cf. Stainton 2006, p. 111.

Even though such radical criticism may be misplaced, we should keep in mind that the vision of the semantic web is built around the notion of knowledge. And – difficult philosophical problems with non-propositional forms of knowledge like knowing-how or foundational perceptual beliefs notwithstanding – an ontology can only codify knowledge that can be explicated in propositional form. So even if we do not subscribe to the strong psychological thesis that subsentential expressions as such are basically meaningless, we could still accept the methodological guideline that ontology design is concerned with knowledge that *can* be expressed or explicated in propositional form. Implicit awareness of the meaning of subsentential expressions thus should always be explicated in propositional form, regardless of whether competent speakers can use or understand such expressions in isolation.¹⁸

So even those ontology designers who would subscribe to compositionality as an essential constituent of their self-understanding still face interesting problems: should we presume that ontologies mirror cross-linguistic universals or is their usefulness limited to speakers within a given linguistic community? Do ontologies mostly track extensions, i. e. the reference of terms, or should we give them an intensional interpretation as well, taking into account their meaning? Do we accept the notion that statements in an ontology are fully devoid of context, so that their meaning really consists of nothing but the sum total of subsentential meanings and the contribution made by syntax?

In thinking about these questions we should never lose track of the fact that ontologies are no end in themselves: they are technological instruments, so that their scope and utility is determined first and foremost by pragmatic considerations. It is therefore imprudent to assume that in order to build an ontology it is necessary to choose one side in these complex and unresolved philosophical debates. We should rather ask ourselves to what extent our understanding of ontology design is determined by unacknowledged biases in our implicit theories of meaning and whether it is possible to build ontologies in a way that is not committed to any explicit stance.

Such a minimal understanding, at least on the heuristic level, of coming to terms with a given domain would consist in two decisive moves:

1. The meaning of propositions (i. e. their ›propositional content‹) is taken to be opaque, it is only referred to by a name. This allows us to avoid any commitment whether or not in a particular case the meaning of a proposition can in fact be analysed compositionally.
2. Propositional contents are only allowed, if they can be connected to a spatio-temporal entity (mostly a person or a document) that articulates a propositional attitude towards this content, i. e. asserts, denies, or reflects upon the content in question. In other words, the propositional content designated by the name ›wine is made of grapes‹ is not to be analysed into a subject term designating a drink, an object term designating fruit, and a relation term designating the process of turning fruit into a drink. And it is

¹⁸ Shirky has not helped the debate along by presenting ontologies and tagging as alternative, but equivalent modes of representing online resources. On the most fundamental level, both technologies solve different problems. Tags are not statements; therefore, they do not encode knowledge. It may be granted that tags are an efficient means for crowd sourced classification and categorisation. But these are not the only – and probably not the most important – use cases for ontologies. Cf. Shirky, *passim*.

not to be admitted, unless we can trace this content to a person or document that either asserts, denies, or reflects on the propositional content ›wine is made from grapes‹.¹⁹

It should be noted that the opacity of propositional meanings is not taken to be absolute. We still can and should talk about subject terms, object terms, and relation terms contributing to the constitution of propositional content. But we can do so without preconceived notions about *how* single terms contribute to the meaning of the propositions they appear in. In order to elucidate this point, I will compare the approach proposed here to standard procedures of modelling knowledge about statements, i. e. reification. But first it will be helpful to discuss a simplified example that is meant to demonstrate that using the approach proposed here we can deal in a simple and transparent manner with inconsistent statements within a domain as well as with statements that may prove to be troublesome when related to other domains.²⁰

3. Legal Wine

§175-2-2 of the Legislative Rule 175CSR 2 governing the activities of the West Virginia Alcohol Beverage Control Commission stipulates that wine in the sense of West Virginia state law is

›any beverage obtained by the fermentation of the natural content of fruits, or other agricultural products, containing sugar and includes, but is not limited to, still wines, champagne and other sparkling wines, carbonated wines, imitation wines, vermouth, cider, perry, sake, or other similar beverages offered for sale or sold as wines containing not less than seven percent (7%) nor more than twenty-four percent (24%) alcohol by volume.«²¹

Beer is defined as ›any beverage obtained by the fermentation of barley, malt, hops, or any other similar product or substitute, and containing more alcohol than that of nonintoxicating beer or nonintoxicating craft beer«. ²² Alcoholic liquors are defined as ›alcohol, beer, including barley beer, wine, *including barley wine* [my emphasis] and distilled spirits, [...]«. ²³ So in West Virginia state law, the concept ›wine‹ includes products based on pears, apples, and rice as long as they contain more than 7% and less than 24% ethanol, i. e. apparently all alcoholic beverages between these limits that are not beer, since beer is discussed under a different heading. But then again ›barley wine‹ is identified as a kind of wine. However, it shares all relevant properties with beer except its alcoholic strength.

¹⁹ So what we model are intentional objects and attitudes towards these objects rather than the objects themselves.

²⁰ This second requirement is particularly relevant, because a central promise of ontologies is interoperability, which is why it employs generic technologies like ›URIs (a generic means to identify entities or concepts in the world), HTTP (a simple yet universal mechanism for retrieving resources, or descriptions of resources), and RDF (a generic graph-based data model with which to structure and link data that describes things in the world)«. Cf. Heath, section ›What is Linked Data?«.

²¹ West Virginia Alcohol Beverage Control Administration, Definition 2.21.

²² West Virginia Alcohol Beverage Control Administration, Definition 2.3.

²³ West Virginia Alcohol Beverage Control Administration, Definition 2.1.

So the law is self-contradictory. If we wanted to model the taxonomy of beverages in West Virginia state law, we would have to settle either for a concept of wine that does include beverages made from barley and does not require that its fermentation is based on sugar. Or we could disregard the subsumption of barley ›wine‹ under the concept of wine, so that our model remains incomplete.

Concepts in law are necessarily vague: courts must have the freedom to apply the law to new beverages that were unknown when the legislation was written. Stipulated meanings in a law can contradict our common-sense notions, so that they cannot easily be mapped on existing ontologies that, like the wine ontology, understand wine as potable liquid that is made from grape. The occurrence of contradictions, vagueness, and tensions between concepts in different domains can lead sceptics to the conclusion that, since concepts are nothing but social constructions that do not follow the strict requirements of the ontology engineer, the whole endeavour of modelling knowledge in a machine-readable way is doomed. Conversely, realists would probably point out that the legal meaning of ›wine‹ in West Virginia could be reconstructed in principle, if we had functioning ontologies of artefacts and social institutions. The resulting determination may be incredibly complex, but feasible in principle.

Or we may wonder whether ›reification‹ can be a solution. RDF, an XML dialect for describing web resources semantically, offers support for this technique, so I will use its syntax to explain the notion.²⁴

RDF represents a reified statement as four statements with particular RDF properties and objects: the statement (S, P, O), reified by resource R, is represented by:

- R rdf:type rdf:Statement
- R rdf:subject S
- R rdf:predicate
- R rdf:object O²⁵

The first triple (R rdf:type rdf:Statement) can be used to refer to the statement that is composed of the tree terms S, P, and O. We could thus refer to the concept »Sake is legal wine in West Virginia« by simply naming the statement »Sake_is_legal_wine_in_WV« (or SLWWV) and composing it out of the subject term »Sake«, the relation »is subclass of«, and the object term »legal_wine_in_West_Virginia«. Such a ›quadlet‹²⁶ does allow us to refer to a statement as a whole. So we could express the intentions of West Virginia legislators in formulating the Legislative Rule by forming a second statement with »West Virginia legislators« as subject term, »stipulate« as relation term, and »RLWWV« as object term: West Virginia legislators stipulate that Sake is legally wine in West Virginia.

²⁴ I discuss reification in RDF for the sake of simplicity. Similar techniques are available for OWL. Cf. Stevens / Lord 2010, *passim*.

²⁵ The Apache Foundation 2011–2014, section »Introduction«.

²⁶ Cf. Stevens / Lord 2010, *passim*.

But reification helps us only as long as the domain to be modelled is not characterised by self-contradictory notions. If we wanted to reify the statement »barley wine is wine«, we would run into problems. Since barley wine is in fact stipulated to be beer and since the stipulation for beer contradicts the stipulations for wine (beer is based on fermentation of grain, wine is based on fermentation of sugar), any ontology trying to capture the intentions of West Virginia legislators is bound to fail, because these intentions contradict each other: a coherent model is impossible. This is different from a situation in which we are merely unsure about the factual truth or falsity of a statement: a reified statement can be false as long as its falsehood is purely factual.²⁷ But reification cannot salvage us from logical or conceptual incoherence.

The way out of this quandary is to deny »barley wine is wine« the status of a RDF statement. »Barley wine is wine« is just the name of a statement containing a subject term, a relation term, and an object term, but none of these terms is part of a RDF triple.²⁸ Hence their aggregation in a statement does not constitute a RDF statement. This expresses the fact that the status of this triple of terms as the description of a resource (i. e. something ›out there‹) is uncertain. Since the reference of the statement is unclear, the same must be presumed for its meaning (or lack thereof). The meaning of the statement is opaque, even though we can specify the terms it contains. But it is equally important to describe the content, whatever it may be, as a propositional content that can be ascribed to the creators of this statement, i. e. presumably legislators in the state of West Virginia.

The main advantage of such an approach over proper reification is that it can be used heuristically: we do not need a full blown ontology for capturing the content of a given discourse in a form that is amenable to further refinement and development. This heuristic approach is particularly useful when we are interested in the connection between what has been said and who said it, i. e. in all domains in which we capture opinions of people, i. e. in all domains that proceed ›doxographically‹. And it can accommodate the development of the intension of a concept over time and thus be helpful to track the history of concepts, beliefs, and theories.

4. A Use Case: Capturing a Philosophical Discourse Doxographically

If we want to condense the approach sketched in this paper into a handy slogan, we could say that it focuses on what people say about the world rather than on what there is in the world. It records opinions rather than facts. In the history of philosophy, doxography, the recording of opinions, is a venerable tradition going back to ancient times.²⁹ So the

²⁷ It is then covered by the ›open world assumption‹ (OWA), at least as long as its truth or falsity is unknown. Cf. Drummond / Shearer 2006, slide 9: »The OWA assumes incomplete information by default.«

²⁸ The ›abstract syntax‹ for RDF (Cyganiak et al. 2014, section 3.1) prescribes that the subject term and the relation term of a RDF triple contain IRIs and not just strings (›literals‹). For the notion of an IRI cf. Duerst / Suignard 2006, section 1.1.

²⁹ However, it should be noted that assessments of its relevance differ. Dillon 2006 calls it »a pretty low form of literature«, Mansfeld 2013 is less skeptical and compares ancient doxographical writings to present-day secondary literature.

minimal ontology for capturing opinions of others can be said to proceed ›doxographically‹. It comprises abstract and spatiotemporal entities, namely persons holding or texts articulating a certain belief and the propositional content of the belief. Propositional attitudes can be understood as properties of spatiotemporal entities, namely the property of asserting, denying, or merely reflecting upon a given propositional content.

So the minimal ontology for capturing opinions of others can be said to proceed ›doxographically‹. It comprises abstract and spatiotemporal entities, namely persons holding or texts articulating a certain belief and the propositional content of the belief. Propositional attitudes can be understood as properties of spatiotemporal entities, namely the property of asserting, denying, or merely reflecting upon a given propositional content.

Such a minimal doxographical ontology can be used to capture the content of a given discourse without making any assumptions about the conceptual structure of the respective domain. In a proof of concept at [EMTO Nanopub I](#) I have assembled ›doxographical facts‹ about the debate on how to define ›philosophy‹ in early modern Iberian philosophy, collating the viewpoints of [Gaspar Cardillo de Villalpando](#), the [Complutenses](#), the [Conimbricenses](#), [Diego Mas](#), [Vicente Montanes](#), [Antonio Rubio](#), [José Saenz de Aguirre](#), and [Franciscus Toletus](#) as ›nanopublications‹.³⁰ Even without additional conceptual analysis of these propositions, we can gain some interesting insights from this purely doxographical ›record keeping‹.

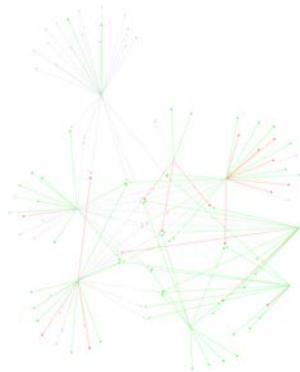


fig. 1: Propositional contents in eight Iberian philosophers debating the proper definition of philosophy. For an interactive version, please open the [SVG-File](#). Javascript must be activated.

Figure 1 shows a network of the eight authors and the propositional content they assert, deny or reflect upon in their texts about the proper definition of philosophy. It has been produced in *gephi*, a very comprehensive tool for the production of network diagrams.³¹ The authors are displayed as ›nodes‹ in this network diagram that only serve as starting points of ›edges‹ (arrows). The edges themselves are coloured according to the propositional attitude that exists between author and propositional content: green arrows signify an assertion, red

³⁰ On the concept of nanopublications and their uses in the humanities cf. Heßbrüggen-Walter 2013, passim.

³¹ A presentation of the guiding principles at work in *gephi* can be found in Bastian et al. 2009, passim.

arrows a negation, grey arrows a neutral stance (e. g. a quotation). Propositional contents are signified by those nodes in which edges end. Since the visualisation is quite complex, zooming and panning is supported via the [SVGpan](#) library. This allows the viewer to explore the structure of the presented network of authors and propositional contents interactively. The points of arrows are linked to URLs of nanopublications on EMTO-Nanopub.

Even a cursory inspection of this visualisation provides some interesting insights into parameters that usually are not in the centre of attention of historians of philosophy. We can discern marked differences in the ratio of grey arrows to coloured arrows in various authors. Antonio Rubio mostly provides theses without taking an explicit stance: most of the arrows starting from this node are grey. In contrast, Cardillo de Villalpando (at the bottom of the figure) prefers a thetical style of writing and expresses only things that are the case: all arrows starting from this node are green. A second surprising result is that, even though many may view Spanish scholasticism as a unified school of thought, many topics come up only in one author. Only a minority of assertions or denials concern more than one or two authors. Reflections like these may lead to a more precise quantitative analysis of argumentative strategies, but would certainly not have come into view by just reading the texts.

But beyond such ›stylometric‹ reflections, visualisations also help us to understand the structure of a debate more precisely. Debates consist in contents that are either asserted or denied. So we can omit all edges that denote mere reflection on a given content and include only assertion and negation as propositional attitudes (edges). And participants in a debate are supposed to endorse at least one thesis that is endorsed or denied by other participants as well.

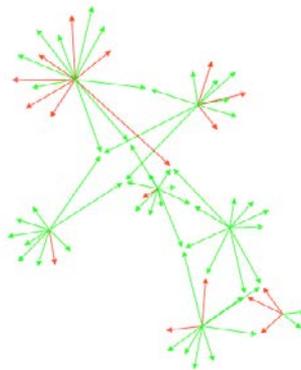


fig. 2: Propositional contents that are asserted or denied in the debate on how to define philosophy. For an interactive version, please open the [SVG-File](#). Javascript must be activated.

To apply these two criteria simplifies the picture considerably. One author drops out of the picture, because he does not fulfil the second criterion: Rubio does not take a stance that is either asserted or denied by another author in the debate. Apparently, there are two camps in the debate, one that seems to focus on philosophy as knowledge of causes (in the upper region

of figure 2) and one that seems to be concerned with the nexus between philosophy and the Divine (in the bottom region of the figure 2). The bridge between both camps is built by Vicente Montañés who asserts contents that can be found in both camps. This impression is reinforced when we simplify further and include only those propositional contents that are asserted by at least two thinkers (figure 3):

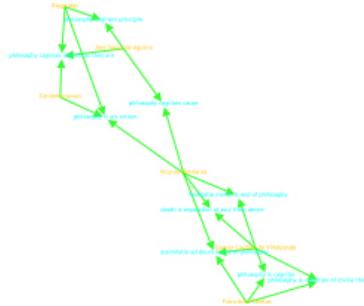


fig. 3: Propositional contents that are asserted by at least two authors in the debate on how to define philosophy. For an interactive version, please open the [SVG-File](#). Javascript must be activated.

Two general points are worth emphasising: first, it should be noted that the visualisations presented here are the result of algorithms for the visualisation of networks implemented in *gephi*. Some minor redactions had to be added manually, but the overall representation of the structure of the debate is not the result of conscious design decisions. Since it is the machine that does the work of structuring the debate, hermeneutic biases are minimised in this step. Second, this approach to visualising excerpts of ›the history of philosophy‹ allows to trace each and every ›visual assertion‹ to the relevant evidence, since every edge that connects an author to a propositional content, i. e. every doxographical statement, is linked to a nanopublication providing the bibliographical data of the source text and the author making the doxographical statement.

›Doxography‹ is an essential, though mostly underrated, element in the workflow of any historian of philosophy. Working with a text, we must first produce summaries, excerpts, or other research notes that help us to fixate its content, before we tackle the more complex task of reconstructing its arguments, comparing them to other sources, and evaluating their validity either in their historical context or in relation to contemporary problems. To deal with this process using digital tools may in itself transform and enhance existing practices in the history of philosophy. But, more importantly, it also opens up new research questions and may change our understanding of the discipline as a whole.

5. Digital Doxography and Heuristic Ontologies: A Vision

In the use case presented here, the ›semantic‹ dimension of semantic web technologies was conspicuously absent. But we can now articulate a broader vision of how the heuristic use

of ontologies could transform not merely the record keeping of the digital doxographer, but transform our strategies of interpreting philosophical sources.

One particularity of (at least some) philosophical theories consists in the way that they try to develop conceptual hierarchies that quite easily could be transformed into statements of an ontology. If we take for example the following two propositional contents:

S:philosophy R:is O:habit

S:part of philosophy R:is O:species of philosophy

For an expert in the domain it is fairly obvious that the first triple expresses a relation of conceptual subordination: It could thus be transformed into a corresponding OWL statement:

```
<owl:Class rdf:ID="philosophy">  
  
<rdfs:subClassOf rdf:resource="habit" />  
  
</owl:Class>
```

The second example expresses the identification of the extensions of two concepts. Everything that is an example of the intension ›is part of philosophy‹ is at the same time an example of the intension ›is a species of philosophy‹. In other words, both concepts are, in the terminology of OWL, equivalent classes:

```
<owl:Class rdf:ID="part of philosophy">  
  
<rdfs:equivalentClassOf rdf:resource="species of philosophy" />  
  
</owl:Class>
```

Of course, it is important to note that again we should not misconstrue such statements as being concerned with philosophy as a thing in the world.³² These statements, too, must be suitably qualified, namely as ›intentional objects‹ of philosophical thought in a given period. And a second important caveat applies: by transforming the isolated statements of a doxographical record into the reconstruction of a conceptual scheme, we leave the domain of ›facts‹ and enter into the realm of interpretation. The more complex a source text is, the more we may expect deviations between different attempts of reconstruction. In this respect, the use of code as a medium of interpretation does not change its fundamental hermeneutic characteristics. But ontologies, understood as the expression of a coherent vision of a given

³² This is a major difference to the goals of the Indiana Philosophy Ontology Project that tries to capture a valid present day understanding of philosophy as a scholarly discipline. Cf. Buckner et al. 2011, passim.

conceptual scheme, could nevertheless develop into powerful tools for the historian of philosophy.

This is particularly true for those domains in which mass digitisation projects have made available large number of previously unknown or inaccessible sources. Since we may expect some progress regarding optical character recognition of historical prints, it is to be hoped that a large number of these texts will at some time in the future be available as machine-readable e-texts. And even though practitioners in the field assert that natural language processing of Latin texts is difficult, because these texts are written in Latin, some progress on this front will hopefully be made too.³³ The upshot of this is that these developments may help us in extracting doxographical triples from a large number of texts, both in Latin and the vernaculars, in order to gain a deeper and more comprehensive understanding of the historical record as it stands. The methodology proposed here may then prove to be a fruitful strategy for turning this content into semantically rich information.

³³ Cf. Passarotti 2010, *passim*.

Bibliographische Angaben

- Apache Jena Documentation. Edited by The Apache Foundation. 2011–2014. [\[online\]](#)
- Mathieu Bastian / Sebastien Heymann / Mathieu Jacomy: Gephi: An Open Source Software for Exploring and Manipulating Networks. In: Proceedings of the Third International ICWSMConference. 2009, S. 361–362. [\[online\]](#)
- Cameron Buckner / Mathias Niepert / Colin Allen: From encyclopedia to ontology: Toward dynamic representation of the discipline of philosophy. In: Synthese 182 (2011), H. 2, p. 205–233. [\[online\]](#)
- John Dillon: Review of Aldo Brancacci (ed.): Philosophy and Doxography in the Imperial Age. Firenze 2005. In: Bryn Mawr Classical Review 12 (2006). [\[online\]](#)
- Nick Drummond / Rob Shearer: The Open World Assumption or Sometimes its nice to know what we don't know. Manchester 2006. [\[online\]](#)
- M. Duerst / M. Suignard: RFC 3987: Internationalized Resource Identifiers (IRIs). 2006. [\[online\]](#)
- Stefan Heßbrüggen-Walter: Tatsachen im semantischen Web: Nanopublikationen in den digitalen Geisteswissenschaften? In: Historyblogosphere: Bloggen in den Geschichtswissenschaften. Edited by Peter Haber / Eva Pfanzer. München 2013, p. 149–160. [\[Nachweis im GBV\]](#)
- Ludger Jansen: Warum sich Artefakte ihrer Marginalisierung widersetzen. In: Deutsche Zeitschrift für Philosophie 61 (2013), H. 2, p. 267–282. [\[Nachweis im OPAC\]](#)
- Theo M.V. Janssen: Frege, Contextuality and Compositionality. In: Journal of Logic, Language and Information 10 (2001), H. 1, p. 115–136. [\[Nachweis im GBV\]](#)
- Max Kemmann: Grasping Technology. Juni 2014. [\[online\]](#)
- Hubertus Kohle: Digitale Bildwissenschaft. Glückstadt 2013. [\[Nachweis im OPAC\]](#)
- Jens Lemanski: Die neuaristotelischen Ursprünge des Kontextprinzips und die Fortführung in der fregeschen Begriffsschrift. In: Zeitschrift für philosophische Forschung 67 (2013), H. 4, p. 566–586. [\[Nachweis im OPAC\]](#)
- Linked Open Data – Frequently Asked Questions. Edited by Tom Heath. [\[online\]](#)
- Jaap Mansfeld: Doxography of Ancient Philosophy. In: The Stanford Encyclopedia of Philosophy (Winter 2013 Edition). Edited by Edward N. Zalta. 2013. [\[online\]](#)
- N. F. Noy / D. L. McGuinness: Ontology Development 101: A Guide to Creating Your First Ontology. 2001. [\[online\]](#)
- OWL Web Ontology Language Reference. Edited by Mike Dean / Guus Schreiber. 2004 [\[online\]](#).
- Marco Carlo Passarotti: Leaving Behind the Less-Resourced Status. The Case of Latin through the Experience of the Index Thomisticus Treebank. In: 7th SaLTmIL Workshop on Creation and Use of Basic Lexical Resources for Less-Resourced Languages, LREC 2010. La Valletta 2010. 2010, p. 27–32. [\[online\]](#)
- RDF 1.1 Concepts and Abstract Syntax. Edited by Richard Cyganiak / David Wood / Markus Lanthaler. 2014. [\[online\]](#)
- Clay Shirky: Ontology is Overrated: Categories, Links, and Tags. [\[online\]](#)
- Barry Smith: Beyond Concepts: Ontology as Reality Representation. In: Proceedings of FOIS 2004. Edited by Achille Varzi / Laure Vieu. Amsterdam, Berlin, Oxford, Tokyo, Washington DC 2014, p. 73–84. [\[Nachweis im OPAC\]](#)
- R. J. Stainton: Context Principle. In: Encyclopedia of Language & Linguistics, vol. 3, Boston 2006, p. 108–115. [\[Nachweis im GBV\]](#)
- Robert Stevens / Phillip Lord: Reification of properties in an ontology. 2010. [\[online\]](#)
- Title 175, Legislative Rule Series 2 Private Club Licensing. Edited by West Virginia Alcohol Beverage Commissioner. [\[online\]](#)
- Zoltán Gendler Szabó: »Compositionality«. In: The Stanford Encyclopedia of Philosophy (Fall 2013 Edition). Edited by Edward N. Zalta. 2013. [\[online\]](#)

Abbildungslegenden und -nachweise

Abb. 1: Propositional contents in eight Iberian philosophers debating the proper definition of philosophy (The rendering of SVG files may vary from browser to browser. Zooming and panning requires the activation of Javascript) (graphic: author).

Abb. 2: Propositional contents that are asserted or denied in the debate on how to define philosophy (graphic: author).

Abb. 3: Propositional contents that are asserted by at least two authors in the debate on how to define philosophy (graphic: author).

Zeitschrift für digitale Geisteswissenschaften

Artikel aus:

Sonderband 1 der ZfdG: Grenzen und Möglichkeiten der Digital Humanities. Hg. von Constanze Baum und Thomas Stäcker. 2015. DOI: [10.17175/sb01](https://doi.org/10.17175/sb01)

Titel:

Computerlinguistische Werkzeuge zur Erschließung und Exploration großer Textsammlungen aus der Perspektive fachspezifischer Theorien

Autor/in:

André Blessing

Kontakt: andre.blessing@ims.uni-stuttgart.de

Institution: Institut für Maschinelle Sprachverarbeitung, Universität Stuttgart

GND: [1058601865](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63862-p0011-9) ORCID:

Autor/in:

Fritz Kliche

Kontakt: kliche@uni-hildesheim.de

Institution: Institut für Informationswissenschaft und Sprachtechnologie, Universität Hildesheim

GND: [1084131943](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63862-p0011-9) ORCID:

Autor/in:

Ulrich Heid

Kontakt: uli@ims.uni-stuttgart.de

Institution: Institut für Informationswissenschaft und Sprachtechnologie, Universität Hildesheim

GND: [111873967](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63862-p0011-9) ORCID:

Autor/in:

Cathleen Kantner

Kontakt: cathleen.kantner@sowi.uni-stuttgart.de

Institution: Institut für Sozialwissenschaften, Universität Stuttgart

GND: [129464872](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63862-p0011-9) ORCID:

Autor/in:

Jonas Kuhn

Kontakt: jonas.kuhn@ims.uni-stuttgart.de

Institution: Institut für Maschinelle Sprachverarbeitung, Universität Stuttgart

GND: [172996090](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63862-p0011-9) ORCID:

DOI des Artikels:

[10.17175/sb001_013](https://doi.org/10.17175/sb001_013)

Nachweis im OPAC der Herzog August Bibliothek:

Erstveröffentlichung:

30.06.2015

Lizenz:

Sofern nicht anders angegeben 

Medienlizenzen:

Medienrechte liegen bei den Autoren

Letzte Überprüfung aller Verweise:

24.05.2016

GND-Verschlagwortung:

[Computerlinguistik](#) | [Textanalyse](#) |

Zitierweise:

André Blessing, Fritz Kliche, Ulrich Heid, Cathleen Kantner, Jonas Kuhn: Computerlinguistische Werkzeuge zur Erschließung und Exploration großer Textsammlungen aus der Perspektive fachspezifischer Theorien. In: Grenzen und Möglichkeiten der Digital Humanities. Hg. von Constanze Baum / Thomas Stäcker. 2015 (= Sonderband der Zeitschrift für digitale Geisteswissenschaften, 1). PDF Format ohne Paginierung. Als text/html abrufbar unter DOI: [10.17175/sb001_013](https://doi.org/10.17175/sb001_013).

André Blessing, Fritz Kliche, Ulrich Heid, Cathleen Kantner, Jonas Kuhn

Computerlinguistische Werkzeuge zur Erschließung und Exploration großer Textsammlungen aus der Perspektive fachspezifischer Theorien

Abstracts

Die Digital Humanities stoßen auf neuartige Probleme, wenn sie Fragen der theoriegeleiteten sozialwissenschaftlichen Forschung bearbeiten. Hier kann die Textanalyse nicht unmittelbar auf linguistischen Merkmalen aufsetzen, sondern sucht den Zugang zu komplexen Begriffen und deren Bedeutungen. Unser Beitrag zur Methodologie der Digital Humanities argumentiert, dass hermeneutisch sensible Korpusanalyseverfahren entwickelt werden können, wenn die eingesetzten Werkzeugkomponenten für die Zieldisziplin transparent bleiben und sich für eine interaktive Exploration des Datenmaterials eignen. Wir stellen interaktive Werkzeuge zur Texterschließung und -analyse vor, die sich flexibel auf fachwissenschaftliche Theorien und Forschungsfragen abstimmen lassen, jedoch gleichzeitig in ihrer Architektur so generisch sind, dass sie breit einsetzbar sind.

Theory-driven research in the social sciences is confronting the Digital Humanities with new challenges. Here, text analysis can not be built directly on linguistic features, but rather the research questions require access to more complex concepts and their meaning. Our methodological contribution to DH argues that it is possible to develop hermeneutically sensitive methods for corpus analysis when the tool components remain transparent for the target discipline and allow for interactive exploration of the underlying data. We provide an overview of interactive tools for text processing and analysis that can be flexibly adjusted to specific disciplinary theories and research questions in the target disciplines, with an underlying architecture that is generic and therefore broadly applicable.

1. Einleitung

Die vielfältige Verfügbarkeit von größeren digitalen Textsammlungen eröffnet grundsätzlich den Zugang zu Fragestellungen für die geistes- und sozialwissenschaftliche Forschung, die mit traditionellen Verfahren nicht oder nur mit unrealistisch großem Zeitaufwand untersucht werden könnten. So kann beispielsweise anhand eines datierten Textkorpus die zeitliche Entwicklung von Textinhalten im größeren Maßstab betrachtet werden – unter zusätzlicher Berücksichtigung von Parametern (wie beispielsweise dem Kontext der untersuchten Texte), zu denen Metadaten wie Land, Autoren o.ä. vorliegen. Das entsprechend datierte und kategorisierte textuelle Datenmaterial scheint über digitale Archive zunächst leicht zugänglich für quantitative Analysen zu sein. Doch dieser erste Eindruck täuscht.

Tatsächlich erweist sich eine systematische, durch geistes- und sozialwissenschaftliche Theoriebildung fundierte Analyse von Textsammlungen als sehr anspruchsvoll. Dies gilt besonders dann, wenn die Textanalyse nicht wie in den sprachwissenschaftlichen Disziplinen unmittelbar auf die Konzepte der linguistischen Theoriebildung aufsetzen kann (beispielsweise durch eine quantitative Analyse der Verwendung von Passivkonstruktionen bei einer Klasse

von Verben), sondern lediglich vermittelnde Funktion beim Zugang zu den Textinhalten bzw. zu nicht-linguistischen Texteigenschaften hat.

Der vorliegende Beitrag stellt die Prozeduren und Tools zur Texterschließung und -analyse im Projekt e-Identity¹ vor, in dem ein großes multinationales und somit mehrsprachiges Zeitungskorpus im zeitlichen Verlauf von über 20 Jahren aus politikwissenschaftlicher Sicht analysiert wird: Wie entwickelten sich die Debatten über Kriege, bewaffnete Konflikte und humanitäre militärische Interventionen in unterschiedlichen Ländern? Welche Rolle spielten kollektive Identitäten (ethnische Zugehörigkeiten, »wir Europäer«, nationale Identitäten, Religionsgemeinschaften etc.) in der Diskussion über internationale Krisensituationen?

Ogleich eine (computer-)linguistische Analyse der Zeitungstexte keinerlei direkte Bedeutung für die untersuchten Fragen hat, kommt ihr als Zwischenschritt eine zentrale Rolle bei der Systematisierung des Vorgehens und der validen Verankerung der quantitativen Untersuchungen in der sozialwissenschaftlichen Theoriebildung zu. Das Ergebnis ist eine sorgfältig abgestimmte interaktive Umgebung für die fachwissenschaftliche Textanalyse, die sich sehr stark auf theoriespezifische Teilfragestellungen zuschneiden lässt, gleichzeitig jedoch in ihrer Werkzeugarchitektur so generisch bleibt, dass sie sich ebenso systematisch auf ganz andere inhaltsanalytische Fragestellungen in großen Textsammlungen anpassen lässt.

Aufgrund dieser über das konkrete Projekt hinausreichenden methodologischen Herausforderungen versteht sich dieser Artikel über die spezifisch technische Vorstellung des entwickelten Werkzeuginventars² hinaus vor allem als Beitrag zur Methodendiskussion in den Digital Humanities und argumentiert für die These, dass die Untersuchung von großen digitalen Sammlungen dann neuartige, hermeneutisch sensible Verfahren hervorbringen kann, wenn die eingesetzten Analyseverfahren und Werkzeuge für die Zieldisziplin transparent und kritisch hinterfragbar bleiben und sich für eine interaktive Exploration des Datenmaterials unter fachwissenschaftlichen Auspizien eignen. Hierfür ist eine intensive Abstimmung der technischen Modellierung mit den fachwissenschaftlichen Fragestellungen und Hypothesen von entscheidender Bedeutung.

Dieser Artikel skizziert in **Abschnitt 2** den fachwissenschaftlichen Hintergrund der Fragestellungen, die im e-Identity-Projekt verfolgt werden, und legt die Grenzen einer textuell-oberflächennahen quantitativen Analyse auf den Artikelsammlungen dar. **Abschnitt 3** stellt den auf die theoretischen Bedürfnisse zugeschnittenen technischen Werkzeugeinsatz im

¹Diese Studie entstand im Rahmen des von Prof. Dr. Cathleen Kantner, Prof. Dr. Jonas Kuhn, Prof. Dr. Manfred Stede und Prof. Dr. Ulrich Heid durchgeführten interdisziplinären Verbundprojekts »Multiple kollektive Identitäten in internationalen Debatten um Krieg und Frieden seit dem Ende des Kalten Krieges. Sprachtechnologische Werkzeuge und Methoden für die Analyse mehrsprachiger Textmengen in den Sozialwissenschaften (e-Identity)«. Wir danken dem Bundesministerium für Bildung und Forschung für die Förderung in den Jahren 2012 bis 2015 im Rahmen der eHumanities-Initiative (Förderkennzeichen: 01UG1234A).

²Der Skopus des vorliegenden Beitrags ist nicht das vollständige Inventar der möglichen computerlinguistischen Werkzeugunterstützung, das im e-Identity-Projekt verwendet und entwickelt wird. Die (halb-)automatische Unterstützung beim Auffinden von Textpassagen, in denen relevante komplexe Konzepte der politikwissenschaftlichen Theoriebildung artikuliert werden, wird hier beispielsweise nicht thematisiert (wenngleich die interaktive Unterstützung der manuellen Textauszeichnung eine wichtige Teilkomponente hierfür darstellt). Weitere Projektbeiträge, u.a. auch von den Projektpartnern an der Universität Potsdam, finden sich auf der [Projektwebsite](#)

Projekt konkret vor: die computerlinguistisch fundierte systematische Textsammlung und -aufbereitung, die in der sogenannten *Explorationswerkbank* bereitgestellt wird (Abschnitt 3.1), die technische Unterstützung einer anspruchsvollen Annotation bzw. Kodierung von Passagen aus der Textsammlung durch Politikwissenschaftler (Abschnitt 3.2) und die Bereitstellung von computerlinguistischen Teilanalysen für weitergehende automatische Analyseschritte (Abschnitt 3.3). Abschnitt 4 schließt mit einem knappen Resümee zur disziplinübergreifenden Zusammenarbeit im Projekt.

2. Inhaltsanalytische Studien auf großen Textsammlungen und das e-Identity-Projekt

Die zunehmende, vielfältige und bequeme Verfügbarkeit großer Mengen digitalisierter Texte über umfangreiche elektronische Textarchive bietet den Sozialwissenschaften im weitesten Sinne viele neue Chancen für die Forschung. Die netzbasierte Verfügbarkeit solchen Textmaterials ermöglicht zudem einen ländervergleichenden Ansatz, der den Bedürfnissen moderner gesellschaftswissenschaftlicher Fragestellungen in einer sich globalisierenden Welt entgegenkommt. Kontinuierlich gesammelte Jahrgänge von Zeitungsinhalten, Parlamentsdebatten, offizielle Dokumente, Rechtstexte, literarische Gesamtausgaben, historische Archive – um nur einige zu nennen – stehen in immer mehr Sprachen umfassend und gut aufgearbeitet bereit. Hinzu kommt eine reiche Auswahl zunehmend anwenderfreundlicherer Textanalysesoftware,³ deren kommerzielle, eher für die qualitativ-interpretierende Inhaltsanalyse geeignete Angebote noch relativ teuer, aber doch meist finanzierbar sind, sowie viele kleinere Anwendungen und umfassende Plattformen (z.B. CLARIN, DARIAH), die computer- oder korpuslinguistische Verfahren der Textanalyse einem breiten Spektrum von akademischen Nutzern zugänglich machen.

Doch warum nutzen die Sozialwissenschaften die vorhandenen neuen Möglichkeiten computer- oder korpuslinguistischer Verfahren der Textanalyse noch so wenig? Warum dominieren immer noch die qualitativ-interpretativen, small-n-Forschungsdesigns über solche, die zumindest auch die Chancen computer- und korpuslinguistischer Methoden für large-scale, big-n, long-time-Forschungsdesigns in den Sozialwissenschaften erschließen? Liegt es an mangelnder Anwenderfreundlichkeit? Oder an einer schlechten Werbung? Wir behaupten, dass die vorhandenen Potentiale aufgrund beträchtlicher Hindernisse nicht ausgeschöpft werden können. Erst wenn diese abgebaut werden, können ebenfalls notwendige Schritte der Verbesserung der Anwenderfreundlichkeit und der Bekanntheit neuer Methoden aus den Digital Humanities den gewünschten Erfolg zeitigen.

Unserer Erfahrung nach scheitert die interdisziplinäre Zusammenarbeit in den Digital Humanities oft an zwei typischen neuralgischen Punkten:

1. Der extreme Aufwand bei der individuellen Forschungsfragen folgenden Korpuserstellung und der damit verbundenen Materialaufbereitung sowie dem

³Für einen Überblick: Krippendorff 2004, S. 281-307; Alexa / Zuell 2000.

Datenmanagement großer Textmengen aus heterogenen und auch in Zukunft nicht standardisierbaren Quellen lässt Sozialwissenschaftler meist schon im Vorfeld der eigentlichen Anwendung computer- und korpuslinguistischer Ansätze scheitern. 2. Sozialwissenschaftler sind daran interessiert, über die Analyse manifester Textinhalte komplexe Sinnzusammenhänge zu rekonstruieren. Sie suchen meist nach *abstrakten* Konzepten, die in der Alltagssprache kaum direkt geäußert werden. Gängige Anwendungen und wörterbuchbasierte Tools werden der Heterogenität der Anforderungen und der Notwendigkeit, von Forschungsprojekt zu Forschungsprojekt neue *Operationalisierungen komplexer fachwissenschaftlicher Begriffe* vorzunehmen, nicht gerecht.

(ad 1) : Die Verfügbarkeit digitaler Textarchive (insbesondere von Zeitungstexten über LexisNexis oder Factiva, politischen Dokumenten und Rechtstexten z.B. über EU-Lex) führte zu einem Boom von textanalytischen Forschungsprojekten in den Sozialwissenschaften – oft mit ländervergleichendem Interesse. Doch meist werden kleine Korpora intensiv mit qualitativen Verfahren analysiert. Korpus- und computerlinguistische Verfahren, mit denen sehr viel größere Fallzahlen bewältigt werden könnten, sind die seltene Ausnahme. Wenige Sozialwissenschaftler nutzen diese Möglichkeiten bisher.⁴ Sie tun dies auch in der Hoffnung, so Daten zu generieren, die Anspruch auf Repräsentativität erheben können und sich mit den gleichen anspruchsvollen statistischen Verfahren auswerten lassen wie die ›harten‹ Daten der quantitativen Sozialforschung, um sich zu diesen in Beziehung zu setzen.⁵

Leider erweist sich die Aussicht auf leichten Zugang zu großen Textmengen oft als Falle. Schon im Vorfeld der Analyse großer Textmengen stellen sich Sozialwissenschaftlern oft unüberwindliche Probleme bei der Erstellung und Aufbereitung des für ihre Fragestellung relevanten Korpus. Manche Archive erlauben das schnelle Download von bis zu 200 Texten *en bloc* (z.B. LexisNexis), andere nur Text für Text. *En bloc* geladene Texte müssen vor der Bearbeitung jedoch wieder in Einzeltexte zerlegt werden. Verschiedene Quellen bieten Texte in unterschiedlichen Text- und Zeichenkodierungen an, gehen mit Sonderzeichen unterschiedlich um und kennzeichnen Metadaten (Datum, Quelle, Autoren etc.) auf unterschiedliche Weise. All dies erschwert das Einlesen des Textmaterials in ein einheitliches Format und erfordert eine Vielzahl von Arbeitsschritten, die nur unter Zuhilfenahme etlicher passgenau programmierter Software-Skripte zu bewältigt sind.

Die großen, leicht zugänglichen Textarchive lassen sich nur mit relativ einfachen Schlagworten und booleschen Schlagwortkombinationen durchsuchen, was zu einer Fülle von semantischen Doppeldeutigkeiten und damit Samplingfehlern⁶ oder – bei zu spezifischer

⁴Große Textmengen wurden für ländervergleichende Untersuchungen politischer Kommunikation z.B. von den folgenden Autoren genutzt: Baker / McEnery 2005; Kantner 2006b; Kantner 2009; Kantner et al. 2008; Koenig et al. 2006; Kutter 2007; Liebert 2007; Renfordt 2009.

⁵ Herrmann 2002, S. 125.

⁶Texte werden im Prozess des Sampling irrtümlich ausgewählt, wenn die Suchworte im Text metaphorisch gebraucht werden. Wird wie in unserem Fall mit Suchworten aus dem Themenfeld Krieg, Militär, Friedenstruppen sowie einer Liste von Krisenstaaten im Untersuchungszeitraum gesucht, können sich durch kriegerische Metaphern z.B. in der Sportberichterstattung Samplingfehler ergeben (z.B. könnte darüber berichtet werden, dass in einem Fußballspiel der Schiedsrichter in der 56. Minute intervenierte ..., ein Spieler aus Bosnien sich verletzte oder auf den Zuschauerrängen teilweise bürgerkriegsartige Zustände herrschten). Sport-Vokabeln auszuschließen, ist jedoch kein geeignetes Gegenmittel, da umgekehrt relevante Artikel oft

Fassung – zu Auslassung von Teilen des relevanten Materials führt. Bei der Auswahl der inhaltlich relevanten Texte (Sampling) aus Archiven fallen zudem Dubletten⁷ an. Dieses ›weiße Rauschen‹ (*noise*)⁸ muss entfernt werden, wenn valide und reliable Forschungsergebnisse erzielt werden sollen.

Diese Probleme sind keine ›Einstiegsprobleme‹ der Umstellung auf digitales Arbeiten, die im Laufe der Zeit verschwinden werden. Ein wesentlicher Zug der Originalität sozialwissenschaftlicher Forschung besteht darin, beständig neue Quellen zu erschließen. Selbst wenn eines Tages ganze Bibliotheken, Zeitungsarchive und Dokumentensammlungen von Rechtstexten in standardisierter elektronischer Form vorlägen, stellten sie doch immer nur einen Teil des möglicherweise relevanten Materials dar. Zur Lösung der genannten Probleme gibt es jedoch noch keine kommerziellen Softwarepakete, die die notwendigen Arbeitsschritte benutzerfreundlich integrieren. Jedes Team schreibt sich daher die nötigen Skripte selbst – was den Teilnehmerkreis automatisch auf solche mit Informatikern beschränkt, Insellösungen hervorbringt und eine erhöhte Gefahr von methodischen Artefakten birgt –, oder aber man kapituliert vor allzu großen Textmengen. Im letzteren Falle sind Sozialwissenschaftler in der Regel auf solche Textmengen verwiesen, die sie noch vollständig lesen, inhaltlich überblicken und manuell bereinigen können. Ein Korpus, das jedoch so klein ist, dass es sich manuell bereinigen lässt, lässt sich in der Folge auch qualitativ-interpretativ auswerten. Computer- und korpuslinguistische Verfahren kommen erst gar nicht zur Anwendung.

Zur Schließung dieser fundamentalen Lücke entwickelt e-Identity eine *Explorationswerkbank* für Korpuserstellung, -bereinigung und -management, welche die Nutzer unterschiedlichsten Quellen, Textformaten und Sprachen individuell anpassen können ([Abschnitt 3.1](#) und [3.2](#)).

(ad 2) : Schwierigkeiten bei der Operationalisierung komplexer fachwissenschaftlicher Begriffe resultieren einerseits aus den Besonderheiten des Gegenstandsbereichs der Sozialwissenschaften. Die Gegenstände dieser Fächer sind zum großen Teil keine materiellen Faktizitäten, die durch die Beobachtung von einem unbeteiligten, ›objektiven‹ Beobachter gemessen werden könnten.⁹ Gerade die interessantesten und am intensivsten debattierten Forschungsgegenstände dieser Fächer sind in den Worten eines der französischen Gründerväter der Soziologie »soziale Fakten«. ¹⁰ Abstrakte Gegenstände wie z.B. »Identitäten«, »Werte«, »Staat«, »Macht«, »soziale Gerechtigkeit« kann man nicht anfassen. Sie existieren nicht außerhalb einer sprachlich verfassten Praxis der Kommunikation und Interaktion, in der die Menschen sich Gedanken über ihr Zusammenleben machen, darüber streiten und in

sportliche Metaphern gebrauchen oder über sportliche Ereignisse berichten (z.B. wenn die multinationalen Truppen in Kundus ein Freundschaftsspiel organisierten, bei dem einheimische, niederländische und deutsche Soldaten gegeneinander antraten).

⁷Typische Ursachen dafür sind z.B. Dopplungen in den Datenbanken (z.B. A-, B- und Online-Versionen von Zeitungstexten) oder unterkomplexe Suchmasken, die eine Zerlegung einer komplexeren Schlagwortkombination nötig machen.

⁸ Gabrielatos 2007, S. 6; Kantner et al. 2011.

⁹Solche ›harten‹ Fakten gibt es natürlich auch. Sie sind z.B. der statistischen Datenerhebung zugänglich. Beispiele hierfür wären demographische und sozio-ökonomische Daten (z.B. Verteilung von Bildung, Einkommen, Armut, Arbeitslosigkeit etc.) oder Daten über politische Präferenzen (z.B. Umfragen, Sonntagsfrage, Wahlergebnisse).

¹⁰ Durkheim 1964 [1895].

Diskursen gemeinschaftlich deuten, sozusagen ›sozial konstruieren‹ und dabei immer wieder neu interpretieren.

Zum anderen resultieren notorische Probleme der Operationalisierung aus der Kluft zwischen theoretischen Konzepten und der Alltagssprache. Sozialwissenschaftler suchen in verschiedenen Arten von Texten Antworten auf Fragen, die vor dem Hintergrund ihrer wissenschaftlichen Theorien relevant sind. Sie wollen nicht einfach beschreiben, *wie* über etwas sprachlich kommuniziert wird, sondern sie beobachten Kommunikationsprozesse durch die Brille theoretischer Begriffe, die gerade *nicht* direkt in der Alltagssprache verwendet werden. Der begriffliche Gehalt ist dabei selbst immer Gegenstand wissenschaftlicher Kontroversen, weil er abhängig ist von der Art und Weise, wie ein Gegenstand oder Gegenstandsbereich im Hinblick auf eine leitende Fragestellung konzeptualisiert wird. Dies hat einen notorischen Pluralismus der sozialwissenschaftlichen Begrifflichkeiten zur Folge. Ein exzellentes Beispiel für diese Schwierigkeiten stellt die Analyse kollektiver Identitäten dar. Kollektive Identität interessiert viele Sozialwissenschaftler. Der Historiker Lutz Niederhammer und andere beklagen eine ausufernde Proliferation von Identitätskonzepten,¹¹ dennoch kommt man ohne diesen Begriff nicht aus,¹² mit dem u.a. erfasst werden soll, was denn eigentlich Gemeinschaften zusammenhält, was den sozialen und kulturellen Kitt ausmacht. Was wir begrifflich unter »kollektiver Identität« verstehen und wie wir es empirisch operationalisieren, ist wesentlich abhängig davon, welchen Identitätsbegriff wir verwenden. Derzeit sind recht unterschiedliche und wohl auch sich wechselseitig ausschließende Identitätsbegriffe auf dem Theoriemarkt vorhanden: zwischen dem sozialpsychologischen, einem hermeneutisch-pragmatischen oder einem poststrukturalistischen Identitätsbegriff liegen Welten. Mit unterschiedlichen theoretischen Brillen achten Forscher auf ganz unterschiedliche Ausdrücke: Aus der Perspektive differenztheoretischer Ansätze (so von Carl Schmidt, Niklas Luhmann, Jacques Derrida und Michel Foucault) zeigt sich der Diskurs einer Ingroup an der Art, wie sie Outgroups (*the other*) benennt, stigmatisiert und sich in Abgrenzung zu ihnen definiert. Sozialpsychologische Ansätze würden nach auffälligen positiven emotionalen Ausdrücken für die eigene Gruppe als Held oder Opfer im untersuchten Textmaterial suchen. Hermeneutisch-pragmatische Ansätze würden nach problemorientierter Kommunikation über das, was aufs Ganze gesehen für das jeweilige Kollektiv – aus Sicht seiner Mitglieder – gut ist, suchen.¹³

Empirisch folgt daraus, dass gerade die interessantesten theoretischen Konzepte in den Sozialwissenschaften von verschiedenen Fachwissenschaftlern höchst unterschiedlich operationalisiert werden. In e-Identity versuchen wir das Mit- oder Gegeneinander von verschiedenen kollektiven Identitäten (z.B. ethnische, nationale, europäische, transatlantische, religiöse Identitäten) in Zeitungsdebatten über Kriege und humanitäre militärische Interventionen zu analysieren. Sprachliche Ausdrücke, die vor dem Hintergrund der einen Theorie als Ausdruck »nationaler Identität« gelten, mögen jedoch vor dem Hintergrund einer anderen Theorie irrelevant sein oder etwas ganz anderes messen. Die Hoffnung, ein für alle Mal universelle linguistische Muster zu identifizieren oder Fachbegriffe über kontextunabhängige Lexika zu operationalisieren, beurteilen wir daher skeptisch. Wir suchen

¹¹ Niethammer 2000; Brubaker / Cooper 2000.

¹² Kantner 2006a; Risse 2010.

¹³ Kantner 2004; Kantner 2006a; Tietz 2002a; Tietz 2002b; Tietz 2010.

nicht nach einfachen sprachlichen Ausdrücken (z.B. Deutschland = Hamburg, Berlin, Elbe, Bodensee, Schwaben etc.), sondern nach Bedeutungen, die in der Alltagssprache selten direkt geäußert werden. Typische Ausdrücke für »deutsche nationale Identität« wären eher Umschreibungen (z.B. »unsere historische Verantwortung«, »Berlin muss endlich...«, »Deutschland in Europa«, »Bündnisverpflichtungen«), die nur in bestimmten Kontexten das Gesuchte meinen, während jedes einzelne Wort in anderen Kontexten womöglich nicht über seine wörtliche Bedeutung hinaus verweist. Denn das, was diesen abstrakten theoretischen Begriffen in der Lebenswelt der Menschen entspricht, ist stark kontextabhängig: nationale Identität äußert sich höchst unterschiedlich in verschiedenen Ländern zu verschiedenen Zeiten, ja selbst zur gleichen Zeit in Bezug auf verschiedene politische Themen – man vergleiche z.B. die Diskussion um deutsche Identität in Fragen der Integration von Migranten mit der Diskussion um deutsche Identität in außenpolitischen Fragen. Sprecher verschiedener sozialer Schichten verleihen ihren kollektiven Zugehörigkeiten auf sehr unterschiedliche Weise Ausdruck. Weitere textanalytische Probleme ergeben sich daraus, dass sich Sprecher sich im gleichen Text als Mitglied verschiedener Gruppen bekennen können oder sich von einer nahen Ebene der Identifikation zu abstrakteren bewegen. »Wir Deutschen«, wird beispielsweise häufig sinngemäß argumentiert, haben aufgrund unserer Geschichte eine besondere historische Verantwortung und sollten uns darum besonders stark dafür einsetzen, dass die »internationale Gemeinschaft« den Konflikt im Land X auf diese oder jene Weise löst. Moderne Menschen haben »multiple Identitäten«¹⁴: gute Europäer zu sein, wurde Bestandteil der deutschen nationalen Identität nach 1945.

Die aus sozialwissenschaftlicher Perspektive interessanten und forschungsleitenden Konzepte sind folglich nicht standardisierbar. Computer- und korpuslinguistische Ansätze sollten daher dem einzelnen Forscherteam Raum für seine *eigene* Operationalisierung lassen und es dabei möglichst umfassend unterstützen (z.B. im Wechselspiel von manueller und automatischer Annotation in »lernenden« Anwendungen).¹⁵ In e-Identity stellt der *Complex Concept Builder*¹⁶ eine solche flexible und interaktive Werkzeugumgebung zur Verfügung (einige Aspekte dazu werden in [Abschnitt 3.3](#) angesprochen).

3. Computerlinguistische Methoden in der Korpusaufbereitung

Dieser Abschnitt soll anhand von zentralen Teilprozeduren in den Abläufen der Dokumentenauswahl, -aufbereitung und der Vorbereitung einer theoretischen Inhaltsanalyse zeigen, wie durch den Einsatz von anspruchsvollen computerlinguistischen Methoden und Werkzeugen für datenintensive sozialwissenschaftliche Korpusstudien einerseits generische Lösungen für wiederkehrende Aufgaben angeboten werden können und andererseits eine sehr spezifische Anpassung des Ablaufs an die theoretischen Annahmen und technischen

¹⁴ Risse 2010.

¹⁵ Dies schließt natürlich nicht aus, dass bewährte Operationalisierungen für die im Umfeld dieser komplexen fachlichen Konzepte ausgedrückten Sachverhalte, Bewertungen und Beziehungen etc. wie üblich analysiert werden können.

¹⁶ Eine Darstellung des Complex Concept Builder findet sich in Blessing et al. 2012.

Randbedingungen der jeweiligen Studie ermöglicht wird. Den Gesamttablauf der Verarbeitung von Dokumenten im *e-Identity*-Projekt skizziert *Abbildung 1* schematisch.

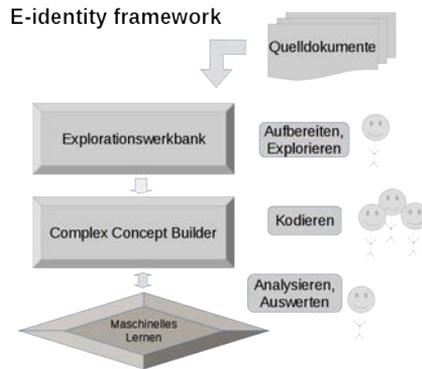


Abb. 1: Schematische Darstellung der Dokumentenverarbeitungskette (Quelle: Eigene Darstellung).

Zunächst (im Schema ganz oben) werden Rohtexte importiert, die im Rahmen der sogenannten *Explorationswerkbank* verwaltet, gesichtet und aufbereitet werden. Ziel ist ein für die weiteren Schritte geeignetes und den methodischen Ansprüchen einer späteren Analyse entsprechendes Korpus. Gleichzeitig bilden die Explorationsmöglichkeiten des entsprechend vorbereiteten Korpus eine wichtige Grundlage für die Hypothesenbildung bzw. Verfeinerung von Ausgangshypothesen. [Abschnitt 3.1](#) widmet sich den Teilaspekten der Datensammlung und -vorbereitung im Detail¹⁷, [Abschnitt 3.2](#) illustriert anhand einiger Beispiele, welche Möglichkeiten sich aus einer computerlinguistischen Aufbereitung für die Arbeit mit dem Korpus ergeben.

Für den anschließenden Arbeitsschritt der Korpusannotation bzw. des Kodierens müssen die theoretischen Konzepte, auf die die späteren Analysen abheben, zunächst entwickelt und operationalisiert und anschließend in der Annotation/Kodierung umgesetzt werden. Zur Unterstützung dieser Prozesse werden in e-Identity Methoden geprüft, die im sogenannten *Complex Concept Builder* subsumiert sind, der nicht im technischen Hauptfokus des vorliegenden Beitrags – den computerlinguistischen Aspekten der Dokumentenaufbereitung – liegt. Die Frage nach den Möglichkeiten der computerlinguistischen Unterstützung des Annotations-Prozesses, der einer sehr gut koordinierten Abstimmung mit den konkreten Fragestellungen aus der Politikwissenschaft bedarf, spielt jedoch zentral in den Kodierungsschritt hinein und wird in [Abschnitt 3.3](#) behandelt.

Im Anschluss an die Annotation/Kodierung stehen die Daten für unterschiedliche Arten der Auswertung zur Verfügung: sie können (i) direkt abgefragt, (ii) für weitere Explorationen geeignet aggregiert und visualisiert oder (iii) einer statistischen Auswertung unterzogen

¹⁷Aus Platzgründen kann hier nur ein Überblick zu den verwendeten Techniken gegeben werden. Eine detaillierte Darstellung der verwendeten Methoden findet sich in Kliche et al. 2014.

werden. Mit Verfahren aus dem maschinellen Lernen können schließlich (iv) auf Basis der vorliegenden Annotationen Modelle für eine automatische Analyse weiterer Daten trainiert und evaluiert werden bzw. (v) in einem interaktiven Zyklus aus vorhandenen Bausteinen weiterführende Analysen erzeugt werden. Aspekte der späteren Auswertung sind von großer Bedeutung bei der Konzeption der Aufbereitung.

3.1. Computerlinguistische Methoden bei der Datensammlung und -vorbereitung

Für die Untersuchung der kollektiven Identitäten wurde ein Sample mit Zeitungsdaten im Umfang von mehr als 800.000 Artikeln erstellt, die von fünf digitalen Medienportalen heruntergeladen wurden. Die Daten stammen aus 12 Zeitungen aus sechs Ländern und enthalten Artikel in drei Sprachen (Deutsch, Englisch, Französisch). Abbildung 2 zeigt rohe Textdaten, wie sie von Medienportalen bereitgestellt werden.¹⁸ Die verschiedenen Medienportale präsentieren die Daten in jeweils eigenen Datenstrukturen. Neben den Fließtexten enthalten die Artikel Metadaten, die für die textanalytische Arbeit wertvoll sind. Die Metadaten liegen zum Teil strukturiert und zum Teil semi-strukturiert vor. Die Beispiele in Abbildung 2 machen deutlich, dass die rohen Textdaten erst erschlossen werden müssen, um die textlichen Inhalte und Metadaten für die Textanalyse verwenden zu können.

```
Document 1 of 145
21 November 2010
(c) 2010
After Lisbon Summit: New priorities for the alliance
BYLINE: By RICHARD HEDU in Lisbon
SECTION: Politics
LENGTH: 734 words
Representatives of the member states of the NATO arrived in
Lisbon ahead of the NATO summit. The alliance's new goals

Elections in Serbia ended with no clear winner. Jeff Brohan.
31 words
PX
English
September 21, 1997
Serbia's presidential elections ended with no clear outcome.
With no candidate receiving over 50% of the votes in the
first round, a second round will be held in October. pg. 43.
Document px00000019970921fwe0024p
```

Abb. 2: Abbildung 2 zeigt Rohe Textdaten, wie sie auf digitalen Medienportalen bereitgestellt werden (Quelle: Eigene Darstellung).

Wir entwickeln mit der *Explorationswerkbank* ein generisches Werkzeug, mit dem Textwissenschaftler rohe Textdaten in unterschiedlichen Datenstrukturen erschließen können. Die Voraussetzung ist, dass die Daten in einer festen Datenstruktur vorliegen. Mit der Werkbank können die Textdaten in textstrukturelle Einheiten (im Folgenden: »Artikel«) segmentiert werden. Aus den Artikeln werden die textlichen Inhalte und Metadaten extrahiert und aus den Daten ein Korpus erstellt. Die Werkbank stellt anschließend Funktionen für die Bereinigung des Korpus und für die textanalytische Arbeit mit den Daten bereit. Ein Wizard führt die Anwender in einer intuitiv nachvollziehbaren Weise durch die Schritte der Datenerschließung. Die Arbeitsschritte des Wizards werden im Folgenden beschrieben.

¹⁸Die Beispiele in diesem Artikel wurden neutralisiert, um Urheberrechtsverletzungen zu vermeiden.

3.1.1 Schritt 1: Die Zerlegung der rohen Textdaten in Artikel

Der Erschließung beginnt mit dem Import der Rohdaten. Textwissenschaftler können Textdaten in verschiedenen Formaten importieren: DOCX, RTF, ODT (Open Office), HTML, TXT (Plaintext ohne Markup). Der Wizard überführt die Daten in Plaintext ohne Markup. Wir konvertieren die unterschiedlichen Datenformate in Plaintext und nehmen den Verlust textstruktureller Merkmale (Fonts, Schriftgröße usw.) in Kauf, um die Texte in einem einheitlichen Format ablegen zu können und um die Textdaten für anschließende computerlinguistische Verarbeitungsschritte zugänglich zu machen. Im speziellen Fall von Zeitungstexten hat das Fehlen von detaillierten auf die Textstruktur bezogenen Metadaten nach unseren Erfahrungen keine negativen Auswirkungen. Die Zeichenkodierungen Latin 1 und Unicode werden in Unicode vereinheitlicht.

Die Anwender lassen sich anschließend Ausschnitte der Rohdaten in einem Vorschaufenster anzeigen. *Abbildung 3* zeigt das Vorschaufenster mit rohen Textdaten.

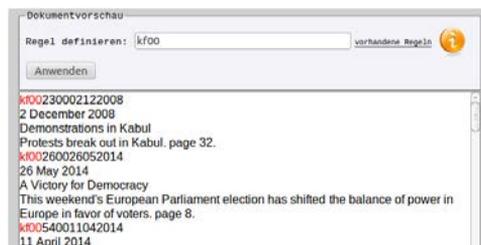


Abb. 3: Ein Ausschnitt von rohen Textdaten wird in einem Vorschaufenster angezeigt. Die Anwender definieren eine Segmentierungsregel, um die Rohdaten in textstrukturelle Einheiten (»Artikel«) zu segmentieren. Im Beispiel trennen Zeilen, die mit der Zeichenfolge kf00 beginnen, die einzelnen Artikel (Quelle: Eigene Darstellung).

Die Anwender suchen in den Rohdaten nach Hinweisen, die die Grenze zwischen Artikeln markieren, und erstellen Segmentierungsregeln. Um eine Segmentierungsregel zu prüfen, werden die von der Regel erfassten Textstellen im Vorschaufenster farblich hervorgehoben, wie es in *Abbildung 1* dargestellt ist. Für die Segmentierungsregeln stehen folgende Funktionen zur Verfügung:

- Definitionen des Segments: Eine feste Zeichenfolge oder ein regulärer Ausdruck;
- Angaben zur minimalen und maximalen Länge des Segments;
- Definitionen für das vorangegangene und das nachfolgende Segment und Angaben zu deren Segmentlängen.

Die Segmentierung ist zeilenbasiert. Der Wizard sucht in den Rohdaten nach den Zeilen (»Segmenten«), die von einer Segmentierungsregel erfasst werden, um die Rohdaten in Artikel aufzutrennen. Die Anwender können wählen, ob eine Zeile, die von einer Segmentierungsregel erkannt wird, zum vorausgehenden oder zum folgenden Artikel gezählt wird. Anschließend wendet der Wizard die Regel auf die importierten Daten an und legt die segmentierten Artikel

in einem Repository ab. Die zugrundeliegenden Segmentierungsregeln und ein Zeitstempel werden als Prozessmetadaten festgehalten. Für jeden importierten Artikel wird außerdem die Anzahl der Segmente (d. h. Zeilen) als einfacher Importcheck festgehalten: Auffallende Abweichungen der Artikellängen können darauf hinweisen, dass Artikelgrenzen nicht richtig erkannt wurden.

3.1.2 Schritt 2: Die Extraktion der textlichen Inhalte und Metadaten aus den Artikeln

In einem zweiten Schritt werden aus den Artikeln die textlichen Inhalte und Metadaten extrahiert. Die Anwender können sich wieder in einem Vorschaufenster Beispielartikel anzeigen lassen und Regeln definieren, nach denen Metadaten und Fließtexte aus den Artikeln extrahiert werden. Die Regeln erkennen wieder nur Segmente, d. h. Zeilen in den Rohdaten der Artikel. Für jede Regel wird ein Bezeichner angegeben. Der Bezeichner ist i. d. R. der Name des extrahierten Metadatum: Datum, Autoren, usw. Für die Erstellung der Extraktionsregeln für Fließtexte und Metadaten stehen mehr Funktionen als für die Segmentierung der Rohdaten in Artikel zur Verfügung. Die Anwender können für die Extraktion von Metadaten und Fließtexten verschiedene Eigenschaften von Segmenten verwenden:

(1) Textmerkmale:

Bestimmte Textmerkmale dienen als positive bzw. negative Indikatoren für Metadaten. Ein positiver Indikator zeigt das Vorhandensein eines Metadatum an. Ein negativer Indikator verhindert die Interpretation eines Segments als Metadaten-Indikator.

(2) Kontextregeln:

Segmente können alternativ

1. im Text Vorgänger oder Nachfolger von anderen Segmenten sein, die ihrerseits Indikatoren enthalten.

Die Segmente können danach beschrieben werden, wo sie in der zu analysierenden textstrukturellen Einheit auftreten (z. B. am Anfang oder am Ende der Einheit, oder nach dem Titel-Segment eines Zeitungsartikels).

(3) Typische Beispiele für Eigenschaften von Segmenten, die solche Segmente zu Indikatoren für Metadaten machen, sind:

1. das Auftreten von bestimmten Schlüsselwörtern (»Anker«);
2. das Auftreten von Mustern von Wörtern oder Zeichenketten einer bestimmten Art (z. B. Jahreszahlen, Datumsmuster, Strukturen aus Vorname und Nachname). Damit

- solche Muster leichter formuliert werden können, können Wortlisten angelegt und eingebunden werden;
3. die minimale bzw. maximale Länge der Segmente.

Die Segmente, die von einer Regel erfasst werden, können wieder im Vorschauartikel farblich hervorgehoben werden, um die Regeln zu testen. Die Anwender können festlegen, ob für eine Regel das gesamte Segment im Repositorium abgelegt werden soll, oder nur der Bereich im Segment, der von der Regel erfasst wird. Für einen Bezeichner können mehrere Regeln erstellt werden. Beispielsweise können Autorenangaben an mehreren Stellen im Artikel auftreten: Am Ende des Fließtextes, unter der Überschrift oder als Teil der strukturierten Metadaten zu Beginn eines Artikels.

Die Anwender erstellen mehrere Regeln und speichern diese als Schablone ab, die in der Folge auf eine bestimmte Datenstruktur appliziert werden kann. Für die Schablone wird ein Name vergeben. In der Schablone werden zusätzliche Prozessmetadaten abgelegt: Beim Speichern jeder Regel wird ein Zeitstempel gesetzt. Weiter wird für jede Regel ein Kommentarfeld bereitgestellt, in dem festgehalten werden kann, für welche Phänomene in den Rohdaten eine Regel erstellt wurde. Schließlich hält die Schablone fest, welche Verarbeitungsschritte die erstellten Regeln verlangen, beispielsweise ob die Tokenisierung der Daten oder die Bestimmung von Wortarten erforderlich ist.

3.1.3 Schritt 3: Die Erschließung der Textdaten

Im dritten Schritt wird die Schablone auf die Artikel gelegt, um die Regeln anzuwenden. Der Wizard prüft, welche Vorverarbeitungsschritte für die Regeln der Schablone notwendig sind, und zeigt sie den Anwendern als Voreinstellungen in einem Formularfeld an. Die Anwender können diese Voreinstellungen erweitern, beispielsweise, wenn Annotationen wie Wortarten für die Volltextsuche zur Verfügung stehen sollen. Anschließend starten die Anwender den Analyseprozess. Die Artikel werden zunächst vorverarbeitet. Um die zeitintensiven Verarbeitungsschritte zu bündeln, erfolgt ohne weitere Interaktion mit den Nutzern nach der Vorverarbeitung die Anwendung der Regeln der Schablone. Die textlichen Inhalte und Metadaten werden erkannt und im Repositorium als Datensätze abgelegt.

3.1.4 Schritt 4: Die Bereinigung des Korpus

Die *Explorationswerkbank* integriert Funktionen, um die Korpusdaten zu bereinigen. Die Anwender können nach verschiedenen Kriterien Datensätze aus dem Korpus ausschließen. Ausgeschlossene Datensätze werden nicht gelöscht, sondern im Repositorium markiert. Eine maximale Länge für leere Artikel und für kurze Artikel kann festgelegt werden. Die Werkbank enthält eine Funktion, um Dubletten und Semi-Dubletten im Korpus aufzudecken. Die Anwender wählen dazu auf einer Skala zwischen 0 und 1 einen Ähnlichkeitswert. Für Textpaare, deren Ähnlichkeit oberhalb dieses Schwellenwerts liegt, wird der kürzere Text aus dem Korpus ausgeschlossen und der längere Text beibehalten. Wenn mehr als zwei Texte identisch oder

ähnlich sind, wird der längste Text beibehalten und die übrigen Texte ausgeschlossen. Die Werkbank bestimmt den Anteil von Zahlen und nicht alphanumerischen Zeichen im Text, um Einheiten auszufiltern, die keinen Fließtext enthalten.

Durch Anwendung einer Topic-Model-Analyse kann jedem Artikel eine Topic-Verteilung zugewiesen werden. Dieser Schritt kann zwar nicht zum automatischen Ausschluss von Artikeln genutzt werden, aber die Exploration der Topics bietet eine effiziente Methode, um eine Kandidatenmenge von Off-Topic-Artikeln manuell zu sichten. Aufbauend auf den manuellen Klassifikationen kann mittels maschinellen Lernens ein Off-Topic-Filter analog zu einem Spam-Filter trainiert werden. In unserem Projekt können dadurch z.B. Artikel zu historischen Kriegen, Sportereignissen und Medienkritiken ausgefiltert werden, die nicht Gegenstand des Forschungskorpus sind.

3.2 Computerlinguistische Methoden für die Korpusexploration

Nach Bereitstellung im Korpus werden die Texte mit computerlinguistischen Werkzeugen aus der CLARIN-Infrastruktur¹⁹ maschinell analysiert. Diese tiefe Analyse ist wichtig, um inhaltliche Untersuchungen zu ermöglichen, die über einfaches Bestimmen von Worthäufigkeiten hinausgehen.

Abbildung 4 skizziert diesen Verarbeitungsschritt. Die Daten werden zur internen Weiterverarbeitung im UIMA-Format in einer Datenbank gespeichert. Diese Art der Repräsentation hat den Vorteil, dass die Annotationsebenen dynamisch erweitert werden können (z.B. können neue Werkzeuge eingebunden werden, die am Projektanfang noch nicht bekannt sind). Des Weiteren erlaubt UIMA auch mehrfache Annotationen und überlappende Annotationen, was eine Grundlage für eine differenzierte Analyse darstellt.

Das Bewusstsein, dass computerlinguistische Verfahren nicht perfekt sind, ist ein wichtiges Kriterium, das den Endanwendern in unterschiedlichster Weise vermittelt werden muss. Der Einsatz unterschiedlicher Analysewerkzeuge auf gleicher Ebene (z.B. mehrere verschiedene Part-of-Speech-Tagger) kann helfen, das System robuster zu machen: bei schwierig zu analysierenden Daten werden unterschiedliche Vorhersagen gemacht, die Wahrscheinlichkeit einer richtigen Lösung steigt also; bei kanonischen Dateninstanzen werden die Werkzeuge übereinstimmen, die Analyse wird somit als relativ verlässlich erkennbar. Je nach Anforderung kann durch die Werkzeugkombination ein Gesamtsystem mit großer Abdeckung/Erkennung (*recall*) oder mit vordringlichem Gewicht auf Güte (*precision*) erzeugt werden.

¹⁹Die verwendeten Stuttgarter Werkzeuge werden detailliert in Mahlow et al. 2014 beschrieben. Eine ausführliche Darstellung, wie die Werkzeuge zu einer Applikation zusammengefasst werden, findet sich in Blessing 2014. Des Weiteren kann eine globale Übersicht der CLARIN-D-Werkzeuge [online](#) abgerufen werden.

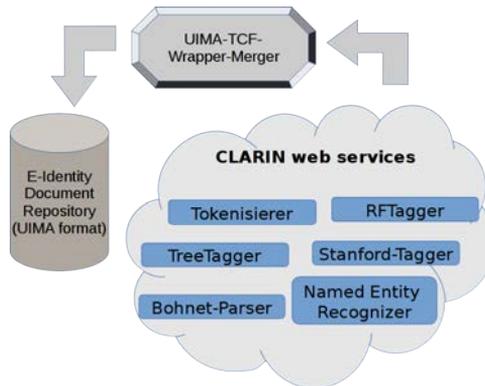


Abb. 4: Computerlinguistische Verarbeitungsschritte, die im Hintergrund der Dokumentenaufbereitung ablaufen (Quelle: Eigene Darstellung).

Folgendes Beispiel (Abbildung 5) zeigt die Analyse des Begriffs »Nato« in englischen Texten durch drei unterschiedliche Part-of-Speech Tagger (TreeTagger²⁰, Stanford-Tagger²¹ und Bohnet-Tagger²²). Man sieht, dass die Werkzeuge in 88% der Fälle zum gleichen Ergebnis kommen, aber in 12% der Fälle (das ist jedes achte Vorkommen) Uneinigkeit besteht. Je nach Anwendung können diese restlichen 12% unterschiedlich interpretiert werden. Die Entscheidung bleibt so dynamisch konfigurierbar und es geht keine Informationen durch ein zu frühes Vereinigen verloren.

Abb. 5: Illustration des redundanten Einsatzes mehrerer Part-of-Speech-Tagger (Quelle: Eigene Darstellung).

Die automatischen Annotationen können je nach Benutzereinstellung in den Texten visualisiert werden, um bei der Exploration Schlüsselstellen im Text schneller zu finden (bei der Kodierung inhaltlicher Kategorien darf diese Funktion nicht verwendet werden, um einen Bias zu vermeiden).

Die verarbeiteten Daten werden indiziert, um eine schnelle Volltextsuche zu ermöglichen. Suchkriterien der Volltextsuche können neben Wortformen auch die Annotationen sein, die in den computerlinguistischen Vorverarbeitungsschritten für die Wörter vergeben worden sind (z. B. Lemmaangaben, Wortarten, etc.); außerdem können die Einträge der Wortlisten mitbenutzt werden, die für die Indikatorensuche angelegt wurden, z. B. Listen von Vornamen, Personen, Orten, Nachrichtenagenturen usw.

Die *Explorationswerkbank* bietet Methoden zur statistischen Auswertung von Artikeln anhand des in den Metadaten hinterlegten Erscheinungsdatums an. So können in unterschiedlichen Zeitintervallen (jährlich, monatlich, wöchentlich, täglich) Zählungen gemacht

²⁰ Schmid 1995.

²¹ Toutanova 2003.

²² Bohnet 2010.

werden und damit die Medienaufmerksamkeit für die Themen des Samples auf der Zeitlinie erfasst werden («Issue Cycles»). Diese Ergebnisse werden in der folgenden inhaltlichen Analyse außerdem zur Normalisierung verwendet.

3.3. Computerlinguistische Unterstützung bei der Annotation

Die genauen textanalytischen Fragestellungen leiten sich in e-Identity aus politikwissenschaftlichen Überlegungen ab und sind mit studienspezifischen Konzepten verbunden. Somit kommen unüberwachte Verfahren, wie sie gern bei der Big-Data-Analyse eingesetzt werden, nur bedingt in Frage. Die Wahrscheinlichkeit, dass ein solches Verfahren genau die gestellte Fragestellung beantwortet, ist sehr gering. Es muss also durch manuelle Kodierung/Annotation ein grundlegendes Wissen erzeugt werden, mithilfe dessen überwachte maschinelle Verfahren arbeiten können.

Im Vorgängerprojekt zu e-Identity wurden komplette Artikel inhaltlich bewertet. Dies ist für die Auswertung der politikwissenschaftlichen Fragestellung völlig ausreichend, aber es hat sich gezeigt, dass diese grobkörnige Kodierung nicht ausreichend ist, um mittels computerlinguistischer Verfahren eine großflächige Analyse vorzunehmen. Unser Ansatz bietet einen neuen Freiheitsgrad, indem jeweils die gewünschte Textstelle zur Kodierung ausgewählt werden muss.

Je nach Projekt verschieben sich die Fragestellungen und somit auch die Variablen der Kodierungen. Dies wird in unserem System durch ein frei rekonfigurierbares Kodierschema ermöglicht. Politikwissenschaftler können dazu ein hierarchisches Kodierbuch inklusive Anmerkungen, die als Tooltipp angezeigt werden, hochladen. Das System generiert daraus eine interaktive Oberfläche, die den Kodierern angezeigt wird. [Abbildung 6](#) zeigt einen Ausschnitt aus der aktuell generierten Schnittstelle für die Kodierung.

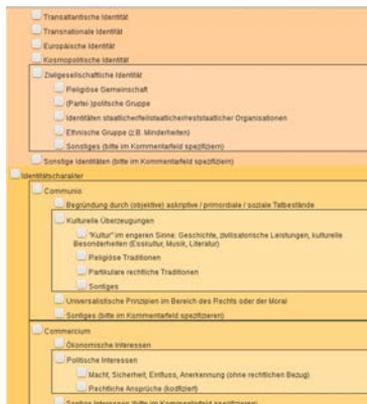


Abb. 6: Die Benutzerschnittstelle für die politikwissenschaftliche Kodierung/Annotation (Quelle: Eigene Darstellung).

3.4 Qualitätssicherung

Qualitätssicherung ist sehr wichtig, um verlässliche Ergebnisse zu erhalten. Unser System bietet durch Mehrfachkodierung ausgewählter Artikel eine Möglichkeit die Qualität der Kodierungen kontinuierlich zu messen.

Dies erlaubt gegebenenfalls ein Eingreifen und Nachschulen von Kodierern, ebenso wird hierdurch die Qualität des Kodierhandbuchs geprüft. Unser System enthält hierfür eine Kodiermanagement-Einheit, welche die Zuteilung und Verwaltung der zu kodierenden Artikel übernimmt. Außerdem kann hiermit auch der Fortschritt der einzelnen Kodierer geprüft werden und je nach Anforderung unterschiedliche Ziele gesetzt werden.

Durch die Komplexität, die durch freie Kodierung beliebiger Textstellen entsteht, ist die Definition einer einfachen Metrik zur Bestimmung der Qualität schwierig. Unser System bietet einen Zwischenschritt mittels graphischer Oberfläche, die es erlaubt, die Kodierung in jedem Artikel visuell zu vergleichen. *Abbildung 7* illustriert eine solche Analyse. Im unteren Teil sind die einzelnen Kodierungen aufgetragen. Jede Zeile repräsentiert einen Kodierer bzw. eine Kodiererin und jedes farbliche Intervall zeigt eine Kodierung an. Die x-Achse beschreibt die Textposition (zeichenweise). Man sieht nun sehr schnell, welche Textstellen eine gute oder weniger gute Übereinstimmung haben. Per Mouse-over können zu jedem Intervall die kodierten Variablen angezeigt werden. Zusätzlich wird im Artikel, der in der oberen Hälfte sichtbar ist, die entsprechende Textstelle hervorgehoben.

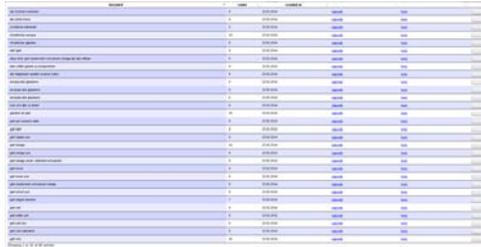


Abb. 7: Vergleichende Darstellung der Kodierungsergebnisse verschiedener Kodierer (Quelle: Eigene Darstellung).

Der Vorteil unserer Implementierung als Webanwendung ist, dass die Benutzer keine spezielle Software installieren müssen. Außerdem kann ein Kodierer standortunabhängig auf das System zugreifen und seine Arbeit jederzeit unterbrechen und wieder aufnehmen.

Die Einbindung bestehender Systeme ist ein weiterer Aspekt unseres Systems. CQPWeb bietet z.B. ein mächtiges Abfragesystem, um mit linguistisch annotierten Daten zu interagieren. Dies kann z.B. genutzt werden, um sich Kollokationen berechnen zu lassen: Tauchen bestimmte Wortkombinationen im Korpus (oder in einem geeignet eingeschränkten Subkorpus) häufiger auf, als dies aufgrund ihrer Basis-Verteilung zu erwarten wäre? *Abbildung 8* zeigt eine Wortliste, für die einfache frequenzbasierte Analysen durchgeführt und dann in CQPWeb weiterverarbeitet werden können. Kollokationsanalysen können abseits des linguistischen

Interesses an idiomatischen Wortverbindungen ausgesprochen hilfreich bei der Exploration eines Korpus sein: So lässt sich beispielsweise fragen, welche Adjektive im Kontext des Lemmas ›Europe‹ in einem Subkorpus britischer Zeitungen systematisch häufiger auftauchen als im US-amerikanischen Vergleichskorpus. Das Potenzial, das sich aus den linguistisch aufbereiteten Texten für weitere Analysen ergibt, ist bei weitem noch nicht erschöpft. Für eine sinnvolle Nutzung ist jedoch eine nicht-triviale Abstimmung der technischen Analysemöglichkeiten auf konkrete nicht-linguistische Fragestellungen nötig.



The screenshot shows a table with multiple columns, likely representing different linguistic or statistical metrics for various words or phrases. The text is small and difficult to read, but it appears to be a data visualization or a list of results from a computational linguistics tool.

Abb. 8: Screenshot zur Kollokationsextraktion (Quelle: Eigene Darstellung).

4. Schluss

Der vorgelegte Beitrag gibt einen Überblick über eine Werkzeugarchitektur zur Unterstützung der Vorbereitung und Erschließung großer Korpora für sozialwissenschaftliche Forschungsfragen. In den Aufbereitungsprozess werden computerlinguistische Methoden so eingebunden, dass linguistische Zwischenrepräsentationen ausgenutzt werden, ohne dass sich die beteiligten Fachwissenschaftler in ihre Details einarbeiten müssen. Der Ansatz fußt grundsätzlich auf generischen Werkzeugketten und ist damit übertragbar auf andere Projekte, unterstützt jedoch gleichzeitig eine sorgfältige Annotation bzw. Kodierung von studienspezifischen Konzepten, die u.a. als Basis für überwachte Lernverfahren genutzt werden kann.

Im Sinne eines transparenten Werkzeugangebots, das eine Sensibilität für mögliche Analysefehler erzeugt und aufrecht erhält, wird den Nutzern bei der Exploration und Hypothesenbildung immer der Bezug zu den tatsächlichen Textdaten und zwischengelagerten Analyseschritten vor Augen geführt.

Die Erfahrung der Projektkooperation zwischen den Computerlinguisten und den Politikwissenschaftlern hat gezeigt, dass Phasen des intensiven Austauschs zum jeweils fachspezifischen Verständnis zentraler Begriffe und Methoden von großer Bedeutung für die Effektivität der gemeinsamen Entwicklungsschritte sind. Ein Einsatz von vordefinierten Werkzeuglösungen kann nicht in der gleichen Weise in einen sozialwissenschaftlichen Erkenntnisprozess eingebunden werden wie die schrittweise Erarbeitung eines textanalytischen Werkzeuginventars, das auf die speziellen Schwierigkeiten und Möglichkeiten der Korpusituation und der verfolgten analytischen Fragestellung angepasst ist.

Bibliographische Angaben

- Melina Alexa / Cornelia Zuell: Text Analysis Software: Commonalities, Differences and Limitations: The Results of a Review. In: *Quality and Quantity* 34 (2000) H. 3, S. 299–321. [[Nachweis im OPAC](#)]
- Paul Baker / Tony McEnery: A Corpus-Based Approach to Discourses of Refugees and Asylum Seekers in UN and Newspaper Texts. In: *Journal of Language and Politics* 4 (2005) H. 2, S. 197–226. [[Nachweis im OPAC](#)]
- André Blessing / Jens Stegmann / Jonas Kuhn (2012a): SOA meets Relation Extraction: Less may be more in Interaction. In: *Proceedings of the Joint CLARIN-D/DARIAH Workshop at Digital Humanities Conference. Hamburg 2013*, S. 6–11. [[online](#)]
- André Blessing / Jonathan Sonntag / Fritz Kliche / Ulrich Heid / Jonas Kuhn / Manfred Stede (2012b): Towards a Tool for Interactive Concept Building for Large Scale Analysis in the Humanities. In: *Proceedings of the 7th Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities Association for Computational Linguistics. Sofia 2012*, S. 55–64. [[online](#)]
- André Blessing / Jonas Kuhn: Textual Emigration Analysis (TEA). In: *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14), European Language Resources Association (ELRA). Reykjavik 2014*, S. 2089–2093. [[Nachweis im GBV](#)]
- Bernd Bohnet: Very High Accuracy and Fast Dependency Parsing is not a Contradiction. In: *The 23rd International Conference on Computational Linguistics (COLING 2010). Beijing 2010*, S. 89–97. [[Nachweis im GBV](#)]
- William R. Brubaker / Frederick Cooper: Beyond »Identity«. In: *Theory and Society* 29 (2000) H.1, S. 1–47. [[Nachweis im GBV](#)]
- Emile Durkheim: *The Rules of Sociological Method*. New York 1964 [zuerst 1895]. [[Nachweis im GBV](#)]
- Constantinos Gabrielatos: Selecting Query Terms to Build a Specialised Corpus from a Restricted-Access Database. In: *ICAME Journal* 31 (2007), S. 5–43. [[Nachweis im GBV](#)]
- Richard K. Herrmann: Linking Theory to Evidence in International Relations. In: *Handbook of International Relations*. Hg. von Walter Carlsnaes / Thomas Risse / Beth A. Simmons. London 2002, S. 119–136. [[Nachweis im GBV](#)]
- Cathleen Kantner: *Kein Modernes Babel. Kommunikative Voraussetzungen Europäischer Öffentlichkeit*. Wiesbaden 2004. [[Nachweis im GBV](#)]
- Cathleen Kantner (2006a): Collective Identity as Shared Ethical Self-Understanding: The Case of the Emerging European Identity. In: *European Journal of Social Theory* 9 (2006) H. 4, S. 501–523. [[Nachweis im GBV](#)]
- Cathleen Kantner (2006b): Die thematische Verschränkung nationaler Öffentlichkeiten in Europa und die Qualität transnationaler politischer Kommunikation. In: *Demokratie in Der Mediengesellschaft*. Hg. von Kurt Imhof / Roger Blum / Heinz Bonfadelli / Otfried Jarren. Wiesbaden 2006, S. 145–160. [[Nachweis im GBV](#)]
- Cathleen Kantner: *Transnational Identity-Discourse in the Mass Media. Humanitarian Military Interventions and the Emergence of a European Identity (1990–2006)*. Berlin 2009.
- Cathleen Kantner / Amelie Kutter / Andreas Hildebrandt / Mark Püttcher: *How to Get Rid of the Noise in the Corpus: Cleaning Large Samples of Digital Newspaper Texts*. Stuttgart 2011. [[online](#)]
- Cathleen Kantner / Amelie Kutter / Swantje Renfordt: The Perception of the EU as an Emerging Security Actor in Media Debates on Humanitarian and Military Interventions (1990–2006). In: *RECON Online Working Paper 19 (2008). Oslo 2008*. [[online](#)]
- Fritz Kliche / André Blessing / Ulrich Heid / Jonathan Sonntag: The *elidentity* Text ExplorationWorkbench. In: *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14) European Language Resources Association (ELRA). Reykjavik 2014*, S. 691–697. [[Nachweis im GBV](#)]
- Thomas Koenig / Sabina Mihelj / John Downey / Mine Gencel Bek: Media Framings of the Issue of Turkish Accession to the EU. In: *Innovation: The European Journal of Social Sciences* 19 (2006) H. 2, S. 149–169. [[Nachweis im GBV](#)]
- Klaus Krippendorff: *Content Analysis: An Introduction to Its Methodology*. London 2004. [[Nachweis im GBV](#)]
- Amelie Kutter: Petitioner or Partner? Constructions of European Integration in Polish Print Media Debates on the EU Constitutional Treaty. In: *Discourse and Contemporary Social Change*. Hg. von Norman Fairclough / Giuseppina Cortese / Patrizia Ardizzone. Bern 2007, S. 433–457. [[Nachweis im GBV](#)]
- Ulrike Liebert: Introduction: Structuring Political Conflict About Europe: National Media in Transnational Discourse Analysis. In: *Perspectives on European Politics and Society* 8 (2007) H. 3, S. 236–260. [[Nachweis im GBV](#)]
- Corstin Mahlow / Kerstin Eckart / Jens Stegmann / André Blessing / Gregor Thiele / Markus Gärtner / Jonas Kuhn: Resources, Tools, and Applications at the CLARIN Center Stuttgart. In: *Proceedings der 12. Konferenz zur Verarbeitung natürlicher Sprache (KONVENS 2014). Hildesheim 2014*, S. 11–21. [[Nachweis im GBV](#)]
- Lutz Niethammer: *»Kollektive Identität«. Heimliche Quellen Einer Unheimlichen Konjunktur*. Reinbek bei Hamburg 2000. [[Nachweis im GBV](#)]

Swantje Renfordt: Framing the use of force : an international rule of law in media reporting ; a comparative analysis of western media debates about military interventions 1990 - 2005. Diss. Berlin 2009. [\[Nachweis im GBV\]](#)

Thomas Risse: A Community of Europeans? Transnational Identities and Public Spheres. Ithaca 2010. [\[Nachweis im GBV\]](#)

Helmut Schmid: Improvements in Part-of-Speech Tagging with an Application to German. In: Proceedings of the ACL SIGDAT-Workshop. Dublin 1995, S. 47-50. [\[online\]](#)

Udo Tietz (2002a): Die Grenzen des »Wir«. Eine Theorie Der Gemeinschaft. Frankfurt 2002. [\[Nachweis im GBV\]](#)

Udo Tietz (2002b): Gemeinwohl, Gemeinwohl Und Die Grenzen Des »Wir«. In: Gemeinwohl und Gemeinwohl. Hg. von Herfried Münkler / Harald Bluhm. Berlin 2002, S. 37-70. [\[Nachweis im GBV\]](#)

Udo Tietz: Les Limites Du »Nous« Libéral. In: Qu'est-Ce Qu'un Collectif? Du Commun À La Politique. Hg. von Laurence Kaufmann / Danny Trom. Paris 2010, S. 173-195. [\[online\]](#)

Kristina Toutanova / Dan Klein / Christopher Manning / Yoram Singer: Feature-Rich Part-of-Speech Tagging with a Cyclic Dependency Network. In: Proceedings of HLT-NAACL. Edmonton, Kanada 2003, S. 252-259. [\[Nachweis im GBV\]](#)

Abbildungslegenden und -nachweise

Abb. 1: Schematische Darstellung der Dokumentenverarbeitungskette (Quelle: Eigene Darstellung).

Abb. 2: Rohe Textdaten, wie sie auf digitalen Medienportalen bereitgestellt werden (Quelle: Eigene Darstellung).

Abb. 3: Ein Ausschnitt von rohen Textdaten wird in einem Vorschaufenster angezeigt. Die Anwender definieren eine Segmentierungsregel, um die Rohdaten in textstrukturelle Einheiten (»Artikel«) zu segmentieren. Im Beispiel trennen Zeilen, die mit der Zeichenfolge kf00 beginnen, die einzelnen Artikel (Quelle: Eigene Darstellung).

Abb. 4: Computerlinguistische Verarbeitungsschritte, die im Hintergrund der Dokumentenaufbereitung ablaufen (Quelle: Eigene Darstellung).

Abb. 5: Illustration des redundanten Einsatzes mehrerer Part-of-Speech-Tagger (Quelle: Eigene Darstellung).

Abb. 6: Die Benutzerschnittstelle für die politikwissenschaftliche Kodierung/Annotation (Quelle: Eigene Darstellung).

Abb. 7: Vergleichende Darstellung der Kodierungsergebnisse verschiedener Kodierer (Quelle: Eigene Darstellung).

Abb. 8: Screenshot zur Kollokationsextraktion (Quelle: Eigene Darstellung).

Zeitschrift für digitale Geisteswissenschaften

Artikel aus:

Sonderband 1 der ZfdG: Grenzen und Möglichkeiten der Digital Humanities. Hg. von Constanze Baum und Thomas Stäcker. 2015. DOI: [10.17175/sb01](https://doi.org/10.17175/sb01)

Titel:

Neue Perspektiven für die Erziehungswissenschaften – Das Potenzial von Mixed Reality für Lehr- und Lernprozesse

Autor/in:

Katharina Schuster

Kontakt: katharina.schuster@ima-zlw-ifu.rwth-aachen.de

Institution: RWTH-Aachen

GND: [1076754309](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63862-p0011-9) ORCID:

Autor/in:

Anja Richert

Kontakt: anja.richert@ima-zlw-ifu.rwth-aachen.de

Institution: RWTH-Aachen

GND: [136938175](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63862-p0011-9) ORCID:

Autor/in:

Sabine Jeschke

Kontakt: sabina.jeschke@ima-zlw-ifu.rwth-aachen.de

Institution: RWTH-Aachen

GND: [128977051](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63862-p0011-9) ORCID:

DOI des Artikels:

[10.17175/sb001_014](https://doi.org/10.17175/sb001_014)

Nachweis im OPAC der Herzog August Bibliothek:

[83020413X](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63862-p0011-9)

Erstveröffentlichung:

19.02.2015

Lizenz:

Sofern nicht anders angegeben 

Medienlizenzen:

Medienrechte liegen bei den Autoren

Letzte Überprüfung aller Verweise:

24.05.2016

GND-Verschlagwortung:

[Pädagogik](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63862-p0011-9) | [Mixed Reality](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63862-p0011-9) | [Mensch-Maschine-Kommunikation](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63862-p0011-9) |

Zitierweise:

Katharina Schuster, Anja Richert, Sabine Jeschke: Neue Perspektiven für die Erziehungswissenschaften – Das Potenzial von Mixed Reality für Lehr- und Lernprozesse. In: Grenzen und Möglichkeiten der Digital Humanities. Hg. von Constanze Baum / Thomas Stäcker. 2015 (= Sonderband der Zeitschrift für digitale Geisteswissenschaften, 1). PDF Format ohne Paginierung. Als text/html abrufbar unter DOI: [10.17175/sb001_014](https://doi.org/10.17175/sb001_014).

Katharina Schuster, Anja Richert, Sabine Jeschke

Neue Perspektiven für die Erziehungswissenschaften – Das Potenzial von Mixed Reality für Lehr- und Lernprozesse

Abstracts

›Mixed Reality‹ beschreibt die Kombination aus virtuellen Umgebungen und natürlichen Benutzerschnittstellen. Das Sichtfeld der nutzenden Person wird hierbei über Head Mounted Displays durch Kopfbewegung gesteuert, Datenhandschuhe ermöglichen die Interaktion mit virtuellen Objekten und omnidirektionale Böden erlauben eine unbegrenzte Fortbewegung durch virtuelle Welten mittels natürlichen Gehens. In einer quasi-experimentellen Untersuchung konnte gezeigt werden, dass ›Mixed Reality‹ sich vor allem auf emotionaler und motivationaler Ebene positiv auswirkt. Die Ergebnisse werden vor dem Hintergrund der Frage interpretiert, inwiefern sich durch derart innovative Technologien grundlegend neue Möglichkeiten für die Erziehungswissenschaften ergeben.

›Mixed Reality‹ describes the combination of virtual environments and natural user interfaces. Here, the user's field of vision is controlled by natural head movements via a head mounted display. Data gloves allow direct interaction with virtual objects and omnidirectional floors enable unrestricted navigation through virtual worlds with natural walking movements. A quasi-experimental study showed that ›mixed reality‹ had a positive effect, particularly on affective and motivational levels. The results are interpreted in the context of the new possibilities for pedagogy offered by such innovative technologies.

1. Einführung – Neue Perspektiven für die Erziehungswissenschaften

»Das Selbstverständnis der deutschsprachigen Erziehungswissenschaft hat sich in den letzten dreißig Jahren geändert und ändert sich weiterhin. Ein solcher Wandel kann – auf den ersten Blick – verstanden werden als ein Hinweis darauf, dass sich die Disziplin neuen Herausforderungen stellt, sich neuen Fragen öffnet, diese mit neuen Methoden der Erkenntnisgewinnung untersucht und neue Lösungen entwickelt.«¹

Diese einleitenden Worte im Vorwort eines Beiheftes der Zeitschrift für Pädagogik mit dem Titel *Das Selbstverständnis der Erziehungswissenschaft: Geschichte und Gegenwart* deuten auf einen vielschichtigen Paradigmenwechsel hin, dem sich die Wissenschaftsdisziplin unterzieht. In den 1960er und 1970er Jahren markierte die Ergänzung des geisteswissenschaftlich-hermeneutisch geprägten Selbstverständnisses um eine empirisch-sozialwissenschaftliche Dimension unter dem Schlagwort der ›realistischen Wende‹ einen ersten methodischen Wendepunkt für die Erziehungswissenschaft.² Eine Öffnung für interdisziplinäre Zusammenarbeit, u.a. mit der pädagogischen Psychologie, der Soziologie, der Ökonomie

¹ Fatke / Oelkers 2014, S. 7.

² Vgl. Fatke / Oelkers 2014, S. 8.

oder den Neurowissenschaften, war die Folge.³ Auch wenn sich die im zitierten Vorwort genannten Herausforderungen und neuen Fragen eher auf eine thematisch-inhaltliche Dimension beziehen⁴, können sie in methodischer Hinsicht auch vor dem Hintergrund der Digital Humanities interpretiert werden.

So geht es im Sinne der Digital Humanities darum, in und für die Geisteswissenschaften »die Prozesse der Gewinnung und Vermittlung neuen Wissens unter den Bedingungen einer digitalen Arbeits- und Medienwelt weiter zu entwickeln«⁵. Der analytische Mehrwert digitaler Werkzeuge für die Geisteswissenschaften liegt dabei nicht nur in neuen Möglichkeiten der Auswertung, wie z.B. Learning Analytics, Data- oder Text-Mining Verfahren. Durch die Entwicklung neuer Informationstechnologien werden völlig neuartige Zugänge zu bekannten Untersuchungsgegenständen ermöglicht – und dem stetigen wissenschaftlichen Fortschritt entsprechend auch neue Untersuchungsgegenstände aufgedeckt. Für die Didaktik und insbesondere die Mediendidaktik spielen dabei auf der Softwareseite multimediale Wissensräume wie virtuelle 3D-Welten eine große Rolle, die sich im Schul- und Hochschulkontext immer größerer Beliebtheit erfreuen.

Seit den 1990er Jahren werden Computerspielsimulationen für Lern- und Trainingszwecke entwickelt und eingesetzt, um komplexe technische Systeme oder den Umgang mit ökonomischen, ökologischen oder sozialen Systemen zu erlernen (z.B. Unternehmensplanung, Klimaentwicklung, Stadtplanung).⁶ Anwendungen reichen von der Simulation prozeduraler Vorgänge im medizinischen Bereich über den Erwerb von Kenntnissen zur Führung eines Fahr- oder Flugzeuges bis hin zur Planung logistischer Prozesse in virtuellen Fabriken. Durch simulierte Verlangsamung oder Beschleunigung von Prozessen oder durch die Simulation von Zukunftsszenarien und historischen Ereignissen lässt sich durch Virtualität nicht nur die Raum- sondern auch die Zeitgrenze durchbrechen.⁷ Unter dem Begriff »Simulation« werden von verschiedenen Autoren unterschiedliche Lern- und Trainingsszenarien subsumiert:

- Visualisierungen komplexer theoretische Konstrukte (z.B. der Mathematik), die aus verschiedenen Perspektiven betrachtet und verändert werden können⁸,
- Situationen, in denen Fehlverhalten zu gesundheitlichen Schäden oder schlimmstenfalls zum eigenen Tod bzw. zum Tod anderer Menschen führt, z.B. durch einen Störfall in einem Atomkraftwerk⁹ oder im medizinischen Bereich während einer Operation¹⁰,
- Situationen, die fast nie eintreten und dennoch möglich sind¹¹, z.B. Evakuierungen bei Massenpaniken,

³ Vgl. Fatke / Oelkers 2014, S. 9.

⁴ Vgl. Fatke / Oelkers 2014, S. 12.

⁵ Sahle et al. 2011, S. 4.

⁶ Vgl. Kerres 2012, S. 12.

⁷ Vgl. Höntzsch 2013, S. 2.

⁸ Vgl. Jeschke 2004, S. 172.

⁹ Vgl. Ewert et al. 2013, S. 5.

¹⁰ Vgl. Flin et al. 2010, S. 40.

¹¹ Vgl. Höntzsch 2013, S. 2.

- Situationen, deren Training mit hohen Kosten verbunden ist, da bspw. teures Equipment zum Einsatz kommt oder hoher Materialverschleiß mit dem Training einhergeht¹²,
- Situationen, die in der Vergangenheit oder in der Zukunft liegen, wie z.B. ein Spaziergang durch das alte Pompeji oder auf dem Mars¹³,
- Situationen, in denen nicht sichtbare Wirkungszusammenhängen und Zustände sichtbar gemacht werden sollen, z.B. Luftströme in einem Gebäude¹⁴.

Für die Anwendung müssen zunächst oben beschriebene Szenarien in virtuellen Umgebungen nachgebaut werden. In Simulationen oder Lernspielen (Serious Games) können die Eigenschaften und Prinzipien mit entsprechenden Inhalten verknüpft und dadurch nachvollziehbar gemacht werden. In den Natur- und Ingenieurwissenschaften dienen virtuelle Labore dem zeit- und ortsunabhängigen Experimentieren und somit dem insbesondere in der Studieneingangsphase so wichtigen Aufbau eines wissenschaftlichen Selbstverständnisses der Studierenden. Wie die obige Aufzählung zeigt, kann der Einsatz virtueller Lernumgebungen verschiedenen Stufen des Lernens dienen. Dies betrifft erstens die Exploration fiktiver Welten, die im physisch Realen nicht umsetzbar sind. Zweitens unterstützen anspruchsvolle Visualisierungstechniken den initialen Verstehensprozess komplexer Lerninhalte und bieten durch Interaktionsmöglichkeiten eine aktive Verarbeitung der Informationen auf unterschiedlichen mentalen Ebenen. Drittens stellt eine virtuelle Lernumgebung häufig die »vorletzte Stufe des Wissenstransfers« dar, die sich in der Übung realer Handlungen im virtuellen Raum manifestiert, bevor diese im physischen realen Kontext angewendet werden müssen. Wie eine bestimmte virtuelle Umgebung erlebt und rezipiert wird, hängt jedoch auch entscheidend von der jeweils eingesetzten Hardware ab.

Ein Nachteil von Simulationen in Bezug auf den Wissenstransfer besteht häufig in der physischen Künstlichkeit der Lernerfahrung. Die Stelle oder Handlung, mit der ein Mensch mit einer Maschine in Kontakt tritt, wird generell als Benutzerschnittstelle (Human Machine Interface, HMI) bezeichnet.¹⁵ Normalerweise interagieren Nutzerinnen und Nutzer mit einer virtuellen Umgebung über einen Personal Computer. Das virtuelle Szenario wird auf einem Monitor angezeigt und das Sichtfeld üblicherweise durch eine Tastatur oder eine Maus gesteuert. Interaktionen mit virtuellen Objekten sowie Fortbewegung erfolgen meist über die gleichen Benutzerschnittstellen. Auf diese Weise überträgt der Nutzer oder die Nutzerin den Kontrollmodus des Computers auf die Aktionen seiner grafischen Repräsentation, auf den Avatar.

Dass Mensch und Computer durch Sprache und Gestik miteinander kommunizieren, ist schon lange nicht mehr der Welt von Science-Fiction-Filmen überlassen. Völlig selbstverständlich werden inzwischen Smartphones und Tablet-Computer mit Wischgesten auf dem Multi-Touch-Display bedient. Das Blättern in Fotoalben, Scrollen in langen Artikeln, Zoomen in bestimmte Bildausschnitte oder Webseiten erfolgt nicht mehr durch Anklicken

¹² Vgl. Ewert et al. 2013, S. 1.

¹³ Vgl. Hoffmann et al. 2013, S. 46.

¹⁴ Vgl. Willenbrock 2012, S. 89.

¹⁵ Vgl. Groner et al. 2008, S. 427.

entsprechender Symbole auf dem Bildschirm sondern durch ›Tippen‹, ›Wischen‹ oder ›Finger auseinanderziehen‹. Computer sind darüber hinaus in der Lage, Körperhaltungen, Bewegungen und Gesten des Menschen zu erkennen, in Befehle zu übersetzen und sie auszuführen. Dies kann sowohl gerätebasiert, z. B. über Datenhandschuhe, oder kamerabasiert geschehen, wie bei der Hardwarekomponente Kinect für die Spielkonsole **Xbox**. Diese Form der Mensch-Computer-Interaktion wird durch das Konzept der natürlichen Benutzerschnittstellen beschrieben (Natural User Interfaces, NUI). In Kombination mit virtuellen Umgebungen werden natürliche Benutzerschnittstellen der sogenannten gemischten Realität zugeordnet, die auch im Deutschen häufig als ›Mixed Reality‹ bezeichnet wird.

Natürliche Benutzerschnittstellen für Visualisierung, Navigation und Interaktion können im Vergleich zu einem feststehenden Rechner oder zu einem Laptop eine weitgehend authentische Lernerfahrung bieten. Bekannte Beispiele sind Flug- und Fahr simulatoren oder bestimmte Spielkonsolen im Unterhaltungsbereich. Ziel der natürlichen Benutzerschnittstellen ist es immer, die Mensch-Computer-Interaktion intuitiver zu gestalten. Bei solchen Geräten stimmt somit die physisch reale Bewegungshandlung mit der virtuell ausgeführten Bewegungshandlung überein.

Der Nutzer oder die Nutzerin soll darüber hinaus den Eindruck haben, regelrecht in die virtuelle Welt eintauchen zu können. In diesem Zusammenhang fällt häufig der Begriff der ›Immersion‹. Als Einstieg sei Immersion zunächst als eine Kategorie der Wirkung von etwas betrachtet, das die Aufmerksamkeit von jemandem beansprucht, der sich auf dieses Etwas konzentriert und einlässt.¹⁶ Etymologisch verweist das lateinische Wort *immersio* auf eine physische Erfahrung des Ein- bzw. Untertauchens in Flüssigkeit. Als Phänomen einer intensiven Grenzerfahrung zwischen zwei Realitätsmodi beschreibt Murray 1999 Immersion als

»a metaphorical term derived from the physical experience of being submerged in water. We seek the same feeling from a psychologically immersive experience that we do from a plunge in the ocean or swimming pool: the sensation of being surrounded by a completely other reality, as different as water is from air, that takes over all of our attention, our whole perceptual apparatus«¹⁷.

Neben der häufig genannten Analogie des Eintauchens bringt die Betrachtung des Immersionsbegriffs in einer Vielzahl wissenschaftlicher Disziplinen entsprechend unterschiedliche Bedeutungsfacetten hervor. Immersive Zustände können dabei sowohl in fiktionalen Welten als auch in nicht-narrativen virtuellen Umgebungen auftreten. Oft ist es das erklärte Ziel von Entwicklern virtueller Welten oder natürlicher Benutzerschnittstellen, den Eindruck des ›Eintauchens‹ durch bestimmte technische Eigenschaften zu unterstützen. Dementsprechend ist in solchen Fällen von immersiven Lernumgebungen oder immersiven Benutzerschnittstellen die Rede. Aus technischer Sicht benötigt der Nutzer oder die Nutzerin für eine erhöhte Immersion eine nahtlose 3D-Sicht der virtuellen Umwelt. Diese wird häufig durch Head Mounted Displays (HMDs) realisiert. Für eine natürliche Navigation in der virtuellen

¹⁶ Vgl. Curtis / Voss 2008, S. 3.

¹⁷ Murray 1999, S. 98f.

Umgebung können omnidirektionale Laufbänder verwendet werden, die eine freie und unbegrenzte Bewegung ermöglichen und die nutzende Person doch an einem physisch klar begrenzten Ort lassen. Durch Datenhandschuhe kann der Nutzer bzw. die Nutzerin intuitiv mit der virtuellen Umgebung sowie den Objekten, die sich in ihr befinden, interagieren.

Für die Untersuchung und die Messbarkeit immersiver Erfahrungen ist jedoch eine Spezifizierung des Konstrukts vonnöten. Der Begriff des Präsenzerlebens verweist auf das Phänomen des räumlichen Eintauchens in eine virtuelle Umgebung und wird relativ einheitlich als ›sense of being there‹ beschrieben. Im Diskurs über den Einsatz virtueller Welten zu Lernzwecken taucht außerdem immer wieder die Frage auf, welche möglicherweise unterstützende Wirkung Immersion auf Lernprozesse haben könnte.¹⁸ In diesem Zusammenhang wird Immersion häufig mit dem Konstrukt des Flow-Erlebens in Verbindung gebracht¹⁹:

»Immersion bezeichnet ein Eintauchen und bezieht sich, anders als der eher auf Fähigkeiten bezogene Flow, auf die Fokussierung von Aufmerksamkeit auf die medialen Inhalte [...] Genau wie beim Spielen ist auch beim Lernen ein solches Eintauchen bzw. Aufgehen in der Aufgabe der Idealzustand zur Erreichung der Lernziele.«²⁰

Sogenannte immersive Lernumgebungen bezeichnen meist selbst erstellte virtuelle Räume, die in kommerzielle Online-3D-Infrastrukturen wie **Second Life** oder entsprechende Open Source-Varianten wie **OpenSim** eingebettet sind.²¹ Die technischen Möglichkeiten, wie z. B. multimodale Informationsaufnahme oder Interaktionsmöglichkeiten mit anderen Nutzern oder Inhalten, werden häufig gleichgesetzt mit dem beabsichtigten Erlebensaspekt der Immersion.²² Positive Effekte von Immersion auf Lernleistung und/oder -motivation gehen oft nicht über Annahmen hinaus.²³ Gebräuchliche Formulierungen wie ›immersives Lernen‹, ›immersive virtuelle Welten‹ oder ›immersive Umgebungen‹ suggerieren, dass der beabsichtigte Effekt automatisch eintritt.

Aufgrund der Möglichkeit, virtuelle Lernumgebungen durch natürliche Bewegungen zu steuern, liegt das große Potenzial natürlicher Benutzerschnittstellen für die Erziehungswissenschaft im praktisch orientierten Teil der Mediendidaktik. Genauer lautet die zu beantwortende Frage, ob und in welcher Form Schüler und/oder Studierende in Mixed Reality-Szenarien besser lernen können als in rein virtueller oder in rein physischer Realität. Bevor natürliche Benutzerschnittstellen flächendeckend zu Lern- und Trainingszwecken zum Einsatz kommen, sind wissenschaftliche Erkenntnisse über die Wahrnehmung der Rezeptionssituation und über den tatsächlichen Lernerfolg notwendig. Deswegen wird die weitere Entwicklung der Hardware von unterschiedlichen mediendidaktischen und

¹⁸ Vgl. Witmer / Singer 1998, S. 238; Höntzsch et al. 2013, S. 3.

¹⁹ Vgl. Höntzsch et al. 2013, S. 3; Breuer 2010, S. 11.

²⁰ Breuer 2010, S. 11.

²¹ Vgl. Höntzsch et al. 2013, S. 2; Hebbel-Seeger 2012, S. 29; Bredl / Groß 2012, S. 37; Zender / Lucke 2012, S. 47; Tarouco et al. 2013, S. 774.

²² Vgl. Höntzsch et al. 2013, S. 4, Bredl / Groß 2012, S. 37.

²³ Vgl. Witmer / Singer 1998, S. 238, Bredl / Groß 2012, S. 37.

medienspsychologischen Studien begleitet, von denen eine im Folgenden näher beschrieben wird. Darüber hinaus stellt sich nach wie vor die Frage, welche Rolle Immersion beim Lernen in virtuellen Welten spielt. Der Grundsatz der Digital Humanities, »die Prozesse der Gewinnung und Vermittlung neuen Wissens unter den Bedingungen einer digitalen Arbeits- und Medienwelt weiter zu entwickeln«²⁴, lässt sich vor diesem Hintergrund auf die Motivation übertragen, das Potenzial von Mixed Reality für den Einsatz in der Lehre zu erforschen.

2. Quasi-experimentelle Untersuchungen zum Potenzial von Mixed Reality für Lehr- und Lernprozesse

2.1 Apparat: Der Mixed Reality-Simulator ›Virtual Theatre‹

Ziel der hier beschriebenen Studie war es zu untersuchen, welchen Einfluss die Benutzerschnittstellen, mit denen eine bestimmte Lernumgebung rezipiert wird, auf Rezeptionsprozesse der Informationsverarbeitung und des Erlebens haben. Um die Effekte natürlicher Benutzerschnittstellen auf Rezeptionsprozesse präzise bestimmen zu können, wurde ein Zweigruppenplan erstellt.²⁵ Der Experimentalgruppe wurde als maßgebliches Treatment die Nutzung eines Mixed-Reality-Simulators (›Virtual Theatre‹) zugewiesen, während die Kontrollgruppe durch Nutzung eines Laptops definiert war. Die Lernumgebung wurde über die Experimental- und die Kontrollgruppe konstant gehalten. Die Daten wurden quantitativ erhoben und statistisch ausgewertet. Bezüglich dieser Auswertung und dementsprechend der Interpretation der Ergebnisse muss jedoch von vornherein eine Einschränkung vorgenommen werden. Der Einfluss der benutzten Hardware auf Rezeptionsprozesse kann nur mit einem Set aus unabhängigen Variablen gemessen werden, welches sich aus der Kombination mehrerer Benutzerschnittstellen zu einem Mixed Reality-Simulator ergibt. Somit ließen sich nicht alle potentiellen Störfaktoren vollständig eliminieren. Weiterhin ließ sich die Zuordnung der Studienteilnehmer zu den beiden Gruppen nicht vollständig randomisieren. Aus diesen Gründen wurde für die Studie die Bezeichnung ›quasi-experimentell‹ gewählt.²⁶

Das in der Studie eingesetzte Virtual Theatre ist ein Mixed Reality-Simulator, der die natürlichen Benutzerschnittstellen eines Head Mounted Displays und eines omnidirektionalen Bodens miteinander vereint. Abbildung 1 zeigt die Nutzung des Virtual Theatres sowie ein exemplarisches Lernszenario aus dem Gebiet der Ingenieurwissenschaften.



Abb. 1: Das Virtual Theatre mit Nutzerin und (vereinfachter) virtueller Umgebung (Foto: Walter Spanjersberg, 2013).

²⁴ Sahle et al. 2011, S. 4.

²⁵ Vgl. Bortz / Döring 2006, S. 523.

²⁶ Vgl. Cook / Campbell 1979, S. 6; Kelle 2007, S. 208.

Das Virtual Theatre kann ausschließlich durch eine einzelne Person genutzt werden, die durch ihre natürlichen Körperbewegungen die Interaktion mit dem System steuert. Die das Virtual Theatre nutzende Person kann sich frei bewegen, indem sie die gewünschte Richtung ansteuert. Der omnidirektionale Boden besteht aus 16 trapezförmigen Elementen. Diese sind aus je 16 festen Rollen zusammengesetzt, die nach außen hin breiter werden. Die Elemente haben an der kurzen Grundseite einen gemeinsamen Ursprung und sind im Mittelpunkt über eine fixierte Plattform miteinander verbunden. Auf der Plattform in der Mitte des Bodens kann sich die nutzende Person bewegen, ohne die Bodenbewegung zu starten. Das Virtual Theatre ermöglicht dem Nutzer außerdem eine nahtlose 3D-Visualisierung der virtuellen Umgebung mit einem stereoskopischen 70° Sichtfeld. Jede Kopfbewegung wird in Echtzeit in der Simulation widergespiegelt.

Das Head Mounted Display und der omnidirektionale Boden können wahlweise mit einem Datenhandschuh oder einem Hand-Tracer kombiniert werden, falls das Szenario keine direkte Objektinteraktion erfordert. An diesem etwa handgroßen Metallkreuz sind vier Infrarotmarker befestigt. *Abbildung 2* zeigt den Datenhandschuh, den Hand-Tracer sowie die entsprechenden Darstellungen der Handbewegung in einer virtuellen Umgebung.



Abb. 2: Datenhandschuh, Hand-Tracer sowie entsprechende Darstellungen in einer virtuellen Umgebung (Fotos: Peer Schlieperskötter, 2014).

Der Hand-Tracer reagiert nicht auf Handgesten, wird jedoch für ein korrektes Positions-Tracking benötigt. Für die im Folgenden beschriebene Studie wurde der Einsatz des Hand-Tracers gewählt. Dieser erlaubt es, eine einzelne mögliche Einflussvariable auf die abhängigen Variablen, die Möglichkeit zur gestengesteuerten Objektinteraktion, zu isolieren. Aufgrund der Wahl des Hand-Tracers als Benutzerschnittstelle wird im Folgenden auch nur noch auf diesen verwiesen.

Um die Position der nutzenden Person in Echtzeit in die Virtualität zu übersetzen, ist das Virtual Theatre mit zehn Infrarotkameras ausgestattet. Die Kameras zeichnen die Positionen der Infrarot-Marker auf, die an dem HMD und dem Hand-Tracer befestigt sind.

2.2 Ablauf der Studie

Im Hinblick auf natürliche Benutzerschnittstellen ist nach wie vor die Frage offen, unter welchen Bedingungen virtuelle Umgebungen zum besten Lernergebnis führen und noch weiter gefragt, inwiefern Immersion und Lernerfolg miteinander zusammenhängen. Der

medienspsychologischen Herangehensweise folgend, ist hierbei die Betrachtung von fünf Faktoren wichtig²⁷:

1. Charakteristika des Medienangebots,
2. der Person,
3. der Situation,
4. des Rezeptionsprozesses, sowie
5. der kurz- und langfristigen Wirkung.

Bei der vorgestellten Studie zum Einfluss natürlicher Benutzerschnittstellen auf Rezeptionsprozesse der Informationsverarbeitung und des Erlebens in virtuellen Lernumgebungen wurden die subjektive Erfahrung des Präsenzerlebens²⁸ und des Flow-Erlebens²⁹ als zentrale Kenngrößen für Immersion gemessen. Um weitere Einblicke in die Rezeptionssituation beim Lernen mit und Erleben von Mixed Reality zu erhalten, wurden situative Emotionen ebenfalls gemessen.

Eine der wichtigsten Fragen innerhalb eines Bildungs- und Trainingskontextes ist, ob die Erfahrung in einer virtuellen Umgebung über natürliche Benutzerschnittstellen zu einem besseren Lernergebnis führt als über einen die klassischen Benutzerschnittstellen wie Maus und Tastatur eines Personal Computers. Leistungsabfrage ist deshalb eine weitere wichtige Komponente des Studiendesigns. Die erwartete Beziehung zwischen Medienangebot, Person, Situation, Rezeptionsprozessen und Wirkung ist in *Abbildung 3* dargestellt.

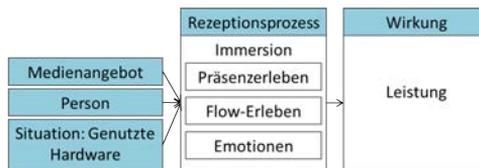


Abb. 3: Erwarteter Zusammenhang zwischen Medienangebot, Person, Situation, Rezeptionsprozessen und Wirkung (Grafik: Katharina Schuster, 2013).

Um den Möglichkeiten des Virtual Theatres gerecht zu werden, wurde als Untersuchungsszenario eine Umgebung ausgewählt, die selbstgesteuertes, exploratives Erlernen räumlicher Strukturen ermöglicht.³⁰ Gewählt wurde ein groß-skaliertes Labyrinth in einer Fabrikhalle. Der virtuelle Schauplatz der Fabrikhalle erschien wichtig, damit die Akustik der physisch realen Halle, in der das Virtual Theatre aufgebaut ist, nicht zu Irritationen in Kombination mit der Darstellung der virtuellen Umgebung führt. Im Labyrinth wurden elf verschiedene Objekte platziert.

²⁷ Vgl. Wirth / Hofer 2008, S. 161.

²⁸ Vgl. Vorderer et al. 2004, S. 3.

²⁹ Vgl. Rheinberg et al. 2003, S. 271.

³⁰ Vgl. Köhler et al. 2008, S. 481.

Der Versuchsablauf bestand aus insgesamt acht Teilen und dauerte etwa eine Stunde. Die Studie wurde von insgesamt drei geschulten Versuchsleiterinnen durchgeführt. Zur Standardisierung des Versuchsablaufs war zuvor eine umfangreiche Versuchsanordnung verfasst worden. **Abbildung 4** zeigt den Gesamt Ablauf am Beispiel der Experimentalbedingung.

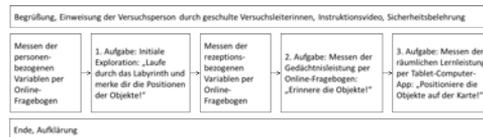


Abb. 4: Versuchsablauf des ›Quasi-Experiments‹ am Beispiel der Experimentalbedingung (Grafik: Katharina Schuster, 2013).

Auch die Aufgaben wurden so gewählt, dass sie der didaktischen Methode des explorativen Lernens entsprechen. Die Anweisung der initialen Explorationsaufgabe lautete, frei durch das Labyrinth zu navigieren und sich die Positionen der Objekte zu merken. Die Probanden erhielten keine Informationen über die Anzahl der aufzufindenden Objekte. So wurde zwar das Lernziel vorgegeben (»Merke dir die Positionen der Objekte«), jedoch nicht die zur Lösung des Problems möglichen Operatoren, wie z. B. verschiedene Orientierungs- oder Merkstrategien. Einer typischen Eigenschaft von E-Learning-Angeboten entsprechend, konnte so jeder Proband und jede Probandin unter Einschränkung der vorgegebenen Maximaldauer im eigenen, für sich am besten geeigneten Lerntempo bleiben und seine oder ihre kognitiven Fähigkeiten selbstgesteuert einsetzen. Beide Gruppen erhielten zuvor die Möglichkeit, für drei Minuten ein Testszenario (eine italienische Piazza) frei zu erkunden. Dies diente dazu, die Probanden an den jeweiligen Steuerungsmodus zu gewöhnen. Zur freien Exploration sowie zum Einprägen der Positionen der Objekte hatten die Probanden acht Minuten Zeit. Um ihnen einen Orientierungspunkt zu geben, wurde ihnen zu Beginn der ersten Aufgabe durch die Versuchsleitung mitgeteilt, dass sie mit Blick nach Norden starten.

Direkt im Anschluss füllten die Probanden am Tablet-Computer den zweiten Teil des Online-Fragebogens aus, mit dem die abhängigen Variablen des Erlebens abgefragt wurden. Zuletzt erfolgte die Leistungsmessung. Die Art der Leistungsmessung wurde über beide Gruppen konstant gehalten. Der Tablet PC (Nexus 10) wurde gewählt, damit sich das Medium der Wissensabfrage sowohl für die Experimental- als auch für die Kontrollgruppe hinreichend vom Medium der Aufgabe unterscheidet.

Für die Leistungsmessung wurden sowohl Aufgaben im Low-Level als auch im High-Level-Bereich gewählt.³¹ Die zweite Aufgabe (Low-Level) lautete, die zuvor im Labyrinth gesehenen Objekte unter verschiedenen, dargebotenen Objekten wiederzuerkennen. Für die dritte Aufgabe (High-Level) mussten die Probanden die Objekte, die sie zuvor im Labyrinth gesehen hatten, auf einem Tablet-Computer über eine direkte Zeigebewegung mit dem Finger per Drag-and-Drop-Funktion an der korrekten Stelle auf einer Karte des Labyrinths positionieren. Für jedes abgelegte Objekt ertönte als auditives Feedback ein Signalton. Während der Anwendung konnten die Objekte immer wieder verschoben werden. Die Applikation (im Folgenden als

³¹ Vgl. Nash et al. 2000, S. 3.

›App‹ bezeichnet) für den Tablet-Computer wurde speziell für diese Studie programmiert. Die Ansicht des Labyrinths in der initialen Explorationsaufgabe und der Zuordnungsaufgabe ist in Abbildung 5 dargestellt.



Abb. 5: Ansicht der virtuellen Umgebung in der ersten und dritten Aufgabe (Foto: Walter Spanjersberg, 2013).

Im Rahmen eines ersten Pre-Tests ($n = 16$) wurde der Gesamtablauf erprobt. Dies diente gleichzeitig als praktische Schulung der Versuchsleiterinnen, nachdem diese zuvor eine theoretische und technische Einführung erhalten hatten. Weiterhin wurde der Fragebogen vom Paper & Pencil-Format auf ein Online-Befragungstool (Sosci-Survey) umgestellt. Im Rahmen eines zweiten Pre-Tests ($n = 18$) wurde der Ablauf erneut getestet. Nach dem zweiten Pre-Test wurde der Fragebogen im Bereich der demografischen Angaben angepasst. Es wurden keine weiteren Änderungen vorgenommen.

2.3 Variablen und Messinstrumente

Im Folgenden werden die einzelnen Variablen des Quasi-Experiments sowie die dafür ausgewählten Messinstrumente näher beschrieben. Die Experimentalbedingung der Nutzung des Virtual Theatres unterschied sich von der Kontrollbedingung der Nutzung eines Laptops hinsichtlich der Ein- und Ausgabegeräte, der Steuerungsmodi und der daraus folgenden physischen Konsequenzen. Die erste unabhängige Variable fasst das Variablenset der Rezeptionssituation zusammen. Abbildung 6 zeigt eine Übersicht der Unterschiede zwischen den Versuchsbedingungen.

Bedingung 1: Virtual Theatre	Bedingung 2: Laptop
Direkte Steuerung des Sichtfelds durch natürliche Kopfbewegung in Interaktion mit HMD.	Indirekte Steuerung des Sichtfelds durch Maus.
Direkte Steuerung der Fortbewegung durch natürliche Gehbewegung in Interaktion mit omnidirektionalem Boden.	Indirekte Steuerung der Fortbewegung durch die Tasten W (vorwärts), A (links), S (rückwärts) und D (rechts) auf der Laptopastatur.
Visuelle Ausgabe: HMD	Visuelle Ausgabe: Bildschirm des Laptops
Zusätzliche Schnittstelle: Hand-Tracer	Keine zusätzliche Schnittstelle
Körperhaltung: Aufrecht stehend bzw. gehend	Körperhaltung: Sitzend
Instruktionsvideo	Kein Instruktionsvideo
Sicherheitshinweise zur Benutzung der Hardware.	Keine Sicherheitshinweise zur Benutzung der Hardware.

Abb. 6: Übersicht: Unterschiede beider Versuchsbedingungen des ›Quasi-Experiments‹ (Grafik: Katharina Schuster, 2014).

Neben demografischen Angaben wurde als personenbezogene, unabhängige Variable das domänenspezifische Interesse mit der Subskala des MEC-Spatial Presence Questionnaire³² gemessen. Das Konstrukt bezieht sich ursprünglich auf die inhaltliche Ebene eines Medienangebots, wie z. B. das Thema eines Buchs oder eines Films. Die Instruktion

³² Vgl. Vorderer et al. 2004, S. 12.

zur Anwendung des MEC-SPQ sieht eine eigenständige Anpassung der Skala durch die Versuchsleitung vor.³³ Im Falle der hier entwickelten virtuellen Umgebung schien die Anpassung auf inhaltlicher Ebene jedoch unangemessen. Da sich Interessen auf bestimmte Themenfelder und Gegenstände, aber auch auf Tätigkeitsformen beziehen können³⁴, wurde die Subskala DSI des MEC-SPQ auf das Interesse an Computerspielen angepasst. Hier lag die Annahme zugrunde, dass Probanden, die häufig Computerspiele spielen, auch an die visuelle Ausgabe einer virtuellen Welt gewöhnt sind, was sich wiederum auf die Rezeption virtueller Lernumgebungen auswirken müsste. Dazu wurden die vier in der Subskala angelegten Items umformuliert, wie beispielsweise »An Computerspielen bin ich generell interessiert«. Zusätzlich zum Interesse an Computerspielen wurde die Spielfrequenz in pro Woche gespielten Stunden abgefragt.

Als abhängige Variablen wurden in dieser Studie Rezeptionsprozesse des Erlebens in Form von Präsenzerleben (erste abhängige Variable) und Flow-Erleben (zweite abhängige Variable) erhoben. Situative Emotionen (dritte abhängige Variable) wurden im Hinblick auf das Potenzial von Mixed Reality zur Steigerung der Lernmotivation ebenfalls abgefragt. Außerdem wurden drei verschiedene Leistungsparameter erfasst. Präsenzerleben wurde mit Subskalen des MEC-SPQ von Vorderer et al.³⁵ gemessen. Zu den Subskalen gehört das mentale Raummodell (Spatial Situation Model, kurz: SSM), mögliche Handlungen (Possible Actions, kurz: PA), Selbstlokalisierung (Self Location, kurz: SL), Hemmung der Rezeption durch bewusste Suche nach Fehlern, z.B. hinsichtlich der Grafik (Suspension of Disbelief, kurz: SOD) und die kognitive Involviertheit (Higher Cognitive Involvement, kurz: HCI).

Das mentale Raummodell (SSM) wurde mit vier Items abgefragt, z. B. »Ich hatte eine genaue Vorstellung von der in der virtuellen Umgebung dargestellten räumlichen Umgebung.« Mögliche Handlungen wurden ebenfalls mit vier Items, wie z. B. »Ich hatte den Eindruck, dass ich selbst in der dargestellten Umgebung aktiv werden konnte.«, gemessen. Auch Selbstlokalisierung wurde mit vier Items erhoben, z. B. »Es war so, als ob sich mein Standort in die präsentierte Umgebung verschoben hätte.« Hemmung durch Fehlersuche wurde mit vier Items, wie z. B. »Ich habe nicht besonders darauf geachtet, ob Fehler bzw. Widersprüche in der virtuellen Umgebung bestehen.«, erfasst. Die kognitive Involviertheit wurde mit acht Items gemessen, wie z. B. »Ich habe gründlich überlegt, inwiefern die dargestellten Dinge etwas miteinander zu tun haben.«

Flow-Erleben wurde mit den Subskalen der gedanklichen Absorbiertheit (vier Items, z. B. »Ich fühlte mich optimal beansprucht.«) sowie glatter automatisierter Verlauf (sechs Items, wie »Meine Gedanken bzw. Aktivitäten liefen flüssig und glatt.«) der Kurzsкала von Rheinberg et al. erhoben (Flow-Kurzsкала, kurz: FKS).³⁶ Situative Emotionen wurden mit der Differentiellen Affekt-Skala zur Erfassung von Emotionen bei der Mediennutzung (M-DAS) gemessen.³⁷ Dazu wurden 13 Subskalen mit jeweils drei Adjektiven (z. B. »amüsiert« für die Subskala

³³ Vgl. Vorderer et. al, S. 5.

³⁴ Vgl. Krapp 2009, S. 53.

³⁵ Vgl. Vorderer et al. 2004, S. 3.

³⁶ Vgl. Rheinberg et al. 2003, S. 271.

³⁷ Vgl. Renaud / Unz 2006, S. 73.

Vergnügen) erfasst. Alle Selbsteinschätzungsvariablen wurden auf einer siebenstufigen Skala mit beschrifteten Extremen von »stimmt überhaupt nicht« bis »stimmt völlig« erfasst. Sie wurden im Rahmen des ersten Pre-Tests mit der Berechnung von Cronbachs Alpha auf interne Konsistenz geprüft und hinsichtlich interner Konsistenz und Ökonomie optimiert.

Als Leistungsparameter wurden die Anzahl korrekt aus dem Labyrinth wieder erkannter Objekte (vierte abhängige Variable), Gesamtdauer (fünfte abhängige Variable) sowie Genauigkeit (sechste abhängige Variable) der Zuordnung von Objekten zu ihren Positionen erhoben. Die Wiedererkennungsaufgabe wurde ebenfalls über das Online-Befragungstool abgefragt. Die Messung der Leistung in dieser Aufgabe wird im Folgenden als Wiedererkennen bezeichnet, gemessen wurde die Anzahl der Fehler im Test. Geschwindigkeit (Millisekunden) und Genauigkeit (Abweichung in Millimetern) wurden automatisiert von der App auf dem Tablet-Computer gemessen und gespeichert.

2.4 Stichprobe und statistische Auswertung

Die Teilnehmer der Studie waren allesamt Studierende der RWTH Aachen sowie der Fachhochschule Aachen. Um eventuellen Enttäuschungen vorzubeugen, wurden die Teilnehmer erst nach der Studie darüber in Kenntnis gesetzt, dass sie randomisiert in einer von zwei verschiedenen Bedingungen getestet wurden. In Bezug auf die Randomisierung muss jedoch einschränkend hinzugefügt werden, dass Studienteilnehmer, die angaben, starke visuelle Beeinträchtigungen (z. B. starke Kurzsichtigkeit) oder Beeinträchtigungen des Bewegungsapparats (z. B. Knieprobleme) zu haben, der Kontrollbedingung zugewiesen wurden.

Im Folgenden wird die Stichprobe genauer beschrieben. Genannt wird der Mittelwert eines gemessenen Wertes (allgemein bezeichnet mit M) und die Standardabweichung (allgemein bezeichnet mit SD , Standard Deviation) als Maß der Streuung. An der Studie nahmen insgesamt $N = 38$ Studierende teil, davon 25 männlich (65,8 %). Zwei Fälle mussten aufgrund von Versuchsabbruch im Virtual Theatre ausgeschlossen werden. Von den Probanden wurden 44,7 % im Mixed Reality-Simulator Virtual Theatre ($n = 17$) und 55,3 % am Laptop ($n = 21$) getestet. Auf Bedingungen bezogen war das Geschlechterverhältnis weitgehend ausgeglichen: Im Virtual Theatre wurden sechs Frauen (35,3%) und am Laptop sieben Frauen (33,3%) getestet. Das Durchschnittsalter betrug 24,71 Jahre ($SD = 3,06$). Auf die Bedingungen bezogen betrug das Durchschnittsalter im Virtual Theatre 25,33 Jahre ($SD = 3,07$) und am Laptop 23,94 Jahre ($SD = 2,95$).

Mit statistischen Methoden wurde untersucht, ob die Ergebnisse der Untersuchungen zufälliger oder systematischer Natur sind. Eine ausführliche Beschreibung über die hier genutzten Auswertungsmethoden und Hinweise darüber, wie die genannten Werte zu lesen sind, findet sich in zahlreichen Lehrbüchern der Statistik.³⁸ Mit univariaten, einfaktoriellem Varianzanalysen (ANOVAs) wird getestet, ob der Faktor der genutzten Hardware statistisch bedeutsame (= signifikante) Unterschiede in Bezug auf eine Zielvariable (z. B. Emotionen)

³⁸ Vgl. Bortz / Döring 2006, S. 196.

hervorrufen. Hierbei wird einerseits zur Überprüfung der Ergebnisse die relevante Teststatistik angegeben (z. B. $F(1,36) = 10.93$) und andererseits, ob es sich bei einem Unterschied um einen statistisch bedeutsamen Befund handelt. Ausschlaggebend hierfür ist die Über- oder Unterschreitung des gewählten Signifikanzniveaus, z. B. 5%, also die akzeptierte Fehlerwahrscheinlichkeit des Tests. Diese Fehlerwahrscheinlichkeit und die Über- oder Unterschreitung dieser Schwelle wird ebenfalls in der Teststatistik genannt (z. B. $p > .05$). Alle statistischen Ergebnisse wurden mit dem Computerprogramm SPSS ausgewertet.

2.5 Ergebnisse

Die Ergebnisse zeigen, dass die Nutzung des Virtual Theatres zu mehr Präsenzerleben führt als die Nutzung des Laptops (Abbildung 7). Signifikante Bedingungsseffekte zeigen sich bei den Subskalen der Selbstlokalisierung (Self Location, $F(1,36) = 15.75$; $p < .001$) und den möglichen Handlungen (Possible Actions, $F(1,36) = 4.90$; $p < .05$).

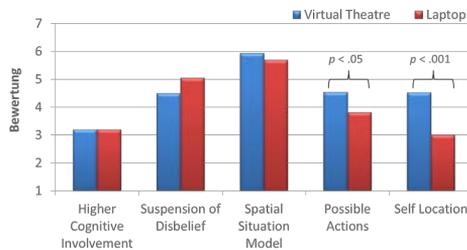


Abb. 7: Präsenzerleben im Vergleich beider Versuchsbedingungen (Grafik: Katharina Schuster, 2014; Screenshot aus Excel, 14.03.2014).

Der positive Einfluss von natürlichen Benutzerschnittstellen auf Flow-Erleben kann nur teilweise bestätigt werden. Hier zeigt sich ein signifikanter Unterschied zwischen den Bedingungen in der intendierten Richtung nur für die Subskala ›Absorbiertheit‹ ($F(1, 36) = 10.63$; $p < .01$, sig.). In der Subskala des glatten automatisierten Verlaufs gibt es keinen Unterschied zwischen den beiden Bedingungen (Abbildung 8).

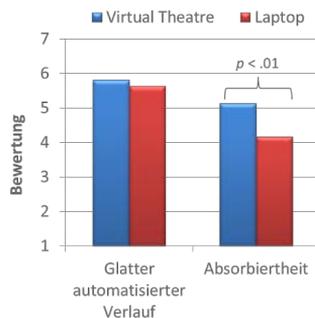


Abb. 8: Flow-Erleben im Vergleich beider Versuchsbedingungen (Grafik: Katharina Schuster, 2014; Screenshot aus Excel, 14.03.2014).

Exemplarisch an Items der jeweiligen Subskalen beschrieben, führt die Nutzung des Mixed Reality-Simulators Virtual Theatre also dazu, dass Probanden vertieft in das sind, was sie gerade machen, selbstvergessen sind und nicht merken, wie die Zeit vergeht (= gedankliche Absorbiertheit). Die Nutzung des Mixed Reality-Simulators führt jedoch nicht dazu, dass die Probanden das Gefühl haben, den Ablauf unter Kontrolle zu haben, dass sie bei jedem Schritt wissen, was zu tun ist oder dass die Gedanken und Aktivitäten flüssig und glatt laufen (= glatter automatisierter Verlauf).³⁹

In Bezug auf die in der Rezeptionssituation erlebten Emotionen lässt sich festhalten, dass insgesamt mehr positive Emotionen im Virtual Theatre erlebt wurden ($F(1, 36) = 6.18; p < .05$; Abbildung 9). Dieser Effekt lässt sich auf die Subskalen ›Faszination‹ ($F(1,36) = 6.22; p < .05$) und ›Ergriffenheit‹ ($F(1,36) = 15.62; p < .001$) zurückführen, welche im Virtual Theatre deutlich ausgeprägter vorkamen.

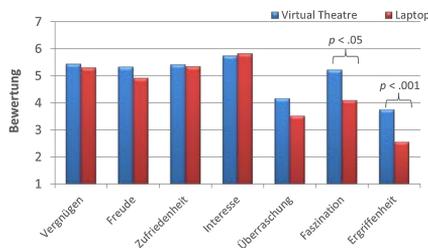


Abb. 9: Positive Emotionen im Vergleich beider Versuchsbedingungen (Grafik: Katharina Schuster, 2014; Screenshot aus Excel, 14.03.2014).

Das Virtual Theatre führt nicht zu besseren Leistungen als die Laptop-Bedingung. In Bezug auf die Anzahl der Fehler im Gedächtnistest ›Labyrinth‹ zeigt sich ein gegenteiliger Effekt: Im Virtual Theatre werden signifikant mehr Fehler gemacht ($F(1,36) = 10.93; p < .01$) als am Laptop (Abbildung 10).



Abb. 10: Leistung im Vergleich beider Versuchsbedingungen (Grafik: Katharina Schuster, 2014; Screenshot aus Excel, 14.03.2014).

³⁹ Vgl. Rheinberg et al. 2003, S. 271.

3. Diskussion

Im Vorfeld wurde angenommen, dass allein die technische Beschaffenheit der natürlichen Benutzerschnittstellen dazu beitragen kann, sich eher in einer virtuellen Umgebung als in der physisch realen Umwelt präsent zu fühlen. Die hochsignifikanten Ergebnisse der Varianzanalyse haben die positive Wirkung der natürlichen Benutzerschnittstelle auf Präsenzerleben bestätigt.

Im Vorfeld wurde außerdem angenommen, dass die technische Beschaffenheit der natürlichen Benutzerschnittstellen dazu beitragen kann, sich fokussiert einer Aufgabe zu widmen und somit Flow-Erleben zu initiieren. Dies haben die Ergebnisse der Datenauswertung bestätigt. In Bezug auf den glatten, automatisierten Verlauf wurde kein Unterschied zwischen den beiden Bedingungen festgestellt, was wiederum darauf zurückzuführen ist, dass Flow ein tätigkeitsspezifisches Konstrukt ist. Der flüssige Verlauf der Gedanken in Bezug auf eine Tätigkeit wird somit anscheinend nach der virtuellen Umgebung beurteilt.

Die positiveren Emotionen im Virtual Theatre lassen sich dadurch erklären, dass alleine das Erlebnis, virtuelle Umgebungen mit natürlichen Benutzerschnittstellen zu erkunden, positive Emotionen initiiert. Das Virtual Theatre als solches vermag insbesondere mehr Faszination und Ergriffenheit auszulösen, als die gleiche Aufgabe an einem Laptop auszuführen. Dies könnte durch den Neuigkeitswert der Technologie begründet sein, welcher wiederum an das Motiv der Neugier der Probanden appelliert haben könnte.⁴⁰

Eines der entscheidendsten Ergebnisse dieser Studie ist, dass die Probanden im Virtual Theatre keine besseren Ergebnisse erzielt haben, als die Probanden am Laptop. Im Gegenteil: in Bezug auf das Wiedererkennen der Objekte lieferte diese Gruppe sogar eine signifikant schlechtere Leistung als die Gruppe am Laptop. Das Ergebnis könnte zum einen schlicht auf zu viele Störfaktoren zurückzuführen sein, die noch mit der Technik des Virtual Theatre einhergehen, wie z. B. der Lautstärkepegel des omnidirektionalen Bodens oder ergonomische Aspekte des Head Mounted Displays. Wären solche technischen Optimierungsaspekte einmal abgeschlossen, so könnte man auch im Virtual Theatre bessere Lernergebnisse erwarten. Zum anderen drängt sich an dieser Stelle jedoch der Verdacht auf, dass die Steuerung virtueller Lernumgebung über natürliche Benutzerschnittstellen zumindest bei Erstbenutzung eine Aufgabe ist, die kognitive Ressourcen benötigt. Mit anderen Worten: So natürlich, wie es der Steuerungsmodus vorgibt zu sein, ist er nicht. Es scheint sich um Eingabemuster zu handeln, die den natürlichen Bewegungen zwar stark ähneln, aber sich so hinreichend von ihnen unterscheiden, dass sie zumindest geübt werden müssen.⁴¹ Hierfür ist die in der Studie gegebene Vorbereitungszeit von drei Minuten nicht ausreichend. Ist das motorische Programm einmal gespeichert und automatisiert, so stünden den Rezeptionsprozessen der Informationsverarbeitung wieder mehr kognitive Ressourcen zur Verfügung.

⁴⁰ Vgl. Edelmann 1996, S. 361.

⁴¹ Vgl. Kiesel / Koch 2012, S. 115.

Eine weitere Interpretation der schlechteren Lernergebnisse im Mixed Reality-Szenario ist die Überlappung von Immersion und Kognition. So ist es allein schon durch die technische Isolation des Head Mounted Displays bedingt, dass mehr Aufmerksamkeit auf die virtuelle Lernumgebung gelenkt wird, als dies beim Vergleichsszenario am Laptop der Fall ist. Die Aufnahme dieser zusätzlichen visuellen Reize aus der Lernumgebung könnte von der Bearbeitung der eigentlichen inhaltlichen Aufgabe ablenken.

4. Ausblick: Mit Mixed Reality zum ›Virtual Turn‹ in den Erziehungswissenschaften?

Insbesondere zu einem späteren Zeitpunkt im Rahmen der Forschung an natürlichen Benutzerschnittstellen in der Hochschullehre wird es notwendig sein, neben Daten der potenziellen Nutzergruppe der Studierenden auch die Anforderungen von Lehrenden zu erheben. Hier geht es vor allem um inhaltliche Aspekte der Gestaltung von virtuellen Lernumgebungen, damit diese auch ein bestimmtes didaktisches Anliegen einlösen.

Das vermeintlich unbefriedigende Ergebnis der schlechteren Leistung im Mixed Reality-Szenario verdeutlicht paradoxerweise das große Potenzial, was sich hinter der neuen Technologie verbirgt: Die realistische Nachbildung einer späteren Anwendungssituation des im schulischen oder hochschulischen Kontext erworbenen Wissens. Denn auch die physische Realität zeichnet sich durch ein hohes Maß an Komplexität und sachfremder Beanspruchung aus. Hierauf kann eventuell ein Lernmodul im Mixed Reality-Simulator besser vorbereiten, als eines, das am Laptop absolviert wird. Aus ethischer Perspektive scheint dies insbesondere für jene Situationen sinnvoll zu sein, in denen Fehlverhalten zu gesundheitlichen Schäden der Person selbst oder anderer beteiligter Personen führen würde. Im Hinblick auf den im Bereich der Hochschullehre häufig diskutierten mangelnden Praxisbezug liefert das junge Forschungsgebiet der Mixed Reality also neue Möglichkeiten.

Die Ergebnisse bestätigen die Notwendigkeit, natürliche Benutzerschnittstellen hinsichtlich ihres technischen Potenzials immer mit den tatsächlich erlebten Emotionen und Wahrnehmungsparametern der Nutzergruppen abzugleichen. Durch eine genaue Überprüfung des Zusammenhangs zwischen Nutzerin bzw. Nutzer, Hardware, virtueller Umgebung und Lernerfolg ist es möglich, die Bedeutung von Immersion für Lernprozesse in virtuellen Umgebungen weiter zu erforschen. Die technische Optimierung natürlicher Benutzerschnittstellen in enger Verzahnung mit der durchdachten und reflektierten Entwicklung passender virtueller Lernszenarien ist für die Erziehungswissenschaften eine mögliche Option, sich mit dieser Art der neuen Methodik der Entwicklung neuer Lösungen für die Vermittlung von Lerninhalten zu widmen.⁴² In Zeiten fortschreitender Digitalisierung aller Gesellschaftsbereiche, demografischen Wandels und hohen Bedarfs nach gut ausgebildeten Fachkräften werden auch die Erziehungswissenschaften nicht umhin kommen, neue Informations- und Kommunikationstechnologien hinsichtlich ihrer Effektivität und ihrer Effizienz zu testen. Ob sich damit auch eine weitere methodische Wende im Sinne eines ›Virtual

⁴² Vgl. Fatke / Oelkers 2014, S. 7.

Turns der Didaktik vollziehen wird, kann nur eine stetige Forschung im Gebiet der virtuellen und gemischten Realität in Bezug auf Lernprozesse zeigen.

Bibliographische Angaben

Jürgen Bortz / Nicola Döring: Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler. 4. überarbeitete Auflage. Heidelberg 2006. [\[Nachweis im GBV\]](#)

Klaus Bredl / Amrei Groß: Gestaltung und Bewertung von Lernszenarien in immersiven virtuellen Welten. In: Lernen und Lehren in virtuellen Welten. Zeitschrift für E-Learning – Lernkultur und Bildungstechnologie 7 (2012), H. 1, S. 36–46. [\[Nachweis im GBV\]](#)

Johannes Breuer: Spielend Lernen? Eine Bestandsaufnahme zum (Digital) Game-Based Learning. Landesanstalt für Medien Nordrhein-Westfalen (LfM). Düsseldorf 2010. [\[online\]](#)

Thomas D. Cook / Donald T. Campbell: Quasi-Experimentation. Design & Analysis Issue for Field Settings. Chicago 1979. [\[Nachweis im GBV\]](#)

Robin Curtis / Christiane Voss: Theorien ästhetischer Immersion. In: Montage AV. Zeitschrift für Theorie und Geschichte audiovisueller Kommunikation (2008), S. 3–10. [\[Nachweis im GBV\]](#)

Walter Edelmann: Lernpsychologie. 5. Auflage. Weinheim 1996. [\[Nachweis im GBV\]](#)

Daniel Ewert / Katharina Schuster / Daniel Schilberg / Sabine Jeschke: Intensifying learner's experience by incorporating the virtual theatre into engineering education. In: Proceedings of the 2013 IEEE Global Engineering Education Conference (EDUCON 2013). Hg. von Institute of Electrical and Electronics Engineers (IEEE). Berlin 2013, S. 207–2012. [\[Nachweis im GBV\]](#)

Reinhard Fatke / Jürgen Oelkers: Das Selbstverständnis der Erziehungswissenschaft: Geschichte und Gegenwart. In: Zeitschrift für Pädagogik 60 (2014), Beiheft 60, S. 7–13. [\[Nachweis im GBV\]](#)

Rohna Flin / Rona Patey / Ronnie Glavin / Nikki Maran: Anaesthetists' non-technical skills, in: British Journal of Anaesthesia 105 (2010), H. 1, S. 38–44. [\[Nachweis im GBV\]](#)

Rudolf Groner / Simon Raess / Phillip Sury: Usability: Systematische Gestaltung und Optimierung von Benutzerschnittstellen. In: Medienpsychologie. Hg. von B. Batinic / M. Appel. Heidelberg 2008, S. 425–446. [\[Nachweis im GBV\]](#)

Andreas Hebbel-Seeger: Motiv: Motivation?! – Warum Lernen in virtuellen Welten (trotzdem) funktioniert. In: Lernen und Lehren in virtuellen Welten. Zeitschrift für E-Learning – Lernkultur und Bildungstechnologie 7 (2012), H.1, S. 23–35. [\[Nachweis im GBV\]](#)

Susan Höntzsch / Uwe Katzky / Klaus Bredl / Frank Kappe / Dirk Krause: Simulationen und simulierte Welten. Lernen in immersiven Lernumgebungen. In: Martin Ebner / Sandra Schön: Lehrbuch für Lernen und Lernen mit Technologien. 2013. [\[online\]](#)

Max Hoffmann / Katharina Schuster / Daniel Schilberg / Sabine Jeschke: Bridging the Gap between Students and Laboratory Experiments. In: Virtual, Augmented and Mixed Reality. Applications of Virtual and Augmented Reality. 6th International Conference, VAMR 2014, Held as Part of HCI International 2014, Heraklion, Greece, June 22–27, 2014, Proceedings, Part II. Hg. von R. Shumaker / S. Lackey. Heidelberg, S. 39–50. [\[Nachweis im GBV\]](#)

Sabina Jeschke: Mathematik in Virtuellen Wissensräumen – IuK-Strukturen und IT-Technologien in Lehre und Forschung. Berlin, Techn. Univ., Diss., 2004. [\[Nachweis im GBV\]](#)

Udo Kelle: Die Integration qualitativer und quantitativer Methoden in der empirischen Sozialforschung. Theoretische Grundlagen und methodologische Konzepte. Wiesbaden 2007. [\[Nachweis im GBV\]](#)

Michael Kerres: Mediendidaktik. Konzeption und Entwicklung mediengestützter Lernangebote. München 2012. [\[Nachweis im GBV\]](#)

Andrea Kiesel / Iring Koch: Lernen. Grundlagen der Lernpsychologie. Wiesbaden 2012. [\[Nachweis im GBV\]](#)

Thomas Köhler / Nina Kahnwald / Martina Reitmaier: Lehren und Lernen mit Multimedia und Internet. In: Medienpsychologie. Hg. von B. Batinic / M. Appel. Heidelberg 2008, S. 477–501. [\[Nachweis im GBV\]](#)

Andreas Krapp: Interesse. In: Handbuch der Allgemeinen Psychologie – Motivation und Emotion. Hg. von V. Brandstätter / J. H. Otto. Göttingen 2009, S. 52–57. [\[Nachweis im OPAC\]](#)

Janet H. Murray: Hamlet on the Holodeck. The Future of Narrative in Cyberspace. Cambridge, Mass. 1999. [\[Nachweis im GBV\]](#)

Eric. B. Nash / Gregory W. Edwards / Jennifer A. Thompson / Woodrow Barfield: A Review of Presence and Performance in Virtual Environments. In: International Journal of Human-Computer Interaction 12 (2000), H. 1, S. 1–41. [\[Nachweis im GBV\]](#)

Dagmar Renaud / Dagmar Unz: Die M-DAS – eine modifizierte Version der Differentiellen Affekt Skala zur Erfassung von Emotionen bei der Mediennutzung. In: Zeitschrift für Medienpsychologie 18 (2006), H. 2, S. 70–75. [\[Nachweis im GBV\]](#)

Falko Rheinberg / Regina Vollmeyer / Stefan Engeser: Die Erfassung des Flow-Erlebens. In: Diagnostik von Motivation und Selbstkonzept. Test und Trends. Hg. von J. Stiensmeier-Pelster / F. Rheinberg. Göttingen 2003, S. 261–279. [\[Nachweis im GBV\]](#)

Patrick Sahle: Digitale Geisteswissenschaften. Hg. vom CCEH, Universität zu Köln. Erstellt im Rahmen des Projekts DARIAH-DE; gefördert vom Bundesministerium für Bildung und Forschung (BMBF) 2011. [\[online\]](#)

Liane Tarouco / Bárbara Gorziza / Ygor Corrêa / Erico M.H. Amaral / Thaísa Müller: Virtual Laboratory for Teaching Calculus: an Immersive Experience. In: Proceedings of the 2012 IEEE Global Engineering Education Conference (EDUCON 2012). Hg. von Institute of Electrical and Electronics Engineers (IEEE). Berlin 2012, S. 774–781. [\[Nachweis im GBV\]](#)

Peter Vorderer / Werner Wirth / Felize Ribeiro Gouveia / Frank Biocca / Timo Saari / Lutz Jäncke / Saskia Böcking / Holger Schramm / Andre Gysbers / Tilo Hartmann / Christoph Klimmt / Jari Laarni / Nicklas Ravaja / Ana Sacau / Thomas Baumgartner / Petra Jäncke: MEC Spatial Presence Questionnaire (MEC-SPQ): Short Documentation and Instructions for Application. Report to the European Community, Project Presence: MEC (IST-2001-37661), 2004. [\[online\]](#)

Harald Willenbrock: Wie im richtigen Leben. In: Brandeins 14 (2012), H. 10, S. 86–91. [\[Nachweis im GBV\]](#)

Werner Wirth / Matthias Hofer: Präsenzerleben. Eine medienpsychologische Modellierung. In: Montage AV. Zeitschrift für Theorie und Geschichte audiovisueller Kommunikation 2008, S. 159–175. [\[Nachweis im GBV\]](#)

Bob G. Witmer / Michael J. Singer: Measuring Presence in Virtual Environments: A Presence Questionnaire. In: Presence: Teleoperators and Virtual Environments 7 (1998), H. 3, S. 225–240. [\[Nachweis im GBV\]](#)

Raphael Zender / Ulrike Lucke: Synchronisationsmechanismen in Mixed Reality-Arrangements. In: Lernen und Lehren in virtuellen Welten. Zeitschrift für E-Learning – Lernkultur und Bildungstechnologie 7 (2012), H. 1, S. 47–60. [\[Nachweis im GBV\]](#)

Abbildungslegenden und -nachweise

Abb. 1: Das Virtual Theatre mit Nutzerin und (vereinfachter) virtueller Umgebung (Foto: Walter Spanjersberg, 2013).

Abb. 2: Datenhandschuh, Hand-Tracer sowie entsprechende Darstellungen in einer virtuellen Umgebung (Fotos: Peer Schlieperskötter, 2014).

Abb. 3: Erwarteter Zusammenhang zwischen Medienangebot, Person, Situation, Rezeptionsprozessen und Wirkung (Grafik: Katharina Schuster, 2013).

Abb. 4: Versuchsablauf des ›Quasi-Experiments‹ am Beispiel der Experimentalbedingung (Grafik: Katharina Schuster, 2013).

Abb. 5: Ansicht der virtuellen Umgebung in der ersten und dritten Aufgabe (Foto: Walter Spanjersberg, 2013).

Abb. 6: Übersicht: Unterschiede beider Versuchsbedingungen des ›Quasi-Experiments‹ (Grafik: Katharina Schuster, 2014).

Abb. 7: Präsenzerleben im Vergleich beider Versuchsbedingungen (Grafik: Katharina Schuster, 2014; Screenshot aus Excel, 14.03.2014).

Abb. 8: Flow-Erleben im Vergleich beider Versuchsbedingungen (Grafik: Katharina Schuster, 2014; Screenshot aus Excel, 14.03.2014).

Abb. 9: Positive Emotionen im Vergleich beider Versuchsbedingungen (Grafik: Katharina Schuster, 2014; Screenshot aus Excel, 14.03.2014).

Abb. 10: Leistung im Vergleich beider Versuchsbedingungen (Grafik: Katharina Schuster, 2014; Screenshot aus Excel, 14.03.2014).

Zeitschrift für digitale Geisteswissenschaften

Artikel aus:

Sonderband 1 der ZfdG: Grenzen und Möglichkeiten der Digital Humanities. Hg. von Constanze Baum und Thomas Stäcker. 2015. DOI: [10.17175/sb01](https://doi.org/10.17175/sb01)

Titel:

Semantisch angereicherte 3D-Messdaten von Kirchenräumen als Quellen für die geschichtswissenschaftliche Forschung

Autor/in:

Felix Lange

Kontakt: flange@mpiwg-berlin.mpg.de

Institution: Max-Planck-Institut für Wissenschaftsgeschichte

GND: [1084013126](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63863-p0011-9) ORCID: [0000-0002-3201-3792](https://orcid.org/0000-0002-3201-3792)

Autor/in:

Martin Unold

Kontakt: martin.unold@hs-mainz.de

Institution: Hochschule Mainz

GND: [1084131374](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63863-p0011-9) ORCID: [0000-0003-2913-2421](https://orcid.org/0000-0003-2913-2421)

DOI des Artikels:

[10.17175/sb001_015](https://doi.org/10.17175/sb001_015)

Nachweis im OPAC der Herzog August Bibliothek:

[830205322](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63863-p0011-9)

Erstveröffentlichung:

19.02.2015

Lizenz:

Sofern nicht anders angegeben 

Medienlizenzen:

Medienrechte liegen bei den Autoren

Letzte Überprüfung aller Verweise:

24.05.2016

GND-Verschlagwortung:

[Sakralbau](#) | [Dreidimensionale Rekonstruktion](#) | [Visualisierung](#) |

Zitierweise:

Felix Lange, Martin Unold: Semantisch angereicherte 3D-Messdaten von Kirchenräumen als Quellen für die geschichtswissenschaftliche Forschung. In: Grenzen und Möglichkeiten der Digital Humanities. Hg. von Constanze Baum / Thomas Stäcker. 2015 (= Sonderband der Zeitschrift für digitale Geisteswissenschaften, 1). PDF Format ohne Paginierung. Als text/html abrufbar unter DOI: [10.17175/sb001_015](https://doi.org/10.17175/sb001_015).

Felix Lange, Martin Unold

Semantisch angereicherte 3D-Messdaten von Kirchenräumen als Quellen für die geschichtswissenschaftliche Forschung

Abstracts

Der vorliegende Beitrag gibt einen Einblick in raumbezogene epigraphische Forschungen von mittelalterlichen und frühneuzeitlichen Kirchen im Rahmen des Forschungsprojektes *Inschriften im Bezugssystem des Raumes* (IBR). Dieses in den Spatial Humanities angesiedelte Vorhaben nutzt mit Hilfe einer speziell für diesen Zweck entwickelten Erfassungssoftware Punktwolken aus terrestrischen Laserscannings als Primärquelle für die Raumanalyse. Zusammen mit epigraphischen Fachdaten und Annotationen entsteht so die Datengrundlage für die Untersuchung vielfältiger Sinn- und Funktionsbezüge von Objekten in Kircheninnenräumen.

This paper provides an overview of spatially oriented epigraphic research in medieval and early modern churches as part of the Spatial Humanities research project *Inschriften im Bezugssystem des Raumes* (IBR). With the help of a custom made capturing software, this field uses geometric point cloud data from terrestrial laser scans as a primary source for spatial analysis. This data, in combination with epigraphic text sources and annotations, forms the basis for research on various symbolic and functional relations between objects in church interiors.

1. Einleitung: Historische Inschriften im Raumkontext

Insbesondere für quellenarme Perioden wie das Altertum oder das Mittelalter ist die Epigraphik eine wichtige historische Hilfswissenschaft. Inschriften als historische Textquellen unterscheiden sich wesentlich von Urkunden und erzählenden Genres, die von vornherein für die Verbreitung durch Kopie und Druck bestimmt waren, durch ihre feste Verbindung zu einem Inschriftenträger und originären Aufstellungsort. Letzterer ist zwar nicht unbedingt statisch gewesen – tatsächlich befinden sich zum Beispiel wenige Memorialinschriften in mittelalterlichen Kirchen noch *in situ*, sondern wurden in Folge religionsgeschichtlicher Ereignisse oder auch im Zuge von Baumaßnahmen zum Teil mehrfach disloziert –, für die meisten in der mittelalterlichen und frühneuzeitlichen Epigraphik untersuchten Inschriftenarten lässt sich aber ein ihnen zugedachter Aufstellungsort annehmen. Dieser Ortsbezug ist im Fall von ortsdeiktischen Ausdrücken wie »hier« sogar bis in den Inschriftentext hinein zu verfolgen. Es kann demnach allgemein gelten, dass sich Inschriften nur im Zusammenhang mit ihrem Inschriftenträger und ihrem dedizierten Anbringungsort verstehen lassen.¹

¹ Vgl. hierzu Fuchs 1991, S. 76: »Inschriftenträger gehören zu den wenigen Quellen mit verbalen Informationen, die anders als Urkunden oder erzählende Quellen immer in der Verklammerung von Text und Aussehen sowie – ganz wichtig – [eig. Hervorh.] zu verstehen sind.«

Doch wie lässt sich dieser Raumbezug methodisch fassen? Wir werden im Folgenden dafür argumentieren, dass geometrische Daten zum Inskriptenträger und seinem Raumkontext eine zuverlässige Quellengrundlage darstellen. Dieser technologiegestützte Ansatz steht im Kontext der *Spatial Humanities*², einem Teilgebiet der digitalen Geisteswissenschaften, zu dem bisher vor allem Untersuchungen von sehr großflächigen Gebieten gerechnet werden: Es geht um weit ausgedehnte Untersuchungsgebiete in der Größe von Stadtteilen³ bis hin zu interkontinentalen Handelsrouten, die naturgemäß als zweidimensionale Flächen betrachtet werden.⁴ Im dreidimensionalen Raum untersucht werden meist eher kleinere Artefakte, etwa bei der Funddokumentation in der Archäologie.⁵ Die zwischen beiden räumlichen Maßstäben liegende Analyse von textuellen, bildlichen und allgemein liturgischen Gegenständen in gebauten Umgebungen wird in mediävistischen und frühneuzeitlichen Forschungen hingegen meist auf eine recht technikferne Weise umgesetzt, so mit Projektionen auf zweidimensionale Grundrisse oder mittels Fotografien.⁶

Im folgenden Abschnitt werden zunächst exemplarisch einige geschichtswissenschaftliche Arbeiten zu Sinn- und Funktionsbezügen in Kircheninnenräumen vorgestellt. Anhand der daran anknüpfenden, aber interpretatorisch weitergehenden Annahme eines räumlichen Systems sinnhafter Bezüge in mittelalterlichen Kirchen sollen dann methodologische Desiderata und ihre Lösung durch geosemantische Software-Tools erläutert werden.⁷

2. Heilige, Gläubige, Bilder, Texte und Liturgie

Schon lange vor dem sogenannten *spatial turn* in den Geisteswissenschaften⁸ sind Wechselbeziehungen und -wirkungen zwischen Objekten, Personen und Handlungen im Kirchenraum ein wichtiges Forschungsfeld in den historischen Disziplinen gewesen. Beispiele sind Gräber, Grabinschriften und Epitaphien. Ein entscheidender Aspekt für die Wahl des Ortes einer Grabstätte innerhalb der Kirche ist die *memoria*, insbesondere die durch Stiftungen ermöglichten liturgischen Vorkehrungen für das Seelenheil eines Verstorbenen. Aber nicht nur die Totenliturgie selbst, auch die Nähe des Toten zu Reliquien und Altären war dabei von großer Bedeutung. Stifter strebten danach, möglichst nahe an entsprechenden Orten bestattet zu werden und so neben ihrem individuellen Totengedenken auch von anderen gottesdienstlichen Handlungen zu profitieren. Dies hat Forscher dazu bewogen, nicht mehr vorhandene Grabmäler (unter anderem) in Abhängigkeit von der Entfernung zu umliegenden Altären und den dort gefeierten Liturgien zu lokalisieren.⁹

² Grundlegend Bruhn et al. 2014.

³ Z. B. Gordon 2008.

⁴ Z. B. Williams 2014.

⁵ Z. B. Karasik 2008.

⁶ Vgl. bspw. Neuheuser 2010. Vgl. hingegen Clark 2007 und Paliou 2014, insbesondere zu St. Vitale in Ravenna, S. 103ff., als Beispiele für technikgestützte raumbezogene Forschungsansätze in den Altertumswissenschaften.

⁷ Die vorliegende Arbeit ist im Rahmen des BMBF-geförderten Verbundprojektes »Inskripten im Bezugssystem des Raumes« entstanden, an dem die Hochschule Mainz und die Akademie der Wissenschaften und der Literatur Mainz beteiligt sind [\[online\]](#).

⁸ Zur Einordnung des Begriffs vgl. Döring / Thielmann 2008, S. 7ff.

⁹ Schwarz 2005, S. 59f. Schwarz verhandelt hier eine symbolisch die eigentliche Grabstätte darstellende Grabtumba für Karl IV. bei der Marienkappelle von St. Veit in Prag.

Auch die Verortung und Ausgestaltung von Nebenaltären ist in den Zusammenhang eines die ganze Kirche umfassenden Bild- und Patrozinienprogramms gestellt worden.¹⁰ Neuheuser wendet dazu allerdings ein, dass die Wahl von Altarheiligen eher privaten Vorlieben folgte und ihre Platzierung von den meist eingeschränkten Möglichkeiten in der Kirche abhing.¹¹

Ähnlich funktionale und sinnhafte Bezüge zur räumlichen Umgebung werden auch für andere bildliche Quellen wie Lehrtafeln und Wandgemälde erforscht. Für St. Bartholomäus in Frankfurt am Main hat Astrid Krüger herausgestellt, dass ein Fries mit einer bestimmten Variante des Bartholomäus-Martyriums die Textauswahl für das Brevier beeinflusst habe, das während der Liturgie gelesen wurde.¹² Text, Bild und liturgische Handlungen stehen hier also in enger Abhängigkeit zueinander.

Für die zeitgenössische Wahrnehmung räumlicher Beziehungen sind aber auch schriftliche Quellen von Bedeutung. Ordinarienbücher, die die Ortsliturgie genau vorschreiben, allgemeine Abhandlungen wie etwa das *Rationale divinatorum officiorum* des Durandus von Mende¹³ oder der Kirchweihbericht des Abtes Suger von Saint-Denis stellen sakrale Objekte und die Kirchenarchitektur in den Kontext der Liturgie.¹⁴ Sie spiegeln und beeinflussen damit die physische Konstruktion und Deutung räumlicher Verhältnisse in Kirchausstattung und Liturgie. So fordert Durandus, Abbildungen der vier Evangelisten in einer bestimmten räumlichen Anordnung zueinander zu zeigen¹⁵, und Suger betont die besondere Bedeutung der Hauptachse des Langhauses (*in medio ecclesiae*).¹⁶

Dieser beispielhafte Überblick zeigt, dass wechselseitige räumliche Bezüge für das Verstehen und ggf. die Rekonstruktion nicht nur von Inschriften, sondern auch von Altären, bildlichen Quellen und Grabmälern von großer Bedeutung sind. Das Projekt »Inschriften im Bezugssystem des Raumes« (IBR) untersucht vor diesem Hintergrund die Frage, inwieweit sich diese semantischen Aspekte von Positionen und Relationen im physischen Raum in der Gesamtanalyse eines Gebäudeinneren zu einem komplexen Geflecht, einem *System* räumlicher Bezüge verdichten lassen.¹⁷ Die Elemente des Systems wären in diesem Fall Personen, die sich im Kontext liturgischer Handlungen im Raum aufhalten und bewegen, und die für die Liturgiefeier relevanten Objekte. Entscheidend ist, dass eine solche sinn- und funktionstragende Anordnung eine historische Relevanz hat, also tatsächlich ein Bedingungsfaktor für Wahrnehmungen und Handlungen historischer Personen war. Es geht hingegen nicht darum, eine Art »Mindmap« zu konstruieren, die eher Assoziationen und Interpretationen heutiger Forscher wiedergibt.

¹⁰ Van der Ploeg 1993, S. 98ff.

¹¹ Neuheuser 2010, S. 190f.

¹² Krüger 2010, S. 149ff.

¹³ Faupel-Dreves 1998, *passim*.

¹⁴ Neuheuser 1998, *passim*.

¹⁵ Vgl. hierzu Faupel-Dreves 1998, S. 671, 675f.

¹⁶ Neuheuser 1998, S. 654. Vgl. allgemein zu diesem Begriff Neuheuser 2010, S. 187f.

¹⁷ Vgl. Schwartz 2014 und Schwartz [in Vorbereitung] für eine ausführliche theoretische und empirische Fundierung.

Dieses angenommene System lässt sich durch einige grundlegende räumliche Relationen strukturieren. Ein Beispiel sind wiederkehrende Eigenschaften von Bildmotiven. Dies soll am Beispiel zur »Wahren Länge Christi« verdeutlicht werden. Dabei handelt es sich im weiteren Sinne um das religionsgeschichtliche Phänomen, von der Größe der Grabstätte Jesu in der Grabeskirche zu Jerusalem auf seine tatsächliche Lebensgröße zu schließen, und das so erhaltene Maß auf Gegenstände wie Amulette sowie auf bildliche Christusdarstellungen anzuwenden.¹⁸ Wenn die Heilige Länge das Thema eines Wandgemäldes in einer Kirche ist, kann man fragen, ob andere Christusdarstellungen im Raum wie Kreuzskulpturen gewissermaßen als Zitat dieser Darstellung zu verstehen und daher ähnlich dimensioniert sind.¹⁹ Wichtig ist auch die relative Positionierung von Objekten und Personen durch Parallelität, Symmetrie und Opposition (d.h. die Platzierung einander gegenüber). Letztere Relation lässt sich etwa in der Zelebrationsrichtung des die Liturgie feiernden Geistlichen *versus populum*, also der Gemeinde zugewandt, erkennen.²⁰ Im Kontext von van der Ploegs Untersuchung zu Nebenaltären, die gewissermaßen in ›Leserichtung‹ angeordnet wurden und einen narrativen Zusammenhang bildeten,²¹ wären die relativen Positionierungen ›links‹ und ›rechts‹ interessant. Wichtig können aber auch vertikale Abfolgen und Himmelsrichtungen sein, was zum Beispiel in der Ausrichtung von Sanktuarien nach Osten bzw. nach Westen deutlich wird.²² Schließlich lassen sich auch durch menschliche Wahrnehmungen und Handlungen determinierte Relationen wie Zugänglichkeit und Sichtbarkeit nennen.²³ Für diese räumlichen Beziehungen muss eine wissenschaftliche Repräsentationsform gefunden werden, die gleichermaßen Dokumentation und Analyse ermöglicht.

3. Der Raum als Quelle

Wie bereits erwähnt, ist die vorwiegende Darstellungsweise für die hier besprochenen Forschungsansätze der wissenschaftliche Text mit schematischen Zeichnungen und Fotografien, was einige methodische Probleme mit sich bringt.

Bei Textquellen wird es im Zuge der Digitalisierung möglich, weitgehend ortsunabhängig zu forschen. Dies stellt für Arbeiten, die auf Archivrecherchen basieren, eine enorme Erleichterung dar.²⁴ Historische Forschungen zu gebauten Umgebungen wie Kircheninnenräumen kommen hingegen bislang im Allgemeinen nicht ohne eine Ortsbesichtigung aus. Vor allem für vergleichende Studien zu einer größeren Anzahl von Kirchen ergeben sich daraus

¹⁸ Grundsätzlich zur Heiligen Länge Herkommer 2008, S. 91 ff.

¹⁹ Frithjof Schwartz, pers. Komm. Diese Frage stellte sich angesichts des Wandgemäldes mit der Länge Christi in der Liebfrauenkirche zu Oberwesel (vgl. Abschnitt 7), wurde aber nach entsprechenden Messungen negativ beantwortet. Vgl. die Inschriftenedition des Wandgemäldes in Nikitsch 2004, S. 139f. (Katalognummer 140).

²⁰ Vgl. hierzu z. B. Jacobsen 2000, S. 68ff., zur religionsgeschichtlichen Bedeutung verschiedener Zelebrationsrichtungen und ihnen entsprechender räumlicher Aufteilungen von Sanktuarien.

²¹ Van der Ploeg 1993, S. 98ff.

²² Jacobsen 2000, S. 72f. und 75ff.

²³ Vgl. Abschnitt 6.

²⁴ Dies ist natürlich nur der Fall, wenn nicht lizenzrechtliche Beschränkungen den ubiquitären Zugriff auf digitale Quellen wenigstens von Bibliotheken aus verhindern. Interessant sind in diesem Zusammenhang die Interviews mit Historikern in Allen / Sieczkiewicz 2010, S. 2, in denen die Befragten Beschränkungen ihrer Arbeit mit digitalisierten historischen Zeitungen durch regionale und nationale Ausrichtungen bei der Lizenznahme beklagen. In US-amerikanischen Bibliotheken etwa finden sich natürlich mehr amerikanische als europäische Zeitschriften und umgekehrt.

naheliegende forschungspraktische Probleme. Noch schwerer wiegt aber, dass genuin raumbezogene Argumentationen auf der Grundlage schematischer Zeichnungen und ausschnittthafter Fotografien für Leser im Allgemeinen nicht vollständig überprüfbar und daher auch nicht falsifizierbar sind. Details der Kirche und ihrer Ausstattung werden lediglich durch Darstellungen des Autors bzw. der Autorin vermittelt, der oder die bei der Auswahl von Fotografien und bei der Reduktion des Vorgefundenen in der vereinfachten Zeichnung möglicherweise Aspekte, die seine Forschungsthese stützen, gegenüber potentiellen Hinweisen auf gegenteilige Annahmen bewusst oder unbewusst bevorzugt hat. Geschichtswissenschaftliche Diskurse werden jedoch idealerweise auf der Grundlage einer allgemein zugänglichen gemeinsamen Quellenbasis geführt. Gerade diese Voraussetzung ist für die raumbezogene Forschung jedoch nicht gegeben. Es bräuchte dafür eine vollständige fotografische Dokumentation etwa der in Rede stehenden Kirche(n). Selbst damit aber hätte man noch keine zuverlässige Forschungsgrundlage; so könnten von den oben angesprochenen Raumrelationen beispielsweise sowohl Sichtbarkeiten als auch Entfernungen und Maße nur bedingt vom Rezipienten selbst bestimmt werden. Es fehlt in diesem Fall eine gemeinsame, umfassende Repräsentation des räumlichen Untersuchungsobjektes.

Die Limitationen des konventionellen zweidimensionalen Ansatzes lassen sich an der im Übrigen sehr überzeugenden Arbeit von Kosch²⁵ zur »Binnentopographie« ehemaliger Kölner Konventkirchen gut festmachen. Kosch will die soziale und liturgische Organisation des Kircheninnenraumes darstellen und zu diesem Zweck unter anderem Zugangsmöglichkeiten und -beschränkungen für Laien und Geistliche in der Kirche kartieren. Dabei sollen auch die verschiedenen historischen Bauzustände berücksichtigt werden. Hierzu nutzt er computergenerierte Längsschnitte und Querschnitte aus verschiedenen Höhen, die auf bestehenden Zeichnungen basieren. Die liturgisch-soziale Raumgliederung wird mit Hilfe von Farbmarkierungen (Pfeilen und Schraffuren) dargestellt. Frühere oder spätere Bauzustände werden durch unterschiedliche Strichstärken gekennzeichnet. Diese Darstellung macht zum Beispiel in unmittelbar nachvollziehbarer Weise deutlich, dass Wege durch die Kirche zu Orten von liturgischer Bedeutung für Laien und Kleriker deutlich voneinander getrennt waren.²⁶ Dass andererseits ein Sarkophag im Kirchenschiff von einer Krypta unter dem Hauptchor aus sichtbar war,²⁷ müssen wir dem Autor glauben. Von welchen Standpunkten aus in der Krypta Sichtbarkeit gegeben war, welchen Teil des Sarkophags man an diesen Stellen sah und ob die Sichtbarkeit auch für die üblichen Körpergrößen im Mittelalter gegeben war, kann der Leser nicht selbst überprüfen. Dies gilt auch für die Quellenbasis der eingezeichneten historischen Raumgliederungen. Kosch räumt selbst ein, dass »die eigentlich unerlässliche, im Einzelnen jeweils durch konkrete Belege gestützte Beweisführung für alle auf den Plänen eingezeichneten Detailangaben in diesem Rahmen nicht oder nur andeutungsweise geleistet werden kann«.²⁸ Schließlich beschränkt es die Aussagekraft der Darstellungen und die (intuitive) Nachvollziehbarkeit der gemachten Aussagen, dass es sich um Zusammenstellungen von zweidimensionalen Zeichnungen handelt.

²⁵ Kosch 2000.

²⁶ Kosch 2000, S. 140.

²⁷ Kosch 2000, S. 139.

²⁸ Kosch 2000, S. 141. Der Autor bezieht sich dabei auf eine Nachfolgepublikation zu den romanischen Altstadtkirchen Kölns, in denen die hier beschriebene Darstellungsweise Verwendung gefunden hat.

4. Dreidimensionale Messdaten als Repräsentationen räumlicher Quellen

Es ist leicht nachvollziehbar, dass die dreidimensionale Darstellung eines Kircheninnenraumes sowohl in seiner Gesamtheit als auch in einem Detailgrad, der Objekte wie Inschriftenträger erkennbar macht, durch digitale Daten erfolgen muss. Aber welche Informationen müssen diese Daten enthalten, und wie sind sie zu visualisieren? Im Hinblick auf die Geschichtlichkeit des Untersuchungsgegenstandes – eine Kirche hat eventuell mehrere Ausbauphasen erlebt, die in ihr enthaltenen Objekte sind zum großen Teil nicht mehr an ihrem ursprünglichen Ort, andere Objekte sind ganz verschwunden – sind hier zunächst externe Annotationen zu nennen. Es muss eine Möglichkeit geben, Hypothesen über die historische räumliche Gestalt gleichsam als ›text‹-kritischen Kommentar in geeigneter Weise getrennt von der Quelle abzubilden, während diese selbst lediglich den ›überlieferten‹, also in der Gegenwart existierenden Raumkörper darstellt.

Eine in der Archäologie häufig verwendete entsprechende Darstellungsform ist das virtuelle Modell, das in den letzten Jahren als Visualisierungs-, aber auch als Analysegrundlage an Popularität gewonnen hat.²⁹ Insbesondere für Gebäude, von denen lediglich Fundamente erhalten sind, ist dies die einzige mögliche Form der Visualisierung. Die Virtualität erschwert jedoch die Trennung des objektiv vorhandenen Baubefundes von der Rekonstruktion: Die Ornamente etwa einer korinthischen Säule, von der in Wirklichkeit nur noch die Basis erhalten ist, werden ohne Relativierung in einem rekonstruierten visualisierten Modell einfach postuliert. Zwar werden in solchen Fällen meist vorgefundene und rekonstruierte Teile grafisch deutlich voneinander getrennt,³⁰ und in wissenschaftlichen Modellierungsumgebungen gibt es die Möglichkeit, die Hintergründe von Rekonstruktionen nach fachspezifischen Standards zu dokumentieren.³¹ Aber auch hier beinhaltet die visuelle Darstellung eine mehr oder weniger implizite Forschungshypothese, eine Modellierungsentscheidung, und ist damit keine Quelle im geschichtswissenschaftlichen Sinn. Bei mittelalterlichen Kirchen ist überdies das Verhältnis von vorgefundener Bausubstanz zu fehlenden Elementen im Allgemeinen vergleichsweise günstig, weshalb eine vollständig virtuelle Darstellung unnötig künstlich erscheint.

Der Gegenentwurf dazu ist die möglichst genaue Erfassung des vorgefundenes Raumes durch moderne Messverfahren. Digitale Aufnahmen des Ist-Zustandes einer historischen Umgebung sind abgesehen von der Wahl der Aufnahmestandpunkte objektiv. Sie sind frei von Interpretation und können daher als Primärquellen angesehen werden, die digital analysierbar sind. Beim terrestrischen Laserscanning (TLS) als Aufnahmemethode, die im Folgenden näher betrachtet wird, ist das digitale Resultat eine Punktwolke. Eine Punktwolke ist eine Liste von Koordinaten, die in diesem Fall aus dem Sichtbereich des Scanners stammt, also die Punkte enthält, an denen der Laserstrahl auf ein festes Objekt – wie eine Wand oder einen Gegenstand – trifft. Durch terrestrische Laserscanner kann also ein Raum durch eine große Zahl von dreidimensionalen Messpunkten abgebildet werden, und zwar je nach Auflösung reduziert,

²⁹ Hermon 2008, S. 36 ff.

³⁰ Vgl. z. B. Hermon 2008, S. 41, Abb. 6.

³¹ Z. B. die prozesshaft dokumentierende Modellierungstechnik in Saldana 2014.

aber ohne dass eine systematische Verfälschung seiner physischen Eigenschaften wie bei einer schlechten Modellierung droht.

TLS-Daten sind objektiv, enthalten allerdings nicht die physischen Zusammenhänge, die für das menschliche Erkennen entscheidend sind; nämlich Objekte wie Altäre oder Säulen. Insbesondere ist die tatsächliche flächenhafte Gestalt physischer Objekte unbekannt, von Informationen über die Geometrie ganz zu schweigen. Daher müssen Teile der Punktwolke ausgeschnitten und mit Fachinformationen verknüpft werden. Für diesen Prozess sind Computerprogramme nötig, die den Menschen bei dieser Aufgabe unterstützen, indem die Koordinaten-Liste geeignet visualisiert und durchsuchbar gemacht wird.³² Weil gängige Softwarelösungen für diese Aufgabe für Ingenieure und Techniker ausgelegt sind und daher ein tiefes Verständnis von analytischer Geometrie und Erfahrung im Umgang mit 3D-Anwendungen voraussetzen, entstand im Rahmen des Projektes IBR eine Eigenentwicklung: der **GenericViewer**. Das Programm bietet durch die Überlagerung von Panorama-Bildern und Punktwolken einen vereinfachten Zugang. Zusätzlich zu der Möglichkeit, Objekte in Punktwolken zu identifizieren, können diese direkt mit Fachdaten angereichert werden.

5. Semantische Anreicherung

In manchen Untersuchungen sind reine Punktwolkendarstellungen hinreichend, um raumbezogene Arbeiten empirisch zu untermauern. Für die datenlogische Abbildung der oben skizzierten raumsemantischen Zusammenhänge ist es aber notwendig, die geometrischen mit wissenschaftlichen Fachdaten anzureichern. Dabei müssen zuerst die händisch markierten geometrischen Objekte mit Klassifizierungen wie ›Altar‹ inhaltlich beschrieben werden. Ein zweiter Schritt ist die Verknüpfung der markierten und benannten Objekte mit möglichen inhaltlichen Bezügen. Auf dem gegenwärtigen Stand der Technik ist es je nach eingesetzter Software zusätzlich nötig, komplexere räumliche Relationen, wie beispielsweise die parallele Anordnung von Objekten, händisch zu markieren. Damit sind die Schritte bereits genannt, wie räumliche Bezüge datentechnisch modelliert und in Datenbanken für vergleichende und quantitative Analysen abfragbar gemacht werden können.

Um vergleichende Forschungen zu betreiben und um die gesammelten Daten öffentlich zugänglich und maschinenlesbar zu machen, sollten sie mit Metadaten beschrieben werden. Diese Metadaten sollten konventionalisiert sein und nicht als einfache Zeichenketten eingegeben werden. Diese Forderungen führen zur Verwendung von Semantic-Web-Technologien. Statt fest definierter Spaltenüberschriften wie in einem relationalen Datenbankschema erhält der Nutzer bzw. die Nutzerin hier die Möglichkeit, für beliebige Subjekte, Objekte und Prädikate Relationen aus ubiquitär gültigen und (maschinen)lesbaren Ontologien und Taxonomien zu wählen. Auf diese Weise können prinzipiell alle markierten Objekte mit anderen Datenbeständen, beispielsweise aus der **Europeana**³³, in Bezug gesetzt werden. Trotzdem ist die nötige Freiheit in der Metadatenauswahl bei entsprechender

³² Für eine technisch vertiefende Darstellung z. B. Wimmer et al. 2006, S. 129 ff.

³³ Das Projekt **3D-Icons** entwickelt Möglichkeiten, wie auch dreidimensionale Daten in Europeana eingespeist werden können.

Wahl des Vokabulars gewährleistet. Semantisch modellierte Objektbezüge erscheinen im Datenmodell als Triple in der Form Subjekt-Prädikat-Objekt (SPO), was die Formulierung von formalisierten Forschungsaussagen, die über bloße Typzuweisungen hinausgehen, auf eine einfache Weise ermöglicht.

Eine Implementierung der hier skizzierten Anwendung für dreidimensionale Objekte ist Arrigo.³⁴ Es erlaubt die dreidimensionale Darstellung und die Klassifizierung von Artefakten mittels semantischer Annotation. Unseres Wissens hingegen existiert aber bis dato keine integrierte Softwareumgebung für die Beschreibung von räumlichen Objekten und ihren Relationen, die für größere gebaute Umgebungen geeignet ist und daher über die Objektdokumentation hinaus komplexere historische Sachverhalte für geisteswissenschaftliche Zwecke abbilden kann.

6. Softwaretechnische Implementierung: GenericViewer

Die von IBR entwickelte Software stellt komplexe geoinformatische Technologien in stark vereinfachter Weise zur Verfügung. Einige räumliche Relationen können (halb-)automatisch berechnet werden: Längen, Flächen, Distanzen und die Lage von Objekten im Raum werden auf Grundlage der 3D-Koordinaten durch den GenericViewer angegeben. Mittels standardisierter Datenbeschreibungen ist ein Export markierter Objekte oder größerer Raumausschnitte zur weiteren Analyse in externen Softwareprogrammen möglich.³⁵ Mit externen Tools lassen sich dann zum Beispiel Viewshed-Analysen durchführen, bei denen für einen gegebenen Punkt im Raum dessen Sichtfeld errechnet wird. Diese Technologie ist eine Grundlage für die komplexen Raumcharakterisierungen der ›Space Syntax‹.³⁶

Für die semantische Anreicherung wurde das Annotationssystem Pundit in einer reduzierten, auf Geo-Objekte spezialisierten Version in den GenericViewer integriert. Pundit ist ursprünglich für die Annotation von HTML-Texten in digitalen Bibliotheken entwickelt worden und ermöglicht die Formulierung von Aussagetripeln.³⁷ Durch das Modul TripleComposer, welches lediglich die Subjekt-Prädikat-Objekt-Slots mittels Drop-Down-Menü ausgeführter Listen zeigt,³⁸ wird eine weitgehende Abstraktion von den im Hintergrund geschriebenen RDF-Graphen erreicht, was einer häufig geäußerten Usability-Forderung an semantische Annotatoren entspricht.³⁹ Gleichzeitig bleibt aber die Möglichkeit, logisch-semantische Zusammenhänge besser abzubilden als durch das bloße semantische Tagging von einzelnen Objekten.⁴⁰ Im GenericViewer können geometrische Objekte, Webseiten, Texte und Elemente aus semantischen Vokabularen zu Klassifizierungen und Aussagen verknüpft werden. Die folgenden Abbildungen (1, 2, 3, 4) zeigen einen zyklischen Workflow der Markierung, Auszeichnung und anschließenden Abfrage entsprechender geosemantischer Daten. Hier

³⁴ Havemann et al. 2009.

³⁵ Vgl. Ganitseva et al. 2014.

³⁶ Griffiths 2012.

³⁷ Nucci et al 2012.

³⁸ Vgl. Abbildung 2.

³⁹ Hinze et al. 2012, S. 170.

⁴⁰ Vgl. z. B. Hinze et al. 2012.

wird ein geometrisches Objekt, in diesem Fall ein Wandgemälde, markiert und im Triple Browser (Abbildung 2) entsprechend klassifiziert. In einer SPARQL-Abfrage werden dann alle Wandgemälde in der Kirche abgefragt (Abbildung 3) und in der Panoramafotografie angezeigt (Abbildung 4).

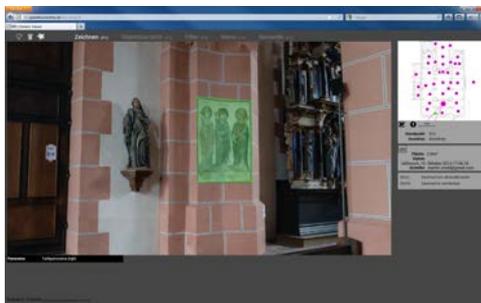


Abb. 1: Markierung einer Geometrie im GenericViewer. Quelle: Martin Unold, i3mainz.

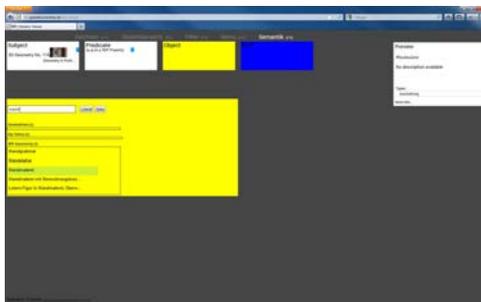


Abb. 2: Klassifizierung. Quelle: Martin Unold, i3mainz.

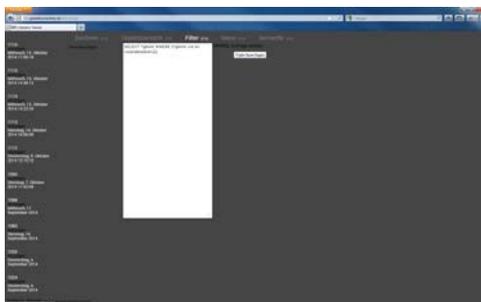


Abb. 3: Abfrage des Triple Stores nach Geometrien. Quelle: Martin Unold, i3mainz.

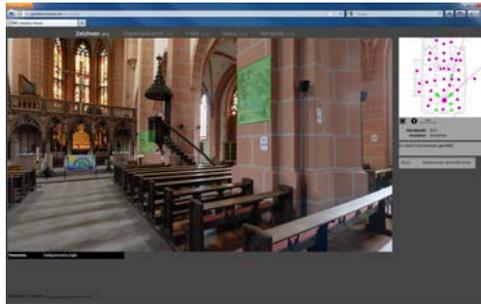


Abb. 4: Anzeige der Abfrage im Panoramabild. Quelle: Martin Unold, i3mainz.

Anhand der annotierten Geometrien im GenericViewer lässt sich erkennen, dass die Dokumentation des vorgefundenen baulichen Zustands als Quelle und ihre geschichtswissenschaftliche Beschreibung in Form von Annotationen strikt voneinander getrennt sind. Letztere trägt nicht zur Visualisierung bei, sondern mündet in Datenbankeinträgen, die Referenzen auf Teile der Punktwolke beinhalten.

Neben händischen Annotationen können Geodaten auch durch ›autoritative‹ wissenschaftliche Texte semantisch angereichert werden. Im Falle von Inschriften sind dies kritische Editionen, wie sie für den deutschen Sprachraum durch das Webportal [Die Deutschen Inschriften Online](#) bereitgestellt werden. Mit Hilfe von XSLT-Transformationen lassen sich diese Daten extrahieren, transformieren und als SPO-Tripel im RDF-Format in den Triple-Store des GenericViewer laden. Im Ergebnis stehen zu jeder Geometrie, sobald sie mit einer Katalognummer aus ‚Deutsche Inschriften Online‘ verbunden ist, eine Vielzahl von zuverlässigen historischen Informationen bereit.⁴¹

Diese Form der Verbindung von strukturierten 3D- und Text-Daten macht den GenericViewer zu einem multimedialen, ›qualitativen‹ GIS-System, wie es Corrigan⁴² vorgeschlagen hat. Während er damit jedoch eine Ergänzung von wissenschaftlichen Geo- und Text-Daten durch eher szenische Begleitinformationen zu meinen scheint,⁴³ geht es hier allein um eine Aggregation analysetauglicher Daten.

Im Folgenden werden zwei im Projektkontext unter Leitung von Frithjof Schwartz durchgeführte Referenzstudien vorgestellt.⁴⁴ In ihnen wird mit Hilfe des Softwaretools der Versuch unternommen, den räumlichen Kontext von Inschriftenträgern adäquat zu modellieren und in empirisch abgesicherter Weise zu analysieren. Beide Arbeiten behandeln die spätgotische Liebfrauenkirche im mittelhessischen Oberwesel.

7. Zwei Referenzstudien

⁴¹ Vgl. den entsprechenden [Webservice](#) dazu, implementiert von Torsten Schrade, AdW Mainz.

⁴² Corrigan 2010.

⁴³ Corrigan 2010, S. 87.

⁴⁴ Schwartz 2014 und Schwartz [in Vorbereitung].

Die an den Baubeginn der Liebfrauenkirche im Jahre 1308 erinnernde Gründungsinschrift befindet sich in der mittleren Maßwerkzone der Chorfenster. An dieser Inschrift ist weniger ihr Text (dieser erwähnt lediglich den Baubeginn⁴⁵), wohl aber die Ausführung innerhalb der Chorfenster und die Anbringung hinter dem Hauptaltar außergewöhnlich. Da dieser Chor erst im Jahre 1331 fertiggestellt wurde, ist die Inschrift offensichtlich auch deutlich nach 1308 dort angebracht worden. Als eine mögliche Erklärung wird in der Literatur auf die Stadtgeschichte verwiesen. Der Baubeginn fiel nämlich in das letzte Jahr, bevor die vormals reichsfreie Stadt 1309 von König Heinrich VII. an den Trierer Erzbischof Balduin verpfändet wurde. Nach Nikitsch handelt es sich um eine an den Erzbischof gerichtete Anspielung auf die Zeit der städtischen Reichsfreiheit. Wichtig für diese Argumentation ist der Anbringungsort der Inschrift, also direkt vor dem Hauptaltar, an dem Balduin selbst möglicherweise die Altarweihe 1331 zelebrierte. Der Erzbischof oder der jeweils am Hauptaltar stehende Priester wären demnach die eigentlichen (Haupt-)Adressaten dieser Inschrift gewesen.⁴⁶ Ein wichtiger Prüfstein für die Plausibilität dieses Argumentes ist die Frage, von welchen Positionen die Inschrift überhaupt lesbar war. Müller gibt hierzu an, dass man »den besten Blick auf die Inschrift vom Hauptaltar aus«⁴⁷ habe. Es stellt sich nun die Frage, ob die über die Chorfenster verteilte Schrift vom Hauptaltar aus vollständig sichtbar war und ob sie wirklich *nur* von dort zu sehen und zu lesen war. Eine exklusive Lesbarkeit für den Kleriker am Hauptaltar wäre ein starkes Argument dafür, dass sie tatsächlich auch vor allem an ihn gerichtet war.

Zur empirischen Prüfung dieser Frage wurde auf der Grundlage der Punktwolke eine Viewshed-Analyse angefertigt. Diese Form der Sichtfeldanalyse berücksichtigt auch den Blickwinkel des Betrachters. Das in *Abbildung 5* in Form in einer Heat-Map dargestellte Ergebnis zeigt, wie viele Stellen in der gesamten Kirche Sichtachsen zu welchen Segmenten der Inschrift aufweisen. Tatsächlich lässt sich an dieser Darstellung ablesen, dass die Inschrift lediglich vom Hauptaltar aus vollständig sichtbar war. Mit Hilfe einer im GenericViewer vorgenommenen Messung der Höhe der Inschrift sowie der Buchstabengrößen konnte zusätzlich abgeschätzt werden, dass der Betrachter bzw. die Betrachterin sie auch tatsächlich *lesen* konnte. Die Interpretation Nikitschs scheint also in diesem Punkt bestätigt zu werden. Diese Untersuchung zeigt damit, wie eine konventionelle, vor allem auf Textquellen gestützte historische Forschungsarbeit durch die Untersuchung des physischen Raumes ergänzt werden kann. Dies ist auch ein weiteres Beispiel dafür, welche große Rolle der Raumkontext von Inschriften für ihre kommunikative Zweckbestimmung spielt.

⁴⁵ »+ IN[C]HOATA [F][V]IT SCE MARIE A[N]O DNI [MC]CC OCATVO, begonnen wurde die Kirche der heiligen Maria im Jahre 1308«, Nikitsch 1996, S. 99.

⁴⁶ Vgl. hierzu Nikitsch 1996, S. 105ff.

⁴⁷ Müller 2008, S. 11.

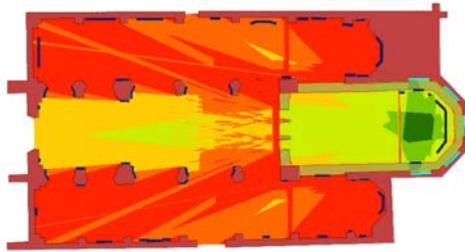


Abb. 5: Viewshed-Analyse für die Lesbarkeit der Bauinschrift. Rötliche Farbtöne signalisieren eine geringe Zahl lesbarer Buchstaben vom jeweiligen Standpunkt, grünliche Farbtöne eine hohe Zahl. Quelle: i3mainz / Julia Ganitseva.

Eine weitere Referenzstudie⁴⁸ untersucht die Stiftungstätigkeit des vermutlich 1515 verstorbenen Kanonikers Petrus Luterns⁴⁹ in der Kirche. Dieser hinterließ unter anderem eine der heiligen Martha gewidmete Kapelle, in der er offenkundig auch beerdigt wurde. Weder das zur Grabstätte gehörende Epitaph mit Inschrift und Grabfigur, noch das Retabel des Martha-Altars sind *in situ* erhalten. Um ein genaueres Bild von den Vorkehrungen Luterns für sein Seelenheil zu erhalten, muss also zunächst der ursprüngliche Ort der Martha-Kapelle und ihrer Bestandteile bestimmt werden. Dazu gibt die Gestaltung der Grabfigur wertvolle Hinweise. Wie eine Auswertung der Messdaten ergab, ist ihr Kopf um ca. 9° nach rechts geneigt. Schwartz nimmt ausgehend von diesem auffälligen Befund wechselseitige visuelle Bezüge zwischen dem Altar als dem Ort der Totenmesse und dem Grab an, die sich laut der Epitaphinschrift beide in dieser Kapelle befanden. Demnach war der Blick der Grabfigur an ihrem bislang unbekanntem ursprünglichen Ort direkt auf den Martha-Altar und damit auf den Liturgen gerichtet, um eine symbolische Nähe zu dem Toten herzustellen.

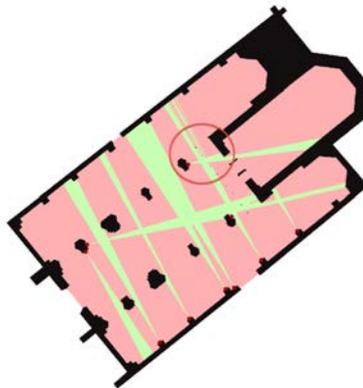


Abb. 6: Viewshed-Analyse für die Lokalisierung des Lutern-Epitaphs mit Sichtfeldern der Grabfigur. Quelle: i3mainz / Julia Ganitseva.

⁴⁸ Schwartz [in Vorbereitung].

⁴⁹ Vgl. Müller 2008, S. 33, zu den aus der Epitaphinschrift hervorgehenden biographischen Angaben.

Ein wichtiger Prüfstein für diese These ist, ob sich für Grabfigur und Altar wirklich ein Paar von Anbringungsorten finden lässt, welches eine Sichtachse zwischen beiden Objekten unter Berücksichtigung der Kopfneigung ermöglicht. Mit Hilfe einer weiteren Viewshed-Analyse wurden mögliche Anbringungsorte entsprechend überprüft. Wichtigstes Auswahlkriterium dabei waren Wandabschnitte, die Spuren einer früheren Anbringung von Gegenständen aufweisen. Für die entsprechende, mit dem GenericViewer auf den digitalen Raumdaten vorgenommene Objekt-Auswahl wurde jeweils eine Sichtfeldanalyse angefertigt, entsprechend der Hypothese, dass sich der Martha-Altar im Sichtfeld der Grabfigur befinden haben müsse. *Abbildung 6* zeigt, dass unter diesen Voraussetzungen als wahrscheinlichster Standort der Kapelle ein Platz innerhalb des Lettners in Frage kommt. Es handelt sich hierbei um den Laubengang des Lettners auf der linken Seite (Markierung in *Abbildung 6*), an dem sich auch heute noch ein Altartisch befindet. Tatsächlich befinden sich in passender Höhe dazu auf dem gegenüberliegenden Pfeiler auch Spuren früherer Halterungen. Offensichtlich ist also lediglich das Retabel an einen anderen Ort gebracht worden, möglicherweise, um einem neuen (und zwischenzeitlich wieder entfernten) Aufsatz auf dem Altartisch Platz zu machen.

Das Vorhandensein dieses Tisches und andere Hinweise haben zwar auch bereits bei der persönlichen Besichtigung des Ortes die Aufmerksamkeit des Forschers geweckt und den Verdacht entstehen lassen, es handele sich dabei um den ursprünglichen Anbringungsort des Martha-Retabels. Das Sichtfeld der Figur als das letztlich entscheidende Argument für diese These jedoch konnte unmöglich ohne technische Hilfsmittel bestimmt werden. Dies wäre allenfalls mit einem Grundriss wie in *Abbildung 6* möglich. Dieser Darstellung fehlt jedoch die dritte Raumdimension, so dass sie nicht zeigen kann, dass die Figur sich auch tatsächlich in Höhe des angenommenen Retabels befand. Schließlich sollte eine solche komplexe Indizienkette für das Fachpublikum nachvollziehbar sein. Das wird aber nur durch eine nachmess- und damit nachprüfbar, aber auch visuell zugängliche Dokumentation der Forschungen und ihres Untersuchungsgegenstandes ermöglicht, wie sie der GenericViewer bietet.

8. Fazit

Die in dieser Arbeit vorgestellten Studien haben gezeigt, dass Zweckbestimmungen und Bedeutungen textueller, aber auch bildlicher und skulpturaler Objekte in mittelalterlichen Kircheninnenräumen im Zusammenhang ihrer historischen räumlichen Umgebung gesehen werden müssen. Für diesen Zusammenhang spielen geometrisch bestimmbare Relationen zu anderen Objekten wie Distanz und Sichtbarkeit häufig eine große Rolle. Aus diesem Grunde sind verlässliche, objektive Geodaten eine hilfreiche Quelle für die Geschichtswissenschaft. Das hier vorgestellte Software-Tool GenericViewer stellt einen Zugang zu dieser Art von Daten dar.

Auch bei Vorliegen geometrischer Analysen ist die Erforschung räumlicher Sinnbezüge jedoch durch die Knappheit historischer Informationen limitiert. Dem kann neben klassischer geschichtswissenschaftlicher Forschungstätigkeit nur durch die systematische Sammlung raumbezogener Daten sowie durch ihre nachhaltige Speicherung und Verfügbarmachung, z. B. über das Semantic Web, begegnet werden.

Bibliographische Angaben

- Avshalom Karasik: Applications of 3D technology as a research tool in archaeological ceramic analysis. In: Beyond Illustration: 2D and 3D Digital Technologies as Tools for Discovery in Archaeology. Hg. von Bernard Frischer / Anastasia Dakouri-Hild. Oxford 2008, S. 103–116. [\[Nachweis im GBV\]](#)
- Robert B. Allen / Robert Sieczkiewicz: How historians use historical newspapers. In: Proceedings of the American Society for Information Science and Technology 47 (2010), H. 1, S. 1–4. DOI 10.1002/meet.14504701131. [\[Nachweis im GBV\]](#)
- Kai-Christian Bruhn / Fredie Kern / Frithjof Schwartz: Spatial Humanities – Eine neue Forschungsdisziplin. In: *gis.Business* 1 (2014), S. 53–55. [\[Nachweis im GBV\]](#)
- John Corrigan: Qualitative GIS and Emergent Semantics. In: The Spatial Humanities: GIS and the Future of Humanities Scholarship. Hg. von David J. Bodenhamer / John Corrigan / Trevor M. Harris Bloomington 2010, S. 76–88. [\[Nachweis im OPAC\]](#)
- David L. Chatford Clark: Viewing the liturgy: a space syntax study of changing visibility and accessibility in the development of the Byzantine church in Jordan. In: *World Archeology* 39 (2007), H. 1, S. 84–104. DOI: 10.1080/00438240601148798. [\[Nachweis im GBV\]](#)
- Jörg Döring / Tristan Thielmann: Einleitung. Der *Spatial Turn* und das geheime Wissen der Geographen. In: *Spatial Turn: das Raumparadigma in den Kultur- und Sozialwissenschaften*. Hg. von Jörg Döring / Tristan Thielmann. Bielefeld 2009, S. 7–45. [\[Nachweis im OPAC\]](#)
- Kirstin Faupel-Dreves: Bildraum oder Kultraum? Symbolische und liturgische Raumgestaltung im Rationale *divinorum officiorum* des Durandus von Mende. In: *Raum und Raumvorstellungen im Mittelalter*. Hg. von Jan Aertens / Andreas Speer. Berlin 1998, S. 665–684. [\[Nachweis im OPAC\]](#)
- Rüdiger Fuchs: Die Inschriften der Stadt Worms. Gesammelt und bearbeitet von Rüdiger Fuchs. Wiesbaden 1991 (= Die Deutschen Inschriften, Bd. 29). [\[Nachweis im OPAC\]](#)
- Julia Ganitseva / Kai-Christian Bruhn / Fredie Kern / Martin Unold: Outlook on TLS - based 3D Building Modeling. In: *Photogrammetrie, Laserscanning, Optische 3D-Messtechnik*. Hg. von Thomas Luhmann / Christina Müller. Heidelberg 2014, S. 224–233. [\[Nachweis im GBV\]](#)
- Colin Gordon: Mapping Decline: St. Louis and the fate of the American city. Philadelphia 2008. [\[Nachweis im GBV\]](#)
- Sam Griffiths: The Use of Space Syntax in Historical Research: current practice and future possibilities. In: Proceedings of the Eighth International Space Syntax Symposium. Santiago 2012. [\[online\]](#).
- Sven Havemann / Volker Settgest / Rene Berndt: The Arrigo Showcase Reloaded – Towards a Sustainable Link between 3D and Semantics. In: *ACM Journal for Computing in Cultural Heritage* 2 (2009), H. 1, Artikel Nr. 4 [unpag.]. DOI 10.1145/1551676.1551679.
- Hubert Herkommer: Heiliges Antlitz und heiliges Maß: Zu Aussehen und Größe des Menschensohnes. In: *Das Heilige Grab, das Heilige Kreuz und die Wahre Länge Christi*. Schwäbisch Gmünd 2008 (= Katalogreihe Museum im Prediger, Bd. 35), S. 67–105. [\[Nachweis im GBV\]](#)
- Sorin Hermon: Reasoning in 3D: a critical appraisal of the role of 3D modelling and virtual reconstructions in archaeology. In: *Beyond Illustration: 2D and 3D Digital Technologies as Tools for Discovery in Archaeology*. Hg. von Bernard Frischer / Anastasia Dakouri-Hild. Oxford 2008, S. 35–44. [\[Nachweis im GBV\]](#)
- Annika Hinze / Ralf Heese / Markus Luczak-Rösch / Adrian Paschke: Semantic enrichment by non-experts: usability of manual annotation tools. In: Proceedings of the 11th international conference on The Semantic Web. Part I. Berlin, Heidelberg 2012, S. 165–181. [\[Nachweis im GBV\]](#)
- Werner Jacobsen: Organisationsformen des Sanktuariums im spätantiken und mittelalterlichen Kirchenbau. Wechselwirkungen von Altarraum und Liturgie aus kunsthistorischer Perspektive. In: *Kölnische Liturgie und ihre Geschichte. Studien zur interdisziplinären Erforschung des Gottesdienstes im Erzbistum Köln*. Hg. von Albert Gerhards / Andreas Odenthal. Münster 2000 (= Liturgiewissenschaftliche Quellen und Forschungen, Bd. 87), S. 67–97. [\[Nachweis im GBV\]](#)
- Clemens Kosch: Zum Projekt einer zeichnerischen Veranschaulichung der sakralen "Binnentopographie" des Hochmittelalters in ehemaligen Konventkirchen Kölns: Methodische Überlegungen am Beispiel von St. Andreas. In: *Kölnische Liturgie und ihre Geschichte, Studien zur interdisziplinären Erforschung des Gottesdienstes im Erzbistum Köln*. Hg. von Albert Gerhards / Andreas Odenthal. Münster 2000, S. 127–142. [\[Nachweis im GBV\]](#)
- Astrid Krüger: Der Heilige im Raum. Bartholomäus-Liturgie und Bartholomäus-Darstellungen im Frankfurter Kaiserdom Sankt Bartholomäus. In: *Heilige, Liturgie, Raum*. Hg. von Dieter R. Bauer / Klaus Herbers / Hedwig Röckelein / Felicitas Schmieder. Stuttgart 2010 (= Beiträge zur Hagiographie, Bd. 8), S. 137–153. [\[Nachweis im OPAC\]](#)
- Dominik Maschek / Michael Schneyder / Marcel Tschanner: Virtual 3D reconstructions – benefit or danger for modern archaeology? Workshop-Paper: 14. Workshop »Cultural Heritage and New Technologies«, 16.–18. November 2009, Wien. [\[online\]](#)
- Sabrina Müller: Die Inschriften der katholischen Pfarrkirchen Unserer Lieben Frau in Oberwesel. Inschriften Mittelrhein-Hunsrück 1: Oberwesel, Liebfrauenkirche. Mainz 2008. [\[Nachweis im GBV\]](#)

Hanns Peter Neuheuser: *Ne lapidum materia apparentium locus vilesceret*. Die Raumvorstellung des Abtes Suger in seiner Kirchweihbeschreibung von St. Denis. In: Raum und Raumvorstellungen im Mittelalter. Hg. von Jan Aertsen / Andreas Speer. Berlin 1998, S. 641–64. [\[Nachweis im OPAC\]](#)

Hanns Peter Neuheuser: Liturgische Raumerschließung und Heiligenverehrung. In: Heilige, Liturgie, Raum. Hg. von Dieter R. Bauer / Klaus Herbers / Hedwig Röckelein / Felicitas Schmieder. Stuttgart 2010, S. 183–216. [\[Nachweis im OPAC\]](#)

Eberhard Josef Nikitsch: Ein Kirchenbau zwischen Bischof und Stadtgemeinde. Zur angeblich verlorenen Bauinschrift von 1308 in der Liebfrauenkirche zu Oberwesel am Rhein, in: Jahrbuch der Westdeutschen Landesgeschichte 22 (1996), S. 95–112. [\[Nachweis im OPAC\]](#)

Eberhard Josef Nikitsch: Die Inschriften des Rhein-Hunsrück-Kreises I. Wiesbaden 2004 (= Die Deutschen Inschriften, Bd. 60). [\[Nachweis im OPAC\]](#)

Michele Nucci / Marco Grassi / Christian Morbidoni / Francesco Piazza: Enriching Digital Libraries Contents with SemLib Semantic Annotation System. Proceedings of the Digital Humanities 2012 Conference. Hamburg 2012, S. 303–306. [\[Nachweis im GBV\]](#)

Eleftheria Paliou: Visibility analysis in 3D spaces: a new dimension to the understanding of social space. In: Spatial analysis and social spaces: interdisciplinary approaches to the interpretation of prehistoric and historic built environments. Topoi. Berlin Studies of the Ancient World. Hg. von Eleftheria Paliou / Undine Lieberwirth / Silvia Polla. Berlin 2014, S. 91–113. [\[Nachweis im GBV\]](#)

Kees van der Ploeg: Art, architecture and liturgy: Siena cathedral in the Middle Ages. Groningen 1993. [\[Nachweis im GBV\]](#)

Marie Saldana: An integrated approach to the procedural modeling of ancient cities and buildings. In: Digital Humanities 2014. Book of Abstracts. Lausanne 2014, S. 351–53. [\[online\]](#)

Michael Viktor Schwarz: Kathedralen zu verstehen. In: Virtuelle Räume. Raumwahrnehmung und Raumvorstellung im Mittelalter. Akten des 10. Symposiums des Mediävistenverbandes, Krems 24.–26. März 2003. Hg. von Elisabeth Vavra. Berlin 2005, S. 47–68. [\[Nachweis im OPAC\]](#)

Frithjof Schwartz: Inschriften im Bezugssystem des Raumes. Forschungsansatz und Verortung im wissenschaftlichen Spektrum. Referenzobjekt und Fallstudien. Projektskizze 2014, unpag. [\[online\]](#)

Frithjof Schwartz: Visuelle Bezüge im Raum. Analyse von Sichtachsen in der Liebfrauenkirche in Oberwesel. In: Zeitschrift für Kunstgeschichte. [in Vorbereitung]

Tim Williams: The Silk Roads: An ICOMOS Thematic Study, Charenton-le-Pont 2014. [online](#)

Michael Wimmer / Claus Scheiblauber: »Instant points: Fast rendering of unprocessed point clouds«. In: Proceedings of the Eurographics Symposium on Point-Based Graphics. Hg. von Mario Botsch / Christine Chen. Zürich 2006, S. 129–136. [\[Nachweis im GBV\]](#)

Abbildungslegenden und -nachweise

Abb. 1: Markierung einer Geometrie im GenericViewer. Quelle: Martin Unold, i3mainz.

Abb. 2: Klassifizierung. Quelle: Martin Unold, i3mainz.

Abb. 3: Abfrage des Triple Stores nach Geometrien. Quelle: Martin Unold, i3mainz.

Abb. 4: Anzeige der Abfrage im Panoramabild. Quelle: Martin Unold, i3mainz.

Abb. 5: Viewshed-Analyse für die Lesbarkeit der Bauinschrift. Rötliche Farbtöne signalisieren eine geringe Zahl lesbarer Buchstaben vom jeweiligen Standpunkt, grüne Farbtöne eine hohe Zahl. Quelle: i3mainz / Julia Ganitseva.

Abb. 6: Viewshed-Analyse für die Lokalisierung des Lutern-Epitaphs mit Sichtfeldern der Grabfigur. Quelle: i3mainz / Julia Ganitseva.

Zeitschrift für digitale Geisteswissenschaften

Artikel aus:

Sonderband 1 der ZfdG: Grenzen und Möglichkeiten der Digital Humanities. Hg. von Constanze Baum und Thomas Stäcker. 2015. DOI: [10.17175/sb01](https://doi.org/10.17175/sb01)

Titel:

Rot rechnen

Autor/in:

Waltraud von Pippich

Kontakt:

von.Pippich@campus.lmu.de

Institution:

Ludwig-Maximilians-Universität, München

GND:

[1084652536](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63864-p0011-9)

ORCID:

[0000-0002-4555-2816](https://orcid.org/0000-0002-4555-2816)

DOI des Artikels:

[10.17175/sb001_016](https://doi.org/10.17175/sb001_016)

Nachweis im OPAC der Herzog August Bibliothek:

[101422873](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63864-p0011-9)

Erstveröffentlichung:

14.04.2016

Lizenz:

Sofern nicht anders angegeben 

Medienlizenzen:

Medienrechte liegen bei den Autoren

Letzte Überprüfung aller Verweise:

24.05.2016

GND-Verschlagwortung:

[Kunstgeschichte \(Fach\)](#) | [Bildverarbeitung](#) | [Farbmessung](#) |

Zitierweise:

Waltraud von Pippich: Rot rechnen. In: Grenzen und Möglichkeiten der Digital Humanities. Hg. von Constanze Baum / Thomas Stäcker. 2015 (= Sonderband der Zeitschrift für digitale Geisteswissenschaften, 1). PDF Format ohne Paginierung. Als text/html abrufbar unter DOI: [10.17175/sb001_016](https://doi.org/10.17175/sb001_016).

Waltraud von Pippich

Rot rechnen

Abstracts

Rot ist eine Farbe und Farbe ist nicht messbar. Ist das so? Das Redcolor-Tool erfasst erstmals Röte in Bildern. Dies geschieht anhand einer Technologie, die jeder Farbe einen Rotwert zumessen kann. Der Beitrag spielt die Implikationen der Farbanalysen für die stilometrische und die historische Bildforschung durch und fragt nach dem Potential der Eigenschaftsmetrik zur Bildung allgemeiner Kategorien für die Geisteswissenschaften. Der stilometrische Teil widmet sich Rotfrequenzen und Farbsummen. Als methodischer Schlüssel wird eine die Stilometrie fundierende ›Theorie der Koordination‹ favorisiert. Der historische Teil erfasst Rot als Symbol der Herrschaft und Macht. Er behandelt exemplarisch Bilder von Tizians Karl V. bis zu Bushs Merkelporträt.

How would you measure color? The software Redcolor-Tool enables redness in images to be quantified for the first time. The technology attributes values of redness to any given color. This paper explores the methodological implications of color analysis for stylometric and historical approaches to visual studies. Furthermore, it examines the potential of new categories for the humanities which could emerge from quantitative feature metrics. The stylometry section is devoted to red color frequencies and color sums. A ›theory of coordination‹ based in stylometry is proposed as the methodological key. The historical section analyses red as the symbol of power and authority. Portraits from Titian´s Charles V to Bush´s portrait of Angela Merkel are analysed, taking the color red as the interpretational basis of traditional and contemporary political iconography.

1. Technologie des Redcolor-Tools

Farbe ist erkenntnistheoretisch der Akzidenz und nicht der Substanz zugeordnet, der Erscheinung und nicht dem Sein. Der Philosophie gilt sie als Paradefall der Qualia-Debatte, sie ist der Stachel im logozentrischen Weltbild. Als Element der Kunst ist sie kaum fassbar. Kunsttheoretisch traditionell dem Materiellen, nicht dem Geistigen zugewiesen,¹ ließe sich gar argumentieren, sie sei kein Gegenstand der geisteswissenschaftlichen Forschung. Beachtet sie die Kunstgeschichte, so geschieht das selten und dann entlang historisch ausgerichteter Fragestellungen.² Es ist möglich, erfolgreich ein Studium der Kunstgeschichte zu absolvieren, ohne jemals das Wort ›rot‹ gesagt zu haben. Über farbige Abbildungen von Bildern ist zu hören: ›Das ist alles falsch.‹ Farbe steht paradigmatisch für das Gebundensein des Forschungsgegenstandes an das wahrnehmende Subjekt und paradigmatisch für Fragiles, das durch jede Art der Reproduktion an ›Echtheit‹ verliert. Dennoch stellt sich die Frage: Ist Farbe messbar? Was ist zu leisten, wenn es Technologie gibt, die Farbe objektiv ermittelt?

¹ Dieser Beurteilung liegt ein starres hierarchisches Wertgefüge zugrunde. Farbtheorien des frühen 20. Jahrhunderts, die Farbe hingegen dem Geistigen zuordnen, zum Beispiel die Theorien von Wassily Kandinsky oder Johannes Itten, gleiten wiederum ins Esoterische ab.

² Max Imdahl untersuchte die Emanzipation der Farbe vom Primat der Linie in der klassischen akademischen Tradition, Imdahl 1988. Über eine historische Ausrichtung hinaus weisen einige Beiträge zum Thema ›Farbe‹ in den Publikationen Busch 2008, Wolfschmidt 2011 und Margrit Vogt / André Karliczek 2013.

Die im Ommer Lab der Ruprecht-Karls-Universität Heidelberg 2012 für die Erforschung von Farbe in Bildern entwickelte Software Redcolor-Tool ermöglicht präzise Analysen der Farbigkeit. Gegenstand der Farbanalysen sind digital vorliegende, entweder digitalisierte oder genuin digitale Bildelemente. Einem **Pixel** wird algorithmisch ein Farbwert zugewiesen. Numerisch wird der Farbwert in Zahlen gefasst, z.B.: Pixel x hat den Farbwert 0,00714. Das Messergebnis lautet 0,00714. Es gibt keine Nominalskala. Das Tool arbeitet ohne die Verwendung von Farbbegriffen und eliminiert mit der durch die Kluft zwischen Gegenstand (Analysandum) und Wort (Farbbegriff) für jede, nicht nur die automatisierte Attribution gegebenen Ungenauigkeit zugleich den methodischen Kardinalfehler. Denn Farbeigenschaften liegen nicht diskret, sondern kontinuierlich vor.

Die Technologie des Redcolor-Tools basiert auf dem **Lab-Farbraum**, der von der **Commission Internationale de l'Éclairage** zum internationalen Standard erhoben wurde. Die Metrik im Lab-Farbraum orientiert sich am menschlichen Sehen. Euklidische Abstände zwischen gleichwertigen Differenzen der Farbempfindung ermöglichen, die menschliche Wahrnehmung mathematisch zu approximieren. Es entsteht ein metrisches Skalenniveau. Die sogenannte Gleichabständigkeit im Lab-Farbraum beinhaltet, dass Farbwerte mit den gleichen metrischen Abständen auch den als gleich wahrgenommenen Differenzen des Farbeindrucks entsprechen. Das intensivste durch das menschliche Auge wahrzunehmende Rot wird durch die informatische Lösung des Redcolor-Tools dem Beginn der Skala (0,00001) subordiniert.

Farbe wird durch die Software also einerseits als Gegebenes (Analysandum) und andererseits als Wahrgenommenes (Metrik) erfasst. Die erarbeiteten Algorithmen führen den Parameter des betrachteten Objekts mit dem Parameter des betrachtenden Subjekts durch Pixelanalyse und Lab-Metrik auf technologischer Ebene zusammen.³ Farben, diese vormals lautlosen Elemente, haben durch den quantifizierenden Zugang zum Farbkosmos und die kalkulierte Subjekt-Objekt-Relation für die wissenschaftliche Forschung gleichsam eine Stimme gewonnen. Die folgenden Fallstudien kontrastieren die Implikationen der digitalen Messtechnologie für die **stilometrische** und die **historische** Bildforschung.

2. Stilometrische Bildforschung

Stilometrie erfasst formale und formalisierte Eigenschaften. Zu diesen Eigenschaften lassen sich neben Größen-, Zahl- und Mengenangaben von Linien, Flächen und weiteren Formen auch Farbeigenschaften zählen. Jedes Bild ereignet sich auf seine eigene Weise, – *individuum est ineffabile* – das Einzige ist unausschöpflich. Dennoch gehen stilometrische Arbeiten davon aus, dass sich Individual- und Epochenstil anhand einer Reihe formaler und formalisierter Merkmale eruieren lassen. Gewiss ist nicht jedes Merkmal in stilometrischen Untersuchungen adäquat erfasst, gewiss ist nicht alles zu berechnen. Doch in der Reduktion der Methoden auf wenige Parameter liegt ihre Effizienz.

³ Die informatische Lösung ist im **Heidelberg Collaboratory for Image Processing** (HCI), in der Computer Vision Group, Ommer Lab, Ruprecht-Karls-Universität Heidelberg, gelungen. Mein Dank gilt Björn Ommer für die Entwicklung des Redcolor-Tools.

Methodisch wird der Farbkreis als aus 16 Farbklassen bestehend verstanden. Die Farbklassen bilden die Farbwerte 1–16. Ein Farbwert wird definiert als diskrete Menge auf der Skala kontinuierlich nebeneinanderliegender feinsten farblicher Nuancen, z.B. ist Farbwert 1 die Menge der Nuancen $\{0,00001, \dots, 1,0\}$. Farbwert 1 entspricht intensivem Rot. Farbwert 16, der tiefem Dunkel gleicht, geht gemäß der spektralen Ordnung im Lab-Farbraum in Farbwert 1 über. Ein Bild wird als Komposition begriffen, die aus den 16 Farbklassen besteht. Das Bilden von Farbklassen birgt eine Reduktion farblicher Vielfalt, zugleich jedoch auch die methodische Erweiterung zur Exploration, Extraktion, Artikulation und möglichen Komparation farbformaler Eigenschaften.

Das Analysandum ist im vorliegenden Fall ein Korpus aus 50 Bildern (m01–m50) von Adolph Menzel.⁴ Digitalisierte Abbildungen der Bilder mit der Auflösung von 300 dpi werden von der Software analysiert und die Prozentzahlen der 16 Farbwerte bemessen auf die Gesamtfläche für jedes Bild berechnet. Bilddaten und Messdaten dieser Analysen sind der Tabelle (Abbildung 1) zu entnehmen. Ein Diagramm (Abbildung 2) der Farbwertkompositionen von 100 Gemälden, links die 50 Korpuswerke, rechts 50 aleatorisch gewählte, verschiedene Werke aus unterschiedlichen Epochen lässt das Vorliegen impliziter *patterns* auf der die homogene Autorschaft repräsentierenden Seite deutlich werden. Die im Diagramm aufscheinenden *patterns* sind durch die Extraktion farbformaler Eigenschaften explizit nachweisbar. Die Messdaten sind die Grundlage der Berechnungen von Farbfrequenzen und Farbwertrelationen. Es gilt, bekannte Muster zu identifizieren und noch namenlose Muster zu benennen.

Abb. 1: Tabelle der Metadaten, Messwerte und Angaben zu statistischen Frequenzen der Farbwerte (Lab-Farbraum, 16 Farbklassen-Modell, Software Redcolor-Tool, HCI) der Korpusbilder © Pippich 2014. Link auf Datei: Waltraud von Pippich: Frequenzen und statistische Dispersion der Farben in 50 Bildern von Adolph Menzel. 2014. Open data LMU Link: <http://dx.doi.org/10.5282/ubm/data.79>

⁴ Die Zuschreibung dieser Werke ist in der Forschung unbestritten, die Autorschaft scheint somit klar. Die Bilder entstammen einer Zeitspanne vom Frühwerk bis hinein ins Spätwerk. Es sind Ölgemälde auf Leinwand, Öl auf Papier (m01, m02, m11, m16), Aquarelle auf Papier (m22, m37, m45) und Gouachen auf Papier (m14, m29, m42). Die Bilder haben unterschiedliche Formate. Thematisch liegt eine Spannweite von Historie, historischem Genre über Genreszenen, einer Landschaftsdarstellung (m15) bis hin zur Wolkenstudie (m30) vor. Unter den Arbeiten sind sowohl Auftragswerke als auch autonom ausgeführte Gemälde. Die Bilder gelten in der Forschung als vollendet, ein Bild (m09) wird als unvollendet verhandelt, wurde vom Künstler aber signiert und datiert, vgl. die Diskussion und die Einordnung als vollendetes Werk bei Busch 2004, S. 87-89.

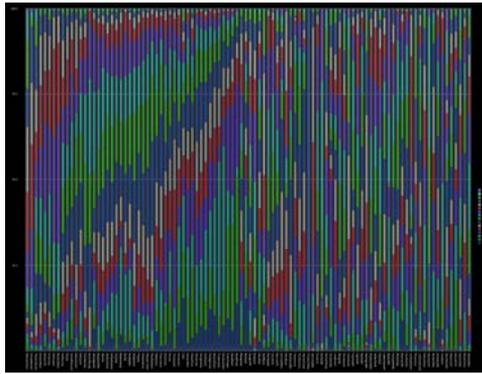


Abb. 2: Stilometrische Bildanalysen, Diagramm der Farbwertkompositionen von 100 Bildern, 50 Bilder von Adolph Menzel (links), 50 Bilder heterogener Autorschaft – von Giotto bis Yves Klein (rechts). Jede Säule im Diagramm repräsentiert ein Bild in seiner farblichen Komposition (Messtechnologie: Lab-Farbraum, 16 Farbklassen-Modell, Software Redcolor-Tool, HCl) © Pippich 2013.

2.1 Statistische Dispersion der Farbwerte

2.1.1 Statistische Variabilität

Die **Range** der Farbigkeit eines Bildes beschreibt die Spannweite der erzielten Messergebnisse auf der verwendeten Skala. Abzulesen ist aus diesem Wert die statistische Streuung und Variabilität der Messergebnisse. 30 der 50 Werke weisen die maximale Range 1–16 auf. 3/5 des Korpus erfüllen das Merkmal extensiver farblicher Variabilität. 2/5 hingegen zeigen eine reduzierte Range. Bei 17 Bildern beträgt diese 2–16, das heißt, die Range ist moderat reduziert. Gemäßigt ist die Variabilität zweier Gemälde mit der Range 3–16. Die kürzeste Range (4–16) wird für Menzels Bild *Berlin-Potsdamer Bahn* gemessen. Spektralanalysen (Abbildung 3),⁵ die die Messergebnisse in Farbe übersetzen und Bildelemente in Farbsummen zusammenführen, vermitteln die unterschiedliche Verteilung der Farbwerte in den Bildern mit minimaler und maximaler Range.

⁵ Der Farbeindruck der Spektralanalysen ist geräteabhängig, den Darstellungen liegt das RGB- oder CYMK-Modell zugrunde, die Lab-Metrik bestimmt die numerischen Werte.

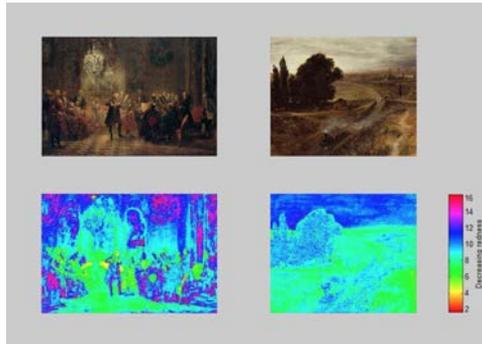


Abb. 3: Oben (von links): Adolph Menzel, *Flötenkonzert Friedrichs des Großen in Sanssouci*, 1852, Öl auf Leinwand, 142 x 205 cm, Berlin, Nationalgalerie, Quelle: Adolph Menzel. *Radikal real*. Kat. Ausst. Kunstthalle der Hypo-Kulturstiftung München und Kupferstichkabinett der staatlichen Museen zu Berlin. Hg. von Bernhard Maaz. München 2008, S. 17, Adolph Menzel, *Berlin-Potsdamer Bahn*, 1847, Öl auf Leinwand, 42 x 52 cm, Berlin, Nationalgalerie, Quelle: Adolph Menzel 1815–1905. *Das Labyrinth der Wirklichkeit*. Kat. Ausst. Nationalgalerie Berlin. Hg. von Claude Keisch und Marie-Ursula Riemann-Reyher. Köln 1996, S. 116, unten: Spektralanalysen, Lab-Farbraum, Software Redcolor-Tool (HCI) – die Korpusbilder mit maximaler (1–16) und minimaler (4–16) Farbränge © Pippich 2014.

Die **Standardabweichung** der Farbverteilung eines Bildes beziffert die durchschnittliche Differenz der Messergebnisse von ihrem Mittelwert. Der Mittelwert (6,25) ist die durchschnittliche Klassenhöhe.⁶ Die durchschnittliche Standardabweichung der Bilder liegt bei 6,4. Das heißt, Klassenhöhen um 12,7 und sehr geringe Klassenhöhen um 0,01 liegen im Korpus im Durchschnitt. Die höchste Standardabweichung erreicht die Farbkomposition des Gemäldes *Frühmesse in einer Salzburger Kirche* (10,79). Gewaltige Klassenhöhen in Kombination mit einer leeren Klasse und äußerst geringen Klassen erzeugen beträchtliche Differenzen vom Mittelwert. Erreicht ist eine niedrige Standardabweichung, die für ähnlich und gleich große Farbmengen steht und somit hinsichtlich der Proportion einzelner Farbklassen zueinander Gleichwertigkeit bezeichnet, tatsächlich einzig in einem *Biergarten* (3,37) betitelten Bild. Der Vergleich der Spektralanalysen der Bilder mit der geringsten und der höchsten Standardabweichung visualisiert die verschiedenen Frequenzen (Abbildung 4).

⁶ Die durchschnittliche Klassenhöhe berechnet sich aus der Summe der Ergebnisse geteilt durch die Anzahl der Ergebnisse. Für jedes Bild liegt die Summe bei 100 (% der Pixel, die Anzahl beträgt 16 (16 Farbklassen), $100:16 = 6,25$). Der Mittelwert jedes Bildes beträgt also 6,25. Ein Bild, das in jeder der 16 Farbklassen das Messergebnis 6,25 (% des Bildes) aufweisen würde, würde die Standardabweichung 0 erzielen. Es gäbe keine Abweichung vom Mittelwert.

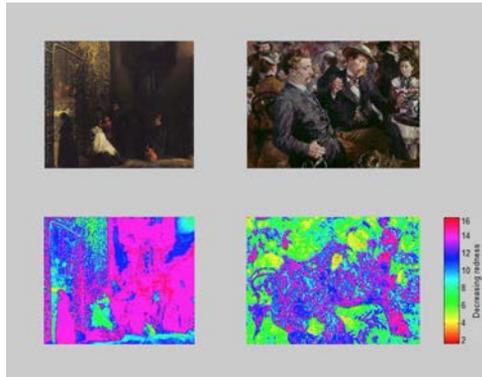


Abb. 4: Oben (von links): Adolph Menzel, *Frühmesse in einer Salzburger Kirche*, 1855, Öl auf Leinwand, 58,4 x 68,3 cm, Wien, Österreichische Galerie im Belvedere, Quelle: Adolph Menzel. Radikal real. Kat. Ausst. Kunsthalle der Hypo-Kulturstiftung München und Kupferstichkabinett der staatlichen Museen zu Berlin. Hg. von Bernhard Maaz. München 2008, S. 167, Adolph Menzel, *Biergarten*, 1883, Pinsel, Tusche, Deckfarben, 13,5 x 17,9 cm, Schweinfurt, Museum Georg Schäfer, Quelle: Adolph Menzel. Radikal real. Kat. Ausst. Kunsthalle der Hypo-Kulturstiftung München und Kupferstichkabinett der staatlichen Museen zu Berlin. Hg. von Bernhard Maaz. München 2008, S. 18, unten: Spektralanalysen, Lab-Farbraum, Software Redcolor-Tool (HCI) – die Korpusbilder mit der höchsten (links: 10,79) und geringsten (rechts: 3,37) statistischen Standardabweichung der Farbwerte (Mittelwert: 6,25) © Pippich 2014.

2.1.2 Berechnungen zentraler Tendenz

Der **Modalwert** ist der Wert, der am häufigsten gemessen wird. Es ist der Wert, dessen Farbe die größte Bildfläche bedeckt. Doch hier ist Vorsicht geboten: denn es liegen nicht nur zusammenhängende Gebiete vor, die Pixel können durch Pixel anderer Farbklassen getrennt sein. Im Korpus lassen sich zehn unterschiedliche Modalwerte zwischen Farbwert 4 und 15 feststellen.⁷ Geprüft werden auch die in einem Bild am zweit- und dritthäufigsten gemessenen Farbwerte. Die Mengen dieser Farbwerte werden mit den anderen Mengen im Bild verglichen. Gewisse Kombinationen fallen dabei auf. Es lassen sich Gruppen an Bildern identifizieren, die Eigenschaften teilen. Die Bestimmungen folgen dem Prinzip der Koordination als einer gleichordnenden, nicht hierarchisierenden Methode der farblichen Stilometrie und werden in [Abschnitt 2.3](#) vorgestellt.

Der **Median** teilt die Anzahl der Messergebnisse in zwei gleich große Hälften. Er wird auch das 50%-Quantil genannt. Die Lage des Median zeigt an, welche Stelle der Skala die Messergebnisse in jeweils 50% und 50% teilt. Der Median der Korpusbilder befindet sich an unterschiedlichen Positionen zwischen 5,5 bis hin zur Stelle 14,3. Dies weist auf ausgesprochene Varianz in der Bedeckung der Bildfläche mit Farben. So sind in der privaten Szene *Wohnzimmer mit der Schwester des Künstlers* 50% der Messergebnisse röter als der Farbwert 5,5. Und in der Nachtszene *Mondschein über der Friedrichsgracht im alten Berlin* sind 50% der Messergebnisse dunkler als der bereits dunkle Median 14,3. Die Spektralanalysen (Abbildung 5) der Gemälde bilden die Unterschiede der Farbdispersion ab.

⁷ Modalwerte sind die Farbwerte 4, 5, 6, 7, 8, 9, 10, 13, 14 und 15.

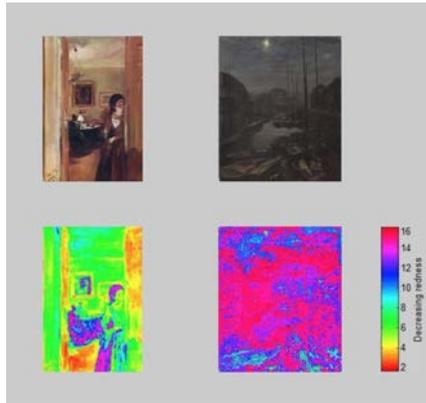


Abb. 5: Oben (von links): Adolph Menzel, *Wohnzimmer mit der Schwester des Künstlers*, 1847, Öl auf Papier, 46,1 x 31,7 cm, München, Neue Pinakothek, Quelle: Adolph Menzel 1815–1905. *Das Labyrinth der Wirklichkeit*. Kat. Ausst. Nationalgalerie Berlin. Hg. von Claude Keisch und Marie-Ursula Riemann-Reyher. Köln 1996, S. 119, Adolph Menzel, *Mondschein über der Friedrichsgracht im alten Berlin*, um 1855, Öl auf Leinwand, 39,5 x 33 cm, Berlin, Nationalgalerie, Quelle: Adolph Menzel 1815–1905. *Das Labyrinth der Wirklichkeit*. Kat. Ausst. Nationalgalerie Berlin. Hg. von Claude Keisch und Marie-Ursula Riemann-Reyher. Köln 1996, S. 177, (unten) Spektralanalysen, Lab-Farbraum, Software Redcolor-Tool (HCl) – die Korpusbilder mit der niedrigsten (links: 5,5) und höchsten (rechts: 14,3) Skalenposition des Median © Pippich 2014.

2.1.3 Berechnungen mittlerer Tendenz

Die **Interquartilsspanne** (IQR für engl. *interquartile range*) erweitert die Aussagekraft des Median. Sie gibt die Spannweite der Messwerte zwischen dem 25%-Quartil und dem 75%-Quartil an. Nun ist zu bestimmen, über wie viele und über welche Skalenwerte verteilt die mittleren, die extremen Lagen aussparenden 50% der Messergebnisse liegen. Bei den untersuchten Werken erstreckt sich die IQR über zwei Farbwerte hinweg bis zu einer Spanne von sechs Farbwerten, sie befindet sich an unterschiedlichen Lagen: Kurz ist sie im Bild *Emilie Menzel schlafend*. Zu den Bildern mit weiter IQR gehört *Kronprinz Friedrich besucht Pesne auf dem Malgerüst in Rheinsberg*. Eine Gegenüberstellung der Spektralanalysen der Bilder mit minimaler und maximaler IQR, bei unterschiedlicher Lage der IQR auf der Skala, vermittelt die Diversität der Farbkompositionen (Abbildung 6).

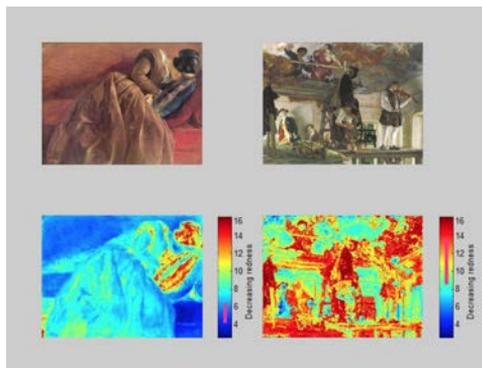


Abb. 6: Oben (von links): Adolph Menzel, *Emilie Menzel schlafend*, um 1848, Öl auf Papier, 46,8 x 80 cm, Hamburg, Hamburger Kunsthalle, Quelle: Adolph Menzel 1815–1905. Das Labyrinth der Wirklichkeit. Kat. Ausst. Nationalgalerie Berlin. Hg. von Claude Keisch und Marie-Ursula Riemann-Reyher. Köln 1996, S. 133, Adolph Menzel, *Kronprinz Friedrich besucht Pesne auf dem Malgerüst in Rheinsberg*, 1861, Gouache auf Papier, 24 x 32 cm, Berlin, Nationalgalerie, Quelle: Adolph Menzel 1815–1905. Das Labyrinth der Wirklichkeit. Kat. Ausst. Nationalgalerie Berlin. Hg. von Claude Keisch und Marie-Ursula Riemann-Reyher. Köln 1996, S. 203, unten: Spektralanalysen, Lab-Farbraum, Software Redcolor-Tool (HCl) – Korpusbilder mit geringer (2) und weiter (5) Interquartilsspanne, die Lage der Interquartilsspanne differiert: 4,5–6,5 (links), 8,5–13,5 (rechts), vgl. die Anzeige der Linie im Farbbalken © Pippich 2014.

2.1.4 Kunst und Durchschnitt

Die Werte der zentralen und mittleren Tendenz sind robust gegenüber **Ausreißern**, sie liefern jedoch hilfreiche Angaben im Rahmen der explorativen Statistik. Die Angaben sind modellbedingt und skalenabhängig, beschreiben aber die Phänomenologie der Bilder. Sie bieten einen analytischen Mehrwert innerhalb der Stilometrie, sie diversifizieren. Anders verhält es sich mit der Angabe eines Durchschnittswertes.⁸ Der Durchschnittswert eines Bildes ist nicht zwangsläufig im Bild selbst zu sehen. Wird der Durchschnittswert auf der Grundlage der Pixelauflösung und der Farbauflösung am Monitor nach dem **RGB-Modell** gebildet, so orientiert der Wert sich nicht an der Phänomenologie der Bilder, er ist auflösungs- und gerätebedingt. Er leistet keinen analytischen Mehrwert, er diversifiziert nicht, er vereinfacht.

Den digitalkünstlerischen Darstellungen Lev Manovichs liegen beispielsweise die durchschnittlichen Grauwerte der Helligkeit und Sättigung gemäß der Pixelauflösung am Monitorgerät zugrunde.⁹ Die vordergründig wie mathematische zweiachsige Koordinatensysteme funktionierenden Darstellungen zeigen auflösungs- und geräteabhängige Durchschnittswerte. Selbst wenn Manovich noch einmal 1 Mio. Bilder nach dieser Methode berechnen ließe, blieben die Ecken der nominell zu Koordinatensystemen deklarierten Abbildungen leer. In mathematischen Koordinatensystemen geben die Ecken die Extreme an. Doch ein System, das *per definitionem* nur Durchschnitte registriert, markiert die Ecken zwangsweise zu blinden Stellen. Zur Extraktion farbformaler Bildeigenschaften ist diese Methode der Medienkunst ungeeignet. Der Durchschnittswert erfasst keine Eigenschaft des Bildes.

Extreme und Ausreißer sind innerhalb des Bildlogos nicht zu unterschätzen. Sie spielen innerhalb des analytisch-stilometrischen Zugangs zu einem Kunstgebilde eben gerade keine marginale Rolle. Ihre eigentümliche Funktion im schöpferischen wie im wirkungsästhetischen Prozess ist kaum zu fassen.¹⁰ Die gegenüber extremen Skalenwerten robusten Werte der zentralen und mittleren Tendenz sind für die farbliche Stilometrie durch das Registrieren extremer und seltener Farbwerte zu ergänzen.

⁸ Der Durchschnittswert berechnet sich z.B. auflösungsbedingt aus der Summe der Werte der Pixel dividiert durch die Pixelanzahl. Ein zu gleichen Teilen aus 5-er und 15-er Werten bestehendes Bild hat den Durchschnittswert 10.

⁹ Manovich 2012, S. 249–278.

¹⁰ Aus stilometrischer Perspektive gilt es insbesondere zu bedenken: Wenn ein Künstler bzw. eine Künstlerin denkt, er oder sie habe mit einer spezifischen Farbwahl, mit einem einzigen silbernen Klecks, etwas Extraordinäres hinzugefügt, vielleicht liegt dies dann statistisch tatsächlich in seinem bzw. ihrem Durchschnitt und er oder sie verlässt somit sein bzw. ihr stilometrisches Muster durch diese Handlung nicht.

2.1.5 Extreme Skalenwerte

Was kann im Bereich Farbe als extrem angesehen werden? Ist es das Glitzernde, der besondere Schein? Statistisch lassen sich Extreme nur im Rahmen einer Skala bestimmen. Die künstliche, durch den Computer programmierte Hierarchisierung des seiner Natur nach nicht hierarchisch organisierten Farbkosmos definiert Farbnuancen um 0,01 und 16 als Extreme. Es ist dabei zwischen extremen und seltenen Skalenwerten zu unterscheiden. Im Menzel-Korpus koinzidieren diese Werte. Dabei liegen die roten Extreme bei den hier untersuchten Gemälden m01–m14 in einem höheren Bereich als bei den Gemälden m15–m50.

Ein Bild sei hinsichtlich seiner Komposition der Rotwerte dargestellt: das Gemälde *Ballszene* (Abbildung 7). Zu sehen ist eine Frau in einem roten Ballkleid als zentrale Figur in einer Szene, die sich am Rande des Parketttreibens abspielt. Dieses Gemälde weist im Verhältnis der Farbwerte $7/6/5/4/3/2$ zueinander kontinuierlich die Zahl Phi (1,61...) auf. Die Werte erzeugen zueinander die goldene Relation.¹¹ Das Bild lässt sich demnach als Farbkomposition begreifen, in der tiefrote Teile und Farben nachlassenden Rots in sich wiederholender Proportion aufeinander abgestimmt sind und sich durch die Zahl aufeinander beziehen. Durch die beschriebene Gleichabständigkeit im Lab-Farbraum betrifft eine Schilderung der Zahlenverhältnisse die Phänomenologie des Bildes. Es ist die Informationstechnologie, die Binnenstrukturen durch Numerik zur Artikulation verhilft und das Bild auf eine zuvor nicht mögliche Weise farbformal zu beschreiben hilft.

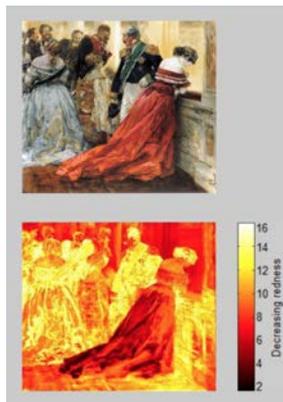


Abb. 7: Oben: Adolph Menzel, *Ballszene*, 1867, Pinsel, Deckfarbe, 26,5 x 30 cm, Schweinfurt, Museum Georg Schäfer, Quelle: Adolph Menzel. Radikal real. Kat. Ausst. Kunsthalle der Hypo-Kulturstiftung München und Kupferstichkabinett der staatlichen Museen zu Berlin. Hg. von Bernhard Maaz. München 2008, S. 199, unten: Spektralanalyse, Lab-Farbraum, Software Redcolor-Tool, HCI © Pippich 2014.

¹¹ Das Verhältnis des **goldenen Schnittes** wird in **Abschnitt 2.2.3** ausführlich besprochen. Die Prozentzahlen der für die *Ballszene* (m13) genannten Farbwerte liegen bei $9,65/6,195/3,768/2,21/1,64/1,0813$ und sind, wie die Zahlen für die im Folgenden dargestellten Rechnungen, der Tabelle (Abbildung 1) zu entnehmen. Weitere Bilder des Korpus, die im Verhältnis der Rotwerte Phi aufweisen, sind: m14 (Farbwerte $7/6/5/4$); m01 ($3/(1+2)$); m02 ($2/1$); m06 ($6/5/3/2$); m11 ($5/4$) und m03 – *Théâtre du Gymnase* –, das in **Abschnitt 2.2.3** analysiert wird.

Die Erfassung der dunklen Extreme vollendet die Berechnungen zur statistischen Dispersion der Farbwerte. Abermals fallen Größenverhältnisse der goldenen Proportion auf. Im Gemälde *Mondschein über der Friedrichsgracht im alten Berlin* bedeckt der Modalwert (15) die Leinwand zu 37,1%. Diese Zahl stellt den Rest des Bildes (62,9%) zum vom Modalwert bedeckten Teil und das gesamte Bild zu diesem Rest in das goldene Verhältnis. Die Bilder *Théâtre du Gymnase* und *Emilie Menzel schlafend* zeigen das goldene Verhältnis der dunkelsten Farbwerte (15/16) zueinander. Bei den Messungen zu Extremen fällt zudem auf, dass zwei Bilder in beiden Skalenextremen hohe Quantitäten erzielen: *Auf der Fahrt durch schöne Natur* und *Blick in einen kleinen Hof*.¹²

2.2 Farbsummenrelationen

2.2.1 Koordination und Bild

Der Logos des Bildes betrifft das Verhältnis der Teile zum Ganzen und das Verhältnis der Teile zueinander. Anders als der Logos der Subordination der Sprache, der aus Verhältnissen begrifflicher Unter- und Überordnung aufbaut, ereignet sich der Logos im Bild, nur im übertragenen Sinne ist von einem Logos zu sprechen.¹³ Es finden sich gleichordnende Verhältnisse der Kombination, des Konsonierens, des Simultanen. Es ist ein ›Logos der Koordination‹, des co-, des cum-, so wie es die lateinische Wortgruppe um *coordinatio* nahelegt. Der Seite einer möglichen Ko-Partizipation von Farben in multiplen Summenverhältnissen entspricht auf der Seite der Wahrnehmung eine ästhetische Kontemplation, die sich, anders als beim Lesen von Worten oder Hören von Sprache, ohne eine durch den Sinn vorgegebene Reihenfolge entfaltet.

Ein Schlüssel zur Erforschung des Farblogos im Bild liegt im Erfassen der Relationen einzelner Farbteile zueinander und zum Bildgesamt. Relationsstudien bieten methodisch entscheidende Vorteile. Sie eliminieren einige skalierungs-, modell-, auflösungs- und aufnahmebedingte Hürden, die Schwierigkeiten konsistenter Formalisierungsprozesse darstellen.¹⁴ Die Relation etwa ist modellbedingt, doch nicht stets skalierungsabhängig. Sie ließe sich bei anderen, die Farbklassen definierenden Schwellwerten im Lab-Farbraum ebenso nachweisen. Die lineare Skala löst sich zudem bei der Beachtung der Relationen im spektralen Farbganzen zum Zyklus, die ursprüngliche, nicht hierarchische Natur des Spektrums entfaltet sich, der Farbkreis schließt sich. Relationen leisten Konnex zwischen Gliedern sowie zwischen Partikeln und Ganzem. So werden die Berechnungen zur Relation zum Instrument der Artikulation von Binnenstrukturen. Auf der Suche nach impliziten *patterns* bietet ein Erfassen

¹² Der Versuch einer Interpretation findet sich in [Abschnitt 2.4](#).

¹³ Wittgenstein fasst in seinen späten *Bemerkungen über die Farben* die systematische Gleichordnung im Ausdruck »das Zusammen der Farbtöne«, Wittgenstein 1984, S. 35. Prägnant hat Gabriel 2010, S. 376, das Andere der Logik des Ästhetischen, seine komplementäre Stellung zur apophantische Aussagen zeugenden Sprache, wiedergegeben.

¹⁴ Zu den aufnahmebedingten Hürden: Bis heute wurden keine Kriterien und Maßstäbe für die professionelle Bildreproduktion formuliert. Es ist die Crux der Arbeit mit Reproduktionen, dass diese einen **Farbstich** aufweisen können. Verschiedene vorgenommene Messungen lassen annehmen, dass bei Bildern mit Farbstich in den Relationen der Farbglieder zueinander die für den erschließenden Vergleich notwendige Stabilität bestehen kann. Die Relation kann in den Bildern mit und ohne Farbstich stabil bleiben.

der Relation die Aussicht, einen Verhältniswert in weitere, komparative Forschung einbinden zu können.

Wie verhalten sich nun im hier untersuchten Korpus einzelne Farbwerte und Summen der Farbwerte zueinander? Einige Zahlen tauchen häufiger auf. Modellbedingte¹⁵ und zahlentheoretische¹⁶ Eigenschaften der Zahlen beleuchten weniger die Natur des individuellen Bildes als diejenige des Modells und der Zahlen. Bekannte mathematische Relationsbeschreibungen reichen für eine systematisch stilometrische Analyse der Farbigkeit in Bildern nicht aus und sind um Prinzipien zu erweitern, die durch den Kunstorganismus Bild entstehen und ihn deshalb beschreiben.¹⁷

2.2.2 Silberne Relation

Das Bild *Eisenwalzwerk* (m31) weist die Farbwerte 8–13 in der folgenden, eigentümlichen Staffellung auf (in %): 7/14/21//15/10/5.¹⁸ Der Modalwert lautet 10 (21%), der zweithäufigste Farbwert ist 11 (15%). Das heißt, die Mengen dieser beiden exponierten Farbwerte ereignen sich zu rötteren und weniger roten Farbwerten in einer 7-er und in einer 5-er Rhythmik. Stabil aber sind dabei die Relationen 2 und 3, die der 7/14/21-Reihe ebenso wie der 5/10/15-Reihe zu Grunde liegen, stabil ist die Relation der Glieder. Hier zeigt sich, dass nicht die absolute Prozentangabe, sondern die Relation der Glieder zueinander einen Anker für den weiteren forschenden Prozess bietet.

Forscht man etwa in m31 an der beschriebenen Stelle weiter, auf die Relation achtend, so zeigt sich, dass die beschriebenen Farbwerte gemeinsam eine Farbsumme von etwas über 70% bilden. Die Quantität approximiert zum Rest des Bildes die **silberne Relation** (2,41...). Es finden sich auch weitere, die silberne Relation annähernde Verhältnisse im Bild.¹⁹ Die Zahlen unterscheiden sich, doch die Zahlenverhältnisse der Farbklassen sind dieselben. Die Relation fungiert als Konstante zwischen den Größen der Glieder und artikuliert ihr Binnenprinzip. Die in rationalen Zahlen nicht vollständig auszudrückende silberne Relation findet durch die Ästhetik zweier Quantitäten vollständigen Ausdruck, ihre Rolle im Farblogos der Bilder ist noch rätselhaft.

2.2.3 Goldene Relation

¹⁵ Hier sind die durch das 16 Farbklassen-Modell bedingten Eigenschaften gemeint.

¹⁶ Z.B. die Primeigenschaft.

¹⁷ Während im Alltag häufig verwendete Verhältnisse (1/2, 1/4) sich meist auf die Teilung eines bereits bestehenden Ganzen mit oft kontingenter Form beziehen, werden einige Muster, die sich aus Artefakten ableiten lassen, auch die Eigenschaft haben, den diesen Erzeugnissen vorangehenden schöpferischen Prozess zu beschreiben.

¹⁸ Der Doppelschrägstrich markiert die Trennung der Prozentzahlen für den Modalwert und den zweithäufigsten Farbwert.

¹⁹ Die Verhältnisse 5/2, 12/5, 29/12, 70/29; Pell-Zahlen (0, 1, 2, 5, 12, 29, 70, ...) die den Divisor der sich logisch fortsetzenden, rekurrenten Reihe bilden, sind als Prozentanteile im Farborganismus von m31 den Messungen zu entnehmen. Pell-Zahl 29 (Farbwerte 11–13), Dividend 41 (Farbwerte 8–10), Pell-Zahl 70 (Farbwerte 8–13), Dividend 99 (Farbwerte 1–16). Und: Die beiden Summen der drei weiteren, auf der Skala jeweils dem Modalwert benachbart liegenden Werte bilden annähernd 29% und stehen jeweils zur beschriebenen Farbsumme der Farbwerte 8–13 in der silbernen Relation.

Warum steht die irrationale Zahl Phi (1,61...) für ein Größenverhältnis, das sich auch *proportio divina* nennt? Das Göttliche hat einen Bezug zum Schöpferischen. Durch einen einzigen Akt der Teilung wird mehr als ein einziger Bezug erreicht. Eine Teilung erzeugt zugleich zweimal dieselbe Relation, das ist eigentlich ein Wunder und jedenfalls eigenartig. Zugleich wird das Gesamt zum größeren Teil und der größere Teil zum kleineren in das von Phi repräsentierte Verhältnis gesetzt. Das Teilungsverhältnis wurde als ein harmonisches beschrieben, die entstehenden Relationen golden und die Teilung der goldene Schnitt genannt.

Zahlen der **Fibonacci-Folge** (1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, ...) sind häufig involviert, wenn das Verhältnis des goldenen Schnittes vorliegt. Sie erreichen im Verhältnis zur in der Fibonacci-Folge vorangehenden Zahl weitgehende Annäherungen an Phi. Jede Zahl der Reihe ist die Summe der beiden vorangehenden Zahlen. Die Fibonacci-Folge repräsentiert ein sich selbst logisch fortsetzendes, selbstursprüngliches Prinzip. Die Kunstproduktion hat Affinität zu dieser Eigenart.

Wegweisend sind die Forschungen Werner Buschs zu Verhältnissen des goldenen Schnittes in der Linien- und Flächenkomposition gerade auch in Menzels Bildern.²⁰ Es gibt demnach eine durch Zahlen zu artikulierende Ordnung in dieser Kunst. Phi findet sich auch im Verhältnis von Farbquantitäten in Menzels Gemälden.²¹ Der Farborganismus nach der goldenen Relation wird hier für drei Gemälde vorgestellt. Die Kreisdiagramme, die die Größenverhältnisse der Farbwerte und zugleich ihren Ort in der spektralen Ordnung abbilden, übernehmen dabei eine mehr als illustrierende Funktion. Denn das durch Phi repräsentierte mathematische Verhältnis ist nicht in rationalen Zahlen darstellbar, mit Worten kaum zu beschreiben. Visuell aber ist es vollständig durch die Perzeption des Größenverhältnisses erfassbar. So wird die Ansicht der Diagramme ein Verständnis für die spezielle Relation erst erzeugen helfen.

Das Bild *Balkonzimmer* (m12, Abbildung 8) gehört zu den besonders bekannten Gemälden Adolph Menzels. Spektralanalysen (Abbildung 9) visualisieren Farbwertgruppen, Kreisdiagramme (Abbildung 10) legen die Proportionsverhältnisse dar. Das ›Menzel-Modalwertpaar‹²² bedeckt die Leinwand zu 37,17%. Der Rest des Bildes steht somit zum Modalwertpaar und das ganze Bild zum Rest im goldenen Verhältnis. Nicht weniger als sechs Mal findet sich diese Farbproportion im Bild.²³ Die Fibonacci-Zahl 55 (Werte 9–12) erzeugt die goldene Relation im Verhältnis zur Fibonacci-Zahl 34 (Werte 5–8).²⁴ Noch einmal ertönt die

²⁰ Busch 1996, S. 457–468 und Busch 2004, zum Größenverhältnis des goldenen Schnittes im Bild *Berlin-Potsdamer Bahn* (m15), S. 75; im Bild *Aufbahrung der Märzgefallenen* (m09), S. 91; über die »abstrakte Ordnung im konkreten Chaos« und den »Flächenorganismus« des Bildes *Pariser Wochentag* (m04), S. 116.

²¹ Einige Verhältnisse wurden im **Abschnitt 2.1.5** »Extreme Skalenwerte« bereits formuliert.

²² Ob ihres prominenten Auftauchens im Korpus kann die Kombination von Modalwert 9 mit zweithäufigstem Wert 10 und von Modalwert 10 mit zweithäufigstem Wert 9 vorläufig ›Menzel-Modalwertpaar‹ genannt werden. 1/5 der untersuchten Bilder, darunter berühmte Bilder Menzels, das *Flötenkonzert in Sanssouci* (m27) und die von Hofmann 1982, S. 31–40, so eindrücklich als »Menzels verstecktes Manifest« geschilderte *Atelierwand* (m19), gehören in diese Gruppe.

²³ Die Farbsummen der Werte 1–8 (37,52%), der Werte 10–12 (37,9%), der Werte 12–6 (39%) und der Werte 13–7 (38%) erzeugen das Verhältnis Phi.

²⁴ Angabe in Prozent. Die Angaben zu den Fibonacci-Zahlen sind allesamt im Folgenden als Prozentangaben für das Bild zu verstehen. Aus Gründen der leichteren Lesbarkeit werden die Prozentzeichen im Text weggelassen, ›Farbwert‹ wird zu ›Wert‹ gekürzt.

Fibonacci-Zahl 55 (Werte 1–9) und erzeugt Phi im Verhältnis zur Fibonacci-Zahl 34 (Werte 1–8), die sie selbst beinhaltet.²⁵



Abb. 8: Adolph Menzel, *Balkonzimmer*, 1845, Öl auf Pappe, 58 x 47 cm, Berlin, Nationalgalerie, Quelle: Adolph Menzel 1815–1905. *Das Labyrinth der Wirklichkeit*. Kat. Ausst. Nationalgalerie Berlin. Hg. von Claude Keisch und Marie-Ursula Riemann-Reyher. Köln 1996, S. 91.

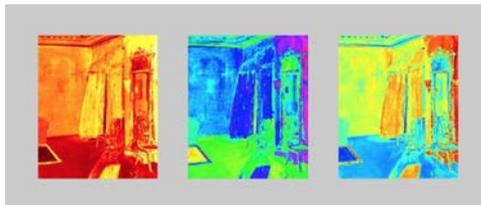


Abb. 9: Spektralanalysen, Adolph Menzel, *Balkonzimmer*, Software Redcolor-Tool, HCI © Pippich 2012.

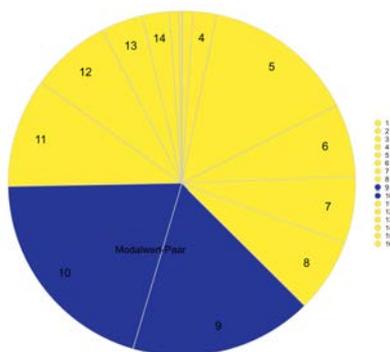


Abb. 10: Diagramm der Farbwerte (Lab-Farbraum, 16-Farbklassen-Modell, Software Redcolor-Tool, HCI), Darstellung der Fibonacci-Relationen: : 100% / gelbes Segment, gelbes Segment / blaues Segment, Adolph Menzel, *Balkonzimmer* © Pippich 2014. Weitere goldene Relationen mit Link auf Datei: Waltraud von Pippich:

²⁵ Adäquatere Visualisierungen der multiplen Farbsummenrelationen befinden sich in der Entwicklung.

Goldene Relationen in Farbkompositionen von Adolph Menzel. Das Gemälde *Balkonzimmer* Open data LMU
Link: <http://dx.doi.org/10.5282/ubm/data.80>

Chaos begegnet dem Betrachter von Menzels Gemälde *Piazza d'Erbe in Verona* (m28 (Abbildung 11), Spektralanalysen (Abbildung 12), Kreisdiagramm (Abbildung 13)), das den ganz gewöhnlichen Alltag auf einem italienischen Wochenmarkt schildert. Werner Busch detektierte die Regeln im Flächenorganismus dieses Bildes.²⁶ Farblich ist ebenfalls Ordnung im Chaos zu entdecken: Eigentümlich färbt das »Menzel-Modalwertpaar« jeweils zu 21% das Bild. Fibonacci-Relationen im Bild sind 8/13/21//21/13/8/5 (Werte 7–12), 55/34 (Werte 10–16, 8–9), abermals 55/34 (Werte 8–10, 11–16), 34/21 (Werte 11–16, 10), 34/21 (Werte 8–9, 10), Wert 9 zu 8, der Wert 8 annähernd zu Wert 7. Die Farbsumme der Werte 9–12 approximiert im Verhältnis zum Leinwandrest den goldenen Schnitt.



Abb. 11: Adolph Menzel, *Piazza d'Erbe in Verona*, 1882–84, Öl auf Leinwand, 74 x 127 cm, Staatliche Kunstsammlungen Dresden, Galerie Neue Meister, Quelle: Adolph Menzel 1815–1905. Das Labyrinth der Wirklichkeit. Kat. Ausst. Nationalgalerie Berlin. Hg. von Claude Keisch und Marie-Ursula Riemann-Reyher. Köln 1996, S. 318.

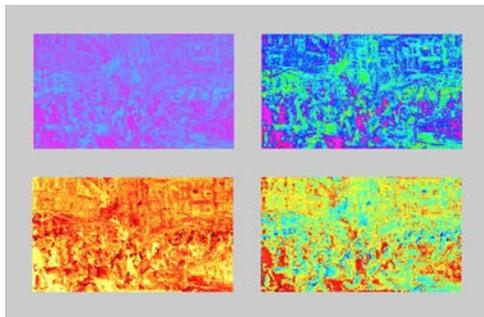


Abb. 12: Spektralanalysen, Adolph Menzel, *Piazza d'Erbe in Verona*, Software Redcolor-Tool, HCI © Pippich 2012.

²⁶ Busch 2004, S. 118–123.

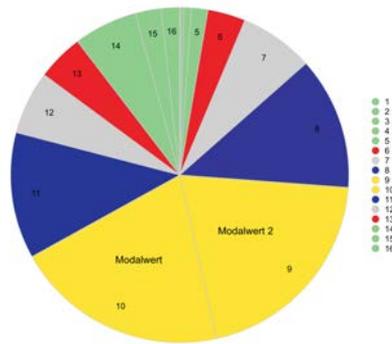


Abb. 13: Diagramm der Farbwerte (Lab-Farbraum, 16 Farbklassen-Modell, Software Redcolor-Tool, HCl), Darstellung der Fibonacci-Relationen: (jeweils) gelbes Segment / blaues Segment, blaues Segment / graues Segment, graues Segment / rotes Segment, rot-grünes Segment / blaues Segment, Adolph Menzel, *Piazza d'Erbe in Verona* © Pippich 2014. Weitere goldene Relationen Link auf Datei: Waltraud von Pippich: Goldene Relationen in Farbkompositionen von Adolph Menzel. Das Gemälde *Piazza d'Erbe in Verona*. 2012. Open data LMU Link: <http://dx.doi.org/10.5282/ubm/data.83>

Das Gemälde *Théâtre du Gymnase* (m03 (Abbildung 14), Spektralanalysen (Abbildung 15), Kreisdiagramm (Abbildung 16)) stellt eine Besonderheit der Menzelsforschung dar. Einzig dieses Bild hielt dem an reiner Farbenwirkung orientierten Urteil Julius Meier-Graefes stand, als dieser in seiner Menzel-Monographie die Werke des Malers vor dem Hintergrund zeitgleicher moderner Bestrebungen der Kunst in Frankreich kritisierte.²⁷ Starke Rotwerte (1–5) bilden die goldene Relation zum Leinwandrest. Fibonacci-Relationen sind 55/34/21 (Werte 1–7, 8–10, 11–3), 21/13 (Werte 3–4, 5). Die Fibonacci-Zahl 89 (Werte 4–16) steht im Bild dem starken Rot (1–3) entgegen. Einer Reprise gleich erklingt dieses Verhältnis im anderen Skalenextrem. Die Fibonacci-Zahl 89 (Werte 1–10) steht dem Dunkel (11–16) entgegen. Gleichzeitig, einer anderen Ordnung als der der Fibonacci-Zahlen entspringend, summieren sich die vier einander im Kreisdiagramm chiasmatisch gegenüberliegenden, jeweils die Fibonacci-Zahl 13 annähernden Farbwerte 4, 5, 9, 10 zu 50% und färben gemeinsam eine Hälfte des Bildes. Das ›Menzel-Modalwertpaar‹ trifft seine farblichen Antagonisten.

²⁷ Meier-Graefe 1906, S. 129–138, 142–144, 167, 182–184, 211. Das Gemälde weist den röttesten Modalwert im Korpus auf, Modalwert 4.



Abb. 14: Adolph Menzel, *Théâtre du Gymnase*, 1856, Öl auf Leinwand, 46 x 62 cm, Berlin, Nationalgalerie. Quelle: Adolph Menzel 1815–1905. *Das Labyrinth der Wirklichkeit*. Kat. Ausst. Nationalgalerie Berlin. Hg. von Claude Keisch und Marie-Ursula Riemann-Reyher. Köln 1996, S. 27.

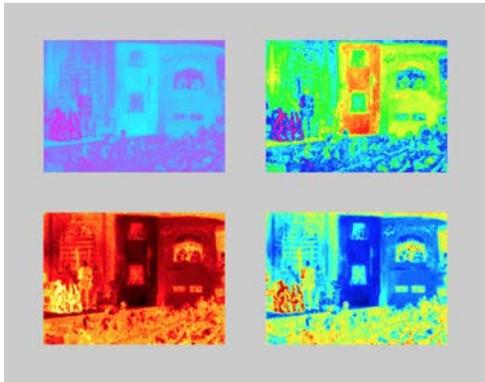


Abb. 15: Spektralanalysen, Adolph Menzel, *Théâtre du Gymnase*, Software Redcolor-Tool, HCI © Pippich 2012.

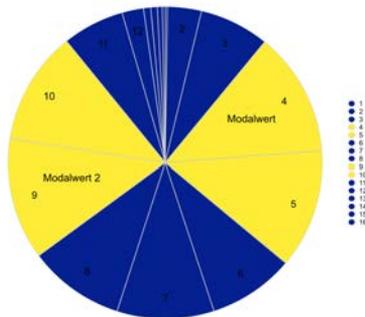


Abb. 16: Diagramm der Farbwerte (Lab-Farbraum, 16 Farbklassen-Modell, Software Redcolor-Tool, HCI), Darstellung der Fibonacci-Relationen, Adolph Menzel, *Théâtre du Gymnase* © Pippich 2014. Link auf Datei: Waltraud von Pippich: Goldene Relationen in Farbkompositionen von Adolph Menzel. Das Gemälde *Théâtre du Gymnase*. 2014. Open data LMU Link: <http://dx.doi.org/10.5282/ubm/data.84>

Wentworth Thompson erforschte den Zusammenhang zwischen dem Wachstum natürlicher Organismen und den Zahlen der Fibonacci-Folge.²⁸ Einzelnen Gliedern eignet durch das goldene Verhältnis zueinander eine besondere Effizienz und Stabilität. Beim Malen entstehen auf der Leinwand künstlich wachsende Farbsummen. Die Farbglieder weisen in Menzels Bildern zueinander häufig die goldene Relation auf, wie gezeigt wurde. Die Bedeutung des goldenen Verhältnisses für jede Farbkreation kann hier nicht erschöpfend bestimmt werden. Der Zusammenhang der Farbkombinationen zur Logik rekurrenter mathematischer Reihen und die in Zahlen zu erfassenden Gemeinsamkeiten zwischen natürlich und künstlich wachsenden Organismen aber ist im Blick zu behalten.

2.3 Eigenschaftsextraktion durch methodische Koordination

Die Messergebnisse zeigen, dass viele Gemälde das Modalwertpaar teilen. Es war nicht vorauszusehen, dass sich diese Bilder auch in der erreichten Range und Standardabweichung, der Position der Quartile und der Interquartilsspanne gleichen oder ähneln würden. Die Eigenschaftsbündel sind wertvoll zur Detektion der *patterns*. Die Verteilung der Farbwerte der Korpusbilder kann in einem kumulativen Frequenzhistogramm (Abbildung 17) dargestellt werden. Die horizontale Achse zeigt die 16 Farbwerte, die vertikale Achse die Prozentzahl bemessen auf die Gesamtfläche des Bildes von 100%. Graphenbündel in der Mitte sowie auf der linken und rechten Seite des Histogramms sind zu determinieren. Das Frequenzhistogramm von 16 Bildern des Korpus zeigt Gruppen, die Eigenschaften teilen. In der Mitte sind die Graphen der prominenten Werke mit dem ›Menzel-Modalwertpaar‹ zu verorten.²⁹ Eine lichtere Farbdispersion weisen die Graphen im Histogramm links auf. Hier gibt es Frequenzen intensiven Rots³⁰ und moderate Rotfrequenzen.³¹ Rechts im Histogramm kumulieren die Graphen der Bilder mit dunklem Modalwert, die wiederum weitere farbformale Eigenschaften teilen.³²

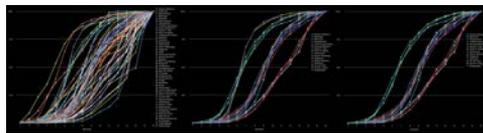


Abb. 17: Kummulative Frequenzhistogramme: Farbverteilung in 50 Bildern von Adolph Menzel (links), Graphenbündel: Farbverteilung in 16 Bildern (Mitte), Graphenbündel: Farbverteilung in 14 Bildern (rechts). Die horizontale Achse der Histogramme zeigt die 16 Farbwerte, die vertikale Achse zeigt die Prozentzahl (Messtechnologie: Lab-Farbraum, 16 Farbklassen-Modell, Software Redcolor-Tool, HCI) © Pippich 2014.

Neben silbernen und goldenen Relationen gehören Modalwerte, Modalwertpaare, Zahlen zu extremen und seltenen Frequenzen, weitere deskriptiv statistische Angaben sowie

²⁸ Wentworth Thompson entdeckte Zahlen der Fibonacci-Folge in den rechts- und linksdrehenden Reihen der Schuppenglieder von Zapfen und im Blütenstand von Blumen, vgl. *Wentworth Thompson 1992*, S. 921–924.

²⁹ Modalwertpaar 9–10 und 10–9 (Modalwert-zweithäufigster Wert), Bilder m19, m25, m26, m27, m28, m29, und Modalwertpaar 8–9, Bild m18.

³⁰ Modalwertpaar 7–6, Bilder m07, m08.

³¹ Modalwertpaar 6–7 und 7–6, Bilder m04, m05, m06.

³² Modalwertpaar 13–10, 13–14, 14–10, 14–16, Bilder m36, m37, m40, m41. Eine Korpuserweiterung ließe zu, die entdeckten *patterns* weiter zu diversifizieren.

Kombinationen dieser Eigenschaften zu *patterns*, die sich gemäß der koordinierenden Methode extrahieren lassen. Im Rahmen der Koordination wird ein Mit- und Nebeneinander, Ähnlichkeit, Unterschied, ein sich Drängen und eine Rhythmik im Aufbau der Elemente erfassbar. Die systematische Koordination ist die methodische Grundlage digitaler farbanalytischer Stilometrie.

Eine Folge der koordinierenden Methode ist, dass fehlende und überschüssige Pixel gleichermaßen zum Vorliegen von *messy data* beitragen. Innerhalb einer als allgemein zu verstehenden ›Demokratie der Pixel‹ trägt jedes Bildelement zum Ergebnis bei, dem Ansatz der ikonographisch identifizierenden, subordinierenden Methode entgegengesetzt.³³ Mit der Ebene der Subordination ist das Feld auch um mögliche Fehler erweitert. Der folgende Versuch der Interpretation einiger detektierte Muster ist in all den Schranken zu verstehen, die eine solche Untersuchung mit sich bringt.

2.4 Semantische Subordination

In Bildern mit sehr hoher Standardabweichung stehen gewaltige Farbmassen geringsten Mengen entgegen und erzeugen den Eindruck des Überdimensionalen. Ikonographisch gehören im untersuchten Korpus Bilder mit kolossaler Standardabweichung zur Vergegenwärtigung des Sakralen (*Frühmesse in einer Salzburger Kirche, Kircheninterieur*). Farbformalistisch diametral entgegengesetzt ist die Komposition gleich großer Farbwertklassen in Schilderungen des Profanen, so im *Biergarten*. Geringen Unterschieden der einzelnen Klassenhöhen entspricht eine geringe Standardabweichung. Die das Vertraute betonende Darstellung *Schlafzimmer des Künstlers in der Ritterstraße* und die das Unkonventionelle und Leichtigkeit verkörpernde Szene *Kronprinz Friedrich besucht Pesne auf dem Malgerüst in Rheinsberg* weisen ebenso sehr geringe Standardabweichungen auf.

Extreme Skalenwerte finden sich jeweils in großen Mengen in den Bildern *Auf der Fahrt durch schöne Natur* und *Blick in einen kleinen Hof* (Abbildung 18). Die inhaltlichen Innen-Außen, Natur-Kultur, Nah-Fern-Dualismen werden von formalen, tief kontrastierenden Farbkombinationen von starker Röte und Unröte wie Rot-Grün begleitet. Die Interquartilsspanne kann bei Gemälden privater, beengter Raumsituation schmal sein wie im Bild *Emilie Menzel schlafend*. Maximal kann sie erscheinen, wenn räumliche Weite und der Kontrast Nah-Fern dargestellt sind, so in den Gemälden *Blick in einen kleinen Hof, Palaisgarten des Prinzen Albrecht, Maskensouper* und *Blick von einem Fenster des Berliner Schlosses*. All diese Beobachtungen bleiben Versuche, eine Analyse der Stadien des Skizzenwerkes und eine Korpuserweiterung ließen die Aussagen präzisieren.

³³ Überpointiert lässt sich formulieren, das praktizierte Abschneiden von Bildrändern beispielsweise in Museums- und Ausstellungskatalogen folge dem Motto ›Das ist in Ordnung, das ist nur der Rand, die Figur ist ja noch zu sehen und bleibt erhalten.‹ Dieser Ansatz ist für die jedes Pixel des Bildes notwendig erfassende Stilometrie ungeeignet.



Abb. 18: Oben (von links) Adolph Menzel, *Auf der Fahrt durch schöne Natur*, 1892, Deckfarben auf Papier, 27,7 x 37,2 cm, Privatsammlung, Quelle: Menzel – der Beobachter. Kat. Ausst. Kunsthalle Hamburg. Hg. von Werner Hofmann. München 1982, S. 280, Adolph Menzel, *Blick in einen kleinen Hof*, 1867, Wasser- und Deckfarben auf Papier, 28 x 22 mm, Berlin, Nationalgalerie, Quelle: Menzel – der Beobachter. Kat. Ausst. Kunsthalle Hamburg. Hg. von Werner Hofmann. München 1982, S. 171, unten: Spektralanalysen (Lab-Farbraum, Software Redcolor-Tool, HCI) – die Bilder im Korpus mit maximalen Messwerten in beiden Skalenextremen © Pippich 2014.

Die vorgenommenen stilometrischen, ohnehin auf Prinzipien der Komparation beruhenden Arbeiten eröffnen die Möglichkeit einer Fülle weiterer, vergleichend vorgehender Studien, etwa zu Farbrelationen in unterschiedlichen Stadien des Werkprozesses,³⁴ im Werk anderer, zeitgenössischer wie nicht zeitgenössischer Künstler aus ähnlichen oder fernen Regionen. Und immer wieder stellt sich die Frage, an welcher Stelle der Schritt auf die semantische Ebene vollzogen werden könnte und die formalen Eigenschaften der Bilder auf Themen und Motive, auf Sinn, zu beziehen sein könnten.

3. Zur Historie von Rot in Herrscherbildern

Mit der Studie zur Historie von Rot in Herrscherbildern und dem Erfassen des Symbolgehalts wird die Farbe nun in eine kategorial verschiedene Fragestellung eingebunden. Es ist die Frage nach Persistenz und Wandel farblicher *patterns* im Laufe der Zeit. Die Analysen folgen dem Prinzip der systematischen Koordination innerhalb des 16 Farbklassen-Modells, die Analyse-Software ist das Redcolor-Tool. Eine Tabelle (Abbildung 19) enthält die Metadaten und Messwerte von Herrscher- und Politikerbildern (h01–h50), die hier in einer notwendig verkürzenden Weise exemplarisch besprochen werden. Die Abbildungen und Rotspektralanalysen sind einer weiteren Tabelle (Abbildung 20) zu entnehmen. Welcher Art sind die Erkenntnisse, die das Rechnen von Rotwerten innerhalb einer historischen Reihenbildung erwarten lässt? Entstehen Momente, in denen das Auge den Algorithmus ergänzt? Sind auch ihre Kategorien für und durch Informatik zu formalisieren?

³⁴ Für Menzel könnte in Richtung der Ölskizzen weiter geforscht werden. Dadurch ließen sich die Thesen zum Protoimpressionismus empirisch stützen. Zur Eigenständigkeit des Malerischen und der Farbe bei Menzel vgl. Kohle 2001, S. 195–207.

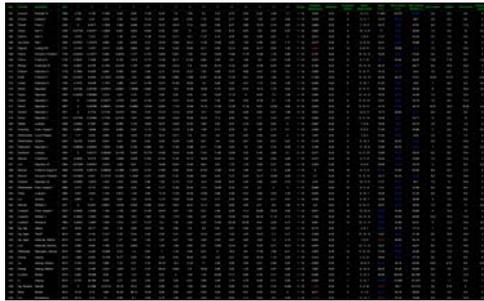


Abb. 19: Tabelle der Metadaten, Messwerte und Angaben zu statistischen Frequenzen der Farbwerte (Lab-Farbraum, 16 Farbklassen-Modell, Software Redcolor-Tool, HCl) in Herrscher- und Politikerbildern © Pippich 2014. Link auf Datei: Waltraud von Pippich: Rotfrequenzen und statistische Farbdispersion in Herrscher- und Politikerbildern (1360-2014). 2014. Open data LMU Link: <http://dx.doi.org/10.5282/ubm/data.81>



Abb. 20: Tabelle der Abbildungen und Rotspektralanalysen (Lab-Farbraum, 16 Farbklassen-Modell, Software Redcolor-Tool, HCl) und Metadaten von Herrscher- und Politikerbildern © Pippich 2015. Link auf Datei: Waltraud von Pippich: Rotspektralanalysen von Herrscher- und Politikerbildern (1360-2014). 2014. Open data LMU Link: <http://dx.doi.org/10.5282/ubm/data.82>

3.1 Rot als Symbol

Rot gilt als die Farbe des Lebens, der Macht, Potenz und Präsenz, der Heiterkeit und des Herrschens. Das Durchrötete, pulsierend Durchblutete wird vom Menschen unmittelbar als belebt verstanden. Kostbar und schwierig herzustellen waren lange Zeit rote Gewebe. Das Tragen roter Gewänder war bis ins Mittelalter hinein streng reglementiert und nur privilegierten Bevölkerungs- und Berufsgruppen, etwa dem Adel und den Juristen, gestattet. Rot ist traditionell die Farbe des Herrschers. Als Farbe der Jakobinermützen wird Rot zugleich unübersehbares Zeichen der Französischen Revolution. Rot ist 1834 die Farbe der Seidenweberaufstände in Lyon und im weiteren Laufe des 19. Jahrhunderts die Farbe unterschiedlicher Emanzipationsbewegungen und der sozialistischen Parteien.³⁵ Das 20. Jahrhundert kennt wie kein anderes die Beeinflussung der Menschen durch die Ästhetik rot wogender Fahnenmeere. Rot wird Farbe der Nationalsozialisten, des Sozialismus und des Kommunismus. Die Farbe durchläuft eine Karriere vom Zeichen für Wenige hin zur Symbolfarbe mit dem Potential zur Identifikation durch Massen.

³⁵ Warnke / Fleckner / Ziegler 2011, Art. *Sozialismus*.

3.2 Rotfrequenzanalysen

3.2.1 Herrscherbilder der klassischen Ikonographie

Der Modalwert zahlreicher Herrscherbilder der klassischen Ikonographie ist der Farbwert 9.³⁶ Die farbliche Range weist sehr oft auf eine maximale Virtuosität der Palette. Auffallend erreicht oder umspielt die Interquartilsspanne vieler Bilder den Wert 4. Eine Standardabweichung zwischen 5 und 6 ist häufig zu messen. Diese Kombination beschreibt ein leitendes farbformales Kontinuum. Die Sonderrolle des *Bildnisses Ludwigs XIV.* von Hyacinthe Rigaud als ikonographisches Vorbild wesentlicher Bereiche der nachfolgenden europäischen Tradition der Herrschermalerei ist auch durch eine Beschreibung seiner exponierten Standardabweichung gewinnbringend zu charakterisieren (Spektralanalyse (Abbildung 21)). Die Farben verteilen sich ebenmäßig entlang der Farbklassen, die Komposition bewirkt die geringe Standardabweichung 3,88 (Mittelwert 6,25). Acht Farbklassen von annähernd gleicher Größe, dazu Skalennachbarn, und das Vorliegen zahlreicher Farbwertkombinationen, die addiert immer wieder die Flächenwerte 50% sowie 75% und 25% approximieren, kennzeichnen die farbformale Eigenart des berühmten monumentalen Gemäldes.³⁷

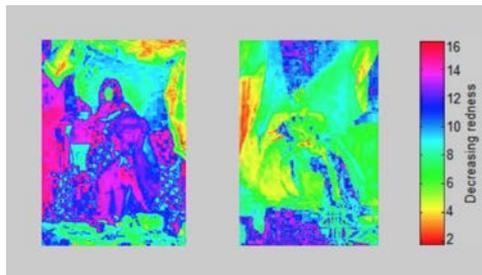


Abb. 21: Rotspektralanalysen, Hyacinthe Rigaud, *Ludwig XIV.*, 1701 (links), Friedrich Amerling, *Franz Joseph I.*, 1832 (rechts), die Herrscherbilder unterscheiden sich in der statistischen Standardabweichung der Farbwerte (Lab-Farbraum, 16 Farbklassen-Modell, Software Redcolor-Tool, HCl), im Bild von Rigaud beträgt die Standardabweichung 3,88, im Bild von Amerling beträgt sie 5,47 (Mittelwert 6,25) © Pippich 2012.

Frequenzen der extremen Skalenwerte variieren epochenspezifisch. Während das soeben charakterisierte barocke Herrscherbildnis in beiden Extremen der Skala massvoll frequentiert ist, lassen sich zahlreiche klassizistische Werke über prozentual reduziertes Rot und stärkere Dunkelfrequenzen beschreiben. Werke des Realismus hingegen sind häufig auch von intensiven Rotwerten, jedoch gering, frequentiert und weisen niedrigere Frequenzen im dunklen Skalenbereich auf.³⁸ Exemplarisch für die Herrscherbildnisse des späteren 19. Jahrhunderts stehend, lassen sich Lenbachs Porträts meist durch schwache Rotfrequenzen

³⁶ In der Tabelle etwa gehören dazu h04, h06, h15, h19, h23, h24, h30, h32. Die in diesem Abschnitt angeführten Angaben sind der Tabelle (Abbildung 19) zu entnehmen, die erwähnten Zahlen sind dort signalrot unterlegt. Die Definitionen statistischer Termini allesamt in Teil 2.1.

³⁷ Die Zahlen sind für m07 in der Tabelle (Abbildung 19) einzusehen. Die Arbeit nach festen Farb- und Formschemata in den Malerateliers Rigauds und beispielsweise auch Davids lässt das Herausstellen der Muster, die aus vielen Händen eine machen, besonders aufschlussreich werden.

³⁸ Die Angaben in diesem Abschnitt sind notwendig verkürzend, es ließen sich jeweils Gegenbeispiele anführen. Zu den beschriebenen *patterns* vgl. in der Tabelle die Messwerte von h14-h19, h21-h24, h26, h28.

und besonders starke Dunkelfrequenzen charakterisieren.³⁹ Ein auffallendes farbformales Kontinuum der traditionellen Herrscherporträts betrifft die silberne Relation zwischen Modalwertpaar und Leinwandrest.⁴⁰ Frappierend ist, wie häufig diese Beziehung zwischen den Farbwerten in den Bildern vorliegt. Dieses Muster gehört zu jenen, die über den Verlauf der Zeit für die politische Ikonographie persistieren.

3.2.2 Politikerbilder heute

Der mediale Wandel von der Malerei über die Photographie hin zur digitalen Bildproduktion und -reproduktion wird begleitet von tiefgreifenden gesellschaftlichen und kulturellen Umwälzungsprozessen. Technologische Innovationen revolutionieren mit der Einführung neuer Drucklege- und Reproduktionsverfahren auch die Darstellungskonventionen. Massenmedien entstehen. Dieser Wandel ist durch die Bildwissenschaft und Kunstgeschichte auch in Erweiterung des methodischen Spektrums der Disziplinen nachzuvollziehen. Die Entwicklung vom offiziellen Auftragsporträt, oft der Ansicht durch wenige Privilegierte vorbehalten, hin zur inoffiziellen, unbeauftragten, durchaus dem Konsens des Dargestellten zuwiderlaufenden Bildveröffentlichung ist auch farbformal zu beschreiben. Es ist zu konstatieren, dass die Rotquantitäten der Darstellungen nach dem Wandel in die Extreme geraten können und sich sehr scharfe Kontraste zwischen Skalenextremen bilden.

Zuvor nicht zu messen ist der intensive Rotton 1 als Modalwert. Nun wird dies ein potentielles Messergebnis. Ein Beispiel sei hierfür angeführt: Im Jahr 2011 bot China der Bundesrepublik Deutschland finanzielle Unterstützung in der Zeit der Finanzkrise an. Das Cover (Abbildung 22) der Frankfurter Allgemeinen Zeitung (FAZ) vom 29.06.2011 zeigt den damaligen chinesischen Staatspräsidenten Wen Jiabao als aufgereckten Zeigefinger vor chinesischer Staatsflagge, der Rotton 1 ist modal. Der Titel zum Bild lautet: »Merkel freut sich über Chinas Hilfsangebot an Europa«.



Abb. 22: Frankfurter Allgemeine Zeitung, Cover vom 29.06.2011, Wen Jiabao © dpa 2011.

³⁹Vgl. in der Tabelle (Abbildung 19) die Bilder h35-h37.

⁴⁰Vgl. in der Tabelle (Abbildung 19) die blauen Ziffern.

Auch der Rotton 2 wird nun Modalwert. Er ist beispielsweise in der Darstellung des FAZ-Covers (Abbildung 23) vom 16.05.2012 zu messen. Hollande war soeben als französischer Ministerpräsident gewählt worden. Er reiste zu seinem ersten Staatsbesuch nach Deutschland. Das Cover zeigt den Staatsakt auf dem roten Teppich vor Schloss Bellevue in Berlin. Kanzlerin Merkel dirigiert den Gast mit Weisegestus ein wenig nach links, – eine Folge dieser Direktion wird sein, dass sie vollständiger auf den roten Teppich passt. Die Bildunterschrift kommentiert: »Einen Schritt nach links. Kanzlerin Merkel weist Präsident François Hollande den rechten Weg«. Affirmativ ist der starke Rotwert 2 als Modalwert eingesetzt in einer offiziellen Darstellung (Abbildung 24) des chinesischen Staatspräsidenten Xi Jinping in der Renmin Ribao, dem täglich erscheinenden Organ der kommunistischen Staatspartei. Der Median der Darstellung erreicht den in Bildern der klassischen Herrscherikonographie vor dem medialen Wandel nicht zu messenden Rotwert 2,5.



Abb. 23: Frankfurter Allgemeine Zeitung, Cover vom 16.05.2012, Angela Merkel und François Hollande © dapd 2012, und Rotspektralanalyse (Lab-Farbraum, Software Redcolor-Tool, HCl) © Ommer/Pippich 2012.



Abb. 24: Renmin Ribao, Cover vom 22.10.2013, Xi Jinping © Xueren Li 2013.

In einer vor dem Medienwandel nicht nachzuweisenden Dimension geraten die Klassenhöhen der Farbwerte in die Extreme. Da sie zusätzlich im Bild erheblich variieren, kann die hohe Standardabweichung als Merkmal großer Teile der modernen Bildproduktion angeführt werden. Enorme Quantität erreicht die Modalfarbe beispielsweise in der FAZ-Coverdarstellung Sigmar Gabriels (Abbildung 25) vor der SPD-Parteifarbe (45,2%). Durch die Akzente auf den Skalenextremen ist die kurze IQR ein weiteres Charakteristikum moderner Bildproduktion. In der traditionellen Herrschermalerei nicht zu messen ist die IQR 2, die das Merkelbildnis (Abbildung 26) von George W. Bush aufweist, auch ist in Werken der klassischen Herrscherikonographie eine signifikante Verkürzung der farblichen Range wie in Bushs Merkelporträt nicht festzustellen.



Abb. 25: Frankfurter Allgemeine Zeitung, Cover vom 28.02.2014, Sigmar Gabriel © Reuters 2014.



Abb. 26: George W. Bush, Angela Merkel, 2014 © Bush/Wade 2014, und Spektralanalyse (Lab-Farbraum, Software Redcolor-Tool, HCI) © Ommer/Pippich 2014.

Persistiert das Muster der silbernen Relation als Abbildungs- und Wahrnehmungskonstante, so besteht die Relation nun nicht mehr zwischen Modalwertpaar und Restbild, sondern zwischen Modalwert und Restbild und erzeugt eine neue ästhetische Prägnanz. Liegen die Extreme rechnerisch demnach im Skalenextrem des Modalwerts und des Median, der Quantität der Modalfarbe, der kurzen IQR, der möglicherweise verkürzten Range und der Bildung der silbernen Relation durch nunmehr einen einzigen Farbwert, so gehören zu zusätzlichen Parametern des bildlichen Wandels der fehlende kompositorische Bezug zwischen Skalennachbarn und der Kontrast.

Anschaulich wird der farbformale Wandel im Vergleich (Abbildung 27) der prozentualen Farbwerte traditioneller Herrscherbilder (h04, h07, h17, h36) und aktueller Photographien von Politikern aus der Tagespresse (h38, h40, h47, h48). Die Diagramme visualisieren farbkompositorische Logoi. Die Darstellungsmethode folgt der Auffassung des bildlichen Logos als bestehend aus dem Verhältnis zwischen Teil und Ganzem. Prozentual zur Gesamtfläche im Uhrzeigersinn dargestellt sind die tiefroten bis hin zu den dunklen Farbmengen. Bilder, die als offizielle Repräsentation einer Staatsmacht Kontinuität und Transzendenz zur ewigen Idee verkörpern sollen, stehen inoffiziellen Bildern im öffentlichen demokratischen Meinungsbildungsprozess gegenüber, die die Kontingenz des Augenblicks einfangen, die den Effekt provozieren, die überraschen und durch das Interessante der Erscheinung fesseln sollen. Die Glieder der ersten Reihe sind komponiert, die Glieder der zweiten Reihe

repräsentieren visuelle Muster für Blickfänge bis hin zum maximalen Kontrast. Zur näheren bildwissenschaftlichen Untersuchung werden nun weitere Kategorien aufschlussreich: Die Orte des Rots im Bild, die Beziehung zwischen Titel/Sujet und farbllichem Muster und die Bildformate drängen sich als weitere Kategorien auf. Es ist der Moment, in dem das Auge den Algorithmus ergänzt.

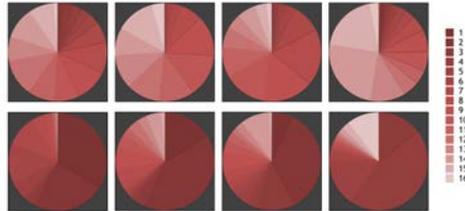


Abb. 27: Diagramme der Farbwertkompositionen von acht Herrscher- und Politikerbildern, obere Reihe (von links): **1** Tizian, Karl V., 1548; **2** Hyacinthe Rigaud, *Ludwig XIV.*, 1701; **3** Jacques Louis David, *Die Krönung Napoleons*, 1805-1807; **4** Franz Lenbach, *Wilhelm I.*, 1887, untere Reihe (von links): **5** FAZ-Cover vom 29.06.2011, Wen Jiabao; **6** FAZ-Cover vom 16.05.2012, Angela Merkel und François Hollande; **7** Renmin Ribao-Cover vom 22.10.2013, Xi Jinping; **8** FAZ-Cover vom 28.02.2014, Sigmar Gabriel. Dargestellt sind die Prozentanteile der 16 Farbwerte im Bild, beginnend im Uhrzeigersinn mit Farbwert 1 (Messtechnologie: Lab-Farbraum, 16-Farbklassen-Modell, Software Redcolor-Tool, HCI) © Pippich 2015.

3.3 Rote Orte

Die traditionelle Ikonographie zeigt den Herrscher im Zentrum. Der Herrscher ist niemals von der Bildkante beschnitten. Der Herrscherkörper ist zentral, frontal, stabil und ganz. Formal koinzidieren mit der Figur meist die Orte des Rots. Die inoffiziellen Darstellungen zeigen, dass auch diese Prinzipien einem Wandel unterliegen. Der Herrscher kann durchaus aus der Bildmitte rücken. Befindet sich der starke Rotwert dennoch im Zentrum, so kann dies etwa in Bildern des historischen Genres im Realismus als Inversion des traditionellen Prinzips auftreten, so in Menzels Gemälde *Friedrich der Große und General Fouqué*. Die rote Figur in der Bildmitte als Spitze der Pyramidalkomposition ist der Diener, nicht der Herrscher. In der Rotspektralanalyse (Abbildung 28) verschwindet der König ganz. Delaroches farbliche Inversion umfasst **Napoleon** kurz vor der Abdankung mit einem Zuviel an Rot. Der Usurpator selbst ist soldatisch gewandet, fast rotlos.⁴¹ Mit den Rotwerten stellen Lenbachs kommerziell erfolgreiche Auftragsporträts auch den Herrscher wieder ins Zentrum (Rotspektralanalysen (Abbildung 29)).

⁴¹ Die suggestive Wirkung des Gemäldes, ein Resultat der lebensgroßen, realistisch gemalten Herrscherfigur im psychologisch schwierigen Moment kurz vor der Abdankung und der in der Darstellung virtuos angewandten farbpsychologischen Gestaltungsweisen Delaroches, wurde von zeitgenössischen Betrachtern beschrieben, vgl. zur zeitgenössischen Rezeption Fleckner 2000, S. 145-167 und S. 156-159.

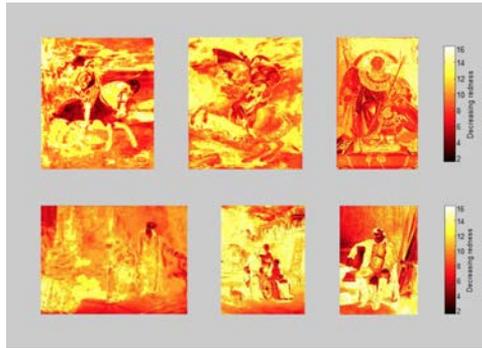


Abb. 28: Rotspektralanalysen, oben (von links): Tizian, *Karl V. nach der Schlacht bei Mülhberg*, 1548; Jacques Louis David, *Napoleon beim Überschreiten des Großen St. Bernhard*, 1802; Auguste Dominique Ingres, *Napoleon auf dem Thron*, 1806; unten (von links): Adolph Menzel, *Platz für den großen Raffael*, 1859; Adolph Menzel, *Friedrich der Große und General Fouqué*, 1852; Paul Delaroche, *Napoleon am 31. März 1814 in Fontainebleau*, 1845 (Messtechnologie: Lab-Farbraum, 16 Farbklassen-Modell, Software Redcolor-Tool, Ommer Lab, HCI) © Ommer/Pippich 2012.



Abb. 29: Rotspektralanalysen (von links): Franz Lenbach, *Franz Joseph I.*, 1873, Franz Lenbach, *Wilhelm I.*, 1887, Franz Lenbach, *Bismarck*, 1890 (Messtechnologie: Lab-Farbraum, 16 Farbklassen-Modell, Software Redcolor-Tool, Ommer Lab, HCI) © Ommer/Pippich 2012.

Bei Herrschertreffen weisen die roten Orte der Darstellung Informationen auf, weil sie zugleich die Stellen der Unröte markieren. Das Bild von Hollande und Sarkozy (h 41) mit klassischer Rotdisposition schafft es nicht auf das FAZ-Cover. Ein Merkmal offizieller chinesischer Bildpropaganda ist das kontinuierliche Ineins von Staatsflagge und -präsident. Bilder von Herrschertreffen verdeutlichen die Penetranz des Prinzips. In den Darstellungen wird die, oft wie die Chinaflagge mit Rot koinzidierende Flagge des anderen Staates entweder nicht oder nur beschnitten gezeigt. Das erste Treffen des damals neu gewählten chinesischen Präsidenten Xi Jinping mit Obama und Merkel beispielsweise fand anlässlich eines G20-Treffens in St. Petersburg im Jahr 2013 statt. Die Renmin Ribao zeigt eine Darstellung von Jinping und Obama (Abbildung 30) prominent auf dem Cover, die US-Flagge ist weitgehend beschnitten. Der Ort der Bildveröffentlichung des Treffens von Jinping und Merkel ist der nicht dem eigentlichen Cover zugehörige untere Teil der Titelseite dieser Renmin Ribao-Ausgabe. Die deutsche Fahne ist nicht im Bild.⁴²

⁴² Veröffentlichte Abbildungen der jeweiligen Szene aus der Sicht US-amerikanischer bzw. deutscher Kamerateams und Reporter waren nicht zu ermitteln.

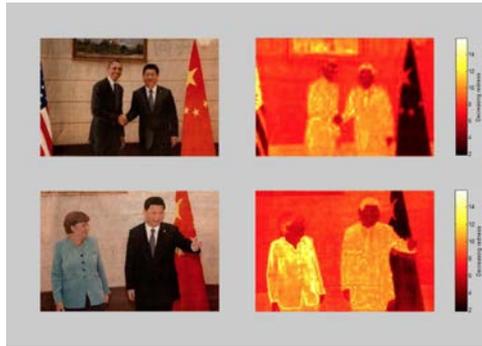


Abb. 30: Renmin Ribao, Cover vom 07.09.2013, Barack Obama und Xi Jinping © Peng Ju 2013, Angela Merkel und Xi Jinping © Jingwen Huang 2013, und Rotspektralanalysen (Lab-Farbraum, Software Redcolor-Tool, HCI) © Pippich 2013.

3.4 Rot und Negation

Das Bild des Staatsmanns entwickelt sich hin bis zur ironischen Repräsentation im Herrscherbild ohne Herrscher. Der Staatsbesuch des damaligen Ministerpräsidenten Wulff in Rom etwa wird auf dem FAZ-Cover vom 14.02.2012 (Abbildung 31) folgendermaßen dokumentiert: Ein ellenlanger, ewiger roter und dabei menschenleerer Teppich rollt sich vor dem Betrachter aus. Der zugehörige Titel lautet: »Wulff in Rom«. Kein Staatsmann ist zu sehen. Wenige Tage später wird Wulff abdanken. Es ist folgerichtig, das übermäßige und nicht gebührende Rot, ein semantisches Zuviel der Ehre, als Fortentwicklung des traditionellen Prinzips der *majestas* und *dignitas* der Herrscherfarbe zu sehen. Der rote Superlativ ist ironisch gemeint.



Abb. 31: Frankfurter Allgemeine Zeitung, Cover vom 14.02.2012, »Wulff in Rom« © dapd 2012.

Die Pixelrhetorik des Zuviel in der Form hyperbolischer Rotquantität wird in Bildern der Tagespresse oftmals von auch ironisch zu verstehenden Textkommentaren begleitet. »Merkel hat ›volles Vertrauen‹ in Gabriel« titelt das angeführte FAZ-Cover. Die Überschrift »Der chinesische Freund« und der Bildkommentar »Die helfende Hand« begleiten das modale Rot

des FAZ-Covers, das den aufgereckten Zeigefinger Jinpings abbildet. Der Titel »Wulff« findet im angeführten Bild des FAZ-Covers seine Negation, »Wulff ohne Wulff«. Für die mit Metadaten arbeitende digitale Bildforschung entsteht ein unauflösliches Kausalitätsparadoxon. Der *semantic gap* entfaltet seine volle Schärfe. Sind Kategorien wie die Orte des Rots als bildliche Parameter informatisch zu formalisieren, so bleibt die Ironie des Bildes als Negation des Titels (und umgekehrt) für automatisierte Verfahren ein schwer zu lösendes Problem.⁴³

3.5 Ästhetische Kategorien

Zu unterschiedlichen Zeiten entstandene Herrscher- und Politikerbilder spiegeln historische Darstellungstraditionen wider. Eine unterschiedliche Staatsform, Religionsauffassung und gesellschaftliche Öffentlichkeit beschreibt ihren Entstehungshorizont. Es ist aufschlussreich, die farbliche Dispersion von fünf Bildern (h07, h19, h24, h36, h39) im Zusammenhang ihres unterschiedlichen Entstehungshintergrundes zu interpretieren: Die Repräsentation absolutistischer Staatsmacht im Bildnis Ludwigs XIV. von Hyacinthe Rigaud entstand 1701 als Geschenk für den spanischen Enfant und sollte diesem als Vorbild dienen. Das klassizistische, offizieller Bildpropaganda entstammende Napoleonbildnis Jacques Louis Davids von 1812 verherrlicht den Usurpator. Das in den Unruhen des Vormärzes 1845, lange nach dem Tod Napoleons entstandene Napoleonbild Paul Delaroches etabliert eine spannungsvolle Szene und zeigt den Nachdenklichen vor einem Zuviel an Rot. Das als offizielles Auftragsporträt 1887 entstandene Bildnis Wilhelms I. von Franz Lenbach repräsentiert den Kaiser in Uniform zentral im Bild, weite, dunkle Farbmassen umfassen die Figur. Nominell »Wulff« darstellend, zeigt das FAZ-Cover 2012 faktisch einen leeren roten Teppich. Diese fünf Bilder unterscheiden sich charakteristisch in der Rotwertkomposition. Ein Diagramm (Abbildung 32) bildet die jeweils eigentümliche Ästhetik der Farbwertmengen ab.

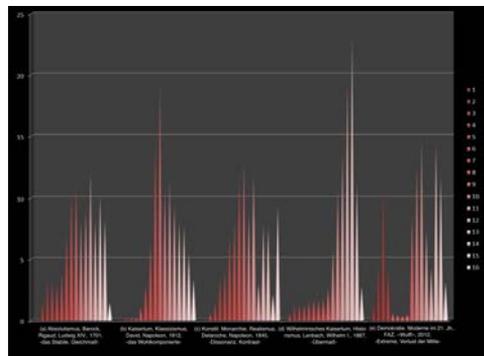


Abb. 32: Diagramm der Farbkomponenten in fünf Herrscher- und Politikerbildern, die horizontale Achse zeigt die Farbwerte für jedes Bild beginnend links mit Farbwert 1 bis Farbwert 16, die vertikale Achse zeigt die Prozentzahl des Farbwerts im Bild (von links): a) Hyacinthe Rigaud, *Ludwig XIV.*, 1701, b) Jacques Louis David, *Kaiser Napoleon in seinem Studierzimmer*, 1812, c) Paul Delaroche, *Napoleon am 31. März 1814 in Fontainebleau*, 1845, d) Franz Lenbach, *Wilhelm I.*, 1887, e) FAZ-Cover vom 14.02.2012, »Wulff in Rom« (Messtechnologie: Lab-Farbraum, 16-Farbklassen-Modell, Software Redcolor-Tool, HCl) © Pippich 2014.

⁴³ Die linguistische Sentimentanalyse jedenfalls kann Ironie nicht erfassen, vgl. Kruschke 2013, S. N5. Zum Superlativ als Ironiesignal Weinrich 2006, S. 64 und Köhler 2004, S. 768–769.

Das Herrscherbild des Barock (a), des Klassizismus (b), des Realismus (c), des Symbolismus oder Historismus (d) und der Moderne im 21. Jahrhundert (e) kennzeichnet eine je unterschiedliche farbliche Dispersion, der eine eigene Ästhetik zu eigen ist. Ästhetische Kategorien der Rotwertkompositionen sind:

- a - das Stabile, Gleichmaß
- b - das Harmonische, Wohlkomponierte
- c - Dissonanz, Kontrast
- d - Überfluss, Übermaß
- e - die Extreme, der Verlust der Mitte

Die Unterschiede dieser ästhetischen Kategorien resultieren aus zutiefst verschiedenen historischen Seh- und Darstellungstraditionen und bildpolitischen, wahrnehmungspsychologischen bis hin zu medientechnischen, produktionsbedingten Faktoren. Sie beschreiben die historische Entwicklung der Herrscherdarstellung aus der Perspektive der Farbkomposition. Nicht alles ist zu allen Zeiten möglich.

4. Fazit

Das digitalisierte/digitale Bild ist der Adressat computergestützter Analysen. Zuvor im Forschungsprozess nicht zu realisierende Fragestellungen lassen sich an das Bild herantragen. Auch Farbe, paradigmatisch für das begrifflich nicht eindeutig zu Fassende, subjektiv Erfahrene stehend, wird durch intelligente informatische Lösungen Gegenstand wissenschaftlicher Forschung. Jedes einzelne Bildelement kann adressiert und in komparative Studien eingebunden werden. Die Möglichkeit, alle Pixel im Bild, viele Bilder und Bilder ganzer kultureller Bereiche und Epochen zu erfassen, zu analysieren und auszuwerten veranlasst zum Staunen. Was lässt sich noch alles herausfinden?⁴⁴ Bestehende Paradigmen lassen sich durch Farbanalysen empirisch prüfen, präzisieren und reformulieren, das Eingehen in Skizzen und frühe Versionen der Bilder lässt die bislang stummen Logoi des Werkprozesses deutlich und markant zu Wort kommen, erlaubt neue Reihen nun entlang farblicher Kriterien der Werkentstehung zu formulieren. Das Lichte, das Dunkle und das Seltene beispielsweise wird vom Computer objektiv erfasst, während der Mensch meist identifizierend verfährt und die Augen ins Kleinteilige schweifen lässt zur inhaltlichen Erfassung des Gesehenen.

Farbformalistische Studien lassen zunächst unberücksichtigt gebliebene Kategorien in neuer Prägnanz aufleben. Dies zeigte die exemplarische Gegenüberstellung einer mithilfe des Redcolor-Tools bearbeiteten stilometrischen und einer historischen Forschungsfrage. Beide Studien konnten an geeigneten Stellen an weitere Kategorien anknüpfen. Diese Kategorien, etwa die Bildformate, die Linearkomposition, die Farborte oder den möglichen Sinn in die Untersuchung integrieren zu können, gehört zur heuristischen Ausrichtung des formalistischen

⁴⁴ Im Rahmen des *big data*-Paradigmas hat Hubertus Kohle die Chancen quantitativer Farbanalysen für die Kunstgeschichte benannt, Kohle 2008, S. 143–144, Kohle 2013, S. 51, und das rein zahlenmäßige Auftreten der Farbe Blau im kunsthistorischen Symbolismus beispielhaft als Untersuchungsfeld der computergestützten Kunstgeschichte formuliert.

Ansatzes. Neue Kategorien, etwa rezeptionsbedingte wie Skandalbilder, thematische wie Altarbilder, formale wie Monumentalbilder, materialorientierte wie »alle Pastelle«, alle Auftragswerke, Werke, die sich nicht verkaufen ließen, können als Faktoren in die Farbstudien ergänzend einfließen und sie umgekehrt als Beginn einer Forschungsfrage erst veranlassen. Hürden für die Homogenität der Daten und die Konsistenz der Formalisierungsprozesse für die digitale Bildanalyse sind modell-, skalierungs- und reproduktionsbedingt. Der gewählte Lab-Farbraum eignet sich für die Analyse der vorgestellten Bilder, für das genuin digitale Bild bietet CIECAM zusätzliche Anhaltspunkte.⁴⁵ Rot rechnen durch ein digitales Tool liefert ein Beispiel für den methodischen Einsatz informatischer Eigenschaftsmetrik für die Geisteswissenschaften. Das durch die Informatik ermöglichte metrische Skalenniveau und die Numerik der Mathematik bieten auch im Rahmen nicht diskret, sondern kontinuierlich vorliegender Eigenschaften Möglichkeiten für systematische und historische Forschungsansätze.

Welche weiteren messbaren Kategorien sind für die informatische direkte Bildadressierung denkbar? Die »Alleviereckenandersfarbigkeit« einer von Constable im Jahr 1822 angefertigten Wolkenstudie im Vergleich (Abbildung 33) zur »Alleviereckengleichfarbigkeit« einer 1823 von Blechen ausgeführten Wolkenstudie zeigt beispielsweise zugleich die formalen Unterschiede der analytischen und synthetisierenden Methode auf, Verfahrensweisen, die wiederum Methoden der Philosophie der Zeit prägnant beschreiben. Asymmetrie wird sich als Prinzip in Gartenanlagen, im Aufbau von Gedichten, der Struktur musikalischer, malerischer und architektonischer Kompositionen des Barock vermutlich auch rechnerisch nachweisen lassen. Die Software kann für die Forschung das Fernste miteinander verbinden, formale Prinzipien in Gebilden in Analogie zueinander berechnen. Für die Geisteswissenschaften liegt in der systematischen Eigenschaftsmetrik durch Software das Potential zur Erforschung unterschiedlicher Artefakte unter einem automatisierten Objektiv. Welche weiteren neuen Technologien zur Bestimmung der Kunst einzusetzen sein werden, wird die Zukunft zeigen.



Abb. 33: Carl Blechen, *Wolken bei Tage mit blauem Himmel*, um 1823, Öl auf Papier, 10,8 x 19,8 cm, Staatliche Museen zu Berlin, Kupferstichkabinett (links) Quelle: *Goethe und die Kunst*, Kat. Ausst. Schirn Kunsthalle Frankfurt, Kunstsammlungen zu Weimar. Hg. von Sabine Schulze. Stuttgart 1994, S. 557, John Constable, *Clouds 5 September 1822*, 1822, Öl auf Papier, 37 x 49 cm, National Gallery of Victoria, Melbourne (rechts) Quelle: *Constable, Impressions of Land, Sea and Sky*. Kat. Ausst. National Gallery of Australia. Hg. von Anne Gray / John Gage. London 2006, S. 177.

⁴⁵ Auch iCAM ist für die Analyse des genuin digitalen Bildes zu nennen. Zur Einführung in die Problematik geeignet Fairchild 2013.

Bibliographische Angaben

- Adolph Menzel. Radikal real: Kat. Ausst. Kunsthalle der Hypo-Kulturstiftung München und Kupferstichkabinett der staatlichen Museen zu Berlin. Hg. von Bernhard Maaz. München 2008. [\[Nachweis im GBV\]](#)
- Adolph Menzel 1815–1905. Das Labyrinth der Wirklichkeit. Kat. Ausst. Nationalgalerie Berlin. Hg. von Claude Keisch / Marie Ursula Riemann-Reyher. Köln 1996, S. 457–468. [\[Nachweis im GBV\]](#)
- Werner Busch: Adolph Menzel. Leben und Werk. München 2004. [\[Nachweis im GBV\]](#)
- Werner Busch: Menzels Landschaften. Bildordnung als Antwort auf die Erfahrung vom Wirklichkeitszerfall. In: Adolph Menzel 1815–1905. Das Labyrinth der Wirklichkeit. Kat. Ausst. Nationalgalerie Berlin. Hg. von Claude Keisch / Marie Ursula Riemann-Reyher. Köln 1996, S. 457–468. [\[Nachweis im GBV\]](#)
- Constable, Impressions of Land, Sea and Sky. Kat. Ausst. National Gallery of Australia. Hg. von Anne Gray / John Gage. London 2006, S. 177. [\[Nachweis im GBV\]](#)
- Erkenntniswert Farbe. Hg. von Margrit Vogt / André Karliczek. Berlin, Jena 2013. [\[Nachweis im GBV\]](#)
- Mark Fairchild: Color Appearance Models. 3. Aufl. 2013. [\[Nachweis im GBV\]](#)
- Farben in Kulturgeschichte und Naturwissenschaft. Hg. von Gudrun Wolfschmidt. Hamburg 2011. [\[Nachweis im GBV\]](#)
- Uwe Fleckner: Napoleon am Scheidewege. In: Kunst der Nationen – Jenseits der Grenzen. Französische und deutsche Kunst vom Ancien Regime bis zur Gegenwart. Hg. von Uwe Fleckner / Martin Schieder / Michael F. Zimmermann. Köln 2000, S. 145–167. [\[Nachweis im GBV\]](#)
- Goethe und die Kunst, Kat. Ausst. Schirn Kunsthalle Frankfurt, Kunstsammlungen zu Weimar. Hg. von Sabine Schulze. Stuttgart 1994, S. 557 [\[Nachweis im GBV\]](#)
- Gottfried Gabriel: Logische Präzision und ästhetische Prägnanz. In: Literaturwissenschaftliches Jahrbuch N.F. Hg. von Volker Kapp / Kurt Müller / Klaus Ridder / Ruprecht Wimmer / Jutta Zimmermann, Bd. 51. Berlin 2010, S. 375–390. [\[Nachweis im GBV\]](#)
- Werner Hofmann: Menzels verschlüsseltes Manifest. In: Menzel – der Beobachter. Kat. Ausst. Hamburger Kunsthalle. Hg. von Werner Hofmann. München 1982, S. 31–40. [\[Nachweis im GBV\]](#)
- Max Imdahl: Farbe. Kunsttheoretische Reflexionen in Frankreich. 2. Aufl. München 1988. [\[Nachweis im GBV\]](#)
- Wilhelm Köhler: Perspektivität und Sprache. Zur Struktur von Objektivierungsformen in Bildern, im Denken und in der Sprache. Berlin 2004. [\[Nachweis im GBV\]](#)
- Hubertus Kohle: Adolph Menzels Friedrich-Bilder. Theorie und Praxis der Geschichtsmalerei im Berlin der 1850er Jahre. Hildesheim / New York 2001. [\[Nachweis im GBV\]](#)
- Hubertus Kohle: Ordnung und Umbruch in der Wissenschaft der Kunstgeschichte: Der Einfluss des Digitalen. In: Reibungspunkte. Ordnung und Umbruch in Architektur und Kunst. Hg. von Hanns Hubach / Barbara von Orelli-Messerli / Tadej Tassinini. Petersberg 2008, S. 141–144. [\[Nachweis im GBV\]](#)
- Hubertus Kohle: Digitale Bildwissenschaft. Glückstadt 2013. [\[Nachweis im GBV\]](#)
- Wolfgang Krischke: *Wer nicht hören kann, muss fühlen. Bazinga! Der neueste sozialtechnologische Schrei, aber noch nicht ironiefest: Die Sentimentanalyse.* In: FAZ vom 10.04.2013, S. N5. [\[Nachweis im GBV\]](#)
- Lev Manovich: How to Compare One Million Images? In: Understanding Digital Humanities. Hg. von David Berry. New York 2012, S. 249–278. [\[Nachweis im GBV\]](#)
- Julius Meier-Graefe: Der junge Menzel. Ein Problem der Kunstökonomie Deutschlands. München 1906. [\[Nachweis im GBV\]](#)
- Menzel – der Beobachter. Kat. Ausst. Hamburger Kunsthalle. Hg. von Werner Hofmann. München 1982, S. 31–40. [\[Nachweis im GBV\]](#)
- Handbuch der politischen Ikonographie. Hg. von Martin Warnke / Uwe Fleckner / Hendrik Ziegler. 2. Aufl. München 2011. [\[Nachweis im GBV\]](#)
- Verfeinertes Sehen. Optik und Farbe im 18. und frühen 19. Jahrhundert. Hg. von Werner Busch. München 2008. [\[Nachweis im GBV\]](#)
- Harald Weinrich: Linguistik der Lüge. 7. Aufl. München 2006. [\[Nachweis im GBV\]](#)
- Ludwig Wittgenstein: Bemerkungen über die Farben. Werkausgabe Bd. 8. Frankfurt a.M. 1984. [\[Nachweis im GBV\]](#)
- D´Arcy Wentworth Thompson: On Growth and Form. New York 1992 [zuerst 1917]. [\[Nachweis im GBV\]](#)

Abbildungslegenden und -nachweise

Abb. 1: Tabelle der Metadaten, Messwerte und Angaben zu statistischen Frequenzen der Farbwerte (Lab-Farbraum, 16 Farbklassen-Modell, Software Redcolor-Tool, HCl) der Korpusbilder © Pippich 2014. Link auf Datei: Waltraud von Pippich: Frequenzen und statistische Dispersion der Farben in 50 Bildern von Adolph Menzel. 2014. Open data LMU Link: <http://dx.doi.org/10.5282/ubm/data.79>

Abb. 2: Stilometrische Bildanalysen, Diagramm der Farbwertkompositionen von 100 Bildern, 50 Bilder von Adolph Menzel (links), 50 Bilder heterogener Autorschaft – von Giotto bis Yves Klein (rechts). Jede Säule im Diagramm repräsentiert ein Bild in seiner farblichen Komposition (Messtechnologie: Lab-Farbraum, 16 Farbklassen-Modell, Software Redcolor-Tool, HCl) © Pippich 2013.

Abb. 3: Oben (von links): Adolph Menzel, *Flötenkonzert Friedrichs des Großen in Sanssouci*, 1852, Öl auf Leinwand, 142 x 205 cm, Berlin, Nationalgalerie, Quelle: Adolph Menzel. Radikal real. Kat. Ausst. Kunsthalle der Hypo-Kulturstiftung München und Kupferstichkabinett der staatlichen Museen zu Berlin. Hg. von Bernhard Maaz. München 2008, S. 17, Adolph Menzel, *Berlin-Potsdamer Bahn*, 1847, Öl auf Leinwand, 42 x 52 cm, Berlin, Nationalgalerie, Quelle: Adolph Menzel 1815–1905. Das Labyrinth der Wirklichkeit. Kat. Ausst. Nationalgalerie Berlin. Hg. von Claude Keisch und Marie-Ursula Riemann-Reyher. Köln 1996, S. 116, unten: Spektralanalysen, Lab-Farbraum, Software Redcolor-Tool (HCl) – die Korpusbilder mit maximaler (1–16) und minimaler (4–16) Farbränge © Pippich 2014.

Abb. 4: Oben (von links): Adolph Menzel, *Frühmesse in einer Salzburger Kirche*, 1855, Öl auf Leinwand, 58,4 x 68,3 cm, Wien, Österreichische Galerie im Belvedere, Quelle: Adolph Menzel. Radikal real. Kat. Ausst. Kunsthalle der Hypo-Kulturstiftung München und Kupferstichkabinett der staatlichen Museen zu Berlin. Hg. von Bernhard Maaz. München 2008, S. 167, Adolph Menzel, *Biergarten*, 1883, Pinsel, Tusche, Deckfarben, 13,5 x 17,9 cm, Schweinfurt, Museum Georg Schäfer, Quelle: Adolph Menzel. Radikal real. Kat. Ausst. Kunsthalle der Hypo-Kulturstiftung München und Kupferstichkabinett der staatlichen Museen zu Berlin. Hg. von Bernhard Maaz. München 2008, S. 18, unten: Spektralanalysen, Lab-Farbraum, Software Redcolor-Tool (HCl) – die Korpusbilder mit der höchsten (links: 10,79) und geringsten (rechts: 3,37) statistischen Standardabweichung der Farbwerte (Mittelwert: 6,25) © Pippich 2014.

Abb. 5: Oben (von links): Adolph Menzel, *Wohnzimmer mit der Schwester des Künstlers*, 1847, Öl auf Papier, 46,1 x 31,7 cm, München, Neue Pinakothek, Quelle: Adolph Menzel 1815–1905. Das Labyrinth der Wirklichkeit. Kat. Ausst. Nationalgalerie Berlin. Hg. von Claude Keisch und Marie-Ursula Riemann-Reyher. Köln 1996, S. 119, Adolph Menzel, *Mondschein über der Friedrichsgracht im alten Berlin*, um 1855, Öl auf Leinwand, 39,5 x 33 cm, Berlin, Nationalgalerie, Quelle: Adolph Menzel 1815–1905. Das Labyrinth der Wirklichkeit. Kat. Ausst. Nationalgalerie Berlin. Hg. von Claude Keisch und Marie-Ursula Riemann-Reyher. Köln 1996, S. 177, (unten) Spektralanalysen, Lab-Farbraum, Software Redcolor-Tool (HCl) – die Korpusbilder mit der niedrigsten (links: 5,5) und höchsten (rechts: 14,3) Skalenposition des Median © Pippich 2014.

Abb. 6: Oben (von links): Adolph Menzel, *Emilie Menzel schlafend*, um 1848, Öl auf Papier, 46,8 x 80 cm, Hamburg, Hamburger Kunsthalle, Quelle: Adolph Menzel 1815–1905. Das Labyrinth der Wirklichkeit. Kat. Ausst. Nationalgalerie Berlin. Hg. von Claude Keisch und Marie-Ursula Riemann-Reyher. Köln 1996, S. 133, Adolph Menzel, *Kronprinz Friedrich besucht Pesne auf dem Malgerüst in Rheinsberg*, 1861, Gouache auf Papier, 24 x 32 cm, Berlin, Nationalgalerie, Quelle: Adolph Menzel 1815–1905. Das Labyrinth der Wirklichkeit. Kat. Ausst. Nationalgalerie Berlin. Hg. von Claude Keisch und Marie-Ursula Riemann-Reyher. Köln 1996, S. 203, unten: Spektralanalysen, Lab-Farbraum, Software Redcolor-Tool (HCl) – Korpusbilder mit geringer (2) und weiter (5) Interquartilsspanne, die Lage der Interquartilsspanne differiert: 4,5–6,5 (links), 8,5–13,5 (rechts), vgl. die Anzeige der Linie im Farbbalken © Pippich 2014.

Abb. 7: Oben: Adolph Menzel, *Ballszene*, 1867, Pinsel, Deckfarbe, 26,5 x 30 cm, Schweinfurt, Museum Georg Schäfer, Quelle: Adolph Menzel. Radikal real. Kat. Ausst. Kunsthalle der Hypo-Kulturstiftung München und Kupferstichkabinett der staatlichen Museen zu Berlin. Hg. von Bernhard Maaz. München 2008, S. 199, unten: Spektralanalyse, Lab-Farbraum, Software Redcolor-Tool, HCl © Pippich 2014.

Abb. 8: Adolph Menzel, *Balkonzimmer*, 1845, Öl auf Pappe, 58 x 47 cm, Berlin, Nationalgalerie, Quelle: Adolph Menzel 1815–1905. Das Labyrinth der Wirklichkeit. Kat. Ausst. Nationalgalerie Berlin. Hg. von Claude Keisch und Marie-Ursula Riemann-Reyher. Köln 1996, S. 91.

Abb. 9: Spektralanalysen, Adolph Menzel, *Balkonzimmer*, Software Redcolor-Tool, HCl © Pippich 2012.

Abb. 10: Diagramm der Farbwerte (Lab-Farbraum, 16-Farbklassen-Modell, Software Redcolor-Tool, HCl), Darstellung der Fibonacci-Relationen: : 100% / gelbes Segment / gelbes Segment / blaues Segment, Adolph Menzel, *Balkonzimmer* © Pippich 2014. Weitere goldene Relationen mit Link auf Datei: Waltraud von Pippich: Goldene Relationen in Farbkompositionen von Adolph Menzel. Das Gemälde *Balkonzimmer* Open data LMU Link: <http://dx.doi.org/10.5282/ubm/data.80>

Abb. 11: Adolph Menzel, *Piazza d'Erbe in Verona*, 1882–84, Öl auf Leinwand, 74 x 127 cm, Staatliche Kunstsammlungen Dresden, Galerie Neue Meister, Quelle: Adolph Menzel 1815–1905. Das Labyrinth der Wirklichkeit. Kat. Ausst. Nationalgalerie Berlin. Hg. von Claude Keisch und Marie-Ursula Riemann-Reyher. Köln 1996, S. 318.

Abb. 12: Spektralanalysen, Adolph Menzel, *Piazza d'Erbe in Verona*, Software Redcolor-Tool, HCl © Pippich 2012.

Abb. 13: Diagramm der Farbwerte (Lab-Farbraum, 16 Farbklassen-Modell, Software Redcolor-Tool, HCl), Darstellung der Fibonacci-Relationen: (jeweils) gelbes Segment / blaues Segment / blaues Segment / graues Segment, graues Segment / rotes Segment, rot-grünes Segment / blaues Segment, Adolph Menzel, *Piazza d'Erbe in Verona* © Pippich 2014. Weitere goldene Relationen Link auf Datei: Waltraud von Pippich: Goldene Relationen in Farbkompositionen von Adolph Menzel. Das Gemälde *Piazza d'Erbe in Verona*. 2012. Open data LMU Link: <http://dx.doi.org/10.5282/ubm/data.83>

Abb. 14: Adolph Menzel, *Théâtre du Gymnase*, 1856, Öl auf Leinwand, 46 x 62 cm, Berlin, Nationalgalerie. Quelle: Adolph Menzel 1815–1905. Das Labyrinth der Wirklichkeit. Kat. Ausst. Nationalgalerie Berlin. Hg. von Claude Keisch und Marie-Ursula Riemann-Reyher. Köln 1996, S. 27.

Abb. 15: Spektralanalysen, Adolph Menzel, *Théâtre du Gymnase*, Software Redcolor-Tool, HCl © Pippich 2012.

Abb. 16: Diagramm der Farbwerte (Lab-Farbraum, 16 Farbklassen-Modell, Software Redcolor-Tool, HCl), Darstellung der Fibonacci-Relationen, Adolph Menzel, *Théâtre du Gymnase* © Pippich 2014. Link auf Datei: Waltraud von Pippich: Goldene Relationen in Farbkompositionen von Adolph Menzel. Das Gemälde *Théâtre du Gymnase*. 2014. Open data LMU Link: <http://dx.doi.org/10.5282/ubm/data.84>

Abb. 17: Kumulative Frequenzhistogramme: Farbverteilung in 50 Bildern von Adolph Menzel (links), Graphenbündel: Farbverteilung in 16 Bildern (Mitte), Graphenbündel: Farbverteilung in 14 Bildern (rechts). Die horizontale Achse der Histogramme zeigt die 16 Farbwerte, die vertikale Achse zeigt die Prozentzahl (Messtechnologie: Lab-Farbraum, 16 Farbklassen-Modell, Software Redcolor-Tool, HCl) © Pippich 2014.

Abb. 18: Oben (von links) Adolph Menzel, *Auf der Fahrt durch schöne Natur*, 1892, Deckfarben auf Papier, 27,7 x 37,2 cm, Privatsammlung, Quelle: Menzel – der Beobachter. Kat. Ausst. Kunsthalle Hamburg. Hg. von Werner Hofmann. München 1982, S. 280, Adolph Menzel, *Blick in einen kleinen Hof*, 1867, Wasser- und Deckfarben auf Papier, 28 x 22 mm, Berlin, Nationalgalerie, Quelle: Menzel – der Beobachter. Kat. Ausst. Kunsthalle Hamburg. Hg. von Werner Hofmann. München 1982, S. 171, unten: Spektralanalysen (Lab-Farbraum, Software Redcolor-Tool, HCl) – die Bilder im Korpus mit maximalen Messwerten in beiden Skalenextremen © Pippich 2014.

Abb. 19: Tabelle der Metadaten, Messwerte und Angaben zu statistischen Frequenzen der Farbwerte (Lab-Farbraum, 16 Farbklassen-Modell, Software Redcolor-Tool, HCl) in Herrscher- und Politikerbildern © Pippich 2014. Link auf Datei: Waltraud von Pippich: Rotfrequenzen und statistische Farbdispersion in Herrscher- und Politikerbildern (1360-2014). 2014. Open data LMU Link: <http://dx.doi.org/10.5282/ubm/data.81>

Abb. 20: Tabelle der Abbildungen und Rotspektralanalysen (Lab-Farbraum, 16 Farbklassen-Modell, Software Redcolor-Tool, HCl) und Metadaten von Herrscher- und Politikerbildern © Pippich 2015. Link auf Datei: Waltraud von Pippich: Rotspektralanalysen von Herrscher- und Politikerbildern (1360-2014). 2014. Open data LMU Link: <http://dx.doi.org/10.5282/ubm/data.82>

Abb. 21: Rotspektralanalysen, Hyacinthe Rigaud, *Ludwig XIV.*, 1701 (links), Friedrich Amerling, *Franz Joseph I.*, 1832 (rechts), die Herrscherbilder unterscheiden sich in der statistischen Standardabweichung der Farbwerte (Lab-Farbraum, 16 Farbklassen-Modell, Software Redcolor-Tool, HCl), im Bild von Rigaud beträgt die Standardabweichung 3,88, im Bild von Amerling beträgt sie 5,47 (Mittelwert 6,25) © Pippich 2012.

Abb. 22: Frankfurter Allgemeine Zeitung, Cover vom 29.06.2011, Wen Jiabao © dpa 2011.

Abb. 23: Frankfurter Allgemeine Zeitung, Cover vom 16.05.2012, Angela Merkel und François Hollande © dapd 2012, und Rotspektralanalyse (Lab-Farbraum, Software Redcolor-Tool, HCl) © Ommer/Pippich 2012.

Abb. 24: Renmin Ribao, Cover vom 22.10.2013, Xi Jinping © Xueren Li 2013.

Abb. 25: Frankfurter Allgemeine Zeitung, Cover vom 28.02.2014, Sigmar Gabriel © Reuters 2014.

Abb. 26: George W. Bush, *Angela Merkel*, 2014 © Bush/Wade 2014, und Spektralanalyse (Lab-Farbraum, Software Redcolor-Tool, HCl) © Ommer/Pippich 2014.

Abb. 27: Diagramme der Farbwertkompositionen von acht Herrscher- und Politikerbildern, obere Reihe (von links): 1 Tizian, *Karl V.*, 1548; 2 Hyacinthe Rigaud, *Ludwig XIV.*, 1701; 3 Jacques Louis David, *Die Krönung Napoleons*, 1805-1807; 4 Franz Lenbach, *Wilhelm I.*, 1887, untere Reihe (von links): 5 FAZ-Cover vom 29.06.2011, Wen Jiabao; 6 FAZ-Cover vom 16.05.2012, Angela Merkel und François Hollande; 7 Renmin Ribao-Cover vom 22.10.2013, Xi Jinping; 8 FAZ-Cover vom 28.02.2014, Sigmar Gabriel. Dargestellt sind die Prozentanteile der 16 Farbwerte im Bild, beginnend im Uhrzeigersinn mit Farbwert 1 (Messtechnologie: Lab-Farbraum, 16-Farbklassen-Modell, Software Redcolor-Tool, HCl) © Pippich 2015.

Abb. 28: Rotspektralanalysen, oben (von links): Tizian, *Karl V. nach der Schlacht bei Mühlberg*, 1548; Jacques Louis David, *Napoleon beim Überschreiten des Großen St. Bernhard*, 1802; Auguste Dominique Ingres, *Napoleon auf dem Thron*, 1806; unten (von links): Adolph Menzel, *Platz für den großen Raffael*, 1859; Adolph Menzel, *Friedrich der Große und General Fouqué*, 1852; Paul Delaroche, *Napoleon am 31. März 1814 in Fontainebleau*, 1845 (Messtechnologie: Lab-Farbraum, 16 Farbklassen-Modell, Software Redcolor-Tool, Ommer Lab, HCl) © Ommer/Pippich 2012.

Abb. 29: Rotspektralanalysen (von links): Franz Lenbach, *Franz Joseph I.*, 1873, Franz Lenbach, *Wilhelm I.*, 1887, Franz Lenbach, *Bismarck*, 1890 (Messtechnologie: Lab-Farbraum, 16 Farbklassen-Modell, Software Redcolor-Tool, Ommer Lab, HCl) © Ommer/Pippich 2012.

Abb. 30: Renmin Ribao, Cover vom 07.09.2013, Barack Obama und Xi Jinping © Peng Ju 2013, Angela Merkel und Xi Jinping © Jingwen Huang 2013, und Rotspektralanalysen (Lab-Farbraum, Software Redcolor-Tool, HCl) © Pippich 2013.

Abb. 31: Frankfurter Allgemeine Zeitung, Cover vom 14.02.2012, »Wulff in Rom« © dapd 2012.

Abb. 32: Diagramm der Farbkomponenten in fünf Herrscher- und Politikerbildern, die horizontale Achse zeigt die Farbwerte für jedes Bild beginnend links mit Farbwert 1 bis Farbwert 16, die vertikale Achse zeigt die Prozentzahl des Farbwerts im Bild (von links): a Hyacinthe Rigaud, *Ludwig XIV.*, 1701, b Jacques Louis David, *Kaiser Napoleon in seinem Studierzimmer*, 1812, c Paul Delaroche, *Napoleon am 31. März 1814 in Fontainebleau*, 1845, d Franz Lenbach, *Wilhelm I.*, 1887, e FAZ-Cover vom 14.02.2012, »Wulff in Rom« (Messtechnologie: Lab-Farbraum, 16-Farbklassen-Modell, Software Redcolor-Tool, HCl) © Pippich 2014.

Abb. 33: Carl Blechen, *Wolken bei Tage mit blauem Himmel*, um 1823, Öl auf Papier, 10,8 x 19,8 cm, Staatliche Museen zu Berlin, Kupferstichkabinett (links) Quelle: Goethe und die Kunst. Kat. Ausst. Schirn Kunsthalle Frankfurt, Kunstsammlungen zu Weimar. Hg. von Sabine Schulze. Stuttgart 1994, S. 557, John Constable, *Clouds 5 September 1822*, 1822, Öl auf Papier, 37 x 49 cm, National Gallery of Victoria, Melbourne (rechts) Quelle: Constable, *Impressions of Land, Sea and Sky*. Kat. Ausst. National Gallery of Australia. Hg. von Anne Gray / John Gage. London 2006, S. 177.

Glossar

Ausreißer	Als Ausreißer wird im Rahmen einer statistischen Untersuchung ein seltenes, sich vom Durchschnitt der Messdaten deutlich abhebendes Messergebnis bezeichnet. Signifikant für einige statistisch deskriptive Angaben, wie die Werte mittlerer und zentraler Tendenz, ist, dass sie robust gegenüber statistischen Ausreißern sind. Sind diese Angaben etwa für einige Methoden des data mining hilfreich, da die Ausreißer auch auf Fehler in der Datenerhebung und -analyse zurückzuführen sein können, so entfaltet beispielsweise die Stilometrie ein Interesse der Erfassung auch seltener Messergebnisse.
Commission Internationale de l'Éclairage	Die Commission Internationale de l'Éclairage (kurz CIE, engl. International Commission on Illumination) ist eine gemeinnützige Körperschaft der länderübergreifenden Zusammenarbeit und des Austausches in technologischen und wissenschaftlichen Fragen auf den Fachgebieten von Licht, Beleuchtung und Farbe. Die CIE ist von der International Organization for Standardization (ISO) als offizielle Instanz zur Festsetzung internationaler Standards auf dem Gebiet Licht, Beleuchtung und Farbe bestätigt worden.
dpi	Die Abkürzung dpi steht für (engl.) dots per inch, sie bezeichnet eine Maßeinheit zur Erfassung der Auflösung von Bildern im digitalen Medium. Die Anzahl der Punkte pro Zoll (1 Zoll entspricht 2,54 cm) gibt als Punktdichte die Auflösung des Bildes wieder.
Farbstich	Eine Abweichung der Farbigkeit einer Bildreproduktion von der Farbigkeit der originalen, reproduzierten Vorlage in der Richtung einer bestimmten Farbnuance wird als Farbstich bezeichnet. Die Abweichungen resultieren aus reproduktionstechnischen Faktoren. Zu diesen gehören die Lichtverhältnisse zum Zeitpunkt der photographischen Reproduktion, die Belichtungszeiten und weitere Einstellungen der aufnehmenden Kamera sowie (material-)technische Bedingungen des verwendeten Bildträgers und des Druckverfahrens. Als Digitalisat weist eine Bildreproduktion zusätzliche Variationen der farblichen Erscheinung durch eine fehlende oder inkorrekte Kalibrierung und die individuelle Farbeinstellung und -auflösung des Monitorgeräts auf.

Fibonacci-Folge	Die Zahlen der Fibonacci-Folge (1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, ...) repräsentieren ein rekurrentes mathematisches Prinzip: Jede Zahl der Fibonacci-Folge ist die Summe der beiden in der Zahlenfolge vorangehenden Zahlen. Zudem approximiert jede Fibonacci-Zahl im Verhältnis zur in der Folge vorangehenden Zahl die goldene Zahl Phi (1,61...).
Goldener Schnitt	Als goldener Schnitt wird das bestimmte Größenverhältnis zweier Glieder bezeichnet, bei dem sich das Gesamt zum größeren Teil ebenso verhält wie der größere Teil zum kleineren Teil. Die das Größenverhältnis des goldenen Schnittes repräsentierende Zahl ist die irrationale Zahl Phi (1,61...).
Interquartilsspanne	Die Interquartilsspanne (auch Interquartilsabstand, kurz IQR für engl. interquartile range) bezeichnet in der Statistik den Abstand zwischen dem ersten Quartil (Q25) und dem dritten Quartil (Q75) einer Häufigkeitsverteilung. Sie berechnet sich aus der Differenz der Werte des dritten und ersten Quartils ($Q75 - Q25 = IQR$). Die Interquartilsspanne beschreibt die Verteilung der mittleren 50% der Messergebnisse. Die IQR ist eine Angabe der sogenannten mittleren Tendenz der gesamten Messdaten, d.h. sie gibt als Streumaß die statistische Mitte der Werte wieder. Deskriptiv erfasst die Interquartilsspanne somit genau eine Hälfte der Messergebnisse. Nicht erfasst wird von der Angabe die Verteilung jener 50% der Messdaten, die sich in den extremeren Lagen der Skala befinden.
Lab-Farbraum	Der Lab-Farbraum (auch CIE $L^*a^*b^*$ oder CIELAB) erfasst Farben in der Weise, wie sie ein normaler Betrachter zu alltäglichen Sehbedingungen wahrnehmen würde. Der Lab-Farbraum ist durch die Wahrnehmungsbezogenheit und die Unabhängigkeit gegenüber technischen Geräten und Systemen der Farberzeugung und -wiedergabe gekennzeichnet. Euklidische Abstände zwischen gleichwertigen Differenzen der Farbempfindung ermöglichen, die menschliche Wahrnehmung mathematisch zu approximieren. Die mathematische Definition der Farben gewährt eine Konvertierung der Farbwerte des Lab-Farbraums in andere Farbräume. Erfasst sind vom Lab-Farbraum alle Farben, die der Mensch wahrnehmen kann. Nicht erfasst sind die sogenannten theoretischen Farben. Der Lab-Farbraum weist mit dieser Kapazität weit über das Fassungsvermögen geräte- und systemabhängiger Farbräume hinaus.

Median	<p>Der Median gibt in einer statistischen Untersuchung den Wert an, der die Messdaten in zwei gleich große Hälften teilt. Er ist das 50%-Quantil. Zu ermitteln ist der Median durch eine Anordnung der Messzahlen der Größe nach. Bei einer ungeraden Anzahl von Messdaten ist der Median der in der Mitte liegende Wert der Zahlenreihe. Ist die Anzahl gerade, berechnet sich der Median meist als arithmetisches Mittel zwischen den beiden Werten, die sich in der Mitte der Zahlenreihe befinden. Der Median ist, wie der Modalwert, eine Angabe der zentralen Tendenz. Charakteristisch für Angaben der zentralen Tendenz ist, dass sie robust gegenüber statistischen Ausreißern sind.</p>
Modalwert	<p>Der Modalwert ist in statistischen Berechnungen derjenige Wert, der im Rahmen eines Messexperiments am häufigsten gemessen wird. Stellt sich nach der Auswertung der Daten heraus, dass ein Wert am häufigsten gemessen wurde, so ist dieser als der dominante, tonangebende, den Modus der Berechnung bestimmende Wert aufzufassen. Der Modalwert ist, wie der Median, eine Angabe der zentralen Tendenz. Diese ist robust gegenüber statistischen Ausreißern.</p>
Pixel	<p>Die kleinste Einheit des Bildes im digitalen Medium heißt Pixel (Abkürzung von engl. picture element). Das Pixel ist in einer Rasterung, die das digitalisierte oder digitale Bild erfasst, die kleinste diskrete, algorithmisch lesbare, informationstragende Entität.</p>
Range	<p>Die Range beschreibt in der Statistik die Spannweite der Messergebnisse innerhalb der den Berechnungen zu Grunde gelegten Skala. Berechnet wird die Range als Differenz zwischen dem größten und dem kleinsten erzielten Messwert ($x_{\max} - x_{\min} = R$). Die Range ist ein Wert zur Deskription der Dispersion und der sogenannten Variabilität der Messwerte. Mit der Range liegt eine Angabe zur statistischen Streuung der Messdaten vor. Wie die statistische Standardabweichung und Varianz trifft die Range als Dispersions- und Streumaß Angaben zur Verteilung der Daten, ohne dabei Robustheit gegenüber statistischen Ausreißern aufzuweisen.</p>
RGB-Modell	<p>Das RGB-Modell (R: Rot, G: Grün, B: Blau) ist die Grundlage der Farbberechnung im RGB-Farbraum. Dieser Farbraum definiert Farben als Farbwerte in Abhängigkeit des technischen Systems, das die Farberzeugung und -wiedergabe gewährleistet. Der RGB-Farbraum gehört zu den sogenannten additiven Farbräumen, d.h. ein Farbwert wird über die</p>

Festsetzung seiner einzelnen Komponenten beschrieben. Im RGB-Farbraum wird ein Farbwert als bestehend aus unterschiedlichen Komponenten von drei Farbkanälen (rot, grün, blau) berechnet. Der RGB-Farbraum ist ein geräte- und systemabhängiger Farbraum, seine Kapazität zur Farberzeugung und -wiedergabe ist durch die konkreten technischen Vorgaben bedingt, sein Vermögen zur Erfassung der Farben im Vergleich zu wahrnehmungsorientierten Farbräumen reduziert.

silberne Relation

Die silberne Relation beschreibt in der Mathematik ein bestimmtes Größenverhältnis zweier Quantitäten. Es ist gegeben, wenn die Summe aus der kleineren Quantität und zweimal der größeren Quantität sich zur größeren Quantität ebenso verhält wie die größere Quantität zur kleineren. Die silberne Relation approximiert den Wert $\sqrt{1+\sqrt{2}}$ (entspricht als Dezimalzahl 2,41...).

Standardabweichung

Die Standardabweichung beschreibt in der Statistik die durchschnittliche Abweichung der Messergebnisse von ihrem gemeinsamen Mittelwert. Sie berechnet sich aus der Summe der einzelnen Abweichungen vom Mittelwert geteilt durch die Anzahl der Messungen. Ohne die Angabe des gemeinsamen Mittelwertes der Messergebnisse wird eine Interpretation der statistischen Standardabweichung nur unzureichend gewährt. Die Standardabweichung ist, neben der Range, ein Wert zur Beschreibung der statistischen Streuung, der Dispersion und der sogenannten Variabilität der Messdaten. Wie die Range trifft die Standardabweichung als Dispersions- und Streumaß Angaben zur Verteilung der Messwerte, ohne dabei Robustheit gegenüber statistischen Ausreißern aufzuweisen.

Zeitschrift für digitale Geisteswissenschaften

Artikel aus:

Sonderband 1 der ZfdG: Grenzen und Möglichkeiten der Digital Humanities. Hg. von Constanze Baum und Thomas Stäcker. 2015. DOI: [10.17175/sb01](https://doi.org/10.17175/sb01)

Titel:

MUSICI und MusMig. Kontinuitäten und Diskontinuitäten

Autor/in:

Berthold Over

Kontakt: over@uni-mainz.de

Institution: Universität Mainz

GND: [gnd_135197848](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63864-p0011-9) ORCID: [0000000161666496](https://orcid.org/0000000161666496)

Autor/in:

Torsten Roeder

Kontakt: torsten.roeder@uni-wuerzburg.de

Institution: Universität Würzburg

GND: [1084606364](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63864-p0011-9) ORCID: [0000000170437820](https://orcid.org/0000000170437820)

DOI des Artikels:

[10.17175/sb001_017](https://doi.org/10.17175/sb001_017)

Nachweis im OPAC der Herzog August Bibliothek:

[830205667](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63864-p0011-9)

Erstveröffentlichung:

19.02.2015

Lizenz:

Sofern nicht anders angegeben 

Medienlizenzen:

Medienrechte liegen bei den Autoren

Letzte Überprüfung aller Verweise:

24.05.2016

GND-Verschlagwortung:

[Musikwissenschaft](#) | [Migration](#) |

Zitierweise:

Berthold Over, Torsten Roeder: MUSICI und MusMig. Kontinuitäten und Diskontinuitäten. In: Grenzen und Möglichkeiten der Digital Humanities. Hg. von Constanze Baum / Thomas Stäcker. 2015 (= Sonderband der Zeitschrift für digitale Geisteswissenschaften, 1). PDF Format ohne Paginierung. Als text/html abrufbar unter DOI: [10.17175/sb001_017](https://doi.org/10.17175/sb001_017).

Berthold Over, Torsten Roeder

MUSICI und MusMig. Kontinuitäten und Diskontinuitäten

Abstracts

Musiker weisen insbesondere in der frühen Neuzeit eine hohe Mobilität auf, die sich auf ganz Europa und darüber hinaus erstreckt. Sie stimulieren damit maßgeblich die Formierung einer gemeinsamen europäischen kulturellen Identität. Zwei aufeinanderfolgende, internationale Musikforschungsprojekte verfolgen dieses Thema mit unterschiedlichen Schwerpunkten. Die jeweiligen Projektdatenbanken nutzen das gleiche Datenbank-System auf verschiedene Art und Weise und mit unterschiedlichen Zielsetzungen. Im Zusammenhang mit Fallbeispielen wird die forschungsgerechte Datenvisualisierung thematisiert, welche im Zuge des aktuellen Projektes weiterentwickelt werden soll.

Especially in the early modern period, musicians were a highly mobile group that traveled throughout Europe and beyond. Accordingly, they significantly stimulated the formation of a common European cultural identity. Two successive, international music research projects are exploring this topic with different emphases. Each project database uses the same database system in a different way and with different objectives. Using case studies, appropriate research data visualizations are discussed, which will be further developed in the ongoing projects.

1. Einleitung

Aus heutiger Sicht ist Migration ein Phänomen der letzten 150 Jahre, als in großen Auswanderungswellen Menschen in die USA emigrierten, durch die Nazi-Herrschaft zur Emigration gezwungen wurden oder aufgrund der Anwerbung von Arbeitskräften, Kriegen und wirtschaftlicher Not ihre Heimat verließen und ihr Glück woanders suchten. Dabei wird übersehen, dass auch in der Frühen Neuzeit Migration zur Lebenswirklichkeit dazugehörte, dass Menschen erstaunliche und heutigen Migranten nicht unähnliche Lebenswege gingen und sowohl Länder- und Kontinente übergreifende als auch räumlich beschränkte Migrationen, etwa von einer in die andere Stadt, vollzogen wurden. Kaufmänner reisten Monate oder Jahre und hielten sich teilweise über längere Zeit an verschiedenen Orten auf, um Geschäfte zu machen. Gesellen wanderten, waren ›auf der Walz‹, um ihre Ausbildung zu vervollkommen, und kehrten unter Umständen nie an ihren Heimatort zurück. Religionsgruppen emigrierten und siedelten sich in anderen Regionen an.

Musiker gehörten schon immer zu einer Berufsgruppe, die eine hohe Mobilität aufweist. Insbesondere aus der Frühen Neuzeit sind zahlreiche Musiker bekannt, die aus unterschiedlichsten Gründen ihre Heimat- und Wirkungsorte temporär oder definitiv verließen. Wanderungsbewegungen beschränkten sich dabei nicht auf einzelne Regionen und Länder, sondern erstreckten sich auf ganz Europa und darüber hinaus.¹ Diesem Phänomen und den

¹ Vgl. dazu ausführlich Goulet / zur Nieden 2015; Ehrmann-Herfort 2013; Ottenberg / Zimmermann 2012; Mähling 2011; Müns 2005; Meyer 2003; Strohm 2001; Brusniak / Koch 2000.

mit ihm verbundenen Konsequenzen wird aktuell in interdisziplinären und internationalen Forschungsgruppen nachgegangen. Während das Projekt **MUSICI. Musicisti europei a Venezia, Roma e Napoli** bis 2012 europäische Musiker fokussierte, die zwischen 1650 und 1750 mit verschiedensten Zielsetzungen nach Venedig, Rom und Neapel reisten und dort in unterschiedlichsten Positionen aktiv waren,² erforscht das jüngst gestartete Projekt **MusMig. Music Migrations in the Early Modern Age: the Meeting of the European East, West and South** seit Herbst 2013 Migrationsbewegungen von Musikern im 17. und 18. Jahrhundert vor allem im östlichen Europa.³

›Migration‹ wird dabei nicht nur als (Aus-)›Wanderung‹ verstanden, sondern als jede Bewegung im territorialen Raum. Dies bietet im Unterschied etwa zur topographischen oder kulturgeschichtlichen eine den historischen Gegebenheiten näher kommende, an Herrschaftsräume orientierte Perspektive. Somit fallen auch Reisen von Musikern und Operntruppen zum Zwecke der Aufführung musikalischer Werke, Kavaliereisen, an denen Musiker teilnahmen, Ausbildungsreisen oder Gesellenwanderungen unter die Untersuchungsobjekte des Projekts. Auch der Begriff ›Musiker‹ wird weit gefasst und umfasst neben Instrumentalisten, Komponisten und Sängern u. a. auch Textdichter, Instrumentenbauer und Musiktheoretiker.⁴ Als Projektergebnis werden Erkenntnisse in Bezug auf die These erwartet, dass die Migration von Musikern maßgeblich zur Dynamik und Synergie der europäischen Kulturszene beigetragen, eine Katalysatorfunktion in Bezug auf Innovationen, auf stilistische Veränderungen sowie auf den Wandel von musikalischen und sozialen Mustern ausgeübt und insgesamt die Formierung einer gemeinsamen europäischen kulturellen Identität stimuliert habe.

2. MUSICI und MusMig

Als Fortführung von MUSICI baut MusMig auf den Erfahrungen des Vorgängerprojekts auf und muss sie angesichts neuer Fragestellungen modifizieren. Neben zahlreichen Kontinuitäten stehen substantielle Diskontinuitäten. Während sich z.B. MUSICI auf drei italienische Städte (Venedig, Rom, Neapel) und ihre musikalischen Institutionen konzentriert und die dorthin reisenden ›ausländischen‹ Musiker (zu denen auch Musiker der anderen betrachteten Städte zählen mussten, da sie aus einem anderen Territorium stammten) erfasst,⁵ ist MusMig breiter angelegt und muss Möglichkeiten bieten, jede Art von Lokalität aufnehmen zu können: vom Hof über die Reichsstadt bis zum ländlichen Kloster. Auf systematischer Ebene hat dies eine stärkere Generalisierung zur Folge, die es ermöglicht, eine größere Bandbreite an Orten zu erfassen. Während für Venedig, Rom und Neapel eine relativ genaue Aufschlüsselung der institutionellen Arbeitgeber von Musikern erfolgen konnte, ist dies aufgrund der Fülle von Institutionen im Untersuchungsgebiet von MusMig nicht mehr möglich. Hier müssen

² Goulet / zur Nieden 2015. Das Projekt wurde von der Deutschen Forschungsgemeinschaft (DFG) und der Agence nationale de la recherche (ANR) finanziert und von Anne-Madeleine Goulet (Paris) und Gesa zur Nieden (Mainz) geleitet.

³ Das von der europäischen Forschungsinitiative **HERA** (Humanities in the European Research Area) finanzierte Projekt vereint unter der Leitung von Vjera Katalinić (Zagreb) Forscher aus Kroatien, Polen, Slowenien und Deutschland.

⁴ Selbstverständlich umfasst dieser Musiker-Begriff auch weibliches Personal, insbesondere Sängerinnen.

⁵ Vgl. die Erläuterungen zur Datenbank bei Berti / zur Nieden / Roeder 2013 sowie Berti / Roeder 2015.

auf genereller Ebene Klassifizierungsmöglichkeiten geschaffen werden, die es dennoch ermöglichen, aussagekräftige Auswertungen zu garantieren. Gleichzeitig müssen angesichts veränderter politischer Rahmenbedingungen etwa Regierungsformen stärker ausdifferenziert werden. Während Venedig eine Republik, Rom eine kirchliche Wahlmonarchie und Neapel ein von einer erbmonarchischen Fremdmacht (Spanien, Frankreich oder Österreich) beherrschtes Territorium war, ist die Bandbreite im Alten Reich und darüber hinaus ungleich größer: vom Heiligen Römischen Reich über Kurfürstentümer, Herzogtümer, Fürstbistümer, Reichsstädte zu weiteren Metropolen (etwa der Stadt Leipzig, die zwar zum Herzogtum Sachsen gehörte, aber aufgrund ihres Status als Messestadt eine Sonderstellung einnahm).

3. Die Datenschemata von MUSICI und MusMig

Auf der Basis der Erfahrungen des MUSICI-Projekts ist das Nachfolgeprojekt MusMig damit begonnen worden, ein Datenschema für die Erfassung von Informationen zu migrierenden Musikern zu entwerfen.

Die Grundlage dafür liefert das **Personendaten-Repository** (PDR) der Berlin-Brandenburgischen Akademie der Wissenschaften, ein seit 2009 laufendes DFG-gefördertes Projekt, das sich mit dem Aufbau einer digitalen Infrastruktur für prosopographische Forschungsprojekte befasst und das ein flexibles Datenschema, ein Serversystem, das Erfassungsprogramm **Archiv-Editor**⁶ sowie verschiedene Webservices entwickelt hat.

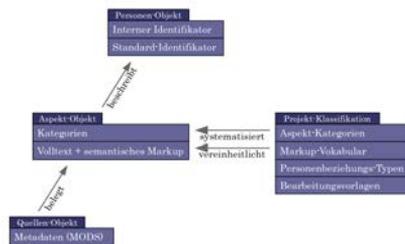


Abb. 1: Das Datenmodell des Personendaten-Repositorys. Grafik: Torsten Roeder, 2014.

Das vom Personendaten-Repository vorgegebene Datenmodell (vgl. Abbildung 1) ist bewusst sehr offen gehalten, da es darauf ausgerichtet ist, sehr unterschiedliche Projekte beherbergen zu können. Den zentralen Angelpunkt des Modells bildet das *Personenobjekt*, dem beliebig viele Informationen zugeordnet werden können. Eine einzelne Information wird *Aspekt* genannt und kann beispielsweise aus einem Namen, einem Beruf oder einem Aufenthaltsort bestehen. Jeder Aspekt muss zudem durch mindestens eine *Quelle* belegt werden.⁷ In Bezug

⁶ Vgl. auch Plutte 2011.

⁷ Für die bibliographische Erfassung der Quellen nutzt das Personendaten-Repository das **Metadata Object Description Schema** (MODS).

auf Musiker könnten dies auch Lehrer- und Schülerverhältnisse, Patronagenetzwerke, Mitwirkung bei Konzerten, komponierte Werke und mehr sein. Insofern besteht jede Person aus einer beliebigen Menge von kleinen, separaten Einzelinformationen. Dies bedeutet, dass sich die Person weniger als klassisches Datenblatt konstituiert, sondern vielmehr in einer dynamischen Form, die sich aus der Menge aller mit ihr verbundenen Informationen definiert.⁸ Nach diesem Prinzip wurden im MUSICI-Projekt ca. 2.200 Aspekte zu über 300 Personen erfasst; in MusMig wird die Datenmenge aufgrund des veränderten Projektumfanges mindestens doppelt so hoch sein.

Um dennoch die sinnvolle, schematische Struktur eines Datenblattes prinzipiell zu erhalten, kann jeder Aspekt einer biographischen *Kategorie* zugeordnet werden. Kategorien konstituieren Gruppen von semantisch vergleichbaren Informationen, z.B. Lebensdaten, Ausbildung, Anstellungen, Reisen, Kompositionen etc. (vgl. *Abbildung 2*). Mithilfe dieser Kategorisierung können die Informationen zu einem Datenblatt bzw. zu einer systematischen Biographie angeordnet werden. Dem generischen Ansatz des Personendaten-Repositoriums entsprechend ist es möglich, die Kategorien durch das Projekt selbst festzulegen. Darüber hinaus erlaubt das System nicht nur die Definition projektinterner Kategorien, sondern auch ein Mapping auf Standards (z.B. CIDOC-CRM⁹), und schafft so eine Vergleichbarkeit und Nutzbarkeit der Daten auch über das einzelne Forschungsprojekt hinaus. Grundsätzlich stehen jedoch die Perspektiven der individuellen Forschungsvorhaben im Vordergrund, so dass die Auswahl an Kategorien – abgesehen von Basisdaten – primär die Herangehensweise von Einzelprojekten unterstützt.



Abb. 2: Beispielaspekte zu Georg Friedrich Händel. In den dunkel schattierten Reitern steht die jeweilige Kategorie. Grafik: Torsten Roeder, 2014.

Das Kategorienschema, welches für das MusMig-Projekt zusammengestellt wurde, ähnelt dem MUSICI-Schema, welches auf der [Datenbank-Website](#) öffentlich verfügbar und dokumentiert ist und außerdem im Abschlussband des Projektes eine ausführliche Besprechung erfährt.¹⁰ Die Kategorien beider Projekte umfassen zum einen übliche Basiskategorien wie Lebensdaten, Namensansetzungen und genealogische Informationen

⁸ Vgl. auch Walkowski 2009, S. 3.

⁹ CIDOC CRM = Comité international pour la documentation / Conceptual Reference Model.

¹⁰ Vgl. Berti / Roeder 2015.

und erlauben zum anderen die Erschließung der speziellen biographischen Strukturen der Lebenswege von migrierenden Musikern und unterstützen auf diese Weise die Untersuchung von Ausbildungswegen, Karrierestationen, Kompositionen, Beziehungsnetzwerken und Rezeptionen. Trotz der ähnlichen Ansätze von MUSICI und MusMig bestehen einige Unterschiede in der praktischen Anwendung, deren Ursache in dem unterschiedlichen Umfang der Projekte liegt. In der Konsequenz finden sich daher auch auf der technischen Ebene sowohl Kontinuitäten als auch Diskontinuitäten von Musici zu MusMig.

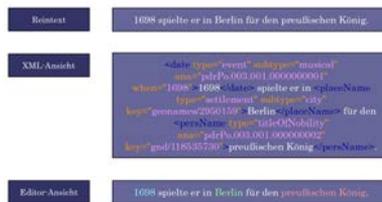


Abb. 3: Semantisches Markup in XML. Das komplexe, semantisch angereicherte XML-Format wird im Editor vereinfacht dargestellt. Grafik: Torsten Roeder, 2014.

Ein Schlüssel zur datenbankgestützten Recherche ist die semantische Anreicherung von Daten. Im Personendaten-Repository besteht jeder einzelne Aspekt aus Freitext, in der Regel aus wenigen Wörtern oder einer Phrase, die zusätzlich semantisch angereichert werden. Abbildung 3 zeigt einen Beispielaspekt: »1698 spielte er in Berlin für den Kurfürsten« (dies beschreibt ein Ereignis im Leben Georg Friedrich Händels). Innerhalb eines solchen Satzes werden mithilfe von XML-Markup inhaltliche Ausdrücke gekennzeichnet wie etwa Datumsangaben, Personennamen, Ortsnamen, Körperschaftsnamen und Sachbegriffe, wodurch eine Basis für die systematische Recherche geschaffen wird. Die XML-Codierung, die zwar grundsätzlich für alle lesbar ist, die einigermaßen mit XML vertraut sind, wird im Archiv-Editor durch eine Bearbeitungsansicht gekapselt, in der die unterschiedlichen Datentypen durch verschiedene Farben hervorgehoben werden. In dem genannten Beispiel kann demnach ein Datum, ein Ort und eine Person erfasst werden. In Attributen werden zusätzliche Informationen festgehalten.

Zusätzlich zu den oben beschriebenen Kategorien enthält die MusMig-Klassifikation auch ein hierarchisches Vokabular, welches bei der Spezifizierung der semantischen Qualität, die in den Ausdrücken steckt, behilflich ist (in den Attributen *type*, und *role*). Das Beispiel liefert einen Ortsnamen (»Berlin«), der als *settlement/city* (Ansiedlung/Stadt) gekennzeichnet werden kann. Das Datum (»1698«) sollte als musikalische Aufführung markiert werden: *event/musical* (Ereignis/Musik). Der »Kurfürst« kann als Person markiert werden und zudem – ohne dass er explizit genannt werden muss – mithilfe des *ana[lysis]*-Attributs an das passende Personenobjekt in der Datenbank angebunden werden. Alternativ kann die Person durch eine

Standard-ID einer Bibliothek, etwa **GND**, **VIAF** oder **LCNAF** im *key*-Attribut identifiziert werden.¹¹ Da das semantische Markup direkt in den Quelltext eingebettet ist, kann auch mehr als eine Person (oder Orte, Daten, Körperschaften) in dem gleichen Aspekt aufgenommen werden (die Identitäten werden dann durch entsprechende *ana[lysis]*- oder *key*-Attribute voneinander unterschieden). Dies ist insbesondere für Migrationsprozesse zwischen mehreren Orten und für Beziehungen zwischen Personen und Organisationen von Bedeutung.

Möchte man später systematisch alle Aufführungen des Jahres 1698, alle Aufführungen in Berlin oder alle Aufführungen in Gegenwart des Kurfürsten eruieren, kann man die gewünschten Daten leicht aus der Datenbank herausfiltern (was durch geeignete Suchmasken unterstützt werden wird). Diese Methode eignet sich auch für die Erzeugung von chronologischen Ansichten, geographischen Übersichten oder statistischen Auswertungen.

Das folgende Beispiel (Abbildung 4) zeigt, wie dies in der MUSICI-Datenbank umgesetzt wird. An der Textstruktur ist erkennbar, dass die Forscher unterschiedliche Praktiken in der Datenerfassung befolgen (der Aspekt zu Johann Adolf Hasse enthält mehr als einen Aufenthalt, während die anderen vier Aspekte genau einen enthalten). Insofern ist das Format nicht immer konsistent. Um ähnliche Situationen in der MusMig-Datenbank zu vermeiden, ist geplant, Vorlagen mit vordefinierten Textschablonen und Basisdatenblätter mit vordefinierten Kategorien zu verwenden.

8	Nome: Antonio Dankeiv
	Risultato: Roma, 1709-1732, almeno.
	Categoria: Permanenza
9	Nome: Urbano Fraus
	Risultato: Roma, 1715-1729, almeno.
	Categoria: Permanenza
10	Nome: Michele Surignach
	Risultato: Roma, 1722-1759.
	Categoria: Permanenza
11	Nome: Euberto Ignazio Loyselet
	Risultato: Roma, prima del 1720 - dopo 1735.
	Categoria: Permanenza
12	Nome: Johann Adolf Hasse
	Risultato: Venezia: 1) Carnevale - luglio 1730. 2) Primavera 1731. 3) Carnevale - maggio 1732. 4) Probabilmente 1733. 5) 1734 - 1736. 6) Autunno 1738 - carnevale 1739. 7) Estate 1744 - estate 1745. 8) Autunno 1746. 9) Probabilmente inverno 1748 - primavera 1749. 10) Probabilmente estate 1752. 11) 1756 - estate 1756. 12) Autunno 1764. 13) 1773 - 1783. Napoli: 1) 1722/ 1724 - 1733. 2) Autunno 1732. 3) Autunno 1758. Roma: 1) Gennaio 1732.
	Categoria: Permanenza

Abb. 4: Abfragebeispiel aus der Musici-Datenbank (Auszug) mit den Ergebnissen zu »Ort = Rom, Zeit = 1725, Kategorie = Aufenthalt«. Der relevante Aspekt ist jeweils unter dem Musikernamen angezeigt. Berti / zur Nieden / Roeder 2013.

Das folgende Beispiel zeigt auf, inwieweit Personen und Orte für das MusMig-Projekt und dessen Datenbank von Bedeutung sind.

¹¹ **GND** = Gemeinsame Normdatei; **VIAF** = Virtual International Authority File; **LCNAF** = Library of Congress Name Authority File.

4. Musik und Dynastie. Musikermigration in dynastischen Kontexten

Dieses Teilprojekt von MusMig untersucht ein Konglomerat an Höfen, die dynastisch miteinander verbunden waren und im Zuge der dynastischen Sukzession miteinander verschmolzen wurden: der Hof der Münchner Wittelsbacher sowie die Wittelsbachischen Nebenlinien Pfalz-Neuburg, Pfalz-Sulzbach und Pfalz-Zweibrücken.¹² Durch die Wittelsbachische Hausunion von 1724 kooperierten diese Linien im ganzen 18. Jahrhundert politisch und militärisch zum Teil eng miteinander.¹³ Die Untersuchung fokussiert hauptsächlich drei Aspekte:

- Erleichtern die auf politischem und militärischem Gebiet sehr engen dynastischen Verbindungen die Migration von Musikern und den damit verbundenen kulturellen Austausch?
- Haben die Sukzessionen Einfluss auf Musikermigrationen, etwa durch Entlassungen oder Residenzverlagerungen?
- Inwiefern dienen Musiker verschiedener lokaler Abstammung der höfischen Profilbildung, wie es landläufiger Meinung nach in Mannheim (böhmische Musiker) und München (italienische Musiker) geschah?

Während in den ersten beiden Aspekten vor allem mögliche Binnenmigrationen untersucht werden, die durch den Austausch von Musikern sowie durch den Transfer des Hofes ausgelöst wurden, stellt der zweite Aspekt Fragen der Musikerrekrutierung in den Mittelpunkt. Die Datengrundlage bilden insbesondere Rechnungs- und Besoldungsbücher, die von den Höfen erhalten sind und aus denen Anstellungszeiten und Gehaltshöhe der Musiker ermittelt werden können. In Einzelfällen finden sich auch weitere Hinweise. Außerdem bieten die seit dem Beginn des 18. Jahrhunderts erscheinenden Hofkalender, die oftmals einen Überblick über den gesamten Hofstaat geben, ergiebige Datenmaterial.

Als Beispiele für migrierende Musiker innerhalb von Dynastien seien im Folgenden der Komponist Johann Christoph Pez (1664–1716), der Sänger Valeriano Pellegrini (ca. 1663–1746) und der Flötist Johann Baptist Wendling (1723–1797) in den Blick genommen.

Johann Christoph Pez wurde in München ausgebildet, begann seine Tätigkeit am Münchner Hof 1688 und wurde im folgenden Jahr vom Kurfürsten Max Emanuel zur weiteren Ausbildung nach Rom geschickt.¹⁴ 1692 bis 1694 hielt er sich wieder in München auf, wo das höfische Musikleben auf ein Minimum reduziert worden war, nachdem Max Emanuel 1692 aufgründ

¹² 1716 folgte Karl Philipp von Pfalz-Neuburg auf seinen Bruder Johann Wilhelm; 1742 trat Karl Theodor von Pfalz-Sulzbach die Erbfolge der Neuburger an; 1777 wurde er als Erbfolger der bayerischen Wittelsbacher Kurfürst von Bayern; nach seinem Tod 1799 ging das Kurfürstentum Bayern an Maximilian Joseph von Pfalz-Zweibrücken über.

¹³ Vgl. Kraus / Schmid 1988, S. 517. Vgl. die Hausverträge und Sukzessionsvereinbarungen von 1724, 1728, 1734, 1746, 1747, 1761, 1766 und 1774 in D-MhSa, Geheimes Hausarchiv, Hausurkunden 1773, 1774, 1787–1790, 1841, 1843–1850, 1852–1853, 1855–1856, 1856–1859, 1872. Dokumente zu Hausverträgen und Sukzessionsvereinbarungen zwischen Pfalz-Sulzbach und Pfalz-Zweibrücken sind in D-MhSa, Geheimes Hausarchiv, Korrespondenzakten 504, 746, 1281–1282, 1689 erhalten.

¹⁴ Zu Pez' Biographie, die hier durch neue Daten ergänzt wird, vgl. grundlegend Rampe / Berben 2005 sowie Roche 2001.

seiner Statthalterschaft der Spanischen Niederlande seine Residenz nach Brüssel verlegt hatte. Dies war wohl ausschlaggebend für die Anstellung Pez' am Hof des Bonner Kurfürsten Joseph Clemens, einem Bruder Max Emanuels.¹⁵ Im Zuge des Spanischen Erbfolgekriegs verließ Pez Bonn und kehrte 1702 an den bayerischen Hof zurück, der durch die Rückkehr Max Emanuels aus den Spanischen Niederlanden wieder an Bedeutung gewann. Dort sollte er so lange ein Wartegeld erhalten, bis eine adäquate Position in der Hofkapelle für ihn bereitstünde.¹⁶ Nach der Flucht des Kurfürsten im Jahre 1704 übernahm er unter der österreichischen Okkupation die musikalische Erziehung der Prinzen¹⁷ und blieb dort bis zu deren Exilierung nach Klagenfurt und Graz im Jahre 1706.¹⁸ Anschließend fand er eine Anstellung am protestantischen Hof des Herzogs von Württemberg in Stuttgart. Pez verließ die katholische Dynastie der Wittelsbacher, um einem protestantischen Regenten zu dienen, was aus konfessioneller Sicht nicht unproblematisch war.¹⁹

Valeriano Pellegrini vertritt den klassischen Typus eines reisenden Sängers (Kastrat), der neben Opernengagements in verschiedenen Städten feste Anstellungen an Höfen hatte. Pellegrini wurde wohl in Bologna geboren und war zeitweise Mitglied der päpstlichen Kapelle.²⁰ Nach Opernengagements in Wien (1699), Mantua (1700), Genua und Piacenza (1701) wurde er 1702 von Max Emanuel fest angestellt, als dieser aus Brüssel nach München zurückkehrte. Dort blieb er bis 1705/06, bevor er aufgrund der Wirren des Spanischen Erbfolgekriegs nach Düsseldorf an den kurfürstlichen Hof Johann Wilhelms von der Pfalz ging, der aus der wittelsbachischen Nebenlinie Pfalz-Neuburg stammte.²¹ Bis zum Tod des Kurfürsten 1716 sang er in Düsseldorf, fungierte als Johann Wilhelms Agent (er kaufte eine Medaillensammlung und Gemälde) und nahm Opernengagements in Venedig und London wahr – vor allem sang er in Opern Georg Friedrich Händels.²² 1716 trat er in den Dienst von Johann Wilhelms Nachfolger, dessen Bruder Karl Philipp, und zog mit dem Hof nach Mannheim. Seinen Lebensabend verbrachte er wieder in Rom.

Der Flötist Johann Baptist Wendling wechselte wohl aufgrund seiner pädagogischen Kompetenz und seines Renommées aus dem Dienst Herzogs Christian IV. von Pfalz-Zweibrücken, dessen Flötenlehrer er 1745 bis 1752 war, in eine Anstellung am Mannheimer Hof Kurfürst Karl Theodors von der Pfalz, wo er die gleiche Funktion übernahm.²³ Wendling unternahm unter beiden Anstellungen zahlreiche Konzertreisen, die ihn u. a. nach Paris, London, Wien und Berlin führten. Als infolge der Sukzession Karl Theodors auf den

¹⁵ D-MhSA, Hofzahlamt 732 (Besoldungsbuch 1694), fol. 61v: »Vermög Sig[na]t] aus Brüssel [...] dato 18. Martij 1695. ist dem Pöze[n] d[a]ß 4. Quartal diss iahrs, weil Er hernach in Chur Cöllnische dienst kome[n], [ver]wiligt word[en].«

¹⁶ D-MhSA, Hofzahlamt 741 (Besoldungsbuch 1702), fol. 77v: »gewest Chur= Cöllnischer Capellmaister, ist vermög ordonanz, in die Churfürstl: dienst aufgenom[m]men = und ihn in dessen Zum Warthgelt, bis er völlige installirt werde, vom .1. [Septem]b[er] diss iahrs angeschafft worden.«

¹⁷ D-MhSA, Hofzahlamt 745 (Besoldungsbuch 6. Juni–31. Dezember 1705), fol. 75v: »musico so Ihro d[er] Churprinzen instruiert; Hofzahlamt 746 (Besoldungsbuch 1706), fol. 59r: »Instructore bei denen ältern durchleichtigen .3. Prinzen«. Vgl. auch Iser 2000, S. 97.

¹⁸ Zur Zeit der Prinzen in Klagenfurt und Graz vgl. Zedler 2012.

¹⁹ Vgl. Owens 2011, S. 167–172.

²⁰ Zur Biographie Pellegrinis vgl. Marx 2008, Bd. 2, S. 762–764; Dean / Rosselli 2001.

²¹ Zu der im New Grove nicht erwähnten Anstellung in München vgl. D-MhSA, Hofzahlamt 741 (Besoldungsbuch 1702), fol. 77r und Hofzahlamt 744 (Hofzahlamt 1705), fol. 75r sowie Over 2007, S. 274. Zur Anstellung in Düsseldorf vgl. Einstein 1908, S. 409.

²² *Agrippina* (Venedig 1709/10), *Il pastor fido*, *Teseo*, evtl. *Lucio Cornelio Silla* (alle London 1712).

²³ Pelker 2007; Gunson 2001; Pelker 2011; Pelker 2002; Gunson 2002.

kurfürstlichen Thron von Bayern der Mannheimer Hof 1778 nach München transferiert wurde, zog Wendling mit.²⁴

5. Visualisierung

Die in den Beispielen gezeigten Sachverhalte sollen zur Veranschaulichung in eine graphisch orientierte Form überführt werden. Das MusMig-Datenmodell (vgl. [Abschnitt 3](#)) erlaubt es, Informationen nach verschiedenen Parametern zu filtern, zu gruppieren und anzuordnen. Üblicherweise läuft eine Suche über den Volltext, und die Darstellung eines Suchergebnisses geschieht dann in Form einer linearen Liste. Jedoch gibt es noch andere und anspruchsvollere Methoden. Durch semantische Filtereinstellungen, die den Typ der eigentlich gesuchten Information bestimmen, und durch die Auswahl einer angemessenen Visualisierungsmethode können Auswahl und Form des Materials auf bestimmte Forschungsinteressen hin zugeschnitten werden.

Einige Visualisierungen sind allgemein verbreitet und in der Lage, einen allgemeinen Überblick des verfügbaren Materials oder dessen Verteilung im Datenbankkorpus zu geben. Um einige Beispiele zu nennen: Zeitleisten ordnen Informationen nach den Proportionen anhand der Verhältnisse ihrer zeitlichen Entfernung. Landkarten verteilen die Informationen (sofern sie mit geospatialen Daten angereichert sind) auf einer zweidimensionalen Kugelprojektion. Tabellen sortieren schachbrettartig nach zwei unabhängigen Parametern. Baumlisten erlauben die Darstellung von hierarchischen Beziehungsstrukturen. Derartige Visualisierungen können dynamisch und generisch erstellt werden, da die Methode im Wesentlichen unabhängig von den semantischen Details in den Daten ist.

In der MUSICI-Datenbank wurde dies mithilfe eines sehr einfachen [Visualisierungs-Interface von Google](#) umgesetzt. Damit ist es möglich, eine gefilterte Sammlung von Aspekten aus der Datenbank automatisch zu verarbeiten und als Ergebnis ein Tortendiagramm, ein Balkendiagramm, eine Landkarte oder auch eine Zeitleiste²⁵ zu erhalten. Das folgende Beispiel ([Abbildung 5](#)) zeigt die Verteilung von allen Aspekten, die in Beziehung zu Aufhalten in Venedig zwischen 1650 und 1750 stehen. Es ist leicht erkennbar, dass in den 1740ern ein Höchststand vorliegt und dass für die 1660er offensichtlich keine Daten vorliegen. Dies sollte eine Einladung für jeden Forscher und jede Forscherin sein, nach möglichen Gründen für dieses Phänomen zu suchen. Es muss erwähnt werden, dass die Datenbank die historische Realität nicht 1:1 widerspiegelt, sondern lediglich die Informationen in der Datenbank, welche einen Auszug aus der riesigen Zahl an Quellen darstellt. Auch wenn das Korpus so repräsentativ wie möglich zusammengestellt ist, wird es niemals exakte Zahlen liefern.

²⁴ Wendling und seine Frau erhielten 1780 eine Umzugskostenerstattung. D-Mhsa, Hofzahlamt 192 (Jahresrechnung 1780); Hofzahlamt 2153 (Belege).

²⁵ Durch Inkludierung des [SIMILE Timeline Widget](#) vom MIT.

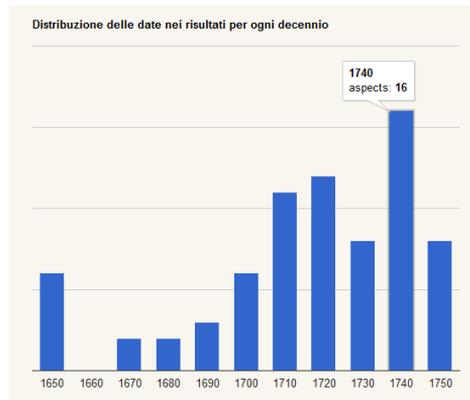


Abb. 5: Zeitliche Verteilung von allen Daten zu Aufenthalten in Venedig in der Musici-Datenbank. Das Hover-Beispiel zeigt die Zahl der verfügbaren Daten auf dem Höchststand in den 1740ern. Berti / zur Nieden / Roeder 2013.

Dieser Visualisierungsansatz zielt darauf, aus der gleichen Datenmenge eine Bandbreite an Ansichten zu generieren, um mehr Perspektiven auf das Material zu erlauben. Das nächste Beispiel (Abbildung 6) zeigt exakt dieselben Informationen (alle Aufenthalte in Venedig), jedoch in einer geographischen Darstellung. Jeder Punkt auf der Karte repräsentiert einen Aspekt, der Informationen zu Venedig und einer weiteren Stadt außerhalb Venedigs enthält, was Relationen zwischen Venedig und jenen anderen Orten impliziert, zum Beispiel Wanderungsbewegungen von Musikern. Die Aufmerksamkeit des Forschers oder der Forscherin könnte insbesondere durch die Punkte außerhalb Italiens angezogen werden (Bergedorf, Grenoble und Šibenik).

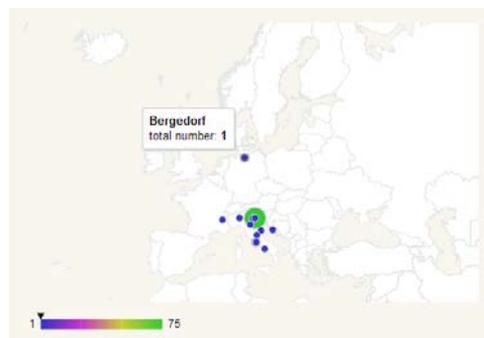


Abb. 6: Räumliche Verteilung aller Aspekte zu Venedig in der Musici-Datenbank. Das Hover-Beispiel (nördlichster Punkt) zeigt, dass es einen Datenbankeintrag zu Bergedorf gibt (der Geburtsort von Johann Adolf Hasse). Berti / zur Nieden / Roeder 2013.

Nachdem generische Visualisierungen in der MUSICI-Datenbank bereits implementiert wurden, zielt das MusMig-Projekt insbesondere darauf, individuellere Visualisierungen zu entwickeln, die die Arbeit des einzelnen Forschers und der einzelnen Forscherin unterstützen. Solche Visualisierungen, die von der gewöhnlichen Gestalt abweichen und mehrere verschiedene, mehrdimensionale Aspekte parallel abbilden, sind weniger verbreitet, aber

potenziell sehr effektiv.²⁶ Sie erlauben es, den Blick auf sehr spezifische Fragen zu richten, benötigen aber auch mehr Ausarbeitung und sind in der Implementierung hochkomplex. Die zwei folgenden Beispiele dazu sind Skizzen auf der Grundlage der zuvor erörterten historischen Fallbeispiele.

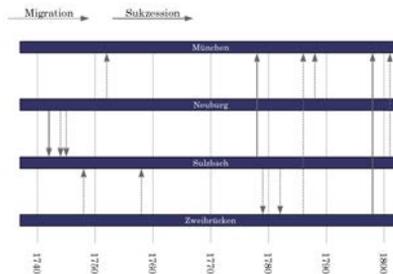


Abb. 7: Parallele Visualisierung von Sukzessionen und Migrationsbewegungen. Grafik: Torsten Roeder, 2014.

Die geographischen Achsen wurden zugunsten der chronologischen Darstellung auf eine Dimension reduziert. Zwischen den horizontalen Ortslinien sind Personendaten in Form von Pfeilen angebracht. Jeder rote Pfeil entspricht einem Musiker, der von Ort A nach Ort B abwanderte. Jeder grüne Pfeil entspricht einer Erbfolge. Es lässt sich damit sichtbar machen, ob bzw. welche Erbfolgen zu Migrationen führten.

Das Grundschema der ersten Skizze (Abbildung 7) entspricht einer einfachen Zeitleiste. Darüber sind die Ortsnamen von vier Höfen als horizontale Linie dargestellt. Ihre geographische Qualität wurde zu einer Linie konvertiert, um die parallel chronologische Perspektive zu ermöglichen. Die Ortslinien sind durch Pfeile verbunden, die Bewegungen einer Person oder eines Hofes repräsentieren. Jede gepunktete Linie beschreibt dabei einen Musiker, der von A nach B wanderte, und jede durchgezogene Linie ein Erbfolge-Ereignis. Dadurch ist es möglich, auf einen Blick Sukzessionen zu identifizieren, die Immigration oder Emigration nach sich ziehen. Es gibt außerdem einen Eindruck von der Attraktivität der Höfe im Vergleich über die Jahrzehnte hinweg. Es obliegt wiederum dem Forscher und der Forscherin, die in der Visualisierung aufgedeckten Phänomene näher zu untersuchen.

²⁶ Dies wird ausführlich dargelegt in Rosenberg / Grafton 2010.

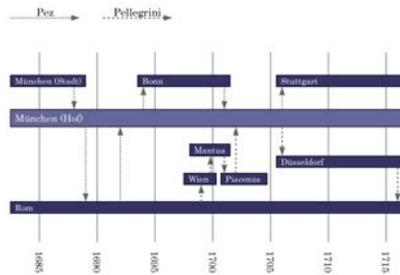


Abb. 8: Lebenswege von Pez und Pellegrini im Vergleich. Grafik: Torsten Roeder, 2014.

Mit einer sehr ähnlichen Methode ist es möglich, die Biographien von Musikern und ihren individuellen Beziehungen zu einem bestimmten Hof zu vergleichen. Diese Skizze (Abbildung 8) setzt den Münchener Hof und die zwei Biographien von Johann Christoph Pez und Valeriano Pellegrini in den Fokus. Migrationsbewegungen sind wiederum durch Pfeile dargestellt, wobei die gepunktete Linie für Pez und die gestrichelte für Pellegrini steht. Es wird sichtbar, dass der Münchener Hof in den frühen 1700ern eine bedeutende Rolle in beiden Biographien spielte: Beide wurden fast zur gleichen Zeit von Max Emanuel engagiert (1701 und 1702) und verließen den Hof nach seiner Flucht (1706). Es wäre von großem Interesse, weitere Musikerbiographien in den Vergleich aufzunehmen und entsprechend weitere musikwissenschaftliche Recherchen zur Zeit der Spanischen Erbfolgekriege durchzuführen.

Ausgehend von diesen Beispielen ist es geplant, eine Reihe von inspirierenden Visualisierungen zu entwickeln, die in der Lage sind, die Herangehensweise der individuellen Projekte zu unterstützen. Dies wird den Forschern und allen anderen potenziellen Datenbanknutzern sowohl einen fokussierten Blick als auch einen exploratorischen Zugang erlauben.

6. Zusammenfassung

Bei allen Gemeinsamkeiten, die MUSICI und MusMig auf eine Stufe zu stellen scheinen, sind jedoch auch signifikante Unterschiede festzustellen. MusMig ist somit nicht nur als bloße Erweiterung von MUSICI zu sehen und MUSICI ebenso wenig als Vorstudie des umfangreicheren MusMig, sondern es handelt sich um zwei Projekte mit eigenständigen Forschungsansätzen, welche sich aber in wesentlichen, sowohl methodischen als auch inhaltlichen Punkten berühren und aufeinander Bezug nehmen. Die Kontinuitäten und Diskontinuitäten spiegeln sich auch in den digitalen Repräsentationen der beiden eng verwandten Projekte, die auf einem gemeinsamen Datenmodell basieren, aber individuelle Klassifikations- und Implementierungsansätze verfolgen.²⁷

²⁷ Eine englische Version dieses Beitrags erscheint im Bericht zum Workshop »Music Migrations: from Source Research to Cultural Studies« (Mainz, 24.-25. April 2014), hg. von Gesa zur Nieden / Berthold Over. Bielefeld [voraussichtl. 2015] (= Mainzer Historische Kulturwissenschaften).

Bibliographische Angaben

- Michela Berti / Torsten Roeder: The »Musici« Database. An Interdisciplinary Cooperation. In: Europäische Musiker in Venedig, Rom und Neapel. Les musiciens européens à Venise, Rome et Naples. Musicisti europei a Venezia, Roma e Napoli. Deutsch-italienische Round-Table-Gespräche. Hg. von Anne-Madeleine Goulet / Gesa zur Nieden. Kassel [u.a.] 2015. (= *Analecta musicologica* 52), S. 633–645. [\[Nachweis im GBV\]](#)
- Winton Dean / John Rosselli: Pellegrini, Valeriano. In: *The New Grove Dictionary of Music and Musicians*. Bd. 19. Hg. von Stanley Sadie. London 2001, S. 299. [\[Nachweis im GBV\]](#)
- The Eighteenth-Century Diaspora of Italian Music and Musicians. Hg. von Reinhard Strohm. Turnhout 2001 (= *Speculum Musicae* 8). [\[Nachweis im OPAC\]](#)
- Alfred Einstein: Italienische Musiker am Hofe der Neuburger Wittelsbacher. 1614–1716. Neue Beiträge zur Geschichte der Musik am Neuburg-Düsseldorfer Hof im 17. Jahrhundert. In: *Sammelbände der Internationalen Musikgesellschaft* 9 (1908), S. 336–424. [\[Nachweis im OPAC\]](#)
- Europäische Musiker in Venedig, Rom und Neapel. Les musiciens européens à Venise, à Rome et à Naples. Musicisti europei a Venezia, Roma e Napoli. Deutsch-italienische Round-Table-Gespräche. Hg. von Anne-Madeleine Goulet / Gesa zur Nieden, 2 Bde. Kassel [u.a.] 2015. (= *Analecta musicologica* 52). [\[Nachweis im OPAC\]](#)
- Emily Gunson: Wendling. In: *The New Grove Dictionary of Music and Musicians*. Hg. von Stanley Sadie. London 2001, Bd. 27, S. 280–282. [\[Nachweis im GBV\]](#)
- Emily Jill Gunson: The Court of Carl Theodor: »A Paradise for Flautists«. In: *Mannheim – Ein Paradies der Tonkünstler?* Hg. von Ludwig Finscher / Bärbel Pelker / Rüdiger Thomsen-Fürst. Frankfurt/Main [u.a.] 2002 (= *Quellen und Studien zur Geschichte der Mannheimer Hofkapelle* 8), S. 263–283. [\[Nachweis im GBV\]](#)
- Handbuch der bayerischen Geschichte. Begr. von Max Spindler. Hg. von Andreas Kraus / Alois Schmid in Verbindung mit Dieter Albrecht. Bd. 2: Das alte Bayern. Der Territorialstaat vom Ausgang des 12. Jahrhunderts bis zum Ausgang des 18. Jahrhunderts. 2. Auflage. München 1988. [\[Nachweis im OPAC\]](#)
- Ulrich Iser: »Wie du ein französisches lied vor meiner gesungen«. Zur musikalischen Erziehung der Wittelsbacher Prinzen. In: *Die Bühnen des Rokoko. Theater, Musik und Literatur im Rheinland des 18. Jahrhunderts*. Hg. von Frank Günther Zehnder. Köln 2000 (= *Der Reiß im Himmel. Clemens August und seine Epoche* 7), S. 86–112. [\[Nachweis im OPAC\]](#)
- Hans Joachim Marx: Händel und seine Zeitgenossen. Eine biographische Enzyklopädie. Laaber 2008. [\[Nachweis im OPAC\]](#)
- Migration und Identität. Wanderbewegungen und Kulturkontakte in der Musikgeschichte. Hg. von Sabine Ehrmann-Herfort. Kassel [u.a.] 2013 (= *Analecta musicologica* 49). [\[Nachweis im OPAC\]](#)
- Le musicien et ses voyages. Pratiques, réseaux et représentations. Hg. von Christian Meyer. Berlin 2003 (= *Musical Life in Europe 1600–1900. Circulation, Institutions, Representation*). [\[Nachweis im GBV\]](#)
- Musicisti europei a Venezia, Roma e Napoli (1650-1750). Hg. von Michela Berti / Gesa zur Nieden / Torsten Roeder. Berlin, Rom 2013. [\[online\]](#)
- Musiker auf Reisen: Beiträge zum Kulturtransfer im 18. und 19. Jahrhundert. Hg. von Christoph-Hellmut Mahling. Augsburg 2011. [\[Nachweis im OPAC\]](#)
- Musiker-Migration und Musik-Transfer zwischen Böhmen und Sachsen im 18. Jahrhundert. Bericht über das Internationale Symposium vom 7. bis 9. November 2008. Hg. von Hans-Günter Ottenberg / Reiner Zimmermann. Dresden [2012]. [\[online\]](#)
- Musik und Migration in Ostmitteleuropa. Hg. von Heike Müns. München 2005 (= *Schriften des Bundesinstituts für Kultur und Geschichte der Deutschen im Östlichen Europa* 23). [\[Nachweis im GBV\]](#)
- Berthold Over: »... sotto l'ombra della Regina de' Pennati«. Antonio Vivaldi, Kurfürstin Therese Kunigunde von Bayern und andere Wittelsbacher. In: *Italian Opera in Central Europe 1614–1780. Volume 3: Opera Subjects and European Relationships*. Hg. von Norbert Dubowy / Corinna Herr / Alina Żórawska-Witkowska in Zusammenarbeit mit Dorothea Schröder. Berlin 2007 (= *Musical Life in Europe 1600–1900. Circulation, Institutions, Representation*), S. 251–297. [\[Nachweis im OPAC\]](#)
- Samantha Owens: The Court of Württemberg-Stuttgart. In: *Music at German Courts, 1715–1760. Changing Artistic Priorities*. Hg. von Samantha Owens / Barbara M. Reul / Janice B. Stockigt. Woodbridge 2011, S. 165–195. [\[Nachweis im OPAC\]](#)
- Bärbel Pelker: Ein »Paradies der Tonkünstler«? Die Mannheimer Hofkapelle des Kurfürsten Carl Theodor. In: *Mannheim – Ein Paradies der Tonkünstler?* Hg. von Ludwig Finscher / Bärbel Pelker / Rüdiger Thomsen-Fürst. Frankfurt/Main [u.a.] 2002 (= *Quellen und Studien zur Geschichte der Mannheimer Hofkapelle* 8), S. 9–33. [\[Nachweis im GBV\]](#)
- Bärbel Pelker: Wendling. In: *Die Musik in Geschichte und Gegenwart. Allgemeine Enzyklopädie der Musik*. Hg. von Ludwig Finscher. Personenteil 17. Kassel [u.a.] 2007, Sp. 765–769. [\[Nachweis im GBV\]](#)
- Bärbel Pelker: The Palatine Court in Mannheim. In: *Music at German Courts, 1715–1760. Changing Artistic Priorities*. Hg. von Samantha Owens / Barbara M. Reul / Janice B. Stockigt. Woodbridge 2011, S. 131–162. [\[Nachweis im OPAC\]](#)
- Christoph Plutte: Archiv-Editor – Software for Personal Data. In: *Research and advanced technology for libraries*. Hg. von Stefan Gradmann. Heidelberg 2011 (= *Lecture Notes in Computer Science* 6966), S. 446–448. [\[Nachweis im GBV\]](#)

Probleme der Migration von Musik und Musikern in Europa im Zeitalter des Barock. Tagungsbericht der 15. Arolser Barock-Festspiele 2000. Hg. von Friedhelm Brusniak / Klaus-Peter Koch. Sinzig 2000 (= Arolser Beiträge zur Musikforschung 9). [\[Nachweis im GBV\]](#)

Siegbert Rampe / Léon Berben: Pez, Johann Christoph. In: Die Musik in Geschichte und Gegenwart. Allgemeine Enzyklopädie der Musik, Personenteil 13. Hg. von Ludwig Finscher. Kassel [u.a.] 2005, Sp. 454–456. [\[Nachweis im GBV\]](#)

Elizabeth Roche: Pez, Johann Christoph. In: The New Grove Dictionary of Music and Musicians. Hg. von Stanley Sadie. London 2001, Bd. 19, S. 533–534. [\[Nachweis im GBV\]](#)

Daniel Rosenberg / Anthony Grafton: Cartographies of Time. Princeton 2010. [\[Nachweis im OPAC\]](#)

Niels-Oliver Walkowski: Zur Problematik der Strukturierung und Abbildung von Personendaten in digitalen Systemen. Vortrag 2009. [\[online\]](#)

Andrea Zedler: Alle Glückseligkeit seiner Education dem allermildesten Ertz-Hause Oesterreich zu danken. Hofstaat, Bildung und musikalische Unterweisung des bayerischen Kurprinzen Karl Albrecht in Graz (1712–1715). In: Historisches Jahrbuch der Stadt Graz 42 (2012), S. 337–366. [\[Nachweis im GBV\]](#)

Abbildungslegenden und -nachweise

Abb. 1: Das Datenmodell des Personendaten-Repositoriums. Grafik: Torsten Roeder, 2014.

Abb. 2: Beispielaspekte zu Georg Friedrich Händel. In den dunkel schattierten Reitern steht die jeweilige Kategorie. Grafik: Torsten Roeder, 2014.

Abb. 3: Semantisches Markup in XML. Das komplexe, semantisch angereicherte XML-Format wird im Editor vereinfacht dargestellt. Grafik: Torsten Roeder, 2014.

Abb. 4: Abfragebeispiel aus der Musici-Datenbank (Auszug) mit den Ergebnissen zu »Ort = Rom, Zeit = 1725, Kategorie = Aufenthalt«. Der relevante Aspekt ist jeweils unter dem Musikernamen angezeigt. Berti / zur Nieden / Roeder 2013.

Abb. 5: Zeitliche Verteilung von allen Daten zu Aufhalten in Venedig in der Musici-Datenbank. Das Hover-Beispiel zeigt die Zahl der verfügbaren Daten auf dem Höchststand in den 1740ern. Berti / zur Nieden / Roeder 2013.

Abb. 6: Räumliche Verteilung aller Aspekte zu Venedig in der Musici-Datenbank. Das Hover-Beispiel (nördlichster Punkt) zeigt, dass es einen Datenbankeintrag zu Bergedorf gibt (der Geburtsort von Johann Adolf Hasse). Berti / zur Nieden / Roeder 2013.

Abb. 7: Parallele Visualisierung von Sukzessionen und Migrationsbewegungen. Grafik: Torsten Roeder, 2014.

Abb. 8: Lebenswege von Pez und Pellegrini im Vergleich. Grafik: Torsten Roeder, 2014.

Zeitschrift für digitale Geisteswissenschaften

Artikel aus:

Sonderband 1 der ZfdG: Grenzen und Möglichkeiten der Digital Humanities. Hg. von Constanze Baum und Thomas Stäcker. 2015. DOI: [10.17175/sb01](https://doi.org/10.17175/sb01)

Titel:

»Losing My Religion« – Einsatz der Videoannotationsdatenbank Pan.do/ra in der kunstgeschichtlichen Analyse von Musikvideos

Autor/in:

Thorsten Wübbena

Kontakt: twuebbena@dt-forum.org

Institution: Deutsches Forum für Kunstgeschichte Paris

GND: [123312396](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-23312396) ORCID: [0000-0001-8172-6097](https://orcid.org/0000-0001-8172-6097)

Autor/in:

Eric Decker

Kontakt: decker@asia-europe.uni-heidelberg.de

Institution: Universität Heidelberg

GND: [1084009560](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-1084009560) ORCID: [0000-0003-3035-2413](https://orcid.org/0000-0003-3035-2413)

Autor/in:

Matthias Arnold

Kontakt: arnold@asia-europe.uni-heidelberg.de

Institution: Universität Heidelberg

GND: [1049623088](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-1049623088) ORCID: [0000-0003-0876-6177](https://orcid.org/0000-0003-0876-6177)

DOI des Artikels:

[10.17175/sb001_018](https://doi.org/10.17175/sb001_018)

Nachweis im OPAC der Herzog August Bibliothek:

[830205977](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-830205977)

Erstveröffentlichung:

19.02.2015

Lizenz:

Sofern nicht anders angegeben 

Medienlizenzen:

Medienrechte liegen bei den Autoren

Letzte Überprüfung aller Verweise:

24.05.2016

GND-Verschlagwortung:

[Musikvideo](#) | [Computergestütztes Verfahren](#) | [Annotation](#) |

Zitierweise:

Thorsten Wübbena, Eric Decker, Matthias Arnold: »Losing My Religion« – Einsatz der Videoannotationsdatenbank Pan.do/ra in der kunstgeschichtlichen Analyse von Musikvideos. In: Grenzen und Möglichkeiten der Digital Humanities. Hg. von Constanze Baum / Thomas Stäcker. 2015 (= Sonderband der Zeitschrift für digitale Geisteswissenschaften, 1). PDF Format ohne Paginierung. Als text/html abrufbar unter DOI: [10.17175/sb001_018](https://doi.org/10.17175/sb001_018).

Thorsten Wübbena, Eric Decker, Matthias Arnold

»Losing My Religion« – Einsatz der Videoannotationsdatenbank Pan.do/ra in der kunstgeschichtlichen Analyse von Musikvideos

Abstracts

In dem vorliegenden Beitrag werden die Möglichkeiten sowie der daraus resultierende Mehrwert der Anwendung einer webbasierten Datenbank für Videoannotationen (Pan.do/ra) in der kulturwissenschaftlichen Untersuchung von Bewegtbildmaterial beleuchtet. Der erste Teil beschäftigt sich am Beispiel eines Musikvideos konkret mit einer Anzahl von Nutzungsmöglichkeiten, während der anschließende zweite Teil das verwendete System in einen allgemeineren Kontext einbettet.

In this paper, we discuss the possibilities as well as the added value of using a web-based databank for video annotation (Pan.do/ra) in cultural studies research on film. The first part demonstrates a variety of potential uses through the example of a music video, and the second part introduces the software on a more general level.

1.

1.1 Intro

Im ersten Teil des Beitrags¹ zur Nutzung der Videoannotationsplattform [Pan.do/ra – Open media archive](#)² im kulturwissenschaftlichen Kontext wird der Schwerpunkt auf den (mit diesem Werkzeug) zu bearbeitenden Bewegtbildinhalten – in diesem Fall Musikvideos – und den damit verbundenen Forschungsfragen liegen, während Matthias Arnold und Eric Decker von [HRA](#), der *Heidelberg Research Architecture*, im sich daran anschließenden zweiten Teil die dahinter liegenden technischen Aspekte und die globaleren Möglichkeiten darstellen und diskutieren werden. Zusammengefasst soll am Ende deutlich werden, dass durch die Verzahnung der Informationstechnologie mit der Fachwissenschaft ein Mehrwert erreicht werden kann, der durch analoge Möglichkeiten nicht hätte erzielt werden können – sozusagen ein geradezu idealtypisches Modell aus dem Gebiet der Digital Humanities.

1.2 Das Musikvideo als Forschungsgegenstand

¹ Autor dieses Teils ist Thorsten Wübbena.

² Pan.do/ra ist eine Online-Applikation, die es – kurz gesagt – mehreren Nutzern gleichzeitig erlaubt, frei gewählte Abschnitte eines Videos mit unterschiedlichen Arten von Annotationen (Layer oder Spuren) zu versehen. Durch diese Annotationen ist es möglich, sowohl einzelne Sequenzen als auch sämtliche Szenen zu einer Thematik direkt über eine URL zu referenzieren. Jede Spur enthält zeitreferenzierte Annotationen eines bestimmten Typs und neben Text können auch georeferenzierte Ortsinformationen und Bilder hinzugefügt werden. Mit seiner integrierten Rechteverwaltung erlaubt es das System, Filme gezielt bestimmten Nutzergruppen zugänglich zu machen oder Filme samt Annotationen frei zugänglich zu machen.

Zahlreiche Musikvideos weisen über ihren genuinen Zweck als Werbeträger hinaus einen ästhetischen Mehrwert auf, der eine eingehendere bildwissenschaftliche Betrachtung herausfordert und aufzeigt, dass sich in diesem Bereich eine eigene Kunstform entwickelt hat. Die künstlerischen Ausdrucksmöglichkeiten, die hier durch die flexible Verknüpfung von Text, Bild und Musik vorhanden sind, eröffnen Szenarien, in denen Text und Musik nicht nur einen kurzen Schwall von heterogenen Bildern zusammenhalten, vielmehr wird ein komplexer Diskurs auf drei Ebenen gestaltet. Das Musikvideo als visuelles Produkt referenziert dabei auch immer wieder auf »verwandte« Medien (Kino- und Fernsehfilm), aber auch auf Werke der bildenden Kunst, andere Kommunikations- und Unterhaltungsformen und auf sich selbst.

Und obgleich das Thema des Musikvideos inzwischen eine verstärkte Aufmerksamkeit innerhalb von Disziplinen wie der Film-, Musik- und Medienwissenschaft sowie der Kunstgeschichte erfahren hat,³ fehlt es in der wissenschaftlichen Bearbeitung des Themas immer noch an einer Auswahl adäquater Werkzeuge zur Unterstützung der Analyse des Untersuchungsgegenstandes und der Präsentation der Ergebnisse.⁴

In der Vergangenheit habe ich gemeinsam mit Henry Keazor zu zeigen versucht, dass eine fruchtbare Betrachtung des Musikvideos nur durch den Blick auf die drei oben erwähnten Ebenen sinnvoll ist.⁵ Ausgehend von dieser Vorgehensweise war das Ansinnen stets, die internen und externen Bezüge einzelner Clips herauszuarbeiten und zu analysieren, wobei bis dato zu bedauern war, dass nicht in stärkerem Maße im Medium selbst gearbeitet werden konnte, sondern jeweils schon bereits zu Beginn – in der Phase der Beobachtungen – auf die Übertragung in eine Textform zurückgegriffen werden musste. Text, Musik und Bild mussten dabei getrennt verhandelt werden und in der Kategorie »Bild« konnte nur mit Standbildern gearbeitet werden.

1.3 Losing My Religion

Am Beispiel des Clips zu R.E.M.'s *Losing My Religion*⁶, der 1991 vom Regisseur Tarsem Singh⁷ umgesetzt wurde, soll nun erläutert und zugleich dargestellt – bzw. in diesem Rahmen

³ Vgl. dazu die Bibliografien in Keazor / Wübbena 2010 sowie die Tatsache, dass auf dem Kunsthistorikertag 2013 in Greifswald dem Musikvideo eine eigene Sektion gewidmet wurde (die Vorträge der Sektion erscheinen in Keazor 2014).

⁴ Erwähnenswert ist an dieser Stelle einzig der Versuch, der an Universität Basel im Bereich der Medienwissenschaften zwischen April 2008 und Juni 2011 unternommen wurde, ein Analyseinstrumentarium zur Interpretation von Inszenierungsformen in Musikvideos zu entwickeln. Das Ergebnis *traVis* (Musikzentriertes Transkriptionsprogramm für audiovisuelle Medienprodukte) wurde im Februar 2013 an der Universität Basel im Rahmen eines Workshops vorgestellt und getestet. Hierbei erwiesen sich bei allen neuen Möglichkeiten schnell die Grenzen von *traVis*: Wie es die Beibezeichnung schon andeutet, ist das Instrument musikzentriert, d.h. die Parameter von Bild und Text kommen eindeutig zu kurz – in letzter Instanz werden sie der Musik doch wieder untergeordnet, wodurch die eigentlich anvisierte Gegenstandsangemessenheit verfehlt wird. Denn in Musikvideos kommen laufende Bilder, gesungener und vorgetragener Text gleichberechtigt zum Zug. Ein weiterer Nachteil ist zudem, dass das Programm dazu führt, dass dem temporalen Verlauf der Clips (bzw. genauer: der ihnen zugrundeliegenden Musik) zu stark gefolgt wird, so dass der Überblick über die Gesamtstruktur des Musikvideos verloren geht, was bei dem Workshop immer wieder deutlich wurde. Damit ist umso ersichtlicher, welches großes Desiderat auf diesem Gebiet weiterhin existiert.

⁵ Vgl. hierzu besonders Keazor / Wübbena 2011.

⁶ Von dem 1991 veröffentlichten Album *Out of Time*.

⁷ Nachdem der 1962 in Indien geborene zukünftige Regisseur ein im Himalaya gelegenes Internat besucht hatte, zog er in die Vereinigten Staaten, um an der Harvard Business School auf einen Master of Business

eher angedeutet – werden, wie stark eine Videoannotationssoftware wie das frei verfügbare Pan.do/ra als webbasiertes Werkzeug in der kunsthistorischen Medienanalyse unterstützend eingesetzt werden kann.⁸

Es ist im Nachhinein schwer zu sagen, inwieweit sich der Erfolg des Songs seiner Präsentation in Clip-Form verdankt, doch allein der Umstand, dass der Clip nicht nur im Erscheinungsjahr Aussichten auf einen Grammy hatte, sondern auch für gleich acht MTV Video Music Awards nominiert war (sechs der Trophäen erhielt das Video dann auch, darunter in den Kategorien »Video of the Year«, »Breakthrough Video« und »Best Direction in a Video«), zeigt, dass man den Anteil der visuellen Interpretation am Erfolg des Stücks keinesfalls unterschätzen darf – zumal das Video den Liedtext in eine Richtung hin interpretiert, die zwar von Texter und Leadsänger Michel Stipe möglicherweise mitgedacht (vgl. auch die Liedzeile »Choosing my confessions«), doch keinesfalls ausschließlich intendiert gewesen ist; und es ist eben diese Deutung, die dem Song den (weite Verbreitung sichernden) Stempel des Skandals und der Anrüchigkeit aufdrückt. Denn die Titelzeile »Losing my religion« ist als Hinweis darauf verstanden worden, dass der Protagonist des Textes seinen Glauben verliert. Tatsächlich aber hat Stipe damit eine vergessene Redewendung aus den Südstaaten der USA aufgegriffen, die soviel bedeutet wie »die Fassung zu verlieren«.⁹

Wie die vielen Hörer nach ihm (und diese dann möglicherweise bereits im Gefolge seiner visuellen Interpretation) konzentriert sich Regisseur Tarsem Singh bei der Konzeption seines Videos auf die mit den Reizworten »religion« und »confessions« angedeutete Sphäre des Glaubens und der Frömmigkeit. Und um diese sowie den im Songtitel angedeuteten Verlust derselben auszugestalten, legt er dem Clip eine Dramaturgie zugrunde, in der es letztendlich um die Durchdringung zweier Ebenen geht: Das ist zum einen die religiöse Sphäre, repräsentiert durch Szenerien, welche dem hinduistischen wie dem christlichen Bilderkosmos entlehnt sind und in letzterem Fall auf Werke der bildenden Kunst des 17. und 20. Jahrhunderts rekurrieren, namentlich auf Schöpfungen des Fotografenduos Pierre et Gilles und des Barockmalers Caravaggio. Ihnen gegenüber gestellt werden nach dem ersten Viertel des Videos und mit Beginn der zweiten Strophe die unter ihnen auf der Erde lebenden Menschen, die in ihrer fahlen Farbigkeit, den starken Hell-Dunkel-Kontrasten und den robusten, eher ärmlich und archaisch anmutenden Typen an Gemälde der Caravaggio-Schule erinnern, die hier gleichwohl nie detailgetreu umgesetzt werden, sondern lediglich einzelne Details liefern (wie z.B. den aus Caravaggios »Berufung des Heiligen Matthäus« bekannten, diagonal die Szene durchziehenden Lichtstrahl.

Administration hin zu studieren. Doch schon bald beschloss er, lieber das Fach »Film« zu belegen und schrieb sich daher am renommierten Art Center College of Design in Pasadena ein. Singh entwickelte sodann seinen oft als opulent und dynamisch bezeichneten visuellen Stil, so u.a. in dem Video zu Suzanne Vegas Song *Tired of Sleeping*. Seinen Durchbruch erzielte er jedoch mit dem hier vorgestellten Musikvideo.

⁸ Dass an dieser Stelle und im Folgenden kein Verweis auf den Videoclip (oder Teile davon) in der Videoannotationssoftware auf dem Server der HRA zu finden ist, liegt an der momentan vorherrschenden großen rechtlichen Unsicherheit in Bezug auf die Online-Bereitstellung von Bildmaterial (statisch oder bewegt), so dass der Serverbetreiber sich in diesem Fall keinem Risiko aussetzen wollte. Deutlich wird daran, dass trotz der bereitstehenden Optionen von freien Online-Werkzeugen eine Einschränkung der Nutzung für bestimmte Themen der Forschung besteht, da letztlich die Frage der Möglichkeit der Verfügbarmachung des Contents entscheidet.

⁹ Vgl. dazu Bruder 2004, S. 85.

1.4 Pan.do/ra

Wie sehen die Möglichkeiten des Umgangs mit dem Musikvideo nun in der Annotationssoftware aus (vgl. Abbildung 1)?¹⁰



Abb. 1: Startseite von Pan.do/ra. Screenshot, erstellt am 17.06.2015, © HRA, CC-Lizenz 0.

Da dieses Musikvideo eine mannigfaltige Anzahl von nachgehenswerten Referenzen beinhaltet, müssen an dieser Stelle einige wenige Beispiele ausreichen. Es soll dabei versucht werden, die Möglichkeiten der Software aus Anwendersicht dicht an den Inhalten zu beleuchten – bei diesem Vorgehen möge man die teils eher deskriptive Art verzeihen.¹¹

Der Leadsänger von R.E.M, Michael Stipe, erscheint von Anfang an immer wieder mit riesigen Engelsflügeln, was sich sodann leitmotivisch durch das ganze Video zieht (Abbildungen 2a-d).



Abb. 2a-d: Stills aus dem Clip *Losing My Religion* – Sänger Michael Stipe mit Flügeln und weitere Szenen (beispielhaft) mit dem Motiv der Engelsflügel in der Editor-Ansicht von Pan.do/ra, jeweils mit dem Tag »Engelsflügel« versehen (vgl. im rechten Bereich den Punkt »Keywords«). Quelle: HRA, Warner Bros., erstellt am 12.10.2014.

¹⁰ Auch an dieser Stelle noch einmal deutlich der Hinweis auf die unbefriedigende Lösung mit Filmstills arbeiten zu müssen. Ein Vorgehen, welches ja gerade durch den Einsatz der webbasierten Videoannotationssoftware überwunden werden sollte, aber hier wegen der rechtlichen Bedenken der Serverbetreiber notwendig ist.

¹¹ Mehr zur inhaltlichen Auseinandersetzung in Keazor / Wübbena 2011, passim.



Abb. 2b



Abb. 2c



Abb. 2d

Eine dieser Szenen stellt ein Gemälde des Barockmalers Caravaggio mit dem in seiner Verzückung von einem Engel gehaltenen Heiligen Franziskus nach, der hier im Clip allerdings durch einen zweiten, älteren Himmelsboten ersetzt wird (vgl. Abbildung 3 und Abbildung 4).¹²

¹² Open Source-Bilddatenbanksoftware der Firma [Coneda UG](#), einer Ausgründung des Kunstgeschichtlichen Instituts der Goethe-Universität Frankfurt. Der Link an dieser Stelle in Pan.do/ra führt in [die Installation des Frankfurter Kunstgeschichtlichen Instituts](#).

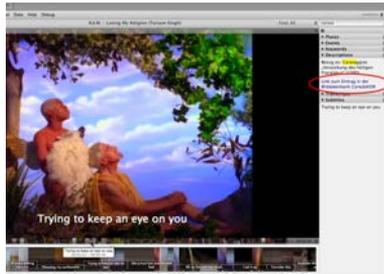


Abb. 3: Link aus Pan.do/ra zum Bilddatenbanksystem ConedaKOR. Quelle: HRA, KGI (Frankfurt/Main), Warner Bros., erstellt am 12.10.2014.

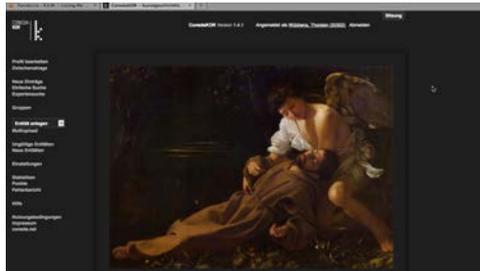


Abb. 4: Entitätsansicht in ConedaKOR: Michelangelo Merisi da Caravaggio, *Verzückung des heiligen Franziskus*, um 1595, Hartford, Wadsworth Atheneum Museum of Art. Quelle: KGI (Frankfurt/Main), [Web Gallery of Art](#), erstellt am 12.10.2014.

In dem Moment, wo der alte Engel versehentlich zur Erde herabstürzt und der andere Engel nach ihm zu greifen versucht, zitiert Singh ein weiteres Gemälde Caravaggios (*Die sieben Werke der Barmherzigkeit*), in dem ebenfalls prägnant das Motiv eines zur Erde hinabgereckten Engels-Armes erscheint (Abbildung 5).



Abb. 5: Interner Link bzw. Einbettung: Michelangelo Merisi da Caravaggio, *Die sieben Werke der Barmherzigkeit*, 1607, Neapel, Chiesa del Pio Monte della Misericordia. Quelle: HRA, KGI (Frankfurt/Main), [Web Gallery of Art](#), Warner Bros., erstellt am 12.10.2014.

Wenn der ältere, gefallene Engel auf der Erde sodann auf neugierige, argwöhnische und grausame Weise von den Menschen untersucht wird, spielt Singh wiederum auf ein Bild Caravaggios (*Der ungläubige Thomas*) an (Abbildung 6).



Abb. 6: Link bzw. Einbettung: Michelangelo Merisi da Caravaggio, *Der ungläubige Thomas*, 1603, Potsdam, Bildergalerie Schloss Sanssouci. Quelle: HRA, KGI (Frankfurt/Main), [Web Gallery of Art](#), Warner Bros., erstellt am 12.10.2014.

Die solcherart nachgestellten oder zitierten Gemälde werden dabei zuweilen in einer poppig-bunten und kitschig-schriellen Ästhetik präsentiert, die sich an den Werken eines französischen Künstlerpaars, des Fotografen Pierre (Commy) und des Malers Gilles (Blanchard)¹³ orientiert, deren *Heiliger Sebastian* dann auch in einer Einstellung des Clips zitiert wird (Abbildung 7).



Abb. 7: Einbettung: Pierre & Gilles, *St. Sebastian*, 1987. Quelle: HRA, KGI (Frankfurt/Main), [Art Forum](#), Warner Bros., erstellt am 12.10.2014.

Abgerundet wird das Personal des Videos durch Männer, die zum einen für die Beleuchtung zuständig zu sein scheinen, zugleich jedoch – als eine eiserne Flügel-Skulptur herstellende Schmiede – als Vertreter der (russisch-kommunistischen) Arbeiterklasse ausgewiesen werden, die dem Jenseitsbezug der Religion das erdverhaftete Vertrauen in die technischen Fähigkeiten des Menschen gegenüberstellen (Abbildung 8a und Abbildung 8b).

¹³ Vgl. zum Werk Marcadé / Cameron 1997.



Abb. 8a-b: Interner Bezug durch Tagging, verdeutlicht durch eine farbige Kennzeichnung auf der unteren Timeline (in diesem Fall »Arbeiter«). Quelle: HRA, Warner Bros., erstellt am 12.10.2014.



Abb. 8b

Der Bezug zu Russland wird ebenfalls in Form des Kostüms und der Choreographie von Stipe hergestellt: Sie sind beide an Aufnahmen vom Auftritt des berühmten Tänzers Vaslav Nijinsky in dem von Claude Debussy geschriebenen Ballett *Jeux* aus dem Jahre 1913 angelehnt (Abbildung 9a und Abbildung 9b).



Abb. 9a-b: Einbettung: Abbildungen des Tanzes von V. Nijinsky zu *Jeux*. Quelle: HRA, KGI (Frankfurt/Main), Warner Bros., erstellt am 12.10.2014.



Abb. 9b

Allein die hier aufgezeigten Bezüge (die z.B. noch um den Verweis auf die 1968 erschienene Kurzgeschichte *Ein sehr alter Mann mit riesengroßen Flügeln* des kolumbianischen Schriftstellers Gabriel García Márquez zu ergänzen wäre, aus der das Motiv des zur Erde gestürzten und dort von Menschen misshandelten Engels stammt) machen deutlich, welche Vielfalt an Verweisen und Zitaten in einen Clip wie demjenigen Singhs eingewoben sein kann.

Im Jahr 2000 hat Tarsem Singh diese im Musikvideo entwickelte Bilderwelt dann auch in seinem Kinodebüt *The Cell* unterzubringen versucht, und so begegnet man dort den bereits bekannten Motiven und Vorbildern, die zuweilen sogar zu identisch anmutenden Szenarien zusammengeführt werden – fast scheint es, als seien einzelne Szenen in der Kulisse des R.E.M.-Videos gedreht worden (vgl. Abbildung 10a und Abbildung 10b).

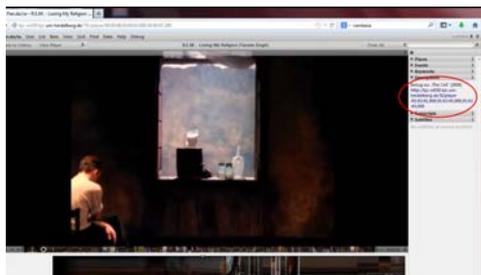


Abb. 10a-b: Interner Link zur korrespondierenden Filmszene aus *The Cell*, 2000 (Director: Tarsem Singh).
Quelle: HRA, KGI (Frankfurt/Main), Warner Bros., New Line Cinema (theatrical), Warner Home Video (DVD), erstellt am 12.10.2014.



Abb. 10b

1.5 Vorteile

Wo liegen aber nun die Vorteile in der Bearbeitung des Materials mit einer – und im Speziellen – mit dieser Videoannotationssoftware?

Für webbasierte Systeme gilt allgemein – nach wie vor erwähnenswert – die zeit- und ortsunabhängige Verfügbarkeit der Inhalte. Für die kollaborative Bearbeitung bei verteilten Forscher- und Studierendengruppen ist dies ein basaler, aber wichtiger Faktor.

Wenn ein Clip mit der Annotationssoftware bearbeitet ist (was aber natürlich nie als »abgeschlossen« betrachtet werden kann) und er für eine spätere Weiterbearbeitung unter einem neuen Aspekt angesehen wird, dann sind die Referenzen an der entsprechenden Stelle im Clip zu finden, ohne dass die entscheidenden Passagen wieder aus einem Text zusammengesucht werden müssen.

Dadurch ist auch für Nichtkenner ein exploratives Sehen des Clips möglich, da sie an den entscheidenden Punkten auf die Bezüge hingewiesen werden, nach denen sie nicht recherchieren konnten, weil sie schlicht nichts davon wussten.

Daneben ist sicherlich aber das Arbeiten im Clip selbst der wichtigste Punkt wissenschaftlicher Auseinandersetzung. Die Bezüge zu bewegten Einheiten, zu Filmausschnitten und nicht nur zu Stills setzen zu können, entspricht dem Medium und bietet die Möglichkeit, sozusagen »in Film zu denken« und zu arbeiten.

1.6 Things-To-Do

Bei der Beschäftigung mit Pan.do/ra konnten eine Reihe von Punkten identifiziert werden, um die das System noch erweitert werden muss, um Musikvideos in ihrem medienübergreifenden Kontext befriedigend analysieren zu können.

Als Beispiele seien hier nur die folgenden genannt:

1. Die direkte Gegenüberstellung mehrerer unterschiedlicher Videos.
2. Die Anpassbarkeit der Benutzeroberfläche an unterschiedliche Analysesituationen.
3. Die Möglichkeit der Annotation reiner Audiodaten ohne Bild.

Wie eingangs gezeigt, stellt sich dieser Forschungsbereich durch seine enge Verbindung von Bild-, Text- und Musik/Audio-Schichten bisher im Hinblick auf Informationsverwaltung und Visualisierung des komplexen Bezugssystems als Problemfall dar, und es erweist sich als Glücksfall, dass im Rahmen des Heidelberger DFG-Projekts **Zur ästhetischen Umsetzung von Musikvideos im Kontext von Handhelds** (Laufzeit: 2011–13) ein möglicher Weg aufgezeigt werden konnte, gemeinsam mit den Kollegen vom Heidelberg Research Architecture an ähnlichen Fragestellungen zu arbeiten.

2.

2.1 Digital Humanities am Heidelberger Exzellenzcluster

Die Heidelberg Research Architecture, kurz HRA, ist die Digital Humanities Unit am Exzellenzcluster **Asien und Europa im Globalen Kontext** der Universität Heidelberg. Zum Kerngeschäft der HRA gehört die Entwicklung des modularen Metadatenframeworks **Tamboti**. Es ermöglicht das gemeinsame Sammeln, Organisieren und Annotieren von digitalisiertem Forschungsmaterial (Text, Bild, Video, Audio). Das System erlaubt das Teilen der Daten und Metadaten innerhalb von Gruppen, von Forschungsgruppen bis hin zu einer weltweiten Veröffentlichung. Dabei setzt die HRA auf eine Kombination hochdeskriptiver internationaler Metadatenstandards (derzeit MODS, TEI und VRA Core) mit einem hohen Anteil kontrollierter multilingualer Vokabulare (wie die Getty Thesauri oder das Virtual International Authority File) und einer modularen Systemarchitektur, die das nachhaltige Erfassen und Annotieren multilingualer Metadaten nach wissenschaftlichen Standards erlaubt.

Tamboti ist für eine Kooperation auf Universitätsebene und darüber hinaus ausgelegt, wobei Kooperationen im Sinne einer gemeinsam nutzbaren Infrastruktur verstanden wird. Mit Tamboti entwickelt die HRA eine auf dem XML-Datenbanksystem **eXist-db** basierende Open Source-Lösung. Tamboti ist als Application für eXist-db weltweit frei nutz- und erweiterbar. Die HRA arbeitet derzeit an zusätzlichen Applikationen (Bildmodul Ziziphus und das Wiki-Modul), die Tamboti um weitere Funktionsbereiche modular erweitern.

Eine größere Nutzergruppe im Exzellenzcluster setzt in der Forschung auf die Arbeit an und mit visuellen Medien, daher liegt hier ein Schwerpunkt der Entwicklungen der HRA. Dabei wurde bereits während der ersten Förderphase des Exzellenzclusters auf die Kombination vorhandener Softwarelösungen mit Eigenentwicklungen gesetzt.

2.2 Videoannotationsdatenbank Pan.do/ra

Im Bereich der kollaborativen Videoannotation hat die HRA bereits in der ersten Förderphase des Exzellenzclusters (2007–2012) in mehreren Projekten Erfahrungen sammeln können. Nach einer Bedarfsanalyse und Tests unterschiedlicher frei zugänglicher Softwarelösungen zur Videoannotation fiel die Wahl auf **pad.ma**, die von einem freien Entwicklerteam¹⁴ als Open Source-Plattform entwickelt wird und jetzt in der überarbeiteten

¹⁴ Die Pan.do/ra-Software wurde ursprünglich im Rahmen des Pad.ma (Public Access Digital Media Archive) Projekts federführend von Jan Gerber, Sebastian Lütgert und Sanjay Bhangar entwickelt. Pad.ma wurde von einer Gruppe von Organisationen aus Indien und Deutschland ins Leben gerufen: CAMP (Mumbai), Alternative Law Forum (Bangalore) und 0x2620 (Berlin). Des Weiteren waren Majlis und Point of View aus Mumbai mit beteiligt. Ausgestattet mit einer Anschubfinanzierung des niederländischen Humanistisch Instituut voor Ontwikkelingssamenwerking (HIVOS) konnte eine erste Version der Videoannotationsdatenbank entwickelt werden. Der weitere Ausbau wird seit 2010 durch die Unterstützung der Bohn Foundation und Foundation for Arts Initiatives ermöglicht. Vgl. [About Pad.ma](#), [About 0xDB](#) und die [faqs](#) der Pan.do/ra Webseite.

Pan.do/ra-Version in Heidelberg installiert ist.¹⁵ Eine Reihe anderer Projekte setzen auch Pan.do/ra ein, darunter 0xDB, Pad.ma und Indiancine.ma.¹⁶

Ausschlaggebend für die Wahl von pad.ma war, dass sie es damals als einzige uns bekannte Software erlaubte, beliebige Sequenzen innerhalb eines Videos mit mehreren Nutzern gleichzeitig zu annotieren. Zudem verfügt sie über eine ausreichende Nutzer- und Rechteverwaltung und kann mit ihrer einfachen Handhabung mehrerer unterschiedlicher Annotationsspuren überzeugen. Der Funktionsumfang ist in der Zwischenzeit, auch vorangetrieben durch gezielte Investitionen der HRA, deutlich gewachsen und in die ebenfalls als Open Source-Software veröffentlichte Nachfolgeversion Pan.do/ra eingeflossen. So wurden beispielsweise eine LDAP-Authentifizierung integriert sowie mehrsprachige Audio- und Annotationsspuren implementiert. Zeitbasierte Annotationen können im .srt und JSON Format im- und exportiert werden. Das erlaubt es zum Beispiel, Audiotranskriptionen in Tools wie [Speedtrans](#)¹⁷ anzufertigen und dann in Pan.do/ra zu importieren.

Für Annotationen stehen in Pan.do/ra unterschiedliche Spuren oder Layer zur Verfügung. So können beispielsweise Annotationen vom Typ ›Subtitle‹ direkt in den Film eingeblendet oder Annotationen vom Typ ›Location‹ auf einem Karteninterface dargestellt werden. Es wird festgehalten, welcher Nutzer oder welche Nutzerin wann welche Annotationen hinzugefügt hat und es ist möglich, die Anzeige von Annotationen nach Autor bzw. Autorin zu filtern – eine Funktion, die sich nicht nur während des Einsatzes im Unterricht bewährt hat.

In den Annotationen ist es grundsätzlich möglich Markup zu verwenden. Die bereits unterstützten HTML5-Tags erlauben Textformatierungen sowie das Einfügen von Hyperlinks und eröffnen auch die interessante Option, Bilder direkt in Annotationen einzubinden (vgl. die Beispiele im ersten Teil). Da sämtliche Annotationen mit URLs versehen sind, lassen sich bequem Szenen in einem oder mehreren Filmen verlinken. Ebenso können Verweise auf einzelne Annotationen von externen Websites gesetzt werden. Dank der Möglichkeit, auch Suchen über eine URL auszulösen, können Anwender beispielsweise sämtliche Stellen im Film aufrufen, die mit demselben Keyword annotiert wurden.¹⁸

Mit dem im Frontend für jede Annotation bereitgestellten *embed code* können darüber hinaus ohne großen Aufwand Filme oder ›Filmschnipsel‹ in andere Webseiten eingebunden werden. Am Exzellenzcluster wird diese Funktion beispielsweise im Open Access e-Journal¹⁹ bereits eingesetzt. Ein derzeit noch eher ungewohnt anmutender Einsatzbereich ist das Einbetten von Videos in Annotationen. So erlaubt der HyperImage Nachfolger Yenda die Nutzung von HTML5-Tags an verschiedenen Stellen. Damit können in Bildannotationen nicht nur Texte und zusätzliche Bilder, sondern auch Videos aus Pan.do/ra eingebettet werden.

¹⁵ [HRA Videoannotationsdatenbank](#).

¹⁶ Vgl. [Demo auf Pan.do/ra](#).

¹⁷ Eine Firefox Erweiterung, die für den Einsatz mit pad.ma entwickelt wurde.

¹⁸ [Beispiel](#).

¹⁹ Vgl. [Transcultural Studies](#).

Die Nutzung von *embed codes* ist nur eine Möglichkeit aus anderen Webseiten heraus auf Daten in Pan.do/ra zuzugreifen. Per API stellt Pan.do/ra eine ganze Reihe von Funktionen für die Anbindung an andere Webapplikationen zur Verfügung.

2.3 Einsatz am Exzellenzcluster Asien und Europa – Global Politics on Screen

Im akademischen Bereich wird die Software am Exzellenzcluster Asien und Europa seit 2010 in unterschiedlichen Projekten und Fachbereichen in Forschung und Lehre eingesetzt. Zusätzlich zu dem in diesem Artikel ausführlich beschriebenen Projekt sei noch das öffentlich zugängliche Global Politics on Screen-Projekt von Madeleine Herren-Oesch et al. erwähnt.²⁰ In diesem Projekt wurde von einer Gruppe von Studierenden und Wissenschaftlern des Historischen Seminars und des Exzellenzclusters Filmmaterial zu einer Mission des Völkerbunds untersucht. Erst während des Projekts wurde immer klarer, dass es sich dabei um eine frühe Version eines japanischen Propagandafilms handelt. Der von der South Manchurian Railway Company (SMR) zu Beginn der 1930er produzierte Film begleitet die sogenannte Lord Lytton Mission des Völkerbundes bei ihrer Untersuchung des Einmarsches der japanischen Streitkräfte in die Mandschurei. Das dabei entstandene Filmmaterial war für die Vorführung in Botschaftskreisen in Europa gedacht. Der Stummfilm wurde für die Vorführung mit Texttafeln in unterschiedlichen Landessprachen aufbereitet.

Die im Rahmen des Heidelberger Projekts vorgenommene Untersuchung des Films konzentriert sich auf eine vermutlich unveröffentlichte Version mit englischen Texttafeln, die eine frühe Fassung des Filmmaterials mit dem Charakter eines Rohschnitts zeigt. Bei der ersten Durchsicht des Filmmaterials waren noch vielen Fragen offen. Es war lediglich klar, dass es sich um ein filmisches Dokument zur Lord Lytton Mission handelte. Sowohl Auftraggeber als auch das anvisierte Zielpublikum waren zunächst nicht bekannt. Mit Hilfe von Pan.do/ra konnte die HRA dem Projekt ein Werkzeug an die Hand geben, diesen und anderen offenen Fragen nachzugehen.

In der ersten Arbeitsphase wurde Pan.do/ra genutzt, um subjektive Auffälligkeiten und Besonderheiten im Film zu finden, zu markieren und zu kommentieren. In den nächsten Phasen hat sich dann das Team Schritt für Schritt dem Film immer stärker und systematisch angenähert. So wurden sämtliche englische Texttafeln transkribiert und dann nach Querverweisen zu Dokumenten in Archiven gesucht. Im [Japan Center for Asian Historical Records](#) wurden mehrere Dokumente mit unterschiedlichen Sprachversionen und in unterschiedlicher Reihenfolge der Texttafeln gefunden.

Da es sich um einen Stummfilm handelt, wurde sämtlichem geschriebenen Text im Film schnell dieselbe Aufmerksamkeit zu Teil wie den Texttafeln. Im Film gezeigte Schriftzüge auf Bannern, Hinweisschildern oder ähnlichem wurden transkribiert und wo nötig ins Englische übersetzt. Dabei hat es sich als sehr vorteilhaft herausgestellt, dass Übersetzungen sofort bei

²⁰ [Global Politics on Screen](#): A Japanese Film on the Lytton Commission in 1932.

Eingabe in die Datenbank für alle Projektmitglieder zur Verfügung standen. Die Banner, die an vielen Stellen im Film den Weg der Kommission begleiten, wirken im Film oft wie Untertitel, die sich je nach verwendeter Sprache an ein unterschiedliches Publikum wenden. Oft stellt sich die Frage, ob die Kommission, das Publikum vor Ort oder das Publikum vor der Leinwand angesprochen werden soll.

Durch das Transkribieren wurde den Projektteilnehmern, die sonst wenig mit Stummfilmen in Berührung gekommen waren, der Stellenwert von geschriebenem Text im Stummfilm deutlich vor Augen geführt. Hinweise, die man ohne die gezielte Verschriftlichung mit den heutigen Sehgewohnheiten vielleicht einfach übersehen hätte, wurden so relevant. Als Beispiel sei hier eine **Szene in einem Hafen** (vermutlich Dairen) angeführt. In einem Kameraschwenk wird gezeigt, wie ein Schiff entladen wird. Dabei ist beiläufig für ca. 3 Sekunden ein Schild, das den Hersteller des Krans identifiziert, zu sehen. Zeit genug, um für einen zeitgenössischen Betrachter aus europäischen Diplomatenkreisen zu bemerken, dass es sich um einen Kran der Demag AG aus Duisburg handelt. Der Rückschluss, dass Japan mit der Besetzung der Mandschurei auch für Sicherheit und Stabilität von Großkunden deutscher Unternehmen sorgt, bleibt den informierten Betrachtern überlassen.

Im weiteren Verlauf des Projekts wurden noch zusätzliche Annotationsebenen verwendet. Mit den Annotationen vom Typ »Location« konnten die gezeigten Schauplätze im Film nachgezeichnet werden und mit **Karten aus dem Kommissionsreport** verglichen werden (vgl. Abbildung 11).



Abb. 11: Gegenüberstellung eines Kartenausschnitts aus dem Kommissionsbericht und eines Screenshots der Location Annotationsspur in Pan.do/ra. Quelle: Route Maps: Itineraries of the Commission in the Far East (League of Nations (ed.): Appeal by the Chinese Government. Report of the Commission of Enquiry, Appendix, Maps No. 13 and 14, Geneva: League of Nations, 1932).

Zur thematischen Aufarbeitung wurden häufig wiederkehrende Argumentationsmuster und Themen des Films mit Keywords versehen. Eines der übergeordneten Ziele des Filmes ist es beispielsweise, den Einmarsch japanischer Truppen in die Mandschurei zu legitimieren. Ein Legitimationsgrund, der immer wieder aufgeführt wird, ist die Zerstörung wichtiger Eisenbahnbrücken durch Anschläge sogenannter »Soldier-Bandits«. Heute geht man davon aus, dass die Anschläge inszeniert waren, um einen Einmarsch vor der Weltöffentlichkeit zu rechtfertigen.

Ebenfalls mit Keywords versehen wurden die Auftritte bestimmter Kommissionsmitglieder im Film. Dabei wurden lediglich die Auftritte von Lord Lytton und des deutschen Abgesandte

Heinrich Schnee über die gesamte Länge des Films transkribiert. Das manuelle Suchen und Markieren hat sich jedoch als sehr zeitaufwändig und nur bedingt aussagekräftig herausgestellt. Um tatsächlich sinnvolle und quantitativ belegbare Aussagen darüber machen zu können, wer wann wie oft und mit wem zusammen gezeigt wird, wäre eine komplette Erfassung aller Gruppenmitglieder nötig. Der dazu nötige Aufwand stand im Rahmen des Projektbudgets aber in keinem vernünftigen Verhältnis zum möglichen Erkenntnisgewinn. Hier wäre eine (semi-)automatisierte Lösung wünschenswert.

2.4 Recherchen – Verfügbare Videoannotationssysteme

Im November 2013 und April 2014 wurde eine größere, ergebnisoffene Recherche der aktuell verfügbaren Anwendungen im Bereich Videodatenbanken und -annotation durchgeführt. Ziel war es, einen genaueren Überblick über den Stand der Entwicklungen auf diesem Gebiet zu erhalten. Darüber hinaus sollte auch geprüft werden, ob neuere oder eventuell für den Einsatz am Exzellenzcluster besser geeignete Lösungen entwickelt wurden, um gegebenenfalls darauf umzusteigen. Die Recherche gestaltete sich schwieriger als erwartet. Die Informationen waren weit verstreut und nur wenige Zusammenstellungen von Tools für Videoanalyse oder Videoannotation bis dato veröffentlicht (vgl. Liste von Video-Annotationssystemen).

Vielversprechend schienen die Einträge in der »Digital Research Tools« (DIRT) Datenbank des Project Bamboo, doch waren sie leider nur bedingt hilfreich. Zum einen waren viele Einträge veraltet und die Liste sehr unvollständig. Auch war es zwar möglich, die Einträge nach Tags zu browsen, doch schon eine kombinierte Suche beispielsweise nach »video« und »annotation« war nicht möglich. Die 24 Einträge beim Tag »video« waren sehr divers und enthielten selbst Bilddatenbanken (Flickr) oder Screen recording Tools (Screenr). Nur 6 Einträge in »video« waren auch mit »annotation« getagged (Advene, Annotator's Workbench, ELAN, Project Pad, Vertov, VideoANT). Darüber hinaus waren die Einträge zwar annotiert, die Inhalte aber meist den Selbstdarstellungen der Webseiten entnommen.²¹

Um die Verlässlichkeit der Informationen und deren Potential als Entscheidungshilfe auf der Suche nach einer geeigneten Software war es auf der »Video«-Seite der »Annotations at Harvard« besser bestellt. Die dort veröffentlichte Liste war zwar mit 18 Einträgen auch überschaubar und die Inhalte stammen aus dem Jahr 2010, aber die Zusammenstellung war inhaltlich deutlich fokussierter und die Anmerkungen oft nicht einfach nur wörtlich von der Website kopiert.²²

Interessant war auch festzustellen, dass das Thema Annotation auch bei Online Guides ausgelassen wurde, so beispielsweise zu beobachten beim ansonsten vorbildlich zusammengestellten »Activists Guide to Archiving Video«.²³

²¹ Im Internet Archive kann noch ein [Snapshot vom 2. Oktober 2013](#) eingesehen werden.

²² [Annotations at Harvard – Video Annotations](#). Harvard Annotation Project, 2010.

²³ [Activists Guide to Archiving Video](#), Version 1.0 (Juli 2013).

Auflistungen in den Buchpublikationen waren hingegen inhaltlich sehr eng gefasst und beschränkten sich auf wenige Tools aus den Bereichen qualitativer Datenanalyse (QDA) bzw. Broadcast Video. Auch lag der Zeitpunkt der Zusammenstellung meist mehrere Jahre vor dem Publikationsdatum. Die folgende Liste zeigt eine Auswahl veröffentlichter Zusammenstellungen von Video-Annotationssystemen:

- Jost et al. 2013: Christofer Jost et al.: *Computergestützte Analyse von audiovisuellen Medienprodukten* (= Qualitative Sozialforschung 22). Springer Fachmedien: Wiesbaden 2013. [\[online\]](#)
- Linked TV: Evlampios Apostolidis et al.: *State of The Art and Requirements Analysis for Hypervideo*. LinkedTV - Television Linked To The Web. 30. September 2012. [\[online\]](#)
- Bamboo: Bamboo DiRT. Digital Research Tools. [2013]. [\[online\]](#)
- Dasiopoulou et al. 2011: Stamatia Dasiopoulou et al.: »A Survey of Semantic Image and Video Annotation Tools.« In: *Knowledge-Driven Multimedia Information Extraction and Ontology Evolution*. Edited by Georgios Paliouras / Constantine D. Spyropoulos / George Tsatsaronis (= Lecture Notes in Computer Science 6050.) Springer: Berlin and Heidelberg 2011, S. 196–239. [\[online\]](#)
- Harvard: Annotations at Harvard – Video. Video Annotation Tools. [2010]. [\[online\]](#)
- UCLA-KB: UCLA Knowledgebase – Video annotation tools. [2009]. [\[online\]](#)

Neben der Verstreutheit der Informationen wurde während der Recherchen sehr schnell deutlich, dass Analyse und Annotation von Videos in sehr vielen Bereichen und mit sehr unterschiedlichen Zielsetzungen eingesetzt werden. So gibt es Anwendungen²⁴, die besonders für linguistische Analysen konzipiert sind, beispielsweise das vom Max Planck Institute for Psycholinguistics in Nijmegen entwickelte ELAN. Dann gibt es Anwendungen, die sich auf die in großen Fernsehanstalten anfallenden bzw. dort archivierten Materialien spezialisieren, wie beispielsweise die Entwicklungen innerhalb des FP7 Projektes LinkedTV. Wieder andere Systeme konzentrieren sich auf den Bereich E-Learning oder sind als Streaming-Plattformen oder Archivportale konzipiert. Darüber hinaus gibt es Lösungen die besonders die Musikanalyse von Filmen unterstützen, wie das bereits erwähnte, in Basel entwickelte trAVIS²⁵, oder sich auf Performances spezialisieren, wie der innerhalb des European Collected Library of Performing Arts (ECLAP) entwickelte MyStoryPlayer. Noch andere Lösungen zielen auf die Erstellung interaktiver Narrationen ab, sind Teil von Softwarepaketen zur qualitativen Datenanalyse, werden in der Lehre eingesetzt oder dienen der Erstellung von Untertiteln. Die folgende Auflistung zeigt Beispiele für Videoanalyse- und -annotationslösungen in verschiedenen Anwendungsbereichen.²⁶

- Linguistische Analyse: ELAN, ANVIL, EXMARaLDA
- Medienanstalten / Archive: LinkedTV, MediaGlobe
- E-Learning / Vortragsportale: Catool, AAV/OpenCast Matterhorn, Alexander Street
- Streaming Plattformen / Archivportale: Kaltura, AV Portal

²⁴ Die Links zu den hier aufgeführten Projekten und Softwarelösungen finden sich zusammen mit weiterführenden Informationen in einem vom Autor zusammengestellten frei verfügbaren Google Docs-Dokument [Overview: Video Annotation Tools](#).

²⁵Vgl. Abschnitt 1.2, Fußnote 4.

²⁶ Vgl. [Overview: Video Annotation Tools](#).

- Musik / Performance: trAVis, MyStoryPlayer
- Video: Mediathread, Pan.do/ra
- Interactive storytelling: Klynt, HyperImage
- Subtitling: Speedtrans, DotSub
- Service based analysis: Semex, Videana

Alle Tools bieten im weiteren Sinne Annotationen an. Viele fanden dabei auch elegante - aber eben doch bereichsspezifische Lösungen. Keins deckt alle Bereiche ab.

Es schien daher sinnvoll, die Ergebnisse der eigenen Recherchen zu systematisieren und einer breiteren Community zur Einsicht und Ergänzung zur Verfügung zu stellen. Darüber hinaus sollten in der Zusammenstellung auch inhaltliche Kriterien erfasst werden, die es Lesern erleichtern, auf der Suche nach einer für sie geeigneten Software eigene Entscheidungen zu treffen. Dazu wurde auch erfasst, ob ein Tool noch aktiv entwickelt wird, ob der Sourcecode veröffentlicht wurde und ob es frei benutzbar ist. Weiterhin war für uns relevant, ob es sich um ein online Tool handelt, es kollaborativ nutzbar ist und ob es potentiell in andere Datenbanken integriert werden kann. Es wurden auch (sehr grobe) Kategorien vergeben, um die jeweilige »Zielgruppe« festzuhalten, beispielsweise »filmstudies«, »linguistic« oder »video platform«. Schließlich wurde auch erfasst, welche Personen oder Institutionen die Entwickler sind und ein kurzer Beschreibungstext aus den online Informationen entnommen. In den etwa fünf Monaten unterschiedlich intensiver Recherchen kamen so knapp 60 Einträge zusammen. Das Dokument ist auf Google Docs veröffentlicht und kann von jedem Nutzer eingesehen, ergänzt und bearbeitet werden.²⁷

2.5 Anforderungen an ein zukünftiges Videoannotationssystem

Bei den Recherchen kamen auch Ressourcen zum Vorschein, in denen Anforderungen für Videoannotationssysteme formuliert werden. Hier sind insbesondere die im Rahmen des Harvard Annotation Projektes zusammengestellten »Requirements of a Video Annotation System«²⁸ zu erwähnen (Abbildung 11), die auch den Interessen und Zielen am Heidelberger Exzellenzcluster Asien und Europa sehr nahe kommen, aber in der Praxis in dieser Art noch in keiner Software verfügbar sind.

²⁷ [Link zum Dokument.](#)

²⁸ [Requirements of a Video Annotation System](#), Harvard Annotation Project, 2010.

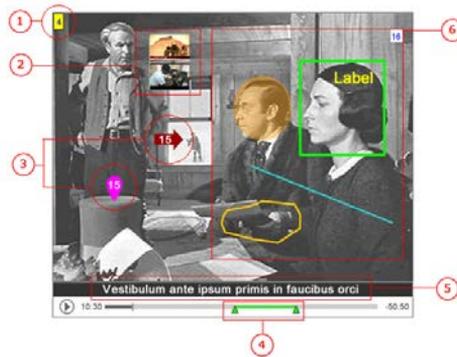


Abb. 12: **Requirements of a Video Annotation System**. Quelle: Harvard Annotation Project, 2010. Mit freundlicher Genehmigung.

Ein solches Annotationssystem sollte lokal oder online verfügbar sein und auf offene Datenstandards setzen. Dabei sollte das Interface erlauben, Videos insgesamt oder in Teilen zu annotieren und Bilder oder andere Videos neben der Visualisierung der Soundspur als Overlays möglich sowie gesprochener Text automatisch als ›baseline annotation‹ verfügbar sein. Außerdem sollten Marker im Bild gesetzt werden können und ihr Aussehen anpassbar sein, wobei einfache geometrische Flächen, Linien und Polygone ebenso erwünscht wären wie kurze, in Farbe, Füllung, Strichstärke usw. anpassbare und möglichst mit der Zoomstufe mitskalierende Text-Tags.²⁹

Viele dieser Anforderungen stimmen mit den Wünschen der Nutzer am Exzellenzcluster Asien und Europa überein. Während die Videoannotationen durch die Funktionalitäten in Pan.do/ra gut abgedeckt sind, ist besonders spannend, wie nah viele der weiterführenden Punkte den bereits im Moment verfügbaren Features der ebenfalls im Cluster eingesetzten virtuellen Forschungs- und Publikationsumgebung HyperImage und dessen durch semantische Annotationen erweiterten Nachfolgers Yenda kommen.³⁰ Zwar sind diese Tools im Moment noch auf Still Images beschränkt, aber die Einbeziehung von Bewegtbildern in die vielfältigen Annotationsmöglichkeiten ist möglich und könnte gemeinsam mit interessierten Kooperationspartnern entwickelt werden.

Ein Bericht über den Weg der Technischen Informationsbibliothek (TIB) von einer Bedarfsanalyse hin zum im Frühjahr 2014 gemeinsam mit dem Hasso-Plattner-Institut für Softwaresystemtechnik an der Universität Potsdam (HPI) eröffneten AV-Portal, einer webbasierten Plattform für audiovisuelle Medien, wurde von Margret Plank auf der LIBER Konferenz 2012 vorgestellt und auf dem LIBER-Blog veröffentlicht.³¹

²⁹ Vgl. **Requirements of a Video Annotation System**, Harvard Annotation Project, 2010.

³⁰ Offizielle Website des Open-Source-Projekts **HyperImage – Eine virtuelle Forschungs- und Publikationsumgebung für den Bilddiskurs**, Yenda – **Picture Knowledge**. Open-Source semantische virtuelle Forschungsumgebung der bitGilde IT Solutions UG.

³¹ Plank 2012 (Der Blogbeitrag war leider im Oktober 2014 online nicht mehr erreichbar.) Vgl. auch Plank / Weichert 2014.

Das AV-Portal ist allerdings nicht für den Einsatz im Forschungsalltag von Geisteswissenschaftlern konzipiert, sondern »optimiert den Zugang zu und die Nutzung von wissenschaftlichen Filmen aus Technik und Naturwissenschaften, wie z.B. Computeranimationen, Vorlesungs- und Konferenzaufzeichnungen.«³² Die in diesen Service integrierten Features wie automatische Sprach- und Texterkennung sowie Anreicherung mit semantischen Informationen sind jedoch auch für die Nutzung über die Naturwissenschaften hinaus spannend und überaus vielversprechend.

2.6 Einbindung von Videoannotationen in Tamboti

Um die in der Datenbank enthaltenen Annotationen der Wissenschaftler über das zentrale Suchinterface in Tamboti verfügbar zu machen wurde als »Proof-of-concept« bereits eine Anbindung der Metadaten implementiert. Dabei werden die Metadaten aus der entsprechenden .srt-Spur als <relatedItem> in den MODS Eintrag des Videos integriert. Der Inhalt der Annotation wird im Subelement <note> und die Zeitspanne in <part> ausgegeben. Dies ist eine Lösung, die für Objektmetadaten inklusive Untertiteln noch sinnvoll erscheint, bei gleichzeitigem mapping von Nutzerannotationen nach MODS aber zu einer Überfrachtung des Eintrags führt. Hier ist in Zukunft eine Trennung von Objektmetadaten (MODS) und Annotation (OADM)³³ angestrebt.

```
<relatedItem type="constituent">
  <titleInfo>
    <title supplied="yes">constituent</title>
  </titleInfo>
  <subject>
    <topic>[TOPIC1, TOPIC2, ...]</topic>
  </subject>
  <place>
    <placeTerm>[NAME-OF-LOCATION]</placeTerm>
  </place>
  <note type="transcription">[CONTENT-OF-TRANSCRIPTION]</note>
  <part>
    <physicalDescription>
      <extent>
        <start>[START-TIME]</start>
        <end>[END-TIME]</end>
        <total>[TOTAL-TIME]</total>
      </extent>
    </physicalDescription>
  </part>
</relatedItem>
```

Abb. 13: Schematische Darstellung der Notation von Annotationen in MODS. © Arnold, 2015, CC-Lizenz 0.

³² Erläuterungen zum AV-Portal der TIB Hannover.

³³ Open Annotation Data Model: Das Datenmodell wurde im Rahmen der Open Annotation Collaboration entwickelt. Zum Einsatz von OADM bei der Annotation von Multimedia Objekten und für einen Vergleich mit anderen Annotationsmodellen vgl. Haslhofer et al. 2014, passim.

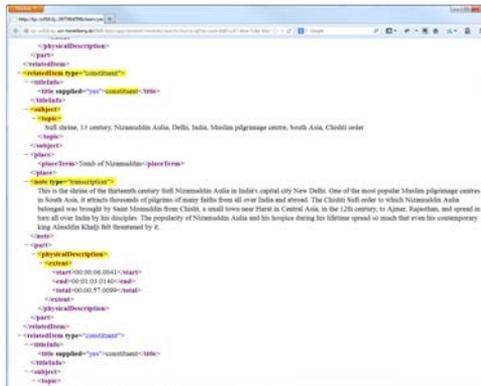


Abb. 14: Beispiel einer Annotation in MODS. Screenshot, erstellt am 30.10.2014, © HRA, CC-Lizenz 0.

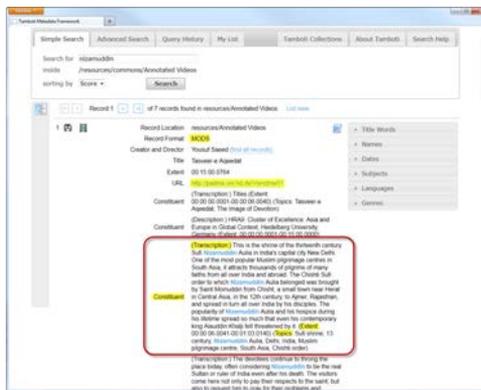


Abb. 15: Anzeige von Video Annotationen in einem Tamboti Datensatz. Screenshot, erstellt am 30.10.2014, © HRA, CC-Lizenz 0.

Dank dieser Erweiterung des Metadatenframeworks ist es möglich, in Tamboti parallel in Metadaten und Annotationen von Videos, Bildern, Texten und bibliographischen Datensätzen zu suchen.



Abb. 16: Eine Tamboti-Suche nach »Krishna« ergibt Treffer in verschiedenen Medien und Datenformaten: Bildmetadaten (VRA Core 4), bibliographische Daten (MODS), Videoannotationen (related items in MODS), Textannotationen (TEI). Screenshot, erstellt am 30.10.2014, © Arnold, [CC-Lizenz 0](#).

2.7 Ausblick

Es gibt noch eine ganze Reihe von Features, die für unterschiedliche Anwendungsszenarien in Forschung und Lehre wünschenswert erscheinen.³⁴ Diese sind natürlich nicht auf den Einsatz in der (digitalen) Kunstgeschichte oder die Geschichtswissenschaft beschränkt, sondern auf viele Bereiche der geisteswissenschaftlichen Forschung anwendbar; ob das nun das (semi-)automatische Erkennen und Verfolgen von Gesichtern, Gegenständen bzw. Mustern im Film oder Standardisierungen bei den Metadaten und Annotationen betrifft.³⁵ Das Geschäftsmodell, das hinter Pan.do/ra steht, macht es sehr wahrscheinlich, dass diese und andere Features implementiert werden können, da die Kosten nicht von einer Institution oder einem Projekt alleine getragen werden müssen.

Was die Weiterentwicklung im Rahmen des Exzellenzclusters angeht, so stehen die Einbindung kontrollierter Vokabulare für Schlagworte und Personen sowie die dynamische Anbindung an das Metadatenframework Tamboti auf der Agenda. Ebenso soll beim Datenaustausch zwischen Videoobjekt-Metadaten (z.B. als MODS Repräsentation) und Nutzerannotationen (im Open Data Annotation Model) unterschieden werden, auch um eine Überfrachtung der Objektmetadaten zu vermeiden.

2.8 Nachtrag

Die Vorstellung des Projekts und der Austausch mit den Kollegen auf der DHd 2014 in Passau erwies sich als überaus fruchtbar. Fabian Cremer beschrieb das Projekt in seinem DHd-Blogbeitrag als »idealtypisch« und »nicht nur leichtgängiger sondern auch nachhaltiger«.³⁶ Anfang Juni 2014 organisierte die Heidelberg Research Architecture einen Round Table *Annotation von digitalen Medien* im Heidelberger Karl Jaspers Centre mit Teilnehmern verschiedener Fachdisziplinen aus Berlin, Darmstadt, Essen, Heidelberg und Rom.³⁷ Ein Folgeworkshop fand im Frühjahr 2015 unter Federführung des DARIAH-DE Cluster 6 *Research Annotations* statt.³⁸ Darüber hinaus wurde die im Artikel vorgestellte Zusammenstellung von Videoannotationstools in die im Rahmen des DARIAH-DE Cluster 6 erarbeitete Liste von Annotationstools eingearbeitet. Diese soll auch an das [DIRT-Verzeichnis](#) weitergegeben und damit einem noch breiteren Nutzerkreis verfügbar gemacht werden.

³⁴ Für einen Einblick in den Bereich Lehre vgl. Lemon et al. 2013, passim.

³⁵ Alles Themen, die derzeit Gegenstand von Forschungsprojekten oder in Websites wie dem AV Portal bereits verfügbar sind. Vgl. Nixon / Troncy 2014, S. 11–15 und 18ff.

³⁶ Cremer 2014.

³⁷ Borek / Reiche 2014.

³⁸ Programm [\[online\]](#)

Bibliographische Angaben

- Activists Guide to Archiving Video, Version 1.0 (July 2013). [\[online\]](#) [\[Nachweis im GBV\]](#)
- Annotations at Harvard – Video Annotations. Harvard Annotation Project 2010. [\[online\]](#)
- Matthias Arnold: Overview: Video Annotation Tools. [\[online\]](#)
- Luise Borek / Ruth Reiche: Round Table »Annotation von digitalen Medien«. In: DHD-Blog vom 29. Juli 2014. [\[online\]](#)
- Frank Bruder: Pop-Splits – Die besten Songs aller Zeiten und ihre Geschichte. Berlin 2004. [\[Nachweis im GBV\]](#)
- Fabian Cremer: Erfolg und Emergenz in den Digital Humanities – ein Tagungseindrucksausschnitt. In: DHD-Blog vom 29. März 2014. [\[online\]](#)
- Gabriel García Márquez: Ein sehr alter Mann mit riesengroßen Flügeln (Un señor muy viejo con unas alas enormes). In: Gabriel García Márquez: Laubsturm. Frankfurt/Main 2004. [\[Nachweis im GBV\]](#)
- Global Politics on Screen: A Japanese Film on the Lytton Commission in 1932. [\[online\]](#)
- Bernhard Haslhofer / Robert Sanderson / Rainer Simon / Herbert van de Sompel: Open Annotations on Multimedia Web Resources. In: Multimedia Tools and Applications 70 (2014) H. 2, S. 847–867. doi:10.1007/s11042-012-1098-9. [\[Nachweis im GBV\]](#)
- Henry Keazor / Thorsten Wübbena: Video thrills the Radio Star: Musikvideos: Geschichte, Themen, Analysen. 3. Auflage. Bielefeld 2011 (zuerst 2005). [\[Nachweis im GBV\]](#)
- Kunstgeschichte, Musikvideo und Bildwissenschaften. Eine Einführung: Come ride with me through the veins of history. Hg. von Henry Keazor. Berlin 2014 [\[online\]](#).
- Narelle Lemon / Meg Colasante / Karen Corneille / Kathy Douglas: Video Annotation for Collaborative Connections to Learning: Case Studies from an Australian Higher Education Context. In: Increasing Student Engagement and Retention Using Multimedia Technologies: Video Annotation, Multimedia Applications, Videoconferencing and Transmedia Storytelling. Hg. von Laura A. Wankel / Patrick Blessinger. 2013, S. 181–214 (= Cutting-Edge Technologies in Higher Education, 6 Part F.). [\[Nachweis im GBV\]](#)
- Bernard Marcadé / Dan Cameron: Pierre et Gilles – The Complete Works, 1976–1996. Köln 1997. [\[Nachweis im GBV\]](#)
- Lyndon Nixon / Raphael Troncy: Survey of Semantic Media Annotation Tools – towards New Media Applications with Linked Media. Paper presented at the Linked Media Workshop, European Semantic Web Conference am 25. Mai 2014. [\[online\]](#)
- Margret Plank: A Portal for Scientific Audiovisual Media: Analysing User Needs. In: LIBER, 3. Oktober 2012. [\[online\]](#)
- Margret Plank / Steffen Weichert: Don't Forget the Users – Developing a Portal for Audiovisual Media with a User-Centred Approach. In: IFLA WLIC 2014. Libraries, Citizens, Societies: Confluence for Knowledge (Session 139 - Audiovisual and Multimedia with Information Technology). Lyon. 16.–22. August 2014. [\[online\]](#)
- Rewind – Play – Fast Forward: The Past, Present and Future of the Music Video. Hg. von Henry Keazor / Thorsten Wübbena. Bielefeld 2010. [\[online\]](#)

Abbildungslegenden und -nachweise

- Abb. 1: Startseite von Pan.do/ra. Screenshot, erstellt am 17.06.2015, © HRA, [CC-Lizenz 0](#).
- Abb. 2: (a-d) Stills aus dem Clip *Losing My Religion* – Sänger Michael Stipe mit Flügeln und weitere Szenen (beispielhaft) mit dem Motiv der Engelsflügel in der Editor-Ansicht von Pan.do/ra, jeweils mit dem Tag »Engelsflügel« versehen (vgl. im rechten Bereich den Punkt »Keywords«). Quelle: HRA, Warner Bros., erstellt am 12.10.2014.
- Abb. 3: Link aus Pan.do/ra zum Bilddatenbanksystem ConedaKOR. Quelle: HRA, KGI (Frankfurt/Main), Warner Bros., erstellt am 12.10.2014.
- Abb. 4: Entitätsansicht in ConedaKOR: Michelangelo Merisi da Caravaggio, *Verzückung des heiligen Franziskus*, um 1595, Hartford, Wadsworth Atheneum Museum of Art. Quelle: KGI (Frankfurt/Main), [Web Gallery of Art](#), erstellt am 12.10.2014.
- Abb. 5: Interner Link bzw. Einbettung: Michelangelo Merisi da Caravaggio, *Die sieben Werke der Barmherzigkeit*, 1607, Neapel, Chiesa del Pio Monte della Misericordia. Quelle: HRA, KGI (Frankfurt/Main), [Web Gallery of Art](#), Warner Bros., erstellt am 12.10.2014.
- Abb. 6: Link bzw. Einbettung: Michelangelo Merisi da Caravaggio, *Der ungläubige Thomas*, 1603, Potsdam, Bildergalerie Schloss Sanssouci. Quelle: HRA, KGI (Frankfurt/Main), [Web Gallery of Art](#), Warner Bros., erstellt am 12.10.2014.
- Abb. 7: Einbettung: Pierre & Gilles, *St. Sebastian*, 1987. Quelle: HRA, KGI (Frankfurt/Main), [Art Forum](#), Warner Bros., erstellt am 12.10.2014.

Abb. 8: Interner Bezug durch Tagging, verdeutlicht durch eine farbige Kennzeichnung auf der unteren Timeline (in diesem Fall »Arbeiter«). Quelle: HRA, Warner Bros., erstellt am 12.10.2014.

Abb. 9: Einbettung: Abbildungen des Tanzes von V. Nijinsky zu *Jeux*. Quelle: HRA, KGI (Frankfurt/Main), Warner Bros., erstellt am 12.10.2014.

Abb. 10: Interner Link zur korrespondierenden Filmszene aus *The Cell*, 2000 (Director: Tarsem Singh). Quelle: HRA, KGI (Frankfurt/Main), Warner Bros., New Line Cinema (theatrical), [Warner Home Video](#) (DVD), erstellt am 12.10.2014.

Abb. 11: Gegenüberstellung eines Kartenausschnitts aus dem Kommissionsbericht und eines Screenshots der Location Annotationsspur in Pan.do/ra. Quelle: Route Maps: Itineraries of the Commission in the Far East (League of Nations (ed.): *Appeal by the Chinese Government. Report of the Commission of Enquiry, Appendix, Maps No. 13 and 14*, Geneva: League of Nations, 1932).

Abb. 12: [Requirements of a Video Annotation System](#). Quelle: Harvard Annotation Project, 2010. Mit freundlicher Genehmigung.

Abb. 13: Schematische Darstellung der Notation von Annotationen in MODS. © Arnold, 2015, [CC-Lizenz 0](#).

Abb. 14: Beispiel einer Annotation in MODS. Screenshot, erstellt am 30.10.2014, © HRA, [CC-Lizenz 0](#).

Abb. 15: Anzeige von Video Annotationen in einem Tamboti Datensatz. Screenshot, erstellt am 30.10.2014, © HRA, [CC-Lizenz 0](#).

Abb. 16: Eine Tamboti-Suche nach »Krishna« ergibt Treffer in verschiedenen Medien und Datenformaten: Bildmetadaten (VRA Core 4), bibliographische Daten (MODS), Videoannotationen (related items in MODS), Textannotationen (TEI). Screenshot, erstellt am 30.10.2014, © Arnold, [CC-Lizenz 0](#).

Zeitschrift für digitale Geisteswissenschaften

Artikel aus:

Sonderband 1 der ZfdG: Grenzen und Möglichkeiten der Digital Humanities. Hg. von Constanze Baum und Thomas Stäcker. 2015. DOI: [10.17175/sb01](https://doi.org/10.17175/sb01)

Titel:

AAC-Fackel. Das Beispiel einer digitalen Musteredition

Autor/in:

Hanno Biber

Kontakt:

hanno.biber@oeaw.ac.at

Institution:

Österreichische Akademie der Wissenschaften

GND:

[133354660](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-33354660)

ORCID:

[0000-0002-0681-771](https://orcid.org/0000-0002-0681-771)

DOI des Artikels:

[10.17175/sb001_019](https://doi.org/10.17175/sb001_019)

Nachweis im OPAC der Herzog August Bibliothek:

[830206329](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-330206329)

Erstveröffentlichung:

19.02.2015

Lizenz:

Sofern nicht anders angegeben 

Medienlizenzen:

Medienrechte liegen bei den Autoren

Letzte Überprüfung aller Verweise:

24.05.2016

GND-Verschlagwortung:

[Edition](#) | [Elektronische Publikation](#) | [Die Fackel \(Zeitschrift\)](#) |

Zitierweise:

Hanno Biber: AAC-Fackel. Das Beispiel einer digitalen Musteredition. In: Grenzen und Möglichkeiten der Digital Humanities. Hg. von Constanze Baum / Thomas Stäcker. 2015 (= Sonderband der Zeitschrift für digitale Geisteswissenschaften, 1). PDF Format ohne Paginierung. Als text/html abrufbar unter DOI: [10.17175/sb001_019](https://doi.org/10.17175/sb001_019).

Hanno Biber

AAC-Fackel. Das Beispiel einer digitalen Musteredition

Abstracts

In diesem Beitrag wird die digitale Musteredition der von Karl Kraus 1899 bis 1936 in Wien herausgegebenen Zeitschrift *Die Fackel* präsentiert. Die *AAC-Fackel* wurde in Zusammenarbeit von Forschern, Wissenschaftlern, Entwicklern und Designern im Rahmen des vom Institut für Corpuslinguistik und Texttechnologie der Österreichischen Akademie der Wissenschaften betriebenen AAC-Austrian Academy Corpus erarbeitet und am 1. Januar 2007 online gestellt. Die digitale Edition der Zeitschrift *Die Fackel* gibt einen kostenfreien Zugang zu 37 Jahrgängen, 415 Heften und 922 Nummern mit 22.586 Seiten und mehr als 6 Millionen Wortformen und bietet eine durchsuchbare Datenbank des gesamten Textes mit Indices, Such- und Navigationsinstrumenten in einem funktionell gestalteten Interface, wobei alle Seiten sowohl als Faksimile als auch als digitaler Text abrufbar sind. Die allgemeinen Prinzipien und die Überlegungen zum Design dieser vielgenutzten Textressource beruhen auf einer corpusbasierten Methodologie und sind bestimmend für die Fragen der wissenschaftlichen Edition und des philologischen Zugangs zu diesem Schlüsseltext der deutschen Literaturgeschichte.

This paper presents the digital edition of an important text in German literary history. The satirical magazine *Die Fackel* (*The Torch*) was published by the satirist and language critic Karl Kraus in Vienna from 1899 until 1936. The digital edition of the *AAC-Fackel* was developed by researchers, scholars, programmers, and designers within the framework of the AAC-Austrian Academy Corpus, operated by the Institute for Corpus Linguistics and Text Technology at the Austrian Academy of Sciences in Vienna, and went online on 1 January 2007. The *AAC-Fackel* offers free online access to the 37 volumes, 415 issues, and 922 issues, comprising 22,586 pages and more than 6 million word forms of *Die Fackel* and contains a fully searchable database of the entire journal with indexes and search and navigation tools in a functionally designed interface, in which all pages of the original journal are available as digital texts and as facsimile images. The general principles of this scholarly digital edition, and its concept and design considerations, are based on a corpus-based methodological approach and are decisive for addressing questions of critical editions and philological access to this key text in German literary history.

1.

»Dennoch: Der Online-Auftritt der ›Fackel‹ entspricht der Programmatik dieser Zeitschrift. Obwohl oder weil man diesen Auftritt in mehrfachem Sinne auch als »Tritt in die Öffentlichkeit« [F 909,12] wird bezeichnen können.«¹ Die *AAC-Fackel*, die digitale Ausgabe der von Karl Kraus vom 1. April 1899 bis Februar 1936 in 922 Nummern in Wien herausgegebenen Zeitschrift *Die Fackel* wurde unter Anwendung computerphilologischer, corpuslinguistischer und texttechnologischer Methoden im Rahmen der an der Österreichischen Akademie der Wissenschaften gegründeten Unternehmung AAC-Austrian Academy Corpus (unter der Leitung von Evelyn Breiteneder) als digitale Musteredition konzipiert, erstellt und am 1. Januar 2007

¹ Welzig 2007, Vorwort [unpag.] [online].

öffentlich und kostenfrei online zugänglich gemacht.² Vorarbeiten für die *AAC-Fackel* beruhen auf den textlexikographischen Forschungen, die an der Österreichischen Akademie der Wissenschaften im Forschungsprojekt *Wörterbuch der Fackel* durchgeführt wurden, für das der digitalisierte Text der Fackel erstmals für Forschungsarbeiten genutzt wurde.³ Das AAC-Austrian Academy Corpus ist ein großes Textcorpus zur deutschen Sprache, »eine umfangreiche und komplex strukturierte Sammlung von digitalen Volltexten zur deutschen Sprache und Literatur im Untersuchungszeitraum 1848 bis 1989«⁴ und somit Basis für die Entwicklung innovativer texttechnologischer Anwendungen im Bereich der Corpusforschung wie auch für philologische, text- und kulturwissenschaftliche Forschungen. Am Institut für Corpuslinguistik und Texttechnologie (ICLTT) der Österreichischen Akademie der Wissenschaften, in dem in der Arbeitsgruppe LIT – Literature in Transition & AAC-Austrian Academy Corpus das AAC betrieben wird, wird mit diesen umfangreichen und gut strukturierten Textressourcen sowie mit anderen Textcorpora philologische Grundlagenforschung in den Bereichen der Corpusforschung und computergestützten Textwissenschaften geleistet.⁵ Das AAC-Programm bildet Schnittstellen zwischen kulturwissenschaftlichen und textwissenschaftlichen Methoden und Prinzipien, philologischer Exaktheit, wissenschaftlichen Editionstechniken, Fragen der Textstrukturierung, Annotation und Kommentierung der Texte einerseits und zeitgemäßer forschungsspezifischer sowie informationstechnologischer Anwendungen andererseits. In diesem dynamischen Aufgabenfeld werden Konvergenzen der Grundlagenforschung im geisteswissenschaftlichen Bereich hin zu deren Umsetzungsmöglichkeiten im Kontext der Digital Humanities erarbeitet. Die *AAC-Fackel* kann neben der digitalen Edition der literarischen Zeitschrift *Der Brenner*⁶ in diesem Kontext als exemplarischer Anwendungsfall einer aus der Corpusforschung und computergestützten Textwissenschaft entstandenen digitalen Musteredition eines literaturgeschichtlich bedeutenden Textes betrachtet werden. Ihr Zustandekommen und die dafür notwendigen Bedingungen resultieren aus einer sich mit Sprache und Fragen des besonderen Sprachgebrauchs auf empirischer Textbasis in einem bestimmten, mit besonderen Eigenschaften ausgestatteten Textcorpus befassenden Forschungsrichtung, wie sie im ICLTT und seinen Vorgängerunternehmungen seit vielen Jahren betrieben wird.

² AAC – Austrian Academy Corpus: AAC-Fackel. Online Version: »*Die Fackel*. Herausgeber: Karl Kraus, Wien 1899–1936.« AAC Digital Edition No. 1. Hg. von Hanno Biber / Evelyn Breiteneder / Heinrich Kabas / Karlheinz Mörth [[online](#)].

³ Vgl. Welzig 1999, S. 1–1056.

⁴ Vgl. Informationen zum AAC auf der [AAC-Website](#) der Österreichischen Akademie der Wissenschaften.

⁵ Vgl. Informationen zum ICLTT auf der Instituts-Website der Österreichischen Akademie der Wissenschaften.

⁶ AAC-Austrian Academy Corpus und Brenner-Archiv: Brenner Online. Online Version: »Der Brenner. Herausgeber: Ludwig Ficker, Innsbruck 1910–1954.« AAC Digital Edition No. 2. Hg. von Hanno Biber / Evelyn Breiteneder / Heinrich Kabas / Karlheinz Mörth [[online](#)].



Abb. 1: AAC-Fackel: *DIE FACKEL* TEXT (© AAC).

Im Laufe der ersten sieben Jahre seit der Veröffentlichung der *AAC-Fackel* haben sich mehr als 27.000 Benutzer auf dem dafür eingerichteten Webserver mit Email-Adresse kostenfrei registriert, um die zu Verfügung gestellten Texte lesen und durchsuchen zu können. Die in der digitalen Edition bereitgestellten Daten werden sowohl von von spezifischen Text- und Sprachinteressen geleiteten Wissenschaftlerinnen und Wissenschaftlern erforscht, als auch von allgemein an der Sprache und Literatur interessierten Leserinnen und Lesern aus aller Welt vielfältig genutzt. Das Interface der *AAC-Fackel*, das nach den für diese Edition entwickelten Prinzipien zur Funktionalität digitaler Textressourcen und von den erforderlichen Überlegungen zum graphischen Design einer elektronischen Version und Neuinterpretation einer historischen literarischen Zeitschrift bestimmt ist, wurde von der in Los Angeles arbeitenden und lehrenden Designerin Anne Burdick in Zusammenarbeit mit dem Projektteam des AAC gestaltet. Das Webinterface der *AAC-Fackel* ermöglicht den Benutzern, die in XML-Formaten und durch die entsprechenden informationstechnologischen Transformationen digital aufbereiteten und darstellbar gemachten Texte sowohl lesen als auch in komplexer Weise die in den Texten präsenten sprachlichen Formen im Detail und in Verwendung corpuslinguistischer Methoden untersuchen und analysieren zu können. Darüber hinaus kann auch in einfacher Weise in den Texten der Zeitschrift nach sprachlichen Einheiten, Wortformen, Wörtern, Namen und anderem gesucht und können dabei deren Eigenschaften erschlossen werden.

Das literarische Werk des wohl bedeutendsten Satiriker deutscher Sprache und darin die 1899 bis 1936 in Wien herausgegebene Zeitschrift *Die Fackel* des 1874 in Jičín in Böhmen geborenen Karl Kraus, der in Wien gelebt, in vielen Städten Europas seine Texte vorgetragen hat und (wörtlich und metaphorisch) weit darüber hinaus rezipiert wurde und wird, ist als überaus bedeutender Beitrag der deutschsprachigen Literatur zur Weltliteratur zu betrachten. Die Überlieferung der satirischen und polemischen Texte nun im digitalen Medium, der Texte mit ihrer thematischen Vielfalt, sprachlichen Komplexität und historischen Relevanz, muss für eine wissenschaftliche Institution, die sich der zeitgemäßen Erforschung des kulturellen Erbes nach neuesten wissenschaftlichen Standards verpflichtet hat, eine überaus wichtige Aufgabe sein. In der digitalen Edition der *AAC-Fackel* wird der gesamte Text dieser Zeitschrift somit einer

neuen, an Literatur, Sprache und Geschichte interessierten Leserschaft zugänglich gemacht, und die einzigartige sprachliche und literarische Qualität dieser Texte wird unter Nutzung computerphilologischer Methoden und texttechnologischer Instrumente durch verschiedene Suchmöglichkeiten und Register in einer neuen und anderen Weise erschlossen.

Die *AAC-Fackel* bietet neben der Volltextsuche und den Registern mit ca. 6 Millionen Wortformen auch ein vollständiges, erstmals publiziertes Inhaltsverzeichnis sämtlicher Texte der Zeitschrift, das für die digitale Edition neu erstellt wurde. Dabei wurden sowohl die Angaben von Karl Kraus in den Überschriften der einzelnen Beiträge bzw. von den Textanfängen und den Inhaltsangaben der Hefte als auch jene Inhaltsverzeichnisse berücksichtigt, die von ihm nachträglich für die jeweiligen Quartalsbände seiner Zeitschrift erstellt wurden. Die vollständige Bildbeigabe aller 22.586 Textseiten (inklusive der Umschlagseiten und Beilagenblätter der Zeitschrift) als digitale Faksimiles ermöglicht den Nutzern die quelleneditorisch korrekte Zitierung aller Texte in den 415 Heften und 922 Nummern der 37 Jahrgänge der Zeitschrift. Es ist für das Interface vorgesehen, weitere neue Funktionen, wie etwa ein auf einer in den letzten Jahren erstellten und bearbeiteten Namensdatenbank der *Fackel* beruhendes Personennamenregister sowie später zusätzlich auch ein Verzeichnis der Varianten der Hefte, in einer solcherart ergänzten erneuerten Version der *AAC-Fackel* zu implementieren.



Abb. 2: *AAC-Fackel*: *DIE FACKEL* CONTENTS und *DIE FACKEL* TEXT (© AAC).

Die *AAC-Fackel* weist einige herausragende Merkmale in ihrer Erscheinungsform auf. Es wurden fünf Frames zu einem stabilen Interface integriert, so dass die die digitale Edition und ihre Prinzipien und Funktionen beschreibenden Paratexte (paratexts) neben dem Suchframe (search/index) und dem dazugehörigen Resultateframe (results) (Abbildung 3) ebenso je nach Bedarf zur Verfügung stehen wie das Frame mit dem Inhaltsverzeichnis (contents) (Abbildung 2) und das Leseframe (texts) (Abbildung 1). Für die im AAC konzipierte digitale Edition wurde im Rahmen der texttechnologischen Forschungen ein spezifisches, funktionell gestaltetes Navigationsmodul konzipiert, mit dessen Hilfe im Leseframe nicht nur von Seite zu Seite, von Heft zu Heft oder von Jahrgang zu Jahrgang navigiert werden kann, sondern man auch (vom Resultateframe ausgehend) zu im jeweiligen Recherchezusammenhang relevanten Textpassagen direkt gelangen kann. »At every point we know where we are, how we arrived,

and how to move around while making use of the analytic features built into the project.»⁷

Die besondere graphische Umsetzung im Bereich des Inhaltsverzeichnisses im Webinterface (contents) kompensiert die nicht unmittelbar erfassbaren physischen Objekteigenschaften der Zeitschriftenhefte und Bände durch symbolische Darstellung in Form von Icons und visuellen Informationen zu den Heften und Bänden der Zeitschrift, so dass diese Größenverhältnisse auf diese Weise unmittelbar erfahrbar gemacht werden. Die digitale Edition bietet somit eine optisch wahrzunehmende Repräsentation der von Papierqualität, Papiervolumen, Druck- und Bindungstechniken bestimmten Eigenschaften eines Druckwerks.

»As both a research tool and a digital record, the AAC allows scholars to explore the pages of Die Fackel without making a trip to the archives. Therefore careful consideration was given to the interpretation of the printed journal's materiality into/onto the screen. While the web version will always come up short in certain aspects when compared to the print originals, a digitized text can perform in ways that a physical archive cannot.«⁸

In der *AAC-Fackel* ist ein linguistisches Suchmodul (search / index) eingerichtet, das es den an den Texten in besonderer Weise interessierten Leserinnen und Lesern ermöglicht, corpusbasierte Abfragen vorzunehmen und die Basisfunktionalität von mit linguistischen Tags versehenen Textressourcen zu nutzen und die mit Part-of-Speech-Tags (nach dem STTS-Tagset) und mit Lemma-Informationen angereicherten Textelemente in großer Detailgenauigkeit zu untersuchen, wofür zweckdienliche technische Erläuterungen zur Abfragesprache und den Abfrageparametern für das Suchinstrument in den Paratexten der digitalen Edition bereitgestellt werden. Neben der Möglichkeit auch komplex verknüpfte Suchen vorzunehmen, beispielsweise nach bestimmten Wortarten und den Konfigurationen im Text nach dem Gebrauch des attributiven Adjektivs vor einem bestimmten Nomen, kann auch mit Abstandsoperatoren gesucht werden, oder können Suchen nach lemmatisierten Formen und in komplexer Weise (unterstützt durch die dahinterliegende MSSQL-Datenbank) vorgenommen werden. Darüber hinaus gibt es auch einen vollständigen Wortformenindex und einen rückläufigen Wortformenindex, der Wortbildungsmuster in den Texten auf einfache Weise nachvollziehbar macht. Für alle Ergebnisse werden die Frequenzangaben ausgegeben, wie auch die jeweilige Stellenangabe mit dem Leseframe direkt verknüpft ist, so dass der Fund mit der Stelle im Kontext der jeweiligen Seite direkt verlinkt ist. Das Suchframe im Webinterface kann je nach Bedarf modifiziert werden, um mehr oder weniger Text in den KWIC-Anzeigen der Resultate auszugeben, die Suchen zeitlich einzuschränken oder die Anzahl der angezeigten Treffer spezifisch anzupassen, um hier einige wichtige Funktionen zu nennen.

⁷ Drucker 2014, S. 159.

⁸ Burdick 2007, unpag.



Abb. 3: AAC-Fackel: SEARCH | INDEX und RESULTS (© AAC).

Komplex organisierte Texte und in diesem besonderen Fall die Texte einer historischen literarischen Zeitschrift werden in der *AAC-Fackel* mit einem Inventar von literarischen Formen und sprachlichen Eigenschaften in einer adäquaten und funktionellen Gestalt im digitalen Medium so wiedergegeben, dass sie unter Nutzung der texttechnologischen Möglichkeiten im neuen Medium, in das die Texte gleichsam übersetzt werden müssen, neu gelesen, interpretiert und analysiert werden können. Aus den Forschungsperspektiven der Literaturwissenschaft (mit Blick auf das Werk von Karl Kraus im Besonderen), der corpusbasierten Textwissenschaft (mit Blick auf die Erforschung von lexikalischen Eigenschaften), der Zeitschriftenforschung (mit Blick auf die Erforschung von Strukturmerkmalen von Zeitschriftentexten im Allgemeinen) kann hier abschließend auch die »Topographie der Texte«⁹ beachtet werden. Dabei geht es um die räumliche Verteilung der Texte im Heft, das einen sprachlichen Raum konstituiert, in dem es »Relationen der in ihm enthaltenen Objekte zueinander gibt, der Nähe oder Ferne sowie Richtungsbeziehungen der aufeinander verweisenden Texte, einen Innenraum der Texte und einen durch die Umschlagseiten entstehenden Außenraum, deren Teile in sich und von außen nach innen auf die Mitte hin strukturiert sind.«¹⁰ Der Textraum der Zeitschrift kann mit dem texttechnologischen Instrument der *AAC-Fackel*, der Navigationsmöglichkeit von Heft zu Heft, den schnellen Auffindungsmöglichkeiten, den Tools zur Darstellung des Inhaltsverzeichnisses aller Hefte, um hier nur ein paar Beispiele zu nennen, nicht nur besonders gut untersucht werden, sondern es werden damit neue, darauf basierende Interpretationen der Texte ermöglicht. Beispielhaft sei hier das zweite im Rahmen des Forschungsprojektes *Wörterbuch der Fackel* erstellte Wörterbuch erwähnt, das für die drei Teilbände nicht nur den digitalisierten Text zur Analyse und Interpretation nutzen konnte, sondern auch für alle angeführten Belegstellen der Fackel direkte, mit URL-Angaben versehene Verweise vom gedruckten Buch in den digitalen Text verzeichnet.¹¹

⁹ Biber 2001, S. 49.

¹⁰ Biber 2014, S. 105ff.

¹¹ Vgl. Welzig 2008, Alpha, S. 1–126 und Chrono, S. 1–555.

Die Antworten auf die Fragestellungen nach den geänderten Sichtweisen und Steuerungsvorgängen durch die digitale Verfügbarkeit der Texte in einer Forschungsressource und der damit verbundenen Reinterpretation und adäquaten Präsentation der Ausgangsmaterialien in einer digitalen Edition liegen in der optimalen Nutzung der durch die informationstechnologische Aufbereitung der strukturellen wie lexikalisch-sprachlichen Einheiten der Texte gegebenen Such- und Indizierungsverfahren sowie den verschiedenen, dadurch eröffneten Zugangsmöglichkeiten zum Textgefüge und zu einzelnen Elementen der Mikro- und Makrostrukturen der Textobjekte. »In der *Fackel* entwickelt Karl Kraus eine eigene Typologie von Texten«,¹² die die semantischen Zuordnungen und Beziehungen der Texte zueinander bestimmt, die mithilfe des digitalen Zugriffs auf die Gesamtheit des Textcorpus ebenso untersucht werden können, wie im sprachlichen Mikrobereich der Texte einzelne sprachliche Formulierungen und ihre Dynamik im Kontext sowie der gesamte Wortformen- und Lemmabestand corpuslinguistisch analysiert werden kann. »Welche Themen in welcher Form behandelt werden, wird im Text an vielen Stellen reflektiert, und viele Themen werden oft in verschiedenen Formen mit unterschiedlichen Intentionen behandelt, in Aufsätzen, in Glossen, Notizen oder Inschriften.«¹³ In einem neuen Index für die digitale Edition der *AAC-Fackel* sollen die Textsorten dieser Zeitschrift genau verzeichnet und dargestellt werden. Eine zentrale Forschungsperspektive liegt von derartigen Fragestellungen ausgehend nicht zuletzt auch in der methodisch-theoretischen Reflexion über die sich durch die Kombination von editionsphilologischen Fragestellungen mit Fragen der Informationstechnologie, der Texttechnologie und der Corpusforschung sowie dem Webinterface Design sich ergebenden Konsequenzen und Forschungsansätze. Die *AAC-Fackel* »creates a humanistic environment that supports question, analysis, and study.«¹⁴

¹² Biber 2001, S. 49.

¹³ Biber 2014, S. 105.

¹⁴ Drucker 2014, S. 160.

Bibliographische Angaben

AAC - Austrian Academy Corpus: AAC-Fackel, Online Version: »Die Fackel. Herausgeber: Karl Kraus, Wien 1899–1936.« AAC Digital Edition No. 1. Hg. von Hanno Biber / Evelyn Breiteneder / Heinrich Kabas / Karlheinz Mörth. [\[online\]](#)

AAC - Austrian Academy Corpus and Brenner-Archiv: Brenner online, Online Version: »Der Brenner. Herausgeber: Ludwig Ficker, Innsbruck 1910–1954«. AAC Digital Edition No. 2. Hg. von Hanno Biber / Evelyn Breiteneder / Heinrich Kabas / Karlheinz Mörth. [\[online\]](#)

Hanno Biber: Die Komposition der *Fackel*. Diss. Universität Wien, Wien 2001. [\[online\]](#)

Hanno Biber: »Ihre höchste Stilleistung ist die graphische Anordnung.« Zur Typographie der satirischen Zeitschrift *Die Fackel* von Karl Kraus. In: Fontes Litterarum – Typographische Gestaltung und literarischer Ausdruck. Hg. von Philipp S. Vanscheidt / Markus F. Polzer. Hildesheim 2014, S. 99–112. [\[Nachweis im GBV\]](#)

Anne Burdick: »Notes on the Design«. In: Paratexts, Notes on the Design [2007], AAC - Austrian Academy Corpus: AAC-Fackel. Online Version: »Die Fackel. Herausgeber: Karl Kraus, Wien 1899–1936« AAC Digital Edition No. 1. Hg. von Hanno Biber / Evelyn Breiteneder / Heinrich Kabas / Karlheinz Mörth. [\[online\]](#)

Johanna Drucker: Graphesis. Visual Forms of Knowledge Production. Cambridge, MA 2014. [\[Nachweis im GBV\]](#)

Werner Welzig: »Wie? Wo? Weswegen?« In: Paratexts, Vorwort (April 2007), AAC - Austrian Academy Corpus: AAC-Fackel. Online Version: *Die Fackel. Herausgeber: Karl Kraus, Wien 1899–1936*. AAC Digital Edition No. 1. Hg. von Hanno Biber / Evelyn Breiteneder / Heinrich Kabas / Karlheinz Mörth. [\[online\]](#)

Werner Welzig: Wörterbuch der Redensarten zu der von Karl Kraus 1899 bis 1936 herausgegebenen Zeitschrift *Die Fackel*. Wien 1999. [\[Nachweis im OPAC\]](#)

Werner Welzig: Schimpfwörterbuch zu der von Karl Kraus 1899 bis 1936 herausgegebenen Zeitschrift *Die Fackel*. Alphabetisches – Chronologisches – Explikatives. Wien 2008. [\[Nachweis im GBV\]](#)

Abbildungslegenden und -nachweise

Abb. 1: AAC-Fackel: *DIE FACKEL* TEXT (© AAC).

Abb. 2: AAC-Fackel: *DIE FACKEL* CONTENTS und *DIE FACKEL* TEXT (© AAC).

Abb. 3: AAC-Fackel: SEARCH | INDEX und RESULTS (© AAC).

Zeitschrift für digitale Geisteswissenschaften

Artikel aus:

Sonderband 1 der ZfdG: Grenzen und Möglichkeiten der Digital Humanities. Hg. von Constanze Baum und Thomas Stäcker. 2015. DOI: [10.17175/sb01](https://doi.org/10.17175/sb01)

Titel:

Heterogene Daten in den Digital Humanities: Eine Architektur zur forschungsorientierten Förderung von Kollektionen

Autor/in:

Tobias Gradl

Kontakt: tobias.gradl@uni-bamberg.de

Institution: Otto-Friedrich-Universität Bamberg

GND: [1084606585](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63862-p0011-9) ORCID: [0000-0002-1392-2464](https://orcid.org/0000-0002-1392-2464)

Autor/in:

Andreas Henrich

Kontakt: andreas.henrich@uni-bamberg.de

Institution: Otto-Friedrich-Universität Bamberg

GND: [111819601](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63862-p0011-9) ORCID:

Autor/in:

Christoph Plutte

Kontakt: plutte@bbaw.de

Institution: Berlin-Brandenburgische Akademie der Wissenschaften

GND: [1084027208](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63862-p0011-9) ORCID:

DOI des Artikels:

[10.17175/sb001_020](https://doi.org/10.17175/sb001_020)

Nachweis im OPAC der Herzog August Bibliothek:

[830207090](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63862-p0011-9)

Erstveröffentlichung:

19.02.2015

Lizenz:

Sofern nicht anders angegeben



Medienlizenzen:

Medienrechte liegen bei den Autoren

Letzte Überprüfung aller Verweise:

24.05.2016

GND-Verschlagwortung:

[Datenintegration](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63862-p0011-9) | [Digital Humanities](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63862-p0011-9) | [Softwaresystem](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63862-p0011-9) | [Architektur \(Informatik\)](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63862-p0011-9) |

Zitierweise:

Tobias Gradl, Andreas Henrich, Christoph Plutte: Heterogene Daten in den Digital Humanities: Eine Architektur zur forschungsorientierten Förderung von Kollektionen. In: Grenzen und Möglichkeiten der Digital Humanities. Hg. von Constanze Baum / Thomas Stäcker. 2015 (= Sonderband der Zeitschrift für digitale Geisteswissenschaften, 1). PDF Format ohne Paginierung. Als text/html abrufbar unter DOI: [10.17175/sb001_020](https://doi.org/10.17175/sb001_020).

Tobias Gradl, Andreas Henrich, Christoph Plutte

Heterogene Daten in den Digital Humanities: Eine Architektur zur forschungsorientierten Föderation von Kollektionen

Abstracts

Traditionelle Ansätze der Datenintegration basieren zumeist auf einer Harmonisierung heterogener Daten im Rahmen einer übergeordneten, integrativen Datenstruktur wie einem globalen Schema oder einer globalen Ontologie. Der vorliegende Beitrag verdeutlicht die Limitationen eines derartigen Harmonisierungsansatzes für den speziellen Kontext der Digital Humanities und zeigt, wie im Rahmen von DARIAH-DE eine forschungsorientierte und fallbasierte Föderation von Daten erreicht werden kann. Mit Hilfe von Collection, Schema und Crosswalk Registry können durch die Registrierung und Assoziation von Kollektionen und darin verwendeten Datenschemata sowohl übergreifende als auch disziplinspezifische Sichten auf Forschungsdaten geschaffen werden.

Traditional approaches to data integration are typically based on the harmonization of heterogeneous data with respect to the constraints of a globally integrative data structure, such as a global schema or ontology. This paper illustrates the limitations of such a harmonization-based approach in the specific context of the Digital Humanities and shows how a research-oriented and case-based data federation can be facilitated by the DARIAH-DE federation architecture. Based on the collection, schema, and crosswalk registries, collections and the data schemata they utilize can be registered and flexibly associated – resulting in the ability to create broad and comprehensive as well as discipline-specific views of research data.

1. Einleitung

Für die kultur- und geisteswissenschaftliche Forschung relevante Ressourcen finden sich zu großen Teilen in den Sammlungen von Museen, Archiven, Bibliotheken, Universitäten und außeruniversitären Forschungseinrichtungen. Mit der Erweiterung des Anwendungsbereiches der Digital Humanities von den Sprachwissenschaften¹ hin zu einer ganzheitlichen Sicht auf die Kultur- und Geisteswissenschaften seit den 1990ern wurden vermehrt Methoden, Anwendungen und Standards für die Digitalisierung, Analyse und Beschreibung von Ressourcen geschaffen.² Die Menge der heute durch öffentliche Netzwerke verfügbaren und für die kultur- und geisteswissenschaftliche Forschung relevanten Kollektionen steigt nicht zuletzt aufgrund der Verwendung von Zugriffs- und Beschreibungsstandards stetig an und bietet Forscherinnen und Forschern einen potenziellen Zugang zu einer Vielzahl heterogener Ressourcen.

In diesem Beitrag stellen wir eine neuartige Föderationsarchitektur vor, die auf eine Erfassung und fallbasierte Zusammenführung von Forschungsdaten nach den individuellen Bedürfnissen von Forschungsprojekten abzielt. Digitale Sammlungen werden zentral

¹ Vgl. die Ausführungen zu Humanities Computing in Schreibman et al. 2004, S. 3ff.

² Grundsätzlich Schreibman et al. 2004, insbesondere S. 564ff.

verzeichnet, zur Vermeidung von Informationsverlusten jedoch nicht harmonisiert, sondern in Form von Beziehungen auf Schemaebene assoziiert, wodurch die Verwendung einer dynamisch föderierten Datenbasis in breiten und interdisziplinären, wie auch in fachspezifischen Anwendungskontexten ermöglicht werden kann.³ Ein übergeordnetes Ziel besteht insbesondere in der Nutzbarmachung des durch Experten hinterlegten Wissens zu Kollektionen und Daten sowie deren Beziehungen für einen weiten Anwenderkreis.

2. Anwendungskontext

Traditionelle Integrationsansätze folgen häufig dem Muster eines physisch harmonisierten Datenbestands auf Basis eines zentralen Schemas.⁴ Verteilte und heterogene, semi-strukturierte Daten werden hierbei in ein gemeinsames Schema übersetzt und stehen für eine einfache Weiterverarbeitung in integrierter Form zur Verfügung. Eine zentrale Aufgabe dieses Ansatzes besteht in der Umsetzung eines hinsichtlich der notwendigen Granularität geeigneten Integrationsschemas. In Bezug auf die Digital Humanities als ganzheitliche Anwendungsdomäne, die sich in Form spezifischer, interdisziplinärer und auch übergreifender Informationsbedürfnisse äußert, führt die Integration aller Disziplinen und Perspektiven jedoch entweder zu Schemata kaum verwaltbarer Komplexität oder – bei der Verwendung eines einfachen Modells, wie z. B. Dublin Core (DCES) – zum Verlust großer Anteile disziplinspezifischer Information.

Für die Konzeption der in **DARIAH-DE** umgesetzten Föderationsarchitektur werden im Folgenden zwei Anwendungsfälle vorgestellt, deren unterschiedliche Anforderungen die Einschränkungen eines solchen zentralistischen Integrationsansatzes verdeutlichen.

2.1 Generische Suche

Mit der generischen Suche verfolgt DARIAH-DE das Ziel, eine übergreifende Suchmöglichkeit zu schaffen, welche die Eigenschaften der Breiten- und Tiefensuche so vereint, dass eine dynamische Anpassung der Suche – z. B. im Hinblick auf eine mögliche Facettierung – erreicht werden kann.⁵ Die übergreifende Suche in eng assoziierten Datenquellen erlaubt unter Anwendung der in der DARIAH-DE Crosswalk Registry definierten Assoziationen und Transformationsregeln – eine detaillierte Auseinandersetzung mit den betrachteten Daten (Tiefensuche). Mit einer wachsenden Zahl einbezogener Kollektionen wird die Granularität der Betrachtung und Facettierung ggf. mangels vorhandener Verbindungen reduziert und nimmt die Form einer Breitensuche ein. Für die dynamische Funktionalität der generischen Suche ist die ad-hoc-Integration ausgewählter Kollektionen basierend auf den für eine konkrete Anfrage relevanten Kollektionen und den zwischen diesen vorliegenden Assoziationen erforderlich, um die jeweils zur Verfügung stehende Granularität von Daten nutzen zu können.

³ Henrich / Gradl 2013, S. 50f.

⁴ Lenzerini 2002, S. 234; Peroni et al. 2013, S. 228f.

⁵ Gradl / Henrich 2013, S. 8ff.

2.2 Datenintegration

Im Gegensatz zu der dynamischen, strukturellen Adaption der generischen Suche an die Zusammensetzung der für eine Anfrage ausgewählten Kollektionen zielen Lösungen der Datenintegration oftmals auf eine Konsolidierung einer a-priori definierten Auswahl von Datenquellen ab.⁶ Anforderungen an eine kollektionsübergreifende Integration sind wesentlich von der verfolgten Forschungsfrage abhängig und können z. B. im Kontext der Ablösung von Systemen durch Neuentwicklungen, aber auch für die Ausweitung der Datenbasis einer bestehenden Analyse- und Visualisierungslösung, wie beispielsweise dem **DARIAH-DE Geobrowser**⁷, auftreten. Die Anwendung eines zentralen Integrationschemas bzw. einer zentralen Ontologie führt im Fall der Datenintegration im Gesamtkontext der Digital Humanities zu Problemen, insbesondere wenn eine spezifische Auswahl von Kollektionen für konkrete Forschungsfragen zusammengefasst werden soll. Werden so beispielsweise Kollektionen aus archäologischen und kunsthistorischen Kontexten integriert, so führt die direkte Integration der spezifischen Datenstrukturen zu einem erhöhten Informationsgehalt gegenüber einer globalen Struktur, die den Fachspezifika nicht gerecht werden kann.

3. Föderationsarchitektur

Die in DARIAH-DE gewählte Architektur (Abbildung 1) besteht aus der Collection Registry zur Verzeichnung von Kollektionen, der Schema Registry zur Verwaltung von Schemata, und der Crosswalk Registry zur Beschreibung von Assoziationen zwischen verschiedenen Schemata. Integrative Dienste wie die generische Suche setzen für die Interpretation und Verarbeitung von Daten der verzeichneten Kollektionen auf den durch die Registries angebotenen Webservices auf.

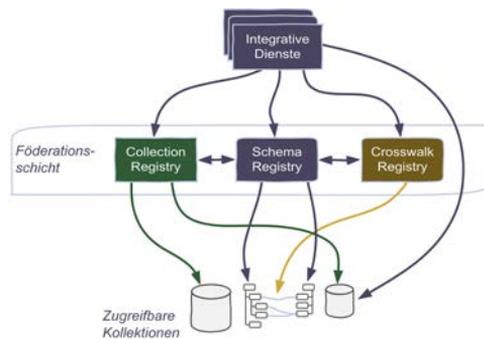


Abb. 1: Komponenten und Zusammenwirken der Föderationsarchitektur [eigene Darstellung].

Für eine Forscherin oder einen Forscher, die oder der eine Sammlung im Rahmen der Föderationsarchitektur registrieren und damit für die Suche, Analyse und den Vergleich mit

⁶ Grundsätzlich Lenzerini 2002, besonders S. 234

⁷ Überblickend Romanello 2013.

anderen Sammlungen zur Verfügung stellen möchte, ergibt sich im Zusammenspiel mit der generischen Suche ein Ablauf in vier Schritten (Abbildung 2):

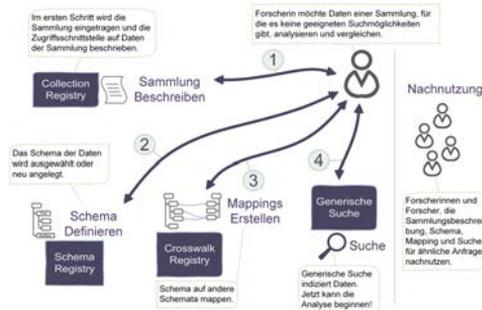


Abb. 2: Schritte der Registrierung von Kollektionen und Schemata [eigene Darstellung].

Die sich aus den einzelnen Schritten ergebenden Informationen stehen zur Nachnutzung für verwandte Forschungsinteressen zur Verfügung und können von integrativen Diensten über Webservices abgefragt werden.

3.1 Collection Registry

Die **Collection Registry** ist ein online zugängliches zentrales Verzeichnis, in dem relevante Sammlungen registriert und durch Fachwissenschaftler beschrieben werden. Das Datenmodell für die Sammlungsbeschreibungen basiert auf dem **Dublin Core Collection Application Profile**, das insbesondere im Hinblick auf die Beschreibung von Zugriffspunkten erweitert wurde. Eine beliebig heterogene Sammlung von Ressourcen wird als Collection bezeichnet und beschreibt ein Konstrukt der Anwendungsdomäne, das zur fachlichen Strukturierung von Archiven und Datenquellen eingesetzt werden kann. Collections können selbst direkt Ressourcen oder weitere untergeordnete Teilcollections beinhalten, und sie können sowohl physische als auch digitale Objekte oder nur Daten aggregieren. Die Sammlungsbeschreibungen decken neben Verschlagwortung, zeitlichen und geografischen Dimensionen auch Sammlungsformate und Informationen zur Datenpflege ab. Für die Auszeichnung mit Schlagworten und Georeferenzierungen werden verschiedene kontrollierte Vokabulare wie **LCHS**, **Dewey Decimal Classification**, **Geonames** u.a. integriert.

Ein Schwerpunkt liegt auf der Beschreibung von Zugriffspunkten wie OAI-PMH-Schnittstellen zur Abfrage der Sammlungselemente für die Weiterverarbeitung durch assoziierte Komponenten. Je Sammlungsbeschreibung können mehrere Zugriffspunkte verzeichnet werden, zu denen neben der URL des Zugriffspunktes auch weitere Angaben wie Zugriffsprotokoll, etwaige Zugriffsbeschränkungen, OAI-PMH subclasses, Dokumentation der Schnittstellen etc. verwaltet werden. Besonders wichtig ist, dass für jeden Zugriffspunkt das von diesem verwendete Schema aus der Schema Registry referenziert wird.

Weiterführende Dienste können alle erforderlichen Informationen für einen Zugriff auf die Sammlungselemente aus der Collection Registry über Webschnittstellen (REST) beziehen.⁸

Neben maschinenlesbaren Schnittstellen für den Zugriff auf die Sammlungsbeschreibungen bietet die Collection Registry ein Benutzerinterface, welches das Anlegen von Sammlungsbeschreibungen und anderen Datenobjekten ebenso unterstützt wie das Suchen, Aktualisieren und Löschen von vorhandenen Beschreibungen. Ausgewählte kontrollierte Vokabulare unterstützen die Eingabe, die Interaktion mit der Schema Registry erlaubt es, eine Sammlungsbeschreibung mit einem bestimmten Schema zu verknüpfen. Für den langfristigen Betrieb wird eine Moderation von DARIAH-DE organisiert, die die Qualität der Daten gewährleisten wird. Die Collection Registry nimmt eine zentrale Rolle in der Datenförderung ein, dient aber zugleich auch als alleinstehender Dienst und Datenquelle für die Suche und Verwaltung von Metadaten zu Datensammlungen.

3.2 Schema- und Crosswalk Registry

In der **Schema- und Crosswalk Registry** werden semi-strukturierte Datenmodelle und Korrelationen zwischen diesen (vgl. Abbildung 3) aus der primären Zielsetzung heraus beschrieben, expliziertes Expertenwissen zu Kollektionen und den darin verwalteten Daten nachnutzen zu können. Die Spezifikationen von Strukturen, z. B. in XML-Schemata, können hierbei in Bezug auf eine Kollektion erweitert und konkretisiert werden, wodurch die Semantik originärer Daten erhalten bleibt und dennoch eine Verfeinerung um zunächst implizites Hintergrundwissen erfolgen kann.

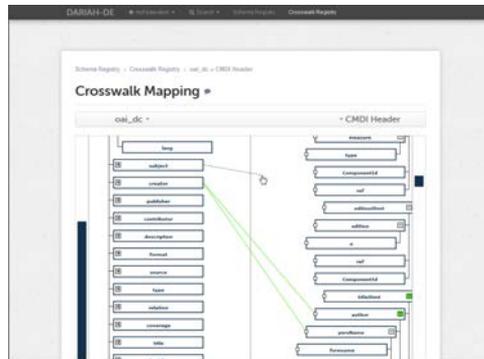


Abb. 3: Assoziation von Schemata in der Crosswalk Registry [eigene Darstellung].

Abbildung 4 zeigt beispielhaft Möglichkeiten zur Verfeinerung von Dublin Core basierend auf dem Wissen zu spezifischen Kollektionen. Manuell modellierte Verarbeitungsregeln führen dabei zu einer erweiterten Version eines Datensatzes, welcher für ein Mapping mit komplexeren Strukturen zur Verfügung steht. Dadurch, dass auch der unveränderte Datensatz

⁸ Plutte / Harms 2012, S. 11f.

weiterhin verwendet werden kann, wird zudem die Kompatibilität zu generischem Dublin Core sichergestellt.



Abb. 4: Beispiele zur kollektionsspezifischen Ergänzung von Dublin Core [eigene Darstellung].

Durch die semantische Assoziation von Kollektionen und darin verwendeten Schemata durch fachwissenschaftliche Experten können integrative Sichten generiert werden, die den Anforderungen konkreter Forschungsfragen entsprechen. Dies wird insbesondere dadurch erreicht, dass nicht technisch motivierte Integrationsziele wie die Vollständigkeit und Korrektheit eines Integrationsschemas im Vordergrund stehen, sondern disziplinspezifische und auch konfliktäre Zusammenhänge modelliert werden können.

Die flexible Anpassbarkeit an übergreifende oder spezifische Fragestellungen wird dabei durch das in Abbildung 5 exemplarisch angedeutete Konzept der forschungsorientierten Föderation digitaler Kollektionen erreicht: Kohärente Bereiche mit eng assoziierten Schemata und Kollektionen werden in der Abbildung durch semantische Cluster widerspiegelt und bilden die Voraussetzung für spezifische Betrachtungen. Für übergreifende Sichten werden wichtige Schemata der einzelnen Cluster (repräsentiert als S3, S5 und S8) mit generischen Schemata, wie z. B. Dublin Core – in der Abbildung symbolisiert durch S10 – assoziiert.

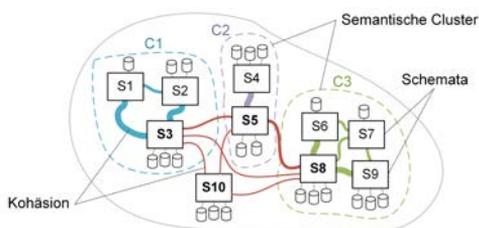


Abb. 5: Prinzip der semantischen Cluster im Beispiel [eigene Darstellung].

3.3 Generische Suche als durchgeführter Use-Case

Mit der generischen Suche wird im Rahmen von DARIAH-DE ein Anwendungsfall der Datenföderation umgesetzt. Hierbei werden Daten aus den in der Collection Registry verzeichneten Kollektionen nach den in der Schema Registry explizierten Strukturen verarbeitet und indiziert. Die Heterogenität der Ressourcen wird zum Zeitpunkt konkreter Suchanfragen basierend auf der zu durchsuchenden Menge von Kollektionen mit Hilfe der Crosswalk Registry aufgelöst.

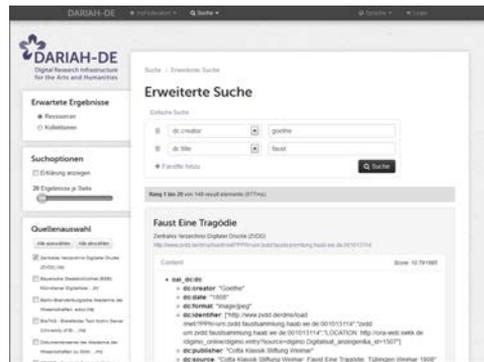


Abb. 6: Anfrageverarbeitung in der generischen Suche [eigene Darstellung].

Abbildung 6 skizziert den Verlauf der Anfrageverarbeitung und die Interaktion mit den Komponenten der Föderationsarchitektur: Am Beginn steht ein Informationsbedürfnis im Rahmen einer Forschungsfrage (1). Zunächst wird nun interaktiv oder automatisch auf Basis der Collection Registry und der von der generischen Suche angebotenen Kollektionssuche die Teilmenge der Kollektionen ermittelt, in denen die Suche durchgeführt werden soll (2). Je feingranularer die Schemata der gewählten Kollektionen in der Crosswalk Registry miteinander verknüpft sind, umso differenzierter können die Anfragen spezifiziert und ausgeführt werden. Der Nutzer kann die Anfrage dabei in einem Schema seiner Wahl formulieren, das als temporäres Integrationsmodell genutzt wird. Die Anfrage wird auf Basis der relevanten Schemainformationen und Transformationsregeln (3) dann so umformuliert, dass sie auf den Indices, die die Daten in ihrem ursprünglichen Schema verwalten, ausgeführt werden kann (4). Ermittelte Ergebnisse werden zusammengefasst und bezüglich ihrer Relevanz für die Anfrage sortiert.

Auf der Basis der eingeführten Föderationsarchitektur und der im Rahmen der generischen Suche implementierten Funktionalität können assoziierte Kollektionen im Rahmen so genannter *Benutzerkollektionen* zusammengestellt und inhaltspezifisch nachgenutzt werden. Neben der Betrachtung der Daten in externen Tools (wie z. B. dem bereits erwähnten Geobrowser), ermöglicht die generische Suche eine unmittelbare Veröffentlichung der Benutzerkollektion in Form einer so genannten Branded Search, einer eigenen Suchoberfläche, welche sowohl optisch als auch inhaltlich an spezifische Bedürfnisse angepasst ist.

Der Bildschirmausschnitt in Abbildung 7 verdeutlicht insbesondere die optische Abgrenzung von der generischen Suche durch eine konfigurierte Farbgebung und die Verwendung von Such- und Organisationslogos. Neben visuellen Aspekten unterscheidet sich die Branded Search auch in Hinblick auf die zu Grunde liegende Datenbasis: Die in einer Branded Search angebotenen Kollektionen spiegeln bei sämtlichen Such-, Analyse- und Visualisierungsaufgaben die von den Erstellern der Suche getroffene Kollektionsauswahl wider.



Abb. 7: Startseite der generischen Suche in Form einer Branded Search [eigene Darstellung].

4. Zusammenfassung

Die vorgestellte Förderationsarchitektur folgt dem Prinzip der dezentralen Integration von Daten.

Mit der generischen Suche kann gezeigt werden, wie durch die Verwendung der einzelnen Förderationskomponenten ein echter Mehrwert für die Recherche über verschiedene heterogene Datensammlungen entstehen und wie eine Alternative zu zentralistischen Ansätzen entwickelt werden kann. Mit einer Ad-hoc-Föderation kann gegenüber einer domänenweiten Harmonisierung die Möglichkeit der individuellen Integrierbarkeit von Daten geschaffen werden, die auf dem Wissen und der Kollaboration von Spezialisten aus verschiedenen Fachwissenschaften basiert und durch ein breites Publikum in Abhängigkeit von konkreten Forschungsfragen eingesetzt werden kann.

Bibliographische Angaben

A companion to Digital Humanities. Hg. von Susan Schreibman / Ray Siemens / John Unsworth. Oxford 2004. [[Nachweis im OPAC](#)]

Tobias Gradl / Andreas Henrich: DARIAH-DE Generische Suche (M 1.4.2.1 - Prototyp): DARIAH-DE Arbeitspapier. 2013. [[online](#), Registrierung erforderlich]

Andreas Henrich / Tobias Gradl: DARIAH(-DE): Digital Research Infrastructure for the Arts and Humanities – Concepts and Perspectives. In: International Journal of Humanities and Arts Computing 7 (2013), S. 47–58. [[Nachweis im GBV](#)]

Maurizio Lenzerini: Data Integration: A Theoretical Perspective. In: Proceedings of the twenty-first ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems. Hg. von der Association for Computing Machinery. New York 2002, S. 233–246. [[Nachweis im GBV](#)]

Silvio Peroni / Francesca Tomasi / Fabio Vitali: Reflecting on the Europeana Data Model. In: Digital Libraries and Archives. Hg. von Maristella Agosti / Floriana Esposito / Stefano Ferilli / Nicola Ferro. Berlin 2013 (= Communications in Computer and Information 354), S. 228–240. [[Nachweis im GBV](#)]

Christoph Plutte / Patrick Harms: Collection Registry (M 1.2.2): DARIAH-DE Arbeitspapier. 2012. [[online](#)]

Matteo Romanello: DARIAH Geo-browser: Exploring Data through Time and Space. 2013. [[online](#)]

Abbildungslegenden und -nachweise

Abb. 1: Komponenten und Zusammenwirken der Förderationsarchitektur [eigene Darstellung].

Abb. 2: Schritte der Registrierung von Kollektionen und Schemata [eigene Darstellung].

Abb. 3: Assoziation von Schemata in der Crosswalk Registry [eigene Darstellung].

Abb. 4: Beispiele zur kollektionsspezifischen Ergänzung von Dublin Core [eigene Darstellung].

Abb. 5: Prinzip der semantischen Cluster im Beispiel [eigene Darstellung].

Abb. 6: Anfrageverarbeitung in der generischen Suche [eigene Darstellung].

Abb. 7: Startseite der generischen Suche in Form einer Branded Search [eigene Darstellung].

ZfdG

Zeitschrift für digitale Geisteswissenschaften

Gesamtbibliographie aus:

Sonderband 1 der ZfdG: Grenzen und Möglichkeiten der Digital Humanities. Hg. von Constanze Baum und Thomas Stäcker. 2015. DOI: [10.17175/sb01](https://doi.org/10.17175/sb01)

Nachweis im OPAC der Herzog August Bibliothek:

Erstveröffentlichung:
19.02.2015

Sofern nicht anders angegeben 

Letzte Überprüfung aller Verweise:

Zitierweise:

Gesamtbibliographie. In: Grenzen und Möglichkeiten der Digital Humanities. Hg. von Constanze Baum / Thomas Stäcker. 2015 (= Sonderband der Zeitschrift für digitale Geisteswissenschaften, 1). PDF Format.

Bibliographie des ersten Sonderbands der ZfdG (2015)

A

- AAC - Austrian Academy Corpus: AAC-Fackel, Online Version: »Die Fackel. Herausgeber: Karl Kraus, Wien 1899–1936.« AAC Digital Edition No. 1. Hg. von Hanno Biber / Evelyn Breiteneder / Heinrich Kabas / Karlheinz Mörth. [online] Biber 2015.
- AAC - Austrian Academy Corpus and Brenner-Archiv: Brenner online, Online Version: »Der Brenner. Herausgeber: Ludwig Ficker, Innsbruck 1910–1954.« AAC Digital Edition No. 2. Hg. von Hanno Biber / Evelyn Breiteneder / Heinrich Kabas / Karlheinz Mörth. [online] Biber 2015.
- Activists Guide to Archiving Video, Version 1.0 (July 2013). [online] [Nachweis im GBV] Wübbena et al 2015.
- The Alcalá account book project, National University of Ireland. Maynooth 2008. [online] Vogeler 2015.
- Melina Alexa / Cornelia Zuell: Text Analysis Software: Commonalities, Differences and Limitations: The Results of a Review. In: *Quality and Quantity* 34 (2000) H. 3, S. 299–321. [Nachweis im OPAC] Blessing et al. 2015.
- Robert B. Allen / Robert Sieczkiewicz: How historians use historical newspapers. In: *Proceedings of the American Society for Information Science and Technology* 47 (2010), H. 1, S. 1–4. DOI 10.1002/meet.14504701131. [Nachweis im GBV] Lange 2015.
- Martin Andert / Jörg Ritter / Paul Molitor: Optimized platform for capturing metadata of historical correspondences. In: *Literary and Linguistic Computing* 6 (2014), S. 1–10. DOI 10.1093/lilc/fqu027. Biehl et al 2015.
- Annotations at Harvard – Video Annotations. Harvard Annotation Project 2010. [online] Wübbena et al 2015.
- Franz-Josef Arlinghaus: Bookkeeping, Double-Entry Bookkeeping. In: *Medieval Italy. An Encyclopedia*. Hg. von Christopher Kleinhenz. New York 2004, S. 147–150. [Nachweis im OPAC] Vogeler 2015.
- Matthias Arnold: Overview: Video Annotation Tools. [online] Wübbena et al 2015.
- Klaus Arnold: Geschichtswissenschaft und elektronische Datenverarbeitung. Methoden, Ergebnisse und Möglichkeiten einer neuen Hilfswissenschaft. In: *Methodenprobleme der Geschichtswissenschaft*. Hg. von Theodor Schieder. München 1974 (= *Historische Zeitschrift, Beiheft N.F. 3*), S. 98–148. [Nachweis im OPAC] Vogeler 2015.
- Apache Jena Documentation. Edited by The Apache Foundation. 2011–2014. [online] Heßbrüggen-Walter 2015.
- Avshalom Karasik: Applications of 3D technology as a research tool in archaeological ceramic analysis. In: *Beyond Illustration: 2D and 3D Digital Technologies as Tools for Discovery in Archaeology*. Hg. von Bernard Frischer / Anastasia Dakouri-Hild. Oxford 2008, S. 103–116. [Nachweis im GBV] Lange / Unold 2015.
- Paul Baker / Tony McEnery: A Corpus-Based Approach to Discourses of Refugees and Asylum Seekers in UN and Newspaper Texts. In: *Journal of Language and Politics* 4 (2005) H. 2, S. 197–226. [Nachweis im OPAC] Blessing et al. 2015.
- Syd Bauman: Transactionography Customized Documentation. In: *Encoding Historical Financial Records Open Access Library*. 20. Juli 2013. [online] Vogeler 2015.
- Itay Bar-Yosef / Issac Beckman / Klara Kedem / Itshak Dinstein: Binarization, character extraction, and writer identification of historical hebrew calligraphy documents. In: *International Journal on Document Analysis and Recognition* 9 (2007), 2, S. 89–99. [Nachweis im GBV] Fecker et al 2015.
- Mathieu Bastian / Sebastian Heymann / Mathieu Jacomy: Gephi: An Open Source Software for Exploring and Manipulating Networks. In: *Proceedings of the Third International ICWSM Conference*. 2009, S. 361–362. [online] Heßbrüggen-Walter 2015.
- Lionel Bently / Laura Biron: Discontinuities between legal conceptions of authorship and social practices. What, if anything, is to be done? In: *The Work of Authorship*. Hg. von Mireille Van Eechoud. Amsterdam 2014, S. 237–276. [online] Ernst 2015.
- Lilo Berg: Rotkäppchen 2.0 (Digital Humanities. Märchenhafte Chance oder Modetrend). In: *Humboldt Kosmos* 102 (2014), S. 12–23. [online] Nerbonne 2015.
- Jean-Francois Bergier / Luigi Solari: Histoire et élaboration statistique. L'exemple de la population de Genève au 15e siècle. In: *Mélanges d'histoire économique et social en hommage au Genève au Professeur Anthony Babel*. Bd. 1. Genf 1963, S. 197–225. [Nachweis im GBV] Vogeler 2015.
- Juliane Betz / Martina Engelbrecht / Christoph Klein / Raphael Rosenberg: Dem Auge auf der Spur: Eine historische und empirische Studie zur Blickbewegung beim Betrachten von Gemälden. In: *Image. Zeitschrift für interdisziplinäre Bildwissenschaft* 11 (2009), S. 29–41. [Nachweis im GBV] Kwastek 2015.
- Hanno Biber: »Ihre höchste Stilleistung ist die graphische Anordnung.« Zur Typographie der satirischen Zeitschrift *Die Fackel* von Karl Kraus. In: *Fontes Litterarum – Typographische Gestaltung und literarischer Ausdruck*. Hg. von Philipp S. Vanscheidt / Markus F. Polzer. Hildesheim 2014, S. 99–112. [Nachweis im GBV] Biber 2015.
- Hanno Biber: Die Komposition der *Fackel*. Diss. Universität Wien. Wien 2001. [online] Biber 2015.

- Douglas Biber: The Multi-Dimensional Approach to Linguistic Analyses of Genre Variation: An Overview of Methodology and Findings. In: *Computers and the Humanities* 26 (1992), H. 5/6, S. 331–345. [Nachweis im GBV] Viehhauser 2015.
- André Blessing / Jonas Kuhn: Textual Emigration Analysis (TEA). In: *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, European Language Resources Association (ELRA). Reykjavik 2014, S. 2089–2093. [Nachweis im GBV] Blessing et al. 2015.
- André Blessing / Jens Stegmann / Jonas Kuhn (2012a): SOA meets Relation Extraction: Less may be more in Interaction. In: *Proceedings of the Joint CLARIN-D/DARIAH Workshop at Digital Humanities Conference*. Hamburg 2013, S. 6–11. [online] Blessing et al. 2015.
- André Blessing / Jonathan Sonntag / Fritz Kliche / Ulrich Heid / Jonas Kuhn / Manfred Stede (2012b): Towards a Tool for Interactive Concept Building for Large Scale Analysis in the Humanities. In: *Proceedings of the 7th Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities Association for Computational Linguistics*. Sofia 2012, S. 55–64. [online] Blessing et al. 2015.
- Hartmut Bleumer / Caroline Emmelius: Generische Transgressionen und Interferenzen. Theoretische Konzepte und historische Phänomene zwischen Lyrik und Narrativ. In: *Lyrische Narrationen – narrative Lyrik. Gattungsinterferenzen in der mittelalterlichen Literatur*. Hg. von Hartmut Bleumer / Caroline Emmelius. Berlin, New York 2011, S. 1–39. [Nachweis im OPAC] Viehhauser 2015.
- Thomas Bögel / Jannik Strötgen / Michael Gertz: Computational Narratology. Extracting Tense Clusters from Narrative Texts, präsentiert auf der 9. Language Resources and Evaluation Conference (LREC'14). Reykjavik 2014. [online] Gius / Jacke 2015.
- Thomas Bögel / Michael Gertz / Evelyn Gius / Janina Jacke / Jan Christoph Meister / Marco Petris / Jannik Strötgen: Collaborative Text Annotation Meets Machine Learning. heureCLÉA, a Digital Heuristics of Narrative. In: *DHCommons Journal* 1 (2015) [online]. Gius / Jacke 2015.
- Bernd Bohnet: Very High Accuracy and Fast Dependency Parsing is not a Contradiction. In: *The 23rd International Conference on Computational Linguistics (COLING 2010)*. Beijing 2010, S. 89–97. [Nachweis im GBV] Blessing et al. 2015.
- Luise Borek / Ruth Reiche: Round Table »Annotation von digitalen Medien«. In: *DHd-Blog* vom 29. Juli 2014. [online] Wübbena et al 2015.
- Jürgen Bortz / Nicola Döring: Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler. 4. überarbeitete Auflage. Heidelberg 2006. [Nachweis im GBV] Jeschke et al 2015.
- Heidrun Elisabeth Boshof: Fest und Alltag in einem spätmittelalterlichen Damenstift. Das Reichsstift Obermünster in Regensburg im Spiegel seiner Rechnungen. In: *Beiträge zur Geschichte des Bistums Regensburg* 42 (2008), S. 187–258. [Nachweis im OPAC] Vogeler 2015.
- Jens Bove / Lutz Heusinger / Angela Kailus: Marburger Informations-, Dokumentations- und Administrationssystem (MIDAS). Handbuch und CD. München 2001. [Nachweis im GBV] Kwastek 2015.
- Manuel Braun: Kodieren, Annotieren, Theoretisieren. Zur Wiederannäherung von Literatur- und Sprachwissenschaft über Korpora. In: *Lili. Zeitschrift für Literaturwissenschaft und Linguistik* 43 (2013), H. 172, S. 83–90. [Nachweis im OPAC] Viehhauser 2015.
- Klaus Bredl / Amrei Groß: Gestaltung und Bewertung von Lernszenarien in immersiven virtuellen Welten. In: *Lernen und Lehren in virtuellen Welten. Zeitschrift für E-Learning – Lernkultur und Bildungstechnologie* 7 (2012), H. 1, S. 36–46. [Nachweis im GBV] Jeschke et al 2015.
- Johannes Breuer: Spielend Lernen? Eine Bestandsaufnahme zum (Digital) Game-Based Learning. Landesanstalt für Medien Nordrhein-Westfalen (LfM). Düsseldorf 2010. [online] Jeschke et al 2015.
- William R. Brubaker / Frederick Cooper: Beyond »Identity«. In: *Theory and Society* 29 (2000) H.1, S. 1–47. [Nachweis im GBV] Blessing et al. 2015.
- Frank Bruder: Pop-Splits – Die besten Songs aller Zeiten und ihre Geschichte. Berlin 2004. [Nachweis im GBV] Wübbena et al 2015.
- Kai-Christian Bruhn / Fredie Kern / Frithjof Schwartz: Spatial Humanities – Eine neue Forschungsdisziplin. In: *gis.Business* 1 (2014), S. 53–55. [Nachweis im GBV] Lange / Unold 2015.
- Cameron Buckner / Mathias Niepert / Colin Allen: From encyclopedia to ontology: Toward dynamic representation of the discipline of philosophy. In: *Synthese* 182 (2011), H. 2, p. 205–233. [online] Heißbrüggen-Walter 2015.
- Hans-Jürgen Bucher: Online-Interaktivität – Ein hybrider Begriff für eine hybride Kommunikationsform. In: *Interaktivität. Ein transdisziplinärer Schlüsselbegriff*. Hg. von Christoph Bieber / Claus Leggewie. Frankfurt/Main, New York 2004, S. 132–167. [Nachweis im GBV] Ernst 2015.
- Axel Bühler: Grundprobleme der Hermeneutik. In: *Hermeneutik. Basistexte zur Einführung in die wissenschaftstheoretischen Grundlagen von Verstehen und Interpretation*. Hg. von Axel Bühler. Heidelberg 2003, S. 3–19. [Nachweis im GBV] Gius / Jacke 2015.
- Marius Bulacu / Lambert Schomaker / Axel Brink: Text-Independent Writer Identification and Verification on Offline Arabic Handwriting. In: *Proceedings of the International Conference of Document Analysis and Recognition (ICDAR)*. Hg. von IEEE. Parana 2007, S. 769–773. DOI 10.1109/ICDAR.2007.4377019. [Nachweis im GBV] Fecker et al 2015.

- Bundesministerium der Justiz und für Verbraucherschutz: Gesetz über Urheberrecht und verwandte Schutzrechte. §1: Allgemeines. [\[online\]](#) Ernst 2015.
- Mercedes Bunz: Die stille Revolution. Wie Algorithmen Wissen, Arbeit, Öffentlichkeit und Politik verändern, ohne dabei viel Lärm zu machen. Berlin 2012. [\[Nachweis im GBV\]](#) Ernst 2015.
- Anne Burdick: »Notes on the Design«. In: Paratexts, Notes on the Design [2007], AAC - Austrian Academy Corpus: AAC-Fackel. Online Version: »Die Fackel. Herausgeber: Karl Kraus, Wien 1899-1936« AAC Digital Edition No. 1. Hg. von Hanno Biber / Evelyn Breiteneder / Heinrich Kabas / Karlheinz Mörth. [\[online\]](#) Biber 2015.
- Manuel Burghardt / Christian Wolff: Digital Humanities: Buzzword oder Strukturwandel der Geisteswissenschaften? In: Blick in die Wissenschaft - Forschungsmagazin der Universität Regensburg 29 (2014), S. 40. [\[online\]](#) Sahle 2015.
- John Burrows: All the Way Through: Testing for Authorship in Different Frequency Strata. In: Literary and Linguistic Computing 22 (2007), H. 1, S. 27–47. [\[Nachweis im GBV\]](#) Viehhauser 2015.
- John Burrows: »Delta: a Measure of Stylistic Difference and a Guide to Likely Authorship. In: Literary and Linguistic Computing 17 (2002), H. 3, S. 267–287. [\[Nachweis im GBV\]](#) Viehhauser 2015.
- Dietrich Busse / Wolfgang Teubert: Ist Diskurs ein sprachwissenschaftliches Objekt? Zur Methodenfrage der historischen Semantik. In: Begriffsgeschichte und Diskursgeschichte. Hg. von Dietrich Busse / Fritz Hermanns / Wolfgang Teubert. Opladen 1994, S. 10–28. [\[Nachweis im OPAC\]](#) Vogeler 2015.
- Hans Butzmann: Die Weissenburger Handschriften. Frankfurt/Main 1964 (= Kataloge der Herzog August Bibliothek Wolfenbüttel: Neue Reihe, Bd. 10). [\[Nachweis im OPAC\]](#) Fecker et al 2015.
- Dino Buzzetti: Digital Representation and the Text Model. In: New Literary History 33 (2002), S. 61–88. DOI 10.1353/nlh.2002.0003.Gius / Jacke 2015.

B

- Die Cameralia des Stadtarchivs Regensburg seit der Mitte des 14. Jahrhunderts. Bearb. von Heidrun Boshof / Regina Brunnhofer. Edition in Progress 2006ff. (= Fontes Civitatis Ratisponensis). [\[online\]](#) Vogeler 2015.
- Bernard Cerquiglini: Eloge de la variante. Histoire critique de la philologie. Paris 1989. [\[Nachweis im OPAC\]](#) Viehhauser 2015.
- Peter Cerwenka / Paul W. Roth: Münzurlaub des 16. Jahrhunderts im Raume des östlichen Österreich. Ein Anwendungsbeispiel der elektronischen Datenverarbeitung in der historischen Forschung. Graz 1972. [\[Nachweis im GBV\]](#) Vogeler 2015.
- Nathanael Chambers / Dan Jurafsky: Unsupervised Learning of Narrative Schemas and Their Participants, präsentiert auf der Association for Computational Linguistics Konferenz (ACL 2009). Singapur 2009. [\[online\]](#) Gius / Jacke 2015.
- David L. Chatford Clark: Viewing the liturgy: a space syntax study of changing visibility and accessibility in the development of the Byzantine church in Jordan. In: World Archeology 39 (2007), H. 1, S. 84–104. DOI: 10.1080/00438240601148798. [\[Nachweis im GBV\]](#) Lange 2015.
- Comédie-Française Register Project. Hg. vom Oxford Brookes und Massachusetts Institute of Technology. 2007ff. [\[online\]](#) Vogeler 2015.
- A Companion to Digital Literary Studies. Hg. von Ray Siemens / Susan Schreibman. Malden 2007. [\[online\]](#) Sahle 2015.
- Comptes de châtelainies savoyardes. Hg. von Christian Guilleré / Jean-Louis Gaulin / Marjorie Burghart. 2003ff. [\[online\]](#) Vogeler 2015.
- Les comptes des consuls de Montferrand (1273–1319). Bearb. von R. Anthony Lodge 2006 (= Éditions en ligne de l'École des Chartes 16). [\[online\]](#) Vogeler 2015.
- Thomas D. Cook / Donald T. Campbell: Quasi-Experimentation. Design & Analysis Issue for Field Settings. Chicago 1979. [\[Nachweis im GBV\]](#) Jeschke et al 2015.
- James H. Coombs / Allen H. Renear / Steven J. DeRose: Markup Systems and the Future of Scholarly Text Processing. In: Communications of the ACM 30 (1987), S. 933–947. DOI 10.1145/32206.32209.Gius / Jacke 2015.
- John Corrigan: Qualitative GIS and Emergent Semantics. In: The Spatial Humanities: GIS and the Future of Humanities Scholarship. Hg. von David J. Bodenhamer / John Corrigan / Trevor M. Harris Bloomington 2010, S. 76–88. [\[Nachweis im OPAC\]](#) Lange 2015.
- Hugh Craig / Arthur Kinney: Shakespeare, Computers, and the Mystery of Authorship. Cambridge 2009. [\[Nachweis im GBV\]](#) Viehhauser 2015.
- Creative Commons: Namensnennung 3.0 Deutschland (CC BY 3.0 DE). [\[online\]](#) Ernst 2015.
- Fabian Cremer: Erfolg und Emergenz in den Digital Humanities – ein Tagungseindruckschnitt. In: DHD-Blog vom 29. März 2014. [\[online\]](#) Wübbena et al 2015.
- Robin Curtis / Christiane Voss: Theorien ästhetischer Immersion. In: Montage AV. Zeitschrift für Theorie und Geschichte audiovisueller Kommunikation (2008), S. 3–10. [\[Nachweis im GBV\]](#) Jeschke et al 2015.

C

- Lutz Danneberg: Zum Autorkonstrukt und zu einer methodologischen Konzeption der Autorintention. In: Rückkehr des Autors. Zur Erneuerung eines umstrittenen Begriffs. Hg. von Fotis Jannidis / Gerhard Lauer / Matías Martínez / Simone Winko. Tübingen 1999, S. 77–105. [\[Nachweis im OPAC\]](#) Gius / Jacke 2015.
- de.hypotheses.org. Wissenschaftliche Blogs: Redaktion. [\[online\]](#) Ernst 2015.
- Debates in the Digital Humanities. Hg. von Matthew K. Gold. Minneapolis 2012. [\[Nachweis im OPAC\]](#) Ernst 2015.
- Defining Digital Humanities. A Reader. Hg. von Melissa Terras / Julianne Nyhan / Edward Vanhoutte. Farnham 2013. [\[Nachweis im OPAC\]](#) Gius / Jacke 2015. Sahle 2015.
- Diverse Autoren: GuttenPlag Wiki. [\[online\]](#) Ernst 2015.
- Diverse Autoren: Twitter-Hashtag #dhd2014. [\[online\]](#) Ernst 2015.
- Diverse Autoren: VroniPlag Wiki. [\[online\]](#) Ernst 2015.
- Defining Digital Humanities: A Reader. Hg. von Edward Vanhoutte / Julianne Nyhan / Melissa Terras. London 2013. [\[Nachweis im GBV\]](#) Nerbonne 2015.
- Digitale Edition der Basler Jahrbuchsbücher (1535/36 bis 1611). Hg. von Susanna Burghartz. Alpha Version 2014. [\[online\]](#) Vogeler 2015.
- Definition of the CIDOC Conceptual Reference Model, Version 5.0.4. Hg. von Patrick Le Boeuf / Martin Doerr / Christian Emil Ore / Stephen Stead. 2011. [\[online\]](#) Vogeler 2015.
- John Dillon: Review of Aldo Brancacci (ed.): Philosophy and Doxography in the Imperial Age. Firenze 2005. In: Bryn Mawr Classical Review 12 (2006). [\[online\]](#) Heßbrüggen-Walter 2015.
- Wilhelm Dilthey: Die Entstehung der Hermeneutik [1900]. In: Wilhelm Dilthey: Gesammelte Schriften. Bd. 5. 8. Auflage. Stuttgart, Göttingen 1990, S. 317–338. [\[Nachweis im GBV\]](#) Gius / Jacke 2015.
- Wilhelm Dilthey: Einleitung in die Geisteswissenschaften. Versuch einer Grundlegung für das Studium der Gesellschaft und der Geschichte. In: Gesammelte Schriften 1. 1875–1900. Stuttgart 1959 (zuerst 1883). [\[online\]](#) Nerbonne 2015.
- Steven J. DeRose / David G. Durand / Elli Mylonas / Allen H. Renear: What is text, really? In: Journal of Computing in Higher Education 1,2 (1990), S. 3–26. [\[Nachweis im GBV\]](#) Vogeler 2015.
- Burcu Dogramaci / Karin Wimmer: Vorwort. In: Netzwerke des Exils. Hg. von Burcu Dogramaci / Karin Wimmer. Berlin 2011, S. 9–11. [\[Nachweis im GBV\]](#) Biehl et al 2015.
- Jörg Döring / Tristan Thielmann: Einleitung. Der *Spatial Turn* und das geheime Wissen der Geographen. In: Spatial Turn: das Raumparadigma in den Kultur- und Sozialwissenschaften. Hg. von Jörg Döring / Tristan Thielmann. Bielefeld 2009, S. 7–45. [\[Nachweis im OPAC\]](#) Lange / Unold 2015.
- Thomas Dreier / Ansgar Ohly: Lehren aus der Vergangenheit – Perspektiven für die Zukunft. In: Plagiate. Wissenschaftsethik und Recht. Hg. von Thomas Dreier / Ansgar Ohly. Tübingen 2013, S. 155–184. [\[Nachweis im GBV\]](#) Ernst 2015.
- Johanna Drucker: Graphesis. Visual Forms of Knowledge Production. Cambridge, MA 2014. [\[Nachweis im GBV\]](#) Biber 2015.
- Nick Drummond / Rob Shearer: The Open World Assumption or Sometimes its nice to know what we don't know. Manchester 2006. [\[online\]](#) Heßbrüggen-Walter 2015.
- Katharina de la Durantaye: Allgemeine Bildungs- und Wissenschaftsschranke. Münster 2014. [\[online\]](#) Ernst 2015.
- M. Duerst / M. Suignard: RFC 3987: Internationalized Resource Identifiers (IRIs). 2006. [\[online\]](#) Heßbrüggen-Walter 2015.
- Emile Durkheim: The Rules of Sociological Method. New York 1964 [zuerst 1895]. [\[Nachweis im GBV\]](#) Blessing et al. 2015.

E

- Early Forward Exchange Markets, 1876–1914. In: EH.net - Economic History Services [\[online\]](#) Vogeler 2015.
- Walter Edelmann: Lernpsychologie. 5. Auflage. Weinheim 1996. [\[Nachweis im GBV\]](#) Jeschke et al 2015.
- Maciej Eder / Jan Rybicki / Mike Kestemont: ›Stylo: a package for stylometric analyses. August 17, 2014. [\[online\]](#) Viehhauser 2015.
- Maciej Eder / Mike Kestemont / Jan Rybicki: Stylometry with R: A Suite of Tools. Digital Humanities 2013: Conference Abstracts. Lincoln 2013, S. 487–89. [\[online\]](#) Nerbonne 2015.
- Maciej Eder / Mike Kestemont / Jan Rybicki: Stylometry with R: a suite of tools. In: Digital Humanities 2013: Conference Abstracts. University of Nebraska-Lincoln, NE, S. 487–489. [\[Nachweis im GBV\]](#) Viehhauser 2015.

- Maciej Eder: Mind Your Corpus: systematic errors in authorship attribution. In: *Literary and Linguistic Computing* 28 (2013), H. 4, S. 603–614. [[Nachweis im GBV](#)] [Viehhauser 2015](#).
- Edition des Steierischen Marchfütterurbars von 1414/1426. Bearb. von Matthias Persting. Testversion [im Erscheinen]. [[online](#)] [Vogeler 2015](#).
- Elizabeth Eisenstein: *The Printing Press as an Agent of Change*. 2 Bde. Cambridge 1979. [[Nachweis im OPAC](#)] [Nerbonne 2015](#).
- Anders Eriksson: Swedia-projektet: dialektforskning i ett jämförande perspektiv. In: *Folkmålsstudier* 43 (2004), S. 11–32. [[online](#)] [Nerbonne 2015](#).
- Thomas Ernst: Jenseits des wissenschaftlichen Werks und des geistigen Eigentums? Die digitale Verbreitung wissenschaftlichen Wissens. Audio-Podcast des Vortrags vom 26. März 2014 auf der Konferenz Digital Humanities – methodischer Brückenschlag oder ›feindliche Übernahme‹? Chancen und Risiken der Begegnung zwischen Geisteswissenschaften und Informatik. [[online](#)] [Ernst 2015](#).
- Thomas Ernst: E-Books, Netzliteratur und das Urheberrecht. In: *SDV – Sprache und Datenverarbeitung. International Journal for Language Data Processing* 36 (2012), H. 2, S. 43–52. [[Nachweis im GBV](#)] [Ernst 2015](#).
- Johannes Evelein: Briefkultur und Exil. In: *Erste Briefe aus dem Exil 1945–1950*. Hg. von Primus-Heinz Kucher / Johannes F. Evelein / Helga Schreckenberger. München 2011, S. 15–31. [[Nachweis im GBV](#)] [Biehl et al 2015](#).
- Daniel Ewert / Katharina Schuster / Daniel Schilberg / Sabine Jeschke: Intensifying learner's experience by incorporating the virtual theatre into engineering education. In: *Proceedings of the 2013 IEEE Global Engineering Education Conference (EDUCON 2013)*. Hg. von Institute of Electrical and Electronics Engineers (IEEE). Berlin 2013, S. 207–212. [[Nachweis im GBV](#)] [Jeschke et al 2015](#).
- eXtensible Business Reporting Language Specifications, version 2.1. Hg. von Philipp Engel et al. 20. Februar 2013. [[online](#)] [Vogeler 2015](#).

F

- Reinhard Fatke / Jürgen Oelkers: Das Selbstverständnis der Erziehungswissenschaft: Geschichte und Gegenwart. In: *Zeitschrift für Pädagogik* 60 (2014), Beiheft 60, S. 7–13. [[Nachweis im GBV](#)] [Jeschke et al 2015](#).
- Kirstin Faupel-Dreves: Bildraum oder Kultraum? Symbolische und liturgische Raumgestaltung im Rationale *divinorum officiorum* des Durandus von Mendel. In: *Raum und Raumvorstellungen im Mittelalter*. Hg. von Jan Aertsen / Andreas Speer. Berlin 1998, S. 665–684. [[Nachweis im OPAC](#)] [Lange / Unold 2015](#).
- Daniel Fecker / Abed Asi / Volker Märgner / Jihad El-Sana / Tim Fingscheidt: Writer Identification for Historical Arabic Documents. In: *Proceedings of the International Conference on Pattern Recognition (ICPR)*. Stockholm 2014, S. 3050–3055. [[Nachweis im GBV](#)] [Fecker et al 2015](#).
- Johann Gottlieb Fichte: Beweis der Unrechtmässigkeit des Büchernachdrucks. Ein Raisonement und eine Parabel. In: *Johann Gottlieb Fichte: Gesamtausgabe der Bayerischen Akademie der Wissenschaften*. Hg. von Reinhard Lauth / Hans Jacob. Band I.1.: Werke 1791–1794. Stuttgart 1964, S. 409–426. [[Nachweis im OPAC](#)] [Ernst 2015](#).
- Kathleen Fitzpatrick: *Planned Obsolescence. Publishing, Technology, and the Future of the Academy*. New York, London 2011. [[Nachweis im OPAC](#)] [Ernst 2015](#).
- Kathleen Fitzpatrick: *Planned Obsolescence. Publishing, Technology, and the Future of the Academy*, auf der Plattform: Media-Commons Press. Open Scholarship in Open Formats. [[online](#)] [Ernst 2015](#).
- Julia Flanders: The Productive Unease of 21st-century Digital Scholarship. In: *Defining Digital Humanities. A Reader*. Hg. von Melissa Terras / Julianne Nyhan / Edward Vanhoutte. Farnham 2013, S. 205–218. [[Nachweis im OPAC](#)] [Ernst 2015](#).
- Marc Flandreau / John Komlos: Target Zones in Theory and History. Credibility, Efficiency, and Policy Autonomy. In: *Journal of Monetary Economics* 53 (2006), S. 1979–1995. DOI: 10.1016/j.jmoneco.2005.10.015. [[Nachweis im GBV](#)] [Vogeler 2015](#).
- Rohna Flin / Rona Patey / Ronnie Glavin / Nikki Maran: Anaesthetists' non-technical skills, in: *British Journal of Anaesthesia* 105 (2010), H. 1, S. 38–44. [[Nachweis im GBV](#)] [Jeschke et al 2015](#).
- Vilém Flusser: *Medienkultur*. Hg. von Stefan Bollmann. 2. Auflage. Frankfurt/Main 1999. [[Nachweis im OPAC](#)] [Ernst 2015](#).
- Dagfinn Føllesdal / Lars Walløe / Jan Elster: Die hypothetisch-deduktive Methode in der Literaturinterpretation. In: *Moderne Interpretationstheorien. Ein Reader*. Hg. von Tom Kindt / Tilmann Köppe. Göttingen 2008, S. 67–78. [[Nachweis im GBV](#)] [Gius / Jacke 2015](#).
- Michel Foucault: Was ist ein Autor? In: *Michel Foucault: Schriften zur Literatur*. Aus dem Französischen von Karin von Hofer / Anneliese Botond. Frankfurt/Main 1988, S. 7–31. [[Nachweis im OPAC](#)] [Ernst 2015](#). [Viehhauser 2015](#).
- Harald Fricke: Methoden? Prämissen? Argumentationsweisen. In: *Vom Umgang mit Literatur und Literaturgeschichte. Positionen und Perspektiven nach der Theoriedebatte*. Hg. von Lutz Danneberg / Friedrich Vollhardt. Stuttgart 1992, S. 211–227. [[Nachweis im GBV](#)] [Gius / Jacke 2015](#).
- Gerd Fritz: *Dynamische Texttheorie*. Gießen 2013. [[online](#)] [Vogeler 2015](#).

Rüdiger Fuchs: Die Inschriften der Stadt Worms. Gesammelt und bearbeitet von Rüdiger Fuchs. Wiesbaden 1991 (= Die Deutschen Inschriften, Bd. 29). [\[Nachweis im OPAC\] Lange 2015.](#)

G

Constantinos Gabrielatos: Selecting Query Terms to Build a Specialised Corpus from a Restricted-Access Database. In: ICAME Journal 31 (2007), S. 5–43. [\[Nachweis im GBV\] Blessing et al. 2015.](#)

Julia Ganitseva / Kai-Christian Bruhn / Fredie Kern / Martin Unold: Outlook on TLS - based 3D Building Modeling. In: Photogrammetrie, Laserscanning, Optische 3D-Messtechnik. Hg. von Thomas Luhmann / Christina Müller. Heidelberg 2014, S. 224–233. [\[Nachweis im GBV\] Lange / Unold 2015.](#)

Hans-Georg Gadamer: Hermeneutik. Wahrheit und Methode [1960]. In: Hans-Georg Gadamer: Gesammelte Werke, Bd. 2. 2., durchgesehene Auflage. Tübingen 1993. [\[Nachweis im GBV\] Gius / Jacke 2015.](#)

Gabriel García Márquez: Ein sehr alter Mann mit riesengroßen Flügeln (Un señor muy viejo con unas alas enormes). In: Gabriel García Márquez: Laubsturm. Frankfurt/Main 2004. [\[Nachweis im GBV\] Wübbena et al 2015.](#)

Gérard Genette: Paratexte. Das Buch vom Beiwerk des Buches. Mit einem Vorwort von Harald Weinrich. Aus dem Französischen von Dieter Hornig. Frankfurt/Main 2001. [\[Nachweis im OPAC\] Ernst 2015.](#)

Gérard Genette: Discours du Récit. In: Gérard Genette: Figures III. Paris 1972, S. 67–282. [\[Nachweis im OPAC\] Gius / Jacke 2015.](#)

Ute Gerhard: Unerhört. Die Geschichte der deutschen Frauenbewegung. Reinbeck 1991. [\[Nachweis im OPAC\] Vogeler 2015.](#)

Alexander Geyken: Methoden bei der Wörterbuchplanung in Zeiten der Internetlexikographie. In: Lexicographica 30 (2014), S. 73–111. [\[Nachweis im GBV\] Vogeler 2015.](#)

Evelyn Gius: Erzählen über Konflikte: ein Beitrag zur digitalen Narratologie. Berlin, Boston 2015. [\[Nachweis im GBV\] Gius / Jacke 2015.](#)

Evelyn Gius / Janina Jacke: Zur Annotation narratologischer Kategorien der Zeit. Guidelines zur Nutzung des CATMA-Tagsets. Hamburg 2014. [\[online\] Gius / Jacke 2015.](#)

Thomas Gloning: Diskursive Praktiken, Textorganisation und Wortgebrauch im Umkreis der ersten Frauenbewegung. In: Historische Pragmatik. Hg. von Peter Ernst. Berlin, Boston 2012, S. 127–146 (= Jahrbuch für germanistische Sprachgeschichte 3). [\[Nachweis im GBV\] Vogeler 2015.](#)

Matthew K. Gold: The Digital Humanities Moment. Die Expanded Open Access-Edition des Buches *Debates in the Digital Humanities*. [\[online\] Ernst 2015.](#)

Jeff Gomez: Print is Dead. Books in Our Digital Age. New York 2008. [\[Nachweis im OPAC\] Ernst 2015.](#)

Jack Goody: What's in a List? In: Jack Goody: The Domestication of the Savage Mind. Cambridge 1977 (= Themes in the social sciences), S. 74–111. [\[Nachweis im GBV\] Vogeler 2015.](#)

Colin Gordon: Mapping Decline: St. Louis and the fate of the American city. Philadelphia 2008. [\[Nachweis im GBV\] Lange / Unold 2015.](#)

Annette Gotscharek / Ulrich Reffle / Christoph Ringlstetter / Klaus Schulz / Andreas Neumann: Towards information retrieval on historical document collections: the role of matching procedures and special lexica. In: International Journal on Document Analysis and Recognition (IJ DAR) 14.2 (2011), S. 159–171. [\[Nachweis im GBV\] Nerbonne 2015.](#)

György Granasztói: Computerized Analysis of a Medieval Tax Roll. In: Acta Historica Academiae scientiarum Hungaricae 17 (1971), S. 13–24. [\[Nachweis im GBV\] Vogeler 2015.](#)

Sam Griffiths: The Use of Space Syntax in Historical Research: current practice and future possibilities. In: Proceedings of the Eighth International Space Syntax Symposium. Santiago 2012. [\[online\] Lange / Unold 2015.](#)

Rudolf Groner / Simon Raess / Phillip Sury: Usability: Systematische Gestaltung und Optimierung von Benutzerschnittstellen. In: Medienpsychologie. Hg. von B. Batinic / M. Appel. Heidelberg 2008, S. 425–446. [\[Nachweis im GBV\] Jeschke et al 2015.](#)

Valentin Groebner: Wissenschaftssprache digital. Die Zukunft von gestern. Konstanz 2014. [\[Nachweis im OPAC\] Ernst 2015.](#)

Klaus Grubmüller: Ich als Rolle. »Subjektivität: als höfische Kategorie im Minnesang? In: Höfische Literatur – Hofgesellschaft – Höfische Lebensformen um 1200. Hg. von Gert Kaiser / Jan-Dirk Müller. Düsseldorf 1986, S. 387–408. [\[Nachweis im OPAC\] Viehhauser 2015.](#)

Klaus Grubmüller: Gattungskonstitution im Mittelalter. In: Mittelalterliche Literatur und Kunst im Spannungsfeld von Hof und Kloster. Ergebnisse der Berliner Tagung, 9.–11. Oktober 1997. Hg. von Nigel F. Palmer / Hans-Jochen Schiewer. Tübingen 1999, S. 193–210. [\[Nachweis im OPAC\] Viehhauser 2015.](#)

Guide to the Software Engineering Body of Knowledge, Version 3.0. Hg. von Pierre Bourque / Richard E. Fairley. IEEE Computer Society, 2014. [\[online\] Nerbonne 2015.](#)

Rolf Gundlach / Carl August Lueckerath: Historische Wissenschaften und Elektronische Datenverarbeitung. Frankfurt/Main u.a. 1976. [\[Nachweis im GBV\] Vogeler 2015.](#)

H

- Susanne Haaf / Alexander Geyken / Frank Wiegand: The DTA-Base Format: A TEI Subset for the Compilation of a Large Reference Corpus of Printed Text from Multiple Sources. In: *Journal of the Text Encoding Initiative (JTEI)*, 8 (2015). DOI: 10.4000/jtei.1114 [online] Geyken / Gloning 2015.
- Hans Van Halteren / Harald Baayen / Fiona Tweedie / Marco Haverkort / Anneke Neijt: New machine learning methods demonstrate the existence of a human stylome. In: *Journal of Quantitative Linguistics* 12.1 (2005), S. 65–77. [Nachweis im GBV] Nerbonne 2015.
- Handbuch der Frauenbewegung. Hg. von Helene Lange / Gertrud Bäumer. Band I–V. Berlin 1901–1906. [Nachweis im GBV] Vogeler 2015.
- Thomas Haigh: We Have Never Been Digital. In: *Communications of the ACM* 57 (2014), H. 9, S. 24–28. DOI: 10.1145/2644148. Sahle 2015.
- Heiko Hartmann: Darstellungsmittel und Darstellungsformen in den erzählenden Werken. In: Wolfram von Eschenbach. Ein Handbuch. Hg. von Joachim Heinze. Berlin, Boston 2011, S. 145–220. [Nachweis im OPAC] Viehhauser 2015.
- Jens Haustein: Mediävistische Stilforschung und die Präsenzkultur des Mittelalters. Mit einem Ausblick auf Gottfried von Straßburg und Konrad von Würzburg. In: *Textprofile stilistisch. Beiträge zur literarischen Evolution*. Hg. von Ulrich Breuer / Bernhard Spies. Bielefeld 2011, S. 43–60. [Nachweis im GBV] Viehhauser 2015.
- Bernhard Haslhofer / Robert Sanderson / Rainer Simon / Herbert van de Sompel: Open Annotations on Multimedia Web Resources. In: *Multimedia Tools and Applications* 70 (2014) H. 2, S. 847–867. doi:10.1007/s11042-012-1098-9. [Nachweis im GBV] Wübbena et al 2015.
- Abdelaali Hassane / Somaya Al-Madeed: ICFHR 2012 Competition on Writer Identification Challenge 2: Arabic Scripts. In: *Proceedings of the International Conference on Frontiers in Handwriting Recognition (ICFHR)*. Hg. von IEEE. Bari 2012, S. 835–840. DOI 10.1109/ICFHR.2012.218. [Nachweis im GBV] Fecker et al 2015.
- Sven Havemann / Volker Settgest / Rene Berndt: The Arrigo Showcase Reloaded – Towards a Sustainable Link between 3D and Semantics. In: *ACM Journal for Computing in Cultural Heritage* 2 (2009), H. 1, Artikel Nr. 4 [unpag.]. DOI 10.1145/1551676.1551679. Lange / Unold 2015.
- Thomas Hays / Stephan Müller: Mittelalter-Philologie im Internet. Digitalisierung mittelalterlicher Handschriften aus Sicht der Forschung. In: *Zeitschrift für deutsches Altertum und deutsche Literatur* 140 (2011), S. 416–420. [Nachweis im OPAC] Fecker et al 2015.
- Friedrich Hebbel: Matteo [1841]. In: *Werke*. Hg. von Gerhard Fricke / Werner Keller / Karl Pörnbacher. Bd. 3. München 1963, S. 465–476. [online] Gius / Jacke 2015.
- Andreas Hebbel-Seeger: Motiv: Motivation?! – Warum Lernen in virtuellen Welten (trotzdem) funktioniert. In: *Lernen und Lehren in virtuellen Welten. Zeitschrift für E-Learning – Lernkultur und Bildungstechnologie* 7 (2012), H.1, S. 23–35. [Nachweis im GBV] Jeschke et al 2015.
- Georg Wilhelm Friedrich Hegel: *Phänomenologie des Geistes*. Hamburg 1988 (zuerst 1806/1807). [Nachweis im GBV] Nerbonne 2015.
- Christiane Heibach: *Literatur im elektronischen Raum*. Frankfurt/Main 2003. [Nachweis im GBV] Ernst 2015.
- Stefan Heidenreich: Form und Filter – Algorithmen der Bilderverarbeitung und Stilanalyse. In: *zeitenblicke* 2 (2003), Nr. 1. [online] Kwastek 2015.
- Wolfgang Ehrhardt Heindold: *Bücher und Buchmacher*. Verlage in der Informationsgesellschaft. 5. Auflage. Heidelberg 2001. [Nachweis im OPAC] Ernst 2015.
- Wolfgang Herborn / Klaus J. Mattheier: Sozialhistorische und sprachgeschichtliche Aspekte eines frühneuzeitlichen Rechnungsbuches der Kölner Kronenburse. In: *Rheinisch-Westfälische Zeitschrift für Volkskunde* 24 (1978), S. 140–182. [Nachweis im GBV] Vogeler 2015.
- Hubert Herkommer: Heiliges Antlitz und heiliges Maß: Zu Aussehen und Größe des Menschensohnes. In: *Das Heilige Grab, das Heilige Kreuz und die Wahre Länge Christi*. Schwäbisch Gmünd 2008 (= Katalogreihe Museum im Prediger, Bd. 35), S. 67–105. [Nachweis im GBV] Lange / Unold 2015.
- Sorin Hermon: Reasoning in 3D: a critical appraisal of the role of 3D modelling and virtual reconstructions in archaeology. In: *Beyond Illustration: 2D and 3D Digital Technologies as Tools for Discovery in Archaeology*. Hg. von Bernard Frischer / Anastasia Dakouri-Hild. Oxford 2008, S. 35–44. [Nachweis im GBV] Lange / Unold 2015.
- David Herlihy / Christiane Klapisch-Zuber: *Les Toscans et leur familles*. Paris 1978. [Nachweis im GBV] Vogeler 2015.
- Richard K. Herrmann: Linking Theory to Evidence in International Relations. In: *Handbook of International Relations*. Hg. von Walter Carlsnaes / Thomas Risse / Beth A. Simmons. London 2002, S. 119–136. [Nachweis im GBV] Blessing et al. 2015.
- Henry III fine rolls Project. Hg. von The National Archives / King's College London. 2009. [online] Vogeler 2015.
- Wolfgang Heß: Rechnung Legen mit Rechenpfennigen. In: *Numismatisches Nachrichtenblatt* 45/4 (1996), S. 11–20. [Nachweis im GBV] Vogeler 2015.

Wolfgang Heß: Rechnung Legen auf Linien. Rechenbrett und Zählisch in der Verwaltungspraxis in Spätmittelalter und Neuzeit. In: Städtisches Haushalts- und Rechnungswesen. Hg. von Erich Maschke / Jürgen Sydow. Sigmaringen 1977 (= Stadt in der Geschichte 2), S. 69–82. [\[Nachweis im OPAC\] Vogeler 2015.](#)

Stefan Heßbrüggen-Walter: Tatsachen im semantischen Web: Nanopublikationen in den digitalen Geisteswissenschaften? In: Historyblogosphere: Bloggen in den Geschichtswissenschaften. Edited by Peter Haber / Eva Pfanzelter. München 2013, p. 149–160. [\[Nachweis im GBV\] Heßbrüggen-Walter 2015.](#)

Annika Hinze / Ralf Heese / Markus Luczak-Rösch / Adrian Paschke: Semantic enrichment by non-experts: usability of manual annotation tools. In: Proceedings of the 11th international conference on The Semantic Web. Part I. Berlin, Heidelberg 2012, S. 165–181. [\[Nachweis im GBV\] Lange / Unold 2015.](#)

Graeme Hirst / Olga Feiguina: Bigrams of syntactic labels for authorship discrimination of short texts. In: Literary and Linguistic Computing 22.4 (2007), S. 405–417. [\[Nachweis im GBV\] Nerbonne 2015.](#)

»Hidden Histories: Computing and the Humanities c. 1965–1985«. Digital Humanities Quarterly-Sonderausgabe 6.3 (2012). [\[online\] Sahle 2015.](#)

Max Hoffmann / Katharina Schuster / Daniel Schilberg / Sabine Jeschke: Bridging the Gap between Students and Laboratory Experiments. In: Virtual, Augmented and Mixed Reality. Applications of Virtual and Augmented Reality, 6th International Conference, VAMR 2014, Held as Part of HCI International 2014, Heraklion, Greece, June 22–27, 2014, Proceedings, Part II. Hg. von R. Shumaker / S. Lackey. Heidelberg, S. 39–50. [\[Nachweis im GBV\] Jeschke et al 2015.](#)

Susan Höntzsch / Uwe Katzky / Klaus Bredl / Frank Kappe / Dirk Krause: Simulationen und simulierte Welten. Lernen in immersiven Lernumgebungen. In: Martin Ebner / Sandra Schön: Lehrbuch für Lernen und Lernen mit Technologien. 2013. [\[online\] Jeschke et al 2015.](#)

Johannes Hohlfeld: Stadtrechnungen als historische Quellen. Ein Beitrag zur Quellenkunde des ausgehenden Mittelalters, dargestellt am Beispiel der Pegauer Stadtrechnungen des 14. und 15. Jahrhundert. Leipzig 1912 (= Bibliothek der sächsischen Geschichte und Landeskunde 4,1). [\[Nachweis im GBV\] Vogeler 2015.](#)

David I. Holmes: Authorship Attribution. In: Computers and the Humanities 28 (1994), H. 2, S. 87–106. [\[Nachweis im GBV\] Viehhauser 2015.](#)

John E. Hopcroft: Computer science: the emergence of a discipline (Turing Award Lecture). In: Communications of the ACM 30 (1987), S. 198–202. [\[Nachweis im GBV\] Nerbonne 2015.](#)

HRK: Empfehlung der 14. Mitgliederversammlung der Hochschulrektorenkonferenz am 14. Mai 2013 in Nürnberg. Gute wissenschaftliche Praxis an deutschen Hochschulen. [\[online\] Ernst 2015.](#)

Paul Huys: Een Deense baljuwsrekening (1377–1379) als spiegeling van het dagelijks leven. In: Bijdragen tot de geschiedenis van Deenze en de Leiestreek 65 (1998), S. 185–218. [\[Nachweis im GBV\] Vogeler 2015.](#)

I

Identifying People in the Past. Hg. von Edward A. Wrigley. London 1973. [\[Nachweis im GBV\] Vogeler 2015.](#)

Informatique et histoire médiévale. Hg. von Lucie Fossier / A. Vauchez / C. Violante. Actes du Colloque de Rome, 20–22 mai 1975. Roma 1977 (= Publications de l'École Française de Rome 31). [\[Nachweis im GBV\] Vogeler 2015.](#)

J

Janina Jacke: Is There a Context-Free Way of Understanding Texts? The Case of Structuralist Narratology. In: Journal of Literary Theory 8 (2014), S. 118–139. [\[Nachweis im GBV\] Gius / Jacke 2015.](#)

Werner Jacobsen: Organisationsformen des Sanktuariums im spätantiken und mittelalterlichen Kirchenbau. Wechselwirkungen von Altarraum und Liturgie aus kunsthistorischer Perspektive. In: Kölnische Liturgie und ihre Geschichte. Studien zur interdisziplinären Erforschung des Gottesdienstes im Erzbistum Köln. Hg. von Albert Gerhards / Andreas Odenthal. Münster 2000 (= Liturgiewissenschaftliche Quellen und Forschungen, Bd. 87), S. 67–97. [\[Nachweis im GBV\] Lange / Unold 2015.](#)

Fotis Jannidis / Gerhard Lauer: Burrow's Delta and Its Use in German Literary History. In: Distant Readings. Topologies of German Culture in the Long Nineteenth Century. Hg. von Matt Erlin / Lynne Tatlock. Rochester 2014, S. 29–54. [\[Nachweis im GBV\] Viehhauser 2015.](#)

Fotis Jannidis: Der Autor ganz nah. Autorstil in Stilistik und Stilometrie. In: Theorien und Praktiken der Autorschaft. Hg. von Matthias Schaffrick / Marcus Willand. Berlin, Boston 2014, S. 169–195. [\[Nachweis im GBV\] Viehhauser 2015.](#)

Fotis Jannidis: Polyvalenz – Konvention – Autonomie. In: Regeln der Bedeutung. Zur Theorie der Bedeutung literarischer Texte. Hg. von Fotis Jannidis / Gerhard Lauer / Matías Martínez / Simone Winko. Berlin 2003, S. 305–328. [\[Nachweis im GBV\] Gius / Jacke 2015.](#)

Fotis Jannidis: Analytische Hermeneutik. Eine vorläufige Skizze. In: Heuristiken der Literaturwissenschaft. Hg. von Uta Klein / Katja Mellmann / Steffanie Metzger. Paderborn 2006, S. 131–144. [\[Nachweis im GBV\] Gius / Jacke 2015.](#)

Fotis Jannidis / Gerhard Lauer / Matías Martínez / Simone Winko: Der Bedeutungsbegriff in der Literaturwissenschaft. Eine historische und systematische Skizze. In: *Regeln der Bedeutung. Zur Theorie der Bedeutung literarischer Texte*. Hg. von Fotis Jannidis / Gerhard Lauer / Matías Martínez / Simone Winko. Berlin 2003, S. 3–30. [\[Nachweis im GBV\]](#) Gius / Jacke 2015.

Ludger Jansen: Warum sich Artefakte ihrer Marginalisierung widersetzen. In: *Deutsche Zeitschrift für Philosophie* 61 (2013), H. 2, p. 267–282. [\[Nachweis im OPAC\]](#) Heßbrüggen-Walter 2015.

Theo M.V. Janssen: Frege, Contextuality and Compositionality. In: *Journal of Logic, Language and Information* 10 (2001), H. 1, p. 115–136. [\[Nachweis im GBV\]](#) Heßbrüggen-Walter 2015.

Wybe Jappe Alberts: Mittelalterliche Stadtrechnungen als Geschichtsquellen. In: *Rheinische Vierteljahrsblätter* 23 (1958), S. 75–96. [\[Nachweis im GBV\]](#) Vogeler 2015.

Gerhard Jaritz: Alltagsgeschichte des Mittelalters. Quellenanalysen mit Hilfe eines fachspezifischen Datenbanksystems. In: *Ludwig Boltzmann Institut für Historische Sozialwissenschaft-Arbeitspapiere Nr. 1*. Salzburg 1984, S. 39–42. [\[online\]](#) Vogeler 2015.

Gerhard Jaritz: Die Reimer Rechnungsbücher (1399–1477) als Quelle zur klösterlichen Sachkultur des Spätmittelalters. In: *Die Funktion der schriftlichen Quellen in der Sachkulturforschung*. Wien 1976 (= Veröffentlichungen des Instituts für mittelalterliche Realienkunde Österreichs 1), S. 145–249. [\[Nachweis im OPAC\]](#) Vogeler 2015.

Hans Robert Jauß: Theorie der Gattungen und Literatur des Mittelalters. In: *Hans Robert Jauß: Alterität und Modernität der mittelalterlichen Literatur. Gesammelte Aufsätze 1956–1976*. München 1977, S. 327–358. [\[Nachweis im OPAC\]](#) Viehhauser 2015.

Sabina Jeschke: *Mathematik in Virtuellen Wissensräumen – IuK-Strukturen und IT-Technologien in Lehre und Forschung*. Berlin, Techn. Univ., Diss., 2004. [\[Nachweis im GBV\]](#) Jeschke et al 2015.

Matthew Jockers: *Macroanalysis. digital methods and literary history*. Champaign 2013. [\[Nachweis im OPAC\]](#) Viehhauser 2015. [Nerbonne 2015](#).

Patrick Juola / R. Harald Baayen: A controlled-corpus experiment in authorship identification by cross-entropy. In: *Literary and Linguistic Computing* 20.1 (2005), S. 59–67. [\[Nachweis im GBV\]](#) Nerbonne 2015.

K

Ben Kaden: Wer übernimmt was? Zum Verhältnis von Digital Humanities und Geisteswissenschaften. In: *LIBREAS. Debatte*, 12. September 2013. [\[online\]](#) Sahle 2015.

Cathleen Kantner / Amelie Kutter / Andreas Hildebrandt / Mark Püttcher: *How to Get Rid of the Noise in the Corpus: Cleaning Large Samples of Digital Newspaper Texts*. Stuttgart 2011. [\[online\]](#) Blessing et al. 2015.

Cathleen Kantner: *Transnational Identity-Discourse in the Mass Media. Humanitarian Military Interventions and the Emergence of a European Identity (1990–2006)*. Berlin 2009. [Blessing et al. 2015](#).

Cathleen Kantner / Amelie Kutter / Swantje Renfordt: *The Perception of the EU as an Emerging Security Actor in Media Debates on Humanitarian and Military Interventions (1990–2006)*. In: *RECON Online Working Paper 19* (2008). Oslo 2008. [\[online\]](#) Blessing et al. 2015.

Cathleen Kantner (2006a): *Collective Identity as Shared Ethical Self-Understanding: The Case of the Emerging European Identity*. In: *European Journal of Social Theory* 9 (2006) H. 4, S. 501–523. [\[Nachweis im GBV\]](#) Blessing et al. 2015.

Cathleen Kantner (2006b): *Die thematische Verschränkung nationaler Öffentlichkeiten in Europa und die Qualität transnationaler politischer Kommunikation*. In: *Demokratie in Der Mediengesellschaft*. Hg. von Kurt Imhof / Roger Blum / Heinz Bonfadelli / Otfried Jarren. Wiesbaden 2006, S. 145–160. [\[Nachweis im GBV\]](#) Blessing et al. 2015.

Cathleen Kantner: *Kein Modernes Babel. Kommunikative Voraussetzungen Europäischer Öffentlichkeit*. Wiesbaden 2004. [\[Nachweis im GBV\]](#) Blessing et al. 2015.

Günter Katzler: *Was erzählen Rechnungsbücher von der Stadt? Das Beispiel des Rechnungsbuchs Bischof Bertholds von Freising*. In: *Pro civitate Austriae*, NF 13 (2008), S. 37–61. [\[Nachweis im OPAC\]](#) Vogeler 2015.

Kunstgeschichte, Musikvideo und Bildwissenschaften. Eine Einführung: Come ride with me through the veins of history. Hg. von Henry Keazor. Berlin 2014 [\[online\]](#) Wübbena et al 2015.

Henry Keazor / Thorsten Wübbena: *Video thrills the Radio Star: Musikvideos: Geschichte, Themen, Analysen*. 3. Auflage. Bielefeld 2011 (zuerst 2005). [\[Nachweis im GBV\]](#) Wübbena et al 2015.

Udo Kelle: *Die Integration qualitativer und quantitativer Methoden in der empirischen Sozialforschung. Theoretische Grundlagen und methodologische Konzepte*. Wiesbaden 2007. [\[Nachweis im GBV\]](#) Jeschke et al 2015.

Max Kemmann: *Grasping Technology*. Juni 2014. [\[online\]](#) Heßbrüggen-Walter 2015.

Michael Kerres: *Mediendidaktik. Konzeption und Entwicklung mediengestützter Lernangebote*. München 2012. [\[Nachweis im GBV\]](#) Jeschke et al 2015.

- Mike Kestemont: Function Words in Authorship Attribution. From Black Magic to Theory? In: *Proceedings of the 3rd Workshop on Computational Linguistics Literature (CLFL)*. Göteborg 2014, S. 59–66. [\[online\]](#) Viehhauser 2015.
- Mike Kestemont: Arthur's Authors. A Quantitative Study of the Rhyme Words in the Middle Dutch Arthurian Epic. In: *Zeitschrift für deutsches Altertum* 142 (2013), H. 1, S. 1–33. [\[Nachweis im OPAC\]](#) Viehhauser 2015.
- Andrea Kiesel / Iring Koch: *Lernen. Grundlagen der Lernpsychologie*. Wiesbaden 2012. [\[Nachweis im GBV\]](#) Jeschke et al 2015.
- Tom Kindt / Tilmann Köppe: *Moderne Interpretationstheorien. Eine Einleitung*. In: *Moderne Interpretationstheorien*. Ein Reader. Hg. von Tom Kindt / Tilmann Köppe. Göttingen 2008, S. 7–26. [\[Nachweis im GBV\]](#) Gius / Jacke 2015.
- Tom Kindt / Hans-Harald Müller: Zum Verhältnis von Deskription und Interpretation. Ein Bestimmungsvorschlag und ein Beispiel. In: *Literatur interpretieren: interdisziplinäre Beiträge zur Theorie und Praxis*. Hg. von Jan Borkowski / Stefan Descher / Felicitas Ferder / Philipp David Heine. Münster 2015. [\[online\]](#). Gius / Jacke 2015.
- Tom Kindt / Hans-Harald Müller: *Wieviel Interpretation enthalten Beschreibungen? Überlegungen zu einer umstrittenen Unterscheidung am Beispiel der Narratologie*. In: *Regeln der Bedeutung. Zur Theorie der Bedeutung literarischer Texte*. Hg. von Fotis Jannidis / Gerhard Lauer / Matias Martínez / Simone Winko. Berlin, New York 2003, S. 286–304. [\[Nachweis im GBV\]](#) Gius / Jacke 2015.
- Adam Kirsch: *Technology Is Taking Over English Departments - The false promise of the digital humanities*, New Republic vom 2. Mai 2014 [\[online\]](#) Sahle 2015.
- Thomas Klein: *Ermittlung, Darstellung und Deutung von Verbreitungstypen in der Handschriftenüberlieferung mittelhochdeutscher Epik*. In: *Deutsche Handschriften 1100–1400. Oxforder Kolloquium 1985*. Hg. von Volker Honemann / Nigel F. Palmer. Tübingen 1988, S. 110–167. [\[Nachweis im OPAC\]](#) Viehhauser 2015.
- Erich Kleinschmidt: *Minnesang als höfisches Zeremonialhandeln*. In: *Archiv für Kulturgeschichte* 58 (1976), S. 35–76. [\[Nachweis im OPAC\]](#) Viehhauser 2015.
- Fritz Kliche / André Blessing et al. 2015 / Ulrich Heid / Jonathan Sonntag: *The eldentity Text Exploration Workbench*. In: *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14) European Language Resources Association (ELRA)*. Reykjavik 2014, S. 691–697. [\[Nachweis im GBV\]](#) Blessing et al. 2015.
- Richard Knipping: *Die Kölner Stadtrechnungen des Mittelalters mit einer Darstellung der Finanzverwaltung*, 2 Bde. Bonn 1897–1898 (= Publikationen der Gesellschaft für rheinische Geschichtskunde 15). [\[Nachweis im OPAC\]](#) Vogeler 2015.
- Herbert Knittler: *Vom Leben auf dem Lande. Die Rechnungen der Herren von Puchheim zu Horn und Göllersdorf 1444–1468*. Edition und Kommentar. St. Pölten 2005 (= NÖ-Schriften 161). [\[Nachweis im GBV\]](#) Vogeler 2015.
- Thomas Koenig / Sabina Mihelj / John Downey / Mine Gencil Bek: *Media Framings of the Issue of Turkish Accession to the EU*. In: *Innovation: The European Journal of Social Sciences* 19 (2006) H. 2, S. 149–169. [\[Nachweis im GBV\]](#) Blessing et al. 2015.
- Hubertus Kohle: *Digitale Bildwissenschaft*. Glückstadt 2013. [\[Nachweis im OPAC\]](#) Heßbrüggen-Walter 2015. Kwastek 2015.
- Hubertus Kohle / Katja Kwastek: *Computer, Kunst und Kunstgeschichte*. Köln 2003. [\[Nachweis im GBV\]](#) Kwastek 2015.
- Thomas Köhler / Nina Kahnwald / Martina Reitmaier: *Lehren und Lernen mit Multimedia und Internet*. In: *Medienpsychologie*. Hg. von B. Batinic / M. Appel. Heidelberg 2008, S. 477–501. [\[Nachweis im GBV\]](#) Jeschke et al 2015.
- Christopher Kolbeck / Paul Rössler: *Aufstieg und Fall des »Item«*. Item Serialität zwischen sprachlicher Innovation und Konstanz - Item in Regensburger Rechnungsbüchern. In: *Stimulus* 19 (2015) [im Erscheinen]. Vogeler 2015.
- Eleftherios E. Koutsofios / Stephen C. North / Russell Truscott / Daniel A. Keim: *Visualizing Large-Scale Telecommunication Networks and Services*. In: *VIS '99 Proceedings of the Conference on Visualization '99*. Los Alamitos 1999, S. 457–461. [\[Nachweis im GBV\]](#) Biehl et al 2015.
- Clemens Kosch: *Zum Projekt einer zeichnerischen Veranschaulichung der sakralen "Binnentopographie" des Hochmittelalters in ehemaligen Konventkirchen Kölns: Methodische Überlegungen am Beispiel von St. Andreas*. In: *Kölnische Liturgie und ihre Geschichte, Studien zur interdisziplinären Erforschung des Gottesdienstes im Erzbistum Köln*. Hg. von Albert Gerhards / Andreas Odenthal. Münster 2000, S. 127–142. [\[Nachweis im GBV\]](#) Lange 2015.
- Siegfried Kracauer: *Schriften*. Band 7. Hg. von Karsten Witte. Frankfurt/Main 1973. [\[Nachweis im OPAC\]](#) Biehl et al 2015.
- Andreas Krapp: *Interesse*. In: *Handbuch der Allgemeinen Psychologie - Motivation und Emotion*. Hg. von V. Brandstätter / J. H. Otto. Göttingen 2009, S. 52–57. [\[Nachweis im OPAC\]](#) Jeschke et al 2015.
- Klaus Krippendorff: *Content Analysis: An Introduction to Its Methodology*. London 2004. [\[Nachweis im GBV\]](#) Blessing et al. 2015.
- Rudolf Kropf: *Zur Auswertung von Urbaren mittels Computer*. In: *Mitteilungen des Instituts für Österreichische Geschichtsforschung* 77 (1969), S. 141–144. [\[Nachweis im OPAC\]](#) Vogeler 2015.
- Astrid Krüger: *Der Heilige im Raum. Bartholomäus-Liturgie und Bartholomäus-Darstellungen im Frankfurter Kaiserdom Sankt Bartholomäus*. In: *Heilige, Liturgie, Raum*. Hg. von Dieter R. Bauer / Klaus Herbers / Hedwig Röckelein / Felicitas Schmieder. Stuttgart 2010 (= Beiträge zur Hagiographie, Bd. 8), S. 137–153. [\[Nachweis im OPAC\]](#) Lange / Unold 2015.

Primus Kucher / Johannes F. Evelein / Helga Schreckenberger: Einleitung. In: Erste Briefe aus dem Exil 1945–1950. Hg. von Primus-Heinz Kucher / Johannes F. Evelein / Helga Schreckenberger. München 2011, S. 9–14. [\[Nachweis im GBV\]](#) [Biehl et al 2015](#).

Rainer Kuhlen: Kollaboratives Schreiben. In: Interaktivität. Ein transdisziplinärer Schlüsselbegriff. Hg. von Christoph Bieber / Claus Leggewie. Frankfurt/Main, New York 2004, S. 216–239. [\[online\]](#) [Ernst 2015](#).

Kunstgeschichte digital. Eine Einführung für Praktiker und Studierende. Hg. von Hubertus Kohle. Berlin 1997. [\[Nachweis im GBV\]](#) [Kwastek 2015](#).

Amelie Kutter: Petitioner or Partner? Constructions of European Integration in Polish Print Media Debates on the EU Constitutional Treaty. In: Discourse and Contemporary Social Change. Hg. von Norman Fairclough / Giuseppina Cortese / Patrizia Ardizzone. Bern 2007, S. 433–457. [\[Nachweis im GBV\]](#) [Blessing et al. 2015](#).

Katja Kwastek: Aesthetics of Interaction in Digital Art. Cambridge, MA 2013. [\[Nachweis im GBV\]](#) [Kwastek 2015](#).

L

Silke Lahn / Jan Christoph Meister: Einführung in die Erzähltextanalyse. Stuttgart, Weimar 2013. [\[Nachweis im OPAC\]](#) [Gius / Jacke 2015](#).

Helene Lange: Kampfzeiten. Aufsätze und Reden aus vier Jahrzehnten. 2 Bände. Berlin 1928. [\[Nachweis im GBV\]](#) [Vogeler 2015](#).

Helene Lange: Die höhere Mädchenschule und ihre Bestimmung. Begleitschrift zu einer Petition an das preußische Unterrichtsministerium und das preußische Abgeordnetenhaus. Berlin 1887. [\[Nachweis im GBV\]](#) [Vogeler 2015](#).

Fritz Laupichler: MIDAS, HIDA, DISKUS – was ist das? In: Arbeitsgemeinschaft der Kunst- und Museumsbibliotheken-News 4 (1998), Nr. 2/3, S. 18–24. [\[online\]](#) [Kwastek 2015](#).

Michiel Leezenberg / Gerard De Vries: Wetenschapsfilosofie voor geesteswetenschappen. Amsterdam 2001. [\[online\]](#) [Nerbonne 2015](#).

Therese N. Leinonen: An acoustic analysis of vowel pronunciation in Swedish dialects. Diss. Universität Groningen 2010. [\[Nachweis im GBV\]](#) [Nerbonne 2015](#).

Jens Lemanski: Die neuaristotelischen Ursprünge des Kontextprinzips und die Fortführung in der fregeschen Begriffsschrift. In: Zeitschrift für philosophische Forschung 67 (2013), H. 4, p. 566–586. [\[Nachweis im OPAC\]](#) [Heßbrüggen-Walter 2015](#).

Narelle Lemon / Meg Colasante / Karen Corneille / Kathy Douglas: Video Annotation for Collaborative Connections to Learning: Case Studies from an Australian Higher Education Context. In: Increasing Student Engagement and Retention Using Multimedia Technologies: Video Annotation, Multimedia Applications, Videoconferencing and Transmedia Storytelling. Hg. von Laura A. Wankel / Patrick Blessing et al. 2015er. 2013, S. 181–214 (= Cutting-Edge Technologies in Higher Education, 6 Part F.). [\[Nachweis im GBV\]](#) [Wübbena et al 2015](#).

Lawrence Lessig: Freie Kultur. Wesen und Zukunft der Kreativität. Aus dem Englischen von Annegret Clausshues / Hartmut Pilch. München 2006. [\[Nachweis im OPAC\]](#) [Ernst 2015](#).

Ulrike Liebert: Introduction: Structuring Political Conflict About Europe: National Media in Transnational Discourse Analysis. In: Perspectives on European Politics and Society 8 (2007) H. 3, S. 236–260. [\[Nachweis im GBV\]](#) [Blessing et al. 2015](#).

Jürgen Link / Ursula Link-Heer: Diskurs/Interdiskurs und Literaturanalyse. In: LiLi. Zeitschrift für Literaturwissenschaft und Linguistik 20 (1990), H. 77, S. 88–99. [\[Nachweis im OPAC\]](#) [Ernst 2015](#).

Peter Lipton: Inference to the Best Explanation. London, New York 1991. [\[Nachweis im GBV\]](#) [Gius / Jacke 2015](#).

Linked Open Data – Frequently Asked Questions. Edited by Tom Heath. [\[online\]](#) [Heßbrüggen-Walter 2015](#).

Jurij M. Lotman: Die Struktur literarischer Texte. 4., unveränderte Auflage. München 1993. [\[Nachweis im GBV\]](#) [Viehhauser 2015](#).

David G. Lowe: Distinctive Image Features from Scale-Invariant Keypoints. In: International Journal of Computer Vision 2004 (60), S. 91–110. [\[Nachweis im GBV\]](#) [Fecker et al 2015](#).

Ana Lučić / Catherine L. Blake: A syntactic characterization of authorship style surrounding proper names. In: LLC: the Journal of Digital Scholarship in the Humanities [im Erscheinen; online seit 29.06.2013]. [\[Nachweis im GBV\]](#) [Nerbonne 2015](#).

M

Sreeraj M / Sumam Mary Idicula: A Survey on Writer Identification Schemes. In: International Journal of Computer Applications 26 (2011), Nr. 2, S. 23–33. DOI 10.5120/3075-4205. [\[online\]](#) [Fecker et al 2015](#).

Mädchenschulgeschichte(n). Die preußische Mädchenschulreform und ihre Folgen. Ariadne 53/54 (2008). [\[Nachweis im GBV\]](#) [Vogeler 2015](#).

- Cerstin Mahlow / Kerstin Eckart / Jens Stegmann / André Blessing et al. 2015 / Gregor Thiele / Markus Gärtner / Jonas Kuhn: Resources, Tools, and Applications at the CLARIN Center Stuttgart. In: Proceedings der 12. Konferenz zur Verarbeitung natürlicher Sprache (KONVENS 2014). Hildesheim 2014, S. 11–21. [\[Nachweis im GBV\]](#) Blessing et al. 2015.
- Inderjeet Mani: Computational Narratology, in: the living handbook of narratology. Hg. von Peter Hühn / John Pier / Wolf Schmid / Jörg Schönert. Hamburg 2013. [\[online\]](#) Gius / Jacke 2015.
- Inderjeet Mani / Barry Schiffman: Temporally Anchoring and Ordering Events in News. In: Time and Event Recognition in Natural Language. Hg. von James Pustejovsky / Robert Gaizauskas. Amsterdam, Philadelphia 2005. [\[online\]](#) Gius / Jacke 2015.
- Thomas Mann: Der Tod. In: Thomas Mann: Große kommentierte Frankfurter Ausgabe. Werke – Briefe – Tagebücher. Hg. von Heinrich Detering. Bd. 2.1. Frankfurt/Main 2004, S. 71–78. [\[Nachweis im OPAC\]](#) Gius / Jacke 2015.
- Christopher D. Manning / Hinrich Schütze: Foundations of Statistical Natural Language Processing. 4. Auflage. Cambridge, Mass. 2001. [\[Nachweis im GBV\]](#) Gius / Jacke 2015.
- Lev Manovich: Museum Without Walls, Art History without Names: Visualization Methods for Humanities and Media Studies. In: Oxford Handbook of Sound and Image in Digital Media. Hg. von Carol Vernallis / Amy Herzog / John Richardson. Oxford 2013, S. 253–278. [\[Nachweis im GBV\]](#) Kwastek 2015.
- Lev Manovich: Wer ist der Autor? Sampling/Remixen/Open Source. In: Lev Manovich: Black Box – White Cube. Aus dem Amerikanischen von Ronald Voullié. Berlin 2005, S. 7–28. [\[Nachweis im GBV\]](#) Ernst 2015.
- Jaap Mansfeld: Doxography of Ancient Philosophy. In: The Stanford Encyclopedia of Philosophy (Winter 2013 Edition). Edited by Edward N. Zalta. 2013. [\[online\]](#) Heßbrüggen-Walter 2015.
- Bernard Marcadé / Dan Cameron: Pierre et Gilles – The Complete Works, 1976–1996. Köln 1997. [\[Nachweis im GBV\]](#) Wübbena et al 2015.
- Dominik Maschek / Michael Schneyder / Marcel Tschanner: Virtual 3D reconstructions – benefit or danger for modern archaeology? Workshop-Paper: 14. Workshop »Cultural Heritage and New Technologies«, 16.–18. November 2009, Wien. [\[online\]](#) Lange / Unold 2015.
- Willard McCarty: Getting there from here. Remembering the future of digital humanities: Roberto Busa Award lecture 2013. In: Literary & Linguistic Computing 29 (2014), H. 3, S. 283–306. DOI:10.1093/lilc/ffqu022. [\[Nachweis im GBV\]](#) Sahle 2015.
- Marshall McLuhan: Culture without Literacy. In: Explorations (1953), H. 1, S. 117–127. [\[Nachweis im GBV\]](#) Ernst 2015.
- Federigo Melis: Storia della Ragioneria. Contributo alla conoscenza e interpretazione delle fonti più significative della storia economica. Bologna 1950. [\[online\]](#) Vogeler 2015.
- Mark Mersiowsky: Die Anfänge territorialer Rechnungslegung im deutschen Nordwesten. Spätmittelalterliche Rechnungen, Verwaltungspraxis, Hof und Territorium (zugl. Diss. phil. Münster 1992). Sigmaringen 2000 (= Residenzenforschung 9), S. 337–340. [\[Nachweis im OPAC\]](#) Vogeler 2015.
- Mark Mersiowsky: Aspekte adligen Lebens um 1400. Frühe westfälische und rheinische Adelsrechnungen im Vergleich. In: Vestigia Monasteriensia. Westfalen – Rheinland – Niederlande. Hg. von Peter Johaneck / Mark Mersiowsky / Ellen Widder. Festschrift für Wilhelm Jansen. Bielefeld 1995 (= Studien zur Regionalgeschichte 5), S. 263–304. [\[Nachweis im OPAC\]](#) Vogeler 2015.
- Jean-Baptiste Michel / Yuan Kui Shen / Aviva Presser Aiden / Adrian Veres / Matthew K. Gray / The Google Books Team / Joseph P. Pickett / Dale Hoiberg / Dan Clancy / Peter Norvig / Jon Orwant / Steven Pinker / Martin A. Nowak / Erez Lieberman Aiden: Quantitative analysis of culture using millions of digitized books. Science 331 (2011), Nr. 6014, S. 176–182. [\[online\]](#) Nerbonne 2015.
- Bettye Miller: Research on Application of Face-recognition Software to Portrait Art Shows Promise. In: UCR Today, 31. Mai 2013. [\[online\]](#) Kwastek 2015.
- Wolfgang Mohr: Hilfe und Rat in Wolframs Parzival. In: Festschrift für Jost Trier zu seinem 60. Geburtstag am 15. Dezember 1954. Hg. von Benno von Wiese / Karl Heinz Borck. Meisenheim 1954, S. 173–197. [\[Nachweis im GBV\]](#) Viehhauser 2015.
- Franco Moretti: Graphs – Maps – Trees. London, New York 2005. [\[Nachweis im GBV\]](#) Viehhauser 2015.
- Franco Moretti: Conjectures on World Literature. In: New Left Review 1 (2000), S. 54–68. [\[Nachweis im GBV\]](#) Viehhauser 2015.
- Nick Montfort / Patsy Baudoin / John Bell / Ian Bogost / Jeremy Douglass / Mark C. Marino / Michael Mateas / Casey Reas / Mark Sample / Noah Watver: 10 PRINT CHR\$(205.5+RND(1)); : GOTO 10. Cambridge MA., London 2013. [\[online\]](#) Ernst 2015.
- Claudine Moulin: Zeichen und ihre Deutung. Zum handschriftennahen Edieren schriftlicher Quellen im interdisziplinären Kontext. In: Die Rechnungsbücher der Stadt Luxemburg. Hg. von Claudine Moulin / Michel Pauly. Luxembourg 2012 (= Schriftenreihe des Stadtarchivs Luxemburg 6), S. 9–17. [\[Nachweis im GBV\]](#) Vogeler 2015.
- Lizzie Muller: Towards an Oral History of Media Art. In: Publications by the Daniel Langlois Foundation. 2008. [\[online\]](#) Kwastek 2015.
- Sabrina Müller: Die Inschriften der katholischen Pfarrkirchen Unserer Lieben Frau in Oberwesel. Inschriften Mittelrhein-Hunsrück 1: Oberwesel, Liebfrauenkirche. Mainz 2008. [\[Nachweis im GBV\]](#) Lange / Unold 2015.

Stefan Münker: Emergenz digitaler Öffentlichkeiten. Die Sozialen Medien im Web 2.0. Frankfurt/Main 2009. [\[Nachweis im GBV\]](#) Ernst 2015.

Janet H. Murray: Hamlet on the Holodeck. The Future of Narrative in Cyberspace. Cambridge, Mass. 1999. [\[Nachweis im GBV\]](#) Jeschke et al 2015.

N

Tobias Nagel: Zur Notwendigkeit einer Ideologiekritik der EDV im Museum. In: Kunstgeschichte digital. Eine Einführung für Praktiker und Studierende. Hg. von Hubertus Kohle. Berlin 1997, S. 84–96. [\[Nachweis im GBV\]](#) Kwastek 2015.

Eric. B. Nash / Gregory W. Edwards / Jennifer A. Thompson / Woodrow Barfield: A Review of Presence and Performance in Virtual Environments. In: International Journal of Human-Computer Interaction 12 (2000), H. 1, S. 1–41. [\[Nachweis im GBV\]](#) Jeschke et al 2015.

Hanns Peter Neuheuser: Liturgische Raumschließung und Heiligenverehrung. In: Heilige, Liturgie, Raum. Hg. von Dieter R. Bauer / Klaus Herbers / Hedwig Röckelein / Felicitas Schmieder. Stuttgart 2010, S. 183–216. [\[Nachweis im OPAC\]](#) Lange 2015.

Hanns Peter Neuheuser: *Ne lapidum materia apparentium locus vilesceret*. Die Raumvorstellung des Abtes Suger in seiner Kirchweihbeschreibung von St. Denis. In: Raum und Raumvorstellungen im Mittelalter. Hg. von Jan Aertsen / Andreas Speer. Berlin 1998, S. 641–64. [\[Nachweis im OPAC\]](#) Lange / Unold 2015.

John Nerbonne: The exact analysis of text. Foreword. In: Inference and Disputed Authorship: The Federalist Papers. Hg. von Frederick Mosteller / David Wallace. 3. Auflage. Stanford 2007. S. XI–XX. [\[online\]](#) Nerbonne 2015.

John Nerbonne: Computational Contributions to the Humanities. In: Literary and linguistic computing 20.1 (2005), S. 25–40. [\[Nachweis im GBV\]](#) Nerbonne 2015.

John Nerbonne: Elektronische Incunabellen. Antrittsvorlesung in Groningen. 1995. [\[online\]](#) Nerbonne 2015.

Joachim Nettelbeck: Abwicklung im Blindflug. In: Frankfurter Allgemeine Zeitung vom 10. September 2015. [\[online\]](#) Sahle 2015.

Andrew J. Newell / Lewis D. Griffin: Natural Image Character Recognition Using Oriented Basic Image Features. In: Proceedings of the International Conference Digital Image Computing Techniques and Applications (DICTA). Hg. von IEEE. Noosa 2011, S. 191–196. DOI 10.1109/DICTA.2011.39. [\[Nachweis im GBV\]](#) Fecker et al 2015.

Lutz Niethammer: »Kollektive Identität«. Heimliche Quellen Einer Unheimlichen Konjunktur. Reinbek bei Hamburg 2000. [\[Nachweis im GBV\]](#) Blessing et al. 2015.

Eberhard Josef Nikitsch: Die Inschriften des Rhein-Hunsrück-Kreises I. Wiesbaden 2004 (= Die Deutschen Inschriften, Bd. 60). [\[Nachweis im OPAC\]](#) Lange 2015.

Eberhard Josef Nikitsch: Ein Kirchenbau zwischen Bischof und Stadtgemeinde. Zur angeblich verlorenen Bauinschrift von 1308 in der Liebfrauenkirche zu Oberwesel am Rhein, in: Jahrbuch der Westdeutschen Landesgeschichte 22 (1996), S. 95–112. [\[Nachweis im OPAC\]](#) Lange / Unold 2015.

Lyndon Nixon / Raphael Troncy: Survey of Semantic Media Annotation Tools – towards New Media Applications with Linked Media. Paper presented at the Linked Media Workshop, European Semantic Web Conference am 25. Mai 2014. [\[online\]](#) Wübbena et al 2015.

Philip Nordmeyer: »Destroy the Digital Humanities! #dighum1213«. Blogbeitrag vom 9.12.2013. [\[online\]](#) Sahle 2015.

Julianne Nyhan: Computation and the Humanities: towards a history of Digital Humanities. Cham 2016. [\[online\]](#) Sahle 2015.

Julianne Nyhan / Oliver Duke-Williams: Joint and multi-authored publication patterns in the Digital Humanities. In: Literary and Linguistic Computing 4 (2014), H. 29, S. 1–13. [\[online\]](#) Ernst 2015.

Julianne Nyhan / Andrew Flinn / Anne Welsh: Oral History and the hidden histories project: towards histories of computing in the Humanities. In: Literary and Linguistic Computing 28 (2013), S. 1–15. DOI: 10.1093/lc/ftq044. [\[Nachweis im GBV\]](#) Sahle 2015.

N. F. Noy / D. L. McGuinness: Ontology Development 101: A Guide to Creating Your First Ontology. 2001. [\[online\]](#) Heßbrüggen-Walter 2015.

Michele Nucci / Marco Grassi / Christian Morbidoni / Francesco Piazza: Enriching Digital Libraries Contents with SemLib Semantic Annotation System. Proceedings of the Digital Humanities 2012 Conference. Hamburg 2012, S. 303–306. [\[Nachweis im GBV\]](#) Lange 2015.

O

Öffentlicher Sprachgebrauch. Praktische, theoretische und historische Perspektiven. Georg Stötzel zum 60. Geburtstag gewidmet. Hg. von Karin Böke / Matthias Jung / Martin Wengeler. Opladen 1996. [\[Nachweis im GBV\]](#) Wolff 2015.

Open Annotation Data Model. Community Draft. Hg. von Robert Sanderson / Paolo Ciccarese / Herbert an de Sompe. 8. Februar 2013. [\[online\]](#) [Vogeler 2015](#).

Open domesday. The first free online copy of Domesday Book. Hg. von Anna Powell-Smith / J.J.N. Palmer. University of Hull. [\[online\]](#) [Vogeler 2015](#).

OWL Web Ontology Language Reference. Edited by Mike Dean / Guus Schreiber. 2004 [\[online\]](#). [Heßbrüggen-Walter 2015](#).

OWL 2 Web Ontology Language. Structural Specification and Functional-Style Syntax. Second Edition. W3C Recommendation. Hg. von Boris Motik / Peter F. Patel-Schneider / Bijan Parsia. 11. Dezember 2012. [\[online\]](#) [Vogeler 2015](#).

P

Otto Pächt: Van Eyck. Die Begründer der altniederländischen Malerei. 1. Auflage 1989. München 1993. [\[Nachweis im GBV\]](#) [Kwastek 2015](#).

Eleftheria Paliou: Visibility analysis in 3D spaces: a new dimension to the understanding of social space. In: Spatial analysis and social spaces: interdisciplinary approaches to the interpretation of prehistoric and historic built environments. Topoi. Berlin Studies of the Ancient World. Hg. von Eleftheria Paliou / Undine Lieberwirth / Silvia Polla. Berlin 2014, S. 91–113. [\[Nachweis im GBV\]](#) [Lange / Unold 2015](#).

Werner Pantke / Arne Haak / Volker Märgner: Color Segmentation for Historical Documents Using Markov Random Fields. In: Proceedings of the 6th International Conference on Soft Computing and Pattern Recognition (SoCPaR 2014). Tunis 2014, S. 151–156. DOI 10.1109/SOCPAR.2014.7007997. [Fecker et al 2015](#).

Werner Pantke / Volker Märgner / Daniel Fecker / Tim Fingscheidt / Abed Asi / Ofer Biller / Jihad El-Sana / Raid Saabni / Mohammed Yehia: HADARA – A Software System for Semi-Automatic Processing of Historical Handwritten Arabic Documents. In: Proceedings of IS&T Archiving 2013, Washington D.C. 2013, S. 161–166 [\[online\]](#). [\[Nachweis im GBV\]](#) [Fecker et al 2015](#).

Hermann Paul: Mittelhochdeutsche Grammatik. 25. Auflage. Tübingen 2007. [\[Nachweis im OPAC\]](#) [Viehhauser 2015](#).

Marco Carlo Passarotti: Leaving Behind the Less-Resourced Status. The Case of Latin through the Experience of the Index Thomisticus Treebank. In: 7th SaLTmIL Workshop on Creation and Use of Basic Lexical Resources for Less-Resourced Languages, LREC 2010. La Valletta 2010. 2010, p. 27–32. [\[online\]](#) [Heßbrüggen-Walter 2015](#).

Kathrin Passig: Standardsituationen der Technologiekritik. Berlin 2013. [\[Nachweis im GBV\]](#) [Ernst 2015](#).

Kat. Ausst. Nana Petzet: System SBF. Inventarisierung der Sammlung. Metainventur. Hg. vom Cuxhavener Kunstverein; Kunstverein und Stiftung Springhornhof. Neuenkirchen, Nürnberg 2003. [\[Nachweis im GBV\]](#) [Kwastek 2015](#).

William Pannapacker: »Big Tent Digital Humanities« – A View From the Edge. In: The Chronicle of Higher Education vom 31. Juli 2011. [\[online\]](#) [Sahle 2015](#).

Matthias Perstling: Multimediale Dokumentation und Edition mehrschichtiger Texte. Das steirisch-landesfürstliche »Marchfütterurbar« von 1414/1426. Diss. phil. Graz 2013. [\[online\]](#) [Vogeler 2015](#).

Alte und neue Philologie. Hg. von Martin-Dietrich Glessgen / Franz Lebsanft. Tübingen 1997. [\[Nachweis im OPAC\]](#) [Viehhauser 2015](#).

John Pier: Narrative Levels. In: the living handbook of narratology. Hg. von Peter Hühn / John Pier / Wolf Schmid / Jörg Schönert. Hamburg 2013. [\[online\]](#) [Gius / Jacke 2015](#).

Wendell Piez: Towards Hermeneutic Markup. An Architectural Outline. In: Digital Humanities 2010. Conference Abstracts, präsentiert auf der Digital Humanities Conference 2010 (DH 2010). London 2010, S. 202–205. [\[online\]](#) [Gius / Jacke 2015](#).

PlagDoc / Martin Kotynek: Schwarmgedanken. Ein Versuch, Lehren aus »GuttenPlag« zu ziehen. Ein Denkanstoß für künftige Recherche-Plattformen im Netz. Ein Diskussionsbeitrag zu »investigativem Crowdsourcing«, In: GuttenPlag Wiki, 8.6.2012. [\[online\]](#) [Ernst 2015](#).

Margret Plank / Steffen Weichert: Don't Forget the Users – Developing a Portal for Audiovisual Media with a User-Centred Approach. In: IFLA WLIC 2014. Libraries, Citizens, Societies: Confluence for Knowledge (Session 139 - Audiovisual and Multimedia with Information Technology). Lyon. 16.–22. August 2014. [\[online\]](#) [Wübbena et al 2015](#).

Margret Plank: A Portal for Scientific Audiovisual Media: Analysing User Needs. In: LIBER, 3. Oktober 2012. [\[online\]](#) [Wübbena et al 2015](#).

Kees van der Ploeg: Art, architecture and liturgy: Siena cathedral in the Middle Ages. Groningen 1993. [\[Nachweis im GBV\]](#) [Lange / Unold 2015](#).

Politische Netzwerkerinnen. Internationale Zusammenarbeit von Frauen 1830–1960. Hg. von Eva Schöck-Quinteros / Anja Schüler / Annika Wilmers / Kerstin Wolff. Berlin 2007. [\[Nachweis im GBV\]](#) [Vogeler 2015](#).

Global Politics on Screen: A Japanese Film on the Lytton Commission in 1932. [\[online\]](#) [Wübbena et al 2015](#).

Bernhard Pörksen / Hanne Detel: Evidenzerfahrungen für alle. Das kontraproduktive Krisenmanagement des Verteidigungsministers und die Logik der Skandalisierung im digitalen Zeitalter. In: Inszenierung als Beruf. Der Fall Guttenberg. Hg. von Oliver Lepsius / Reinhart Meyer-Kalkus. Berlin 2011, S. 56–69. [\[Nachweis im OPAC\]](#) Ernst 2015.

R

R Core Team: A language and environment for statistical computing. R Foundation for Statistical Computing. Wien 2013. [\[online\]](#) Viehhauser 2015.

Ursula Rautenberg: Die historisch-systematische Medienspezifität des Buches. In: Ursula Rautenberg / Dirk Wetzels: Buch. Tübingen 2001, S. 8–10. [\[Nachweis im OPAC\]](#) Ernst 2015.

Fausto Ravida: Graphematisch-phonologische Analyse der Luxemburger Rechnungsbücher (1388–1500). Ein Beitrag zur historischen Stadtsprachenforschung. Heidelberg 2012 (= Germanistische Bibliothek 43). [\[Nachweis im GBV\]](#) Vogeler 2015.

RDF 1.1 Concepts and Abstract Syntax. Edited by Richard Cyganiak / David Wood / Markus Lanthaler. 2014. [\[online\]](#) Heßbrüggen-Walter 2015.

RDF 1.1 Semantics. W3C Recommendation Hg. von Patrick J. Hayes / Peter F. Patel-Schneider. 25. Februar 2014. [\[online\]](#) Vogeler 2015.

Rechnungen des Mittelalters im Digitalen Zeitalter. Ein Modell. Hg. von Georg Vogeler. 2013ff. [\[online\]](#) Vogeler 2015.

Die Rechnungsbücher der Stadt Luxemburg. Bearb. von Claudine Moulin / Michel Pauly. Luxemburg 2007–2012ff. (= Schriftenreihe des Stadtarchivs Luxemburg 1–6). [\[Nachweis im GBV\]](#) Vogeler 2015.

Dagmar Renaud / Dagmar Unz: Die M-DAS – eine modifizierte Version der Differentiellen Affekt Skala zur Erfassung von Emotionen bei der Mediennutzung. In: Zeitschrift für Medienpsychologie 18 (2006), H. 2, S. 70–75. [\[Nachweis im GBV\]](#) Jeschke et al 2015.

Swantje Renfordt: Framing the use of force : an international rule of law in media reporting ; a comparative analysis of western media debates about military interventions 1990 - 2005. Diss. Berlin 2009. [\[Nachweis im GBV\]](#) Blessing et al. 2015.

Rewind – Play – Fast Forward: The Past, Present and Future of the Music Video. Hg. von Henry Keazor / Thorsten Wübbena. Bielefeld 2010. [\[online\]](#) Wübbena et al 2015.

Falko Rheinberg / Regina Vollmeyer / Stefan Engeser: Die Erfassung des Flow-Erlebens. In: Diagnostik von Motivation und Selbstkonzept. Test und Trends. Hg. von J. Stiensmeier-Pelster / F. Rheinberg. Göttingen 2003, S. 261–279. [\[Nachweis im GBV\]](#) Jeschke et al 2015.

Thomas Risse: A Community of Europeans? Transnational Identities and Public Spheres. Ithaca 2010. [\[Nachweis im GBV\]](#) Blessing et al. 2015.

Peter M.W. Robinson: Is There a Text in These Variants? In: The Literary Text in the Digital Age. Hg. von Richard J. Finneran. Ann Arbor (Mi) 1996, S. 99–115. [\[Nachweis im OPAC\]](#) Fecker et al 2015.

Jan Röhner: Feindliche Übernahme? Die Geisteswissenschaften wehren sich gegen falsche Ansprüche der Informatik, aber setzen auf die »Digital Humanities«. In: Frankfurter Allgemeine Zeitung vom 19. Juli 2013. [\[online\]](#) Sahle 2015.

Marie-Laure Ryan: Possible Worlds, Artificial Intelligence, and Narrative Theory. Bloomington 1991. [\[Nachweis im GBV\]](#) Gius / Jacke 2015.

S

Patrick Sahle: Digitale Editionsformen. Zum Umgang mit der Überlieferung unter den Bedingungen des Medienwandels, 3 Bände. Norderstedt 2013 (= Schriften des Instituts für Dokumentologie und Editorik 7–9). [\[Nachweis im OPAC\]](#) Viehhauser 2015. Vogeler 2015.

Patrick Sahle: DH Studieren! Auf dem Weg zu einem Kern- und Referenzcurriculum der Digital Humanities. DARIAH-DE Working Papers Nr. 1. Göttingen 2013. [\[online\]](#) Sahle 2015.

Patrick Sahle: Digitale Geisteswissenschaften. Hg. vom CCEH, Universität zu Köln. Erstellt im Rahmen des Projekts DARIAH-DE; gefördert vom Bundesministerium für Bildung und Forschung (BMBF) 2011. [\[online\]](#) Jeschke et al 2015.

Marie Saldana: An integrated approach to the procedural modeling of ancient cities and buildings. In: Digital Humanities 2014. Book of Abstracts. Lausanne 2014, S. 351–53. [\[online\]](#) Lange / Unold 2015.

Angelika Schaser: Frauenbewegung in Deutschland. Darmstadt 2006. [\[Nachweis im GBV\]](#) Vogeler 2015.

Arno Schirokauer: Studien zur mittelhochdeutschen Reimgrammatik. In: Beiträge zur Geschichte der deutschen Sprache und Literatur (PBB) 47 (1923), S. 1–126. [\[Nachweis im OPAC\]](#) Viehhauser 2015.

Friedrich Schleiermacher: Vorlesungen zur Hermeneutik und Kritik [1805–1833]. In: Friedrich Schleiermacher. Kritische Gesamtausgabe. Bd. 4. Berlin [u.a.] 2012. [\[Nachweis im OPAC\]](#) Gius / Jacke 2015.

- Helmut Schmid: Improvements in Part-of-Speech Tagging with an Application to German. In: Proceedings of the ACL SIGDAT-Workshop. Dublin 1995, S. 47-50. [\[online\]](#) [Blessing et al. 2015](#).
- Helmut Schmid: Probabilistic Part-of-Speech Tagging Using Decision Trees. In: Proceedings of International Conference on New Methods in Language Processing. Manchester 1994, S. 44-49. [\[online\]](#) [Biehl et al 2015](#).
- Wolf Schmid: Elemente der Narratologie. Berlin [u.a.] 2008. [\[Nachweis im GBV\]](#) [Gius / Jacke 2015](#).
- Desmond Schmidt: The Inadequacy of Embedded Markup for Cultural Heritage Texts. In: Literary and Linguistic Computing 25 (2010), S. 337-356. DOI 10.1093/lilc/fqq007. [Gius / Jacke 2015](#).
- Helmut Schmid: Improvements in Part-of-Speech Tagging with an Application to German. In: Natural Language Processing Using Very Large Corpora. Hg. von Susan Armstrong / Kenneth Church / Pierre Isabelle / Sandra Manzi / Evelyne Tzoukermann / David Yarowsky. Dordrecht 1999, S. 13-26. [\[Nachweis im GBV\]](#) [Biehl et al 2015](#).
- Rüdiger Schnell: Minnesang und Singspruch im 13. Jahrhundert. Gattungsdifferenzen und Gattungsinterferenzen. In: Transformationen der Lyrik im 13. Jahrhundert. Wildbader Kolloquium 2008. In Verbindung mit Eckart Conrad Lutz und Klaus Ridder hg. von Susanne Köbele. Berlin 2013 (= Wolfram-Studien 21), S. 287-348. [\[Nachweis im OPAC\]](#) [Viehhauser 2015](#).
- Rüdiger Schnell: »Autor« und »Werk« im deutschen Mittelalter. Forschungskritik und Forschungsperspektiven. In: Neue Wege der Mittelalter-Philologie. Landshuter Kolloquium 1996. Hg. von Joachim Heinzle / L. Peter Johnson / Gisela Vollmann-Profe. Berlin 1998 (= Wolfram-Studien 15), S. 12-73. [\[Nachweis im OPAC\]](#) [Viehhauser 2015](#).
- Arthur Schnitzler: Blumen [1894]. In: Arthur Schnitzler: Ausgewählte Werke in acht Bänden. Bd. 1: Leutnant Gustl. Erzählungen 1892-1907. Frankfurt/Main 2004, S. 98-107. [\[Nachweis im GBV\]](#) [Gius / Jacke 2015](#).
- Christoph Schöch: Corneille, Molière et les autres. Stilometrische Analysen zu Autorschaft und Gattungszugehörigkeit im französischen Theater der Klassik. In: Literaturwissenschaft im digitalen Medienwandel. Hg. von Christof Schöch / Lars Schneider. Philologie im Netz, Beiheft 7 (2014), S. 130-157. [\[online\]](#) [Viehhauser 2015](#).
- Thomas Schröder: Die Handlungsstruktur von Texten. Ein integrativer Beitrag zur Texttheorie. Tübingen 2003. [\[Nachweis im GBV\]](#) [Vogeler 2015](#).
- Schuldbücher und Rechnungen der Großschäffer und Lieger des Deutschen Ordens in Preußen, Bd. 1: Großschäfferei Königsberg I (Ordensfoliant 141). Bearb. von Cordelia Heß / Christina Link / Jürgen Sarnowsky. Köln, Weimar, Wien 2008 (= Veröffentlichungen aus den Archiven Preußischer Kulturbesitz, 62,1; zugleich Quellen und Darstellungen zur Hansischen Geschichte N.F., LIX, 1). [\[Nachweis im OPAC\]](#) [Vogeler 2015](#).
- Schuldbücher und Rechnungen der Großschäffer und Lieger des Deutschen Ordens in Preußen, Bd. 3: Großschäfferei Marienburg. Bearb. von Christina Link / Jürgen Sarnowsky. Köln, Weimar, Wien 2008 (= Veröffentlichungen aus den Archiven Preußischer Kulturbesitz, 62,3; zugleich Quellen und Darstellungen zur Hansischen Geschichte N.F. LIX, 3). [\[Nachweis im OPAC\]](#) [Vogeler 2015](#).
- Die Schul- und Rechnungsbücher des Deutschen Ordens. Eine synoptische Edition im Internet. Hg. von Christina Link / Jürgen Sarnowsky. Hamburg 2007ff. [\[online\]](#) [Vogeler 2015](#).
- Michael Viktor Schwarz: Kathedralen zu verstehen. In: Virtuelle Räume. Raumwahrnehmung und Raumvorstellung im Mittelalter. Akten des 10. Symposiums des Mediävistenverbandes, Krens 24.-26. März 2003. Hg. von Elisabeth Vavra. Berlin 2005, S. 47-68. [\[Nachweis im OPAC\]](#) [Lange / Unold 2015](#).
- Frithjof Schwartz: Visuelle Bezüge im Raum. Analyse von Sichtachsen in der Liebfrauenkirche in Oberwesel. In: Zeitschrift für Kunstgeschichte. [in Vorbereitung] [Lange / Unold 2015](#).
- Frithjof Schwartz: Inschriften im Bezugssystem des Raumes. Forschungsansatz und Verortung im wissenschaftlichen Spektrum. Referenzobjekt und Fallstudien. Projektskizze 2014, unpag. [\[online\]](#) [Lange / Unold 2015](#).
- Kurt Seidel: Klosterrechnungen als Geschichtsquelle. In: Deutsche Geschichtsblätter 12 (1911), S. 291-297. [\[Nachweis im OPAC\]](#) [Vogeler 2015](#).
- Shared Canvas Data Model. Hg. von Robert Sanderson / Benjamin Albritton, Version 1.0. 14. Februar 2013. [\[online\]](#) [Vogeler 2015](#).
- Clay Shirky: Ontology is Overrated: Categories, Links, and Tags. [\[online\]](#) [Heßbrüggen-Walter 2015](#).
- Edward Shorter: The Historian and the Computer. A Practical Guide. Engelwood Cliffs, N.J. 1971. [\[Nachweis im GBV\]](#) [Vogeler 2015](#).
- Ben Shneiderman / Catherine Plaisant: Designing the User Interface: Strategies for Effective Human-Computer Interaction. 5. Auflage. Reading 2010 (zuerst 1987). [\[Nachweis im GBV\]](#) [Nerbonne 2015](#).
- Roberto Simanowski: Interfictions. Vom Schreiben im Netz. Frankfurt/Main 2002. [\[Nachweis im GBV\]](#) [Ernst 2015](#).
- Barry Smith: Beyond Concepts: Ontology as Reality Representation. In: Proceedings of FOIS 2004. Edited by Achille Varzi / Laure Vieu. Amsterdam, Berlin, Oxford, Tokyo, Washington DC 2014, p. 73-84. [\[Nachweis im OPAC\]](#) [Heßbrüggen-Walter 2015](#).
- Early U.S. Securities Prices. Zusammenestellt von Richard E. Sylla / Jack Wilson / Robert E. Wright. In: EH.Net - Economic History Services [\[online\]](#) [Vogeler 2015](#).
- Stéfan Sinclair / Geoffrey Rockwell / the Voyant Tools Team: Voyant Tools (web application). 2012. [\[online\]](#) [Viehhauser 2015](#).

- Karl Stackmann: Neue Philologie? In: *Modernes Mittelalter. Neue Bilder einer populären Epoche*. Hg. von Joachim Heinze. Frankfurt/Main 1994, S. 398–427. [Nachweis im GBV] Viehhauser 2015.
- Thomas Stäcker: Erfahrungsbericht Helmstedter Drucke Online an der Herzog August Bibliothek Wolfenbüttel. In: *Volltext via OCR. Möglichkeiten und Grenzen. Testscenarien zu den Funeralschriften der Staatsbibliothek zu Berlin – Preußischer Kulturbesitz*. Hg. von Maria Federbusch / Christian Polzin. Berlin 2013 (= Beiträge aus der Staatsbibliothek zu Berlin – Preußischer Kulturbesitz Band 43), S.123–135. [online] Fecker et al 2015.
- R. J. Stainton: Context Principle. In: *Encyclopedia of Language & Linguistics*, vol. 3, Boston 2006, p. 108–115. [Nachweis im GBV] Heßbrüggen-Walter 2015.
- Elisabeth Steiner / Johannes Stigler: GAMS and Cirilo Client: policies, documentation and tutorial. 2014. [online] Vogeler 2015.
- Fabian Steiner / Max Trecker: Zeitgemäß publizieren? Veröffentlicht von Lilian Landes am 17. Februar 2014, auf dem Weblog: *Rezensieren – Kommentieren – Bloggen: Wie kommunizieren Geisteswissenschaftler in der digitalen Zukunft?* [online] Ernst 2015.
- Felix Steiner: *Dargestellte Autorschaft. Autorkonzept und Autorsubjekt in wissenschaftlichen Texten*. Tübingen 2009. [Nachweis im GBV] Ernst 2015.
- Angela Steinsiek: Alles Wikipedia? Kommentieren heute am Beispiel der Jean Paul-Brief-Edition. In: *Brief-Edition im digitalen Zeitalter*. Hg. von Anne Bohnenkamp / Elke Richter. Berlin, Boston 2013, S. 229–235. [Nachweis im OPAC] Biehl et al 2015.
- Robert Stevens / Phillip Lord: Reification of properties in an ontology. 2010. [online] Heßbrüggen-Walter 2015.
- Georg Stötzel / Martin Wengeler: *Kontroverse Begriffe. Geschichte des öffentlichen Sprachgebrauchs in der Bundesrepublik Deutschland*. In Zusammenarbeit mit Karin Böke, Hildegard Gorny, Silke Hahn, Matthias Jung, Andreas Musolf, Cornelia Tönnesen. Berlin, New York 1995. [Nachweis im GBV] Vogeler 2015.
- Peter Strohschneider: *Tanzen und Singen. Leichs von Ulrich von Winterstetten, Heinrich von Sax sowie dem Tannhäuser und die Frage nach dem rituellen Status des Minnesangs*. In: *Mittelalterliche Lyrik. Probleme der Poetik*. Hg. von Thomas Cramer / Ingrid Kasten. Berlin 1999, S. 197–231. [Nachweis im OPAC] Viehhauser 2015.
- Jannik Strötgen / Michael Gertz: *HeidelTime. High Quality Rule-Based Extraction and Normalization of Temporal Expressions*. In: *Proceedings of the 5th International Workshop on Semantic Evaluation, präsentiert auf der Association for Computational Linguistics Conference*. Uppsala 2010, S. 321–324. [online] Gius / Jacke 2015.
- Hans Sveistrup / Agnes von Zahn-Harnack: *Die Frauenfrage in Deutschland. Strömungen und Gegenströmungen 1790–1930. Sachlich geordnete und erläuterte Quellenkunde*. Burg 1934. [Nachweis im GBV] Vogeler 2015.
- Patrik Svensson: *The digital humanities as a humanities project*. In: *Arts and Humanities in Higher Education* 11/1-2 (2012), S. 42–60. [Nachweis im GBV] Sahle 2015.
- Zoltán Gendler Szabó: »Compositionality«. In: *The Stanford Encyclopedia of Philosophy* (Fall 2013 Edition). Edited by Edward N. Zalta. 2013. [online] Heßbrüggen-Walter 2015.

T

- Liane Tarouco / Bárbara Gorziza / Ygor Corrêa / Erico M.H. Amaral / Thaísa Müller: *Virtual Laboratory for Teaching Calculus: an Immersive Experience*. In: *Proceedings of the 2012 IEEE Global Engineering Education Conference (EDUCON 2012)*. Hg. von Institute of Electrical and Electronics Engineers (IEEE). Berlin 2012, S. 774–781. [Nachweis im GBV] Jeschke et al 2015.
- TEI P5. *Guidelines for Electronic Text Encoding and Interchange*. Hg. von TEI Consortium. Version 2.6, 2014. [online] Vogeler 2015.
- Helmut Tervooren: *Sangspruchdichtung*. Stuttgart, Weimar 1995. [Nachweis im OPAC] Viehhauser 2015.
- Melissa Terras: *Peering Inside the Big Tent: Digital Humanities and the Crisis of Inclusion*. In: *Blogbeitrag vom 26. Juli 2011*. [online] Sahle 2015.
- Manfred Thaller: *What is a text within the Digital Humanities, or some of them, at least?* In: *digital humanities 2012*. Hamburg 2012. [online] Vogeler 2015.
- Udo Tietz: *Les Limites Du »Nous« Libéral*. In: *Qu'est-Ce Qu'un Collectif? Du Commun À La Politique*. Hg. von Laurence Kaufmann / Danny Trom. Paris 2010, S. 173–195. [online] Blessing et al. 2015.
- Udo Tietz (2002a): *Die Grenzen des »Wir«*. Eine Theorie Der Gemeinschaft. Frankfurt 2002. [Nachweis im GBV] Blessing et al. 2015.
- Udo Tietz (2002b): *Gemeinsinn, Gemeinwohl Und Die Grenzen Des »Wir«*. In: *Gemeinwohl und Gemein Sinn*. Hg. von Herfried Münkler / Harald Bluhm. Berlin 2002, S. 37–70. [Nachweis im GBV] Blessing et al. 2015.
- Charles Tilly: *Computers in Historical Analysis*. In: *Computers and the Humanities* 7 (1973), S. 323–335. [Nachweis im GBV] Vogeler 2015.

Title 175, Legislative Rule Series 2 Private Club Licensing. Edited by West Virginia Alcohol Beverage Commissioner. [\[online\] Heßbrüggen-Walter 2015.](#)

Christoph Titz: Anonymer Schavan-Ankläger: »Ich wollte das nicht unter den Tisch fallen lassen«. In: Spiegel Online, 5. Mai 2012. [\[online\] Ernst 2015.](#)

Horst Thomé: Werk. In: Reallexikon der deutschen Literaturwissenschaft. Neubearbeitung des Reallexikons der deutschen Literaturgeschichte. Hg. von Jan-Dirk Müller, gemeinsam mit Georg Braungart et al. Band III: P-Z. Berlin, New York 2003, S. 832–834. [\[Nachweis im GBV\] Ernst 2015.](#)

Kathryn Tomasek / Syd Bauman: Encoding Financial Records for Historical Research. In: Journal of the Text Encoding Initiative 6 (2013). DOI: 10.4000/jtei.895. [\[online\] Vogeler 2015.](#)

Kathryn Tomasek: Encoding Financial Records for Historical Research. Whitepaper 2012. NEH Ref: HD-51224-11. [\[online\] Vogeler 2015.](#)

Kristina Toutanova / Dan Klein / Christopher Manning / Yoram Singer: Feature-Rich Part-of-Speech Tagging with a Cyclic Dependency Network. In: Proceedings of HLT-NAACL. Edmonton, Kanada 2003, S. 252–259. [\[Nachweis im GBV\] Blessing et al. 2015.](#)

V

William Vaughan: Computergestützte Bildrecherche und Bildanalyse. In: Kunstgeschichte digital. Eine Einführung für Praktiker und Studierende. Hg. von Hubertus Kohle. Berlin 1997, S. 97–106. [\[Nachweis im GBV\] Kwastek 2015.](#)

Georg Vogeler: Digitale Edition von Wirtschafts- und Rechnungsbüchern. In: Wirtschafts- und Rechnungsbücher des Mittelalters und der Frühen Neuzeit. Hg. von Gudrun Gleba / Niels Petersen. Göttingen 2015. [\[Nachweis im GBV\] Vogeler 2015.](#)

Georg Vogeler: Spätmittelalterliche Steuerbücher deutscher Territorien, Teil 1: Überlieferung und hilfswissenschaftliche Analyse. In: Archiv für Diplomatik 49 (2003), S. 165–295; Teil 2: Funktionale Analyse und Typologie. In: Archiv für Diplomatik 50 (2004), S. 57–204. [\[Nachweis im OPAC\] Vogeler 2015.](#)

Peter Vorderer / Werner Wirth / Felize Ribeiro Gouveia / Frank Biocca / Timo Saari / Lutz Jäncke / Saskia Böcking / Holger Schramm / Andre Gysbers / Tilo Hartmann / Christoph Klimmt / Jari Laarni / Nicklas Ravaja / Ana Sacau / Thomas Baumgartner / Petra Jäncke: MEC Spatial Presence Questionnaire (MEC-SPQ): Short Documentation and Instructions for Application. Report to the European Community, Project Presence: MEC (IST-2001-37661), 2004. [\[online\] Jeschke et al 2015.](#)

Wilhelm Voßkamp: Gattungen als literarisch-soziale Institutionen. Zu Problemen sozial- und funktionsgeschichtlich orientierter Gattungstheorie und -historie. In: Textsortenlehre – Gattungsgeschichte. Hg. von Walter Hinck. Heidelberg 1977, S. 27–44. [\[Nachweis im OPAC\] Viehhauser 2015.](#)

W

Philippe Wampfler: »online first«. Geisteswissenschaften als Social Media. In: Zürcher Jahrbuch für Wissenschaftsgeschichte. Digital Humanities. H. 9 (2013), S. 79–102. [\[Nachweis im OPAC\] Ernst 2015.](#)

Scott Weingart: Submissions to Digital Humanities 2015 (mehrere Teile). In: the scottbot irregular. 2013f. [\[online\] Sahle 2015.](#)

Scott Weingart: Acceptances to Digital Humanities 2014. In: the scottbot irregular. 2013f. [\[online\] Sahle 2015.](#)

Scott Weingart: Submissions to Digital Humanities 2014. In: the scottbot irregular. 2013f. [\[online\] Sahle 2015.](#)

Scott Weingart: Acceptances to Digital Humanities 2013. In: the scottbot irregular. 2013f. [\[online\] Sahle 2015.](#)

Werner Welzig: Wörterbuch der Redensarten zu der von Karl Kraus 1899 bis 1936 herausgegebenen Zeitschrift *Die Fackel*. Wien 1999. [\[Nachweis im OPAC\] Biber 2015.](#)

Werner Welzig: Schimpfwörterbuch zu der von Karl Kraus 1899 bis 1936 herausgegebenen Zeitschrift *Die Fackel*. Alphabetisches – Chronologisches – Explikatives. Wien 2008. [\[Nachweis im GBV\] Biber 2015.](#)

Werner Welzig: »Wie? Wo? Weswegen?« In: Paratexts, Vorwort (April 2007), AAC - Austrian Academy Corpus: AAC-Fackel. Online Version: *Die Fackel*. Herausgeber: Karl Kraus, Wien 1899–1936. AAC Digital Edition No. 1. Hg. von Hanno Biber / Evelyn Breiteneder / Heinrich Kabas / Karlheinz Mörth. [\[online\] Biber 2015.](#)

Frank Wende: Briefe aus dem Exil. 1933–1945. In: Der Brief. Eine Kulturgeschichte der schriftlichen Kommunikation. Hg. von Klaus Beyrer / Hans-Christian Täubrich. Frankfurt/Main 1996, S. 172–183. [\[Nachweis im OPAC\] Biehl et al 2015.](#)

Aude Wirth: Les comptes des receveurs et leur intérêt pour l'histoire sociale. L'exemple de la Lorraine. In: Noms de métiers et catégories professionnelles. Acteurs, pratiques, discours (XVe siècle à nos jours). Hg. von Hanne Georges / Judde de Larivière Claire. Toulouse 2010, S. 97–111. [\[Nachweis im GBV\] Vogeler 2015.](#)

Harald Willenbrock: Wie im richtigen Leben. In: Brandeins 14 (2012), H. 10, S. 86–91. [\[Nachweis im GBV\] Jeschke et al 2015.](#)

Tim Williams: The Silk Roads: An ICOMOS Thematic Study, Charenton-le-Pont 2014. [online](#)] [Lange / Unold 2015](#).

Wybo Wiersma / John Nerbonne / Timo Lauttamus: Automatically extracting typical syntactic differences from corpora. In: LLC: the Journal of Digital Scholarship in the Humanities 26.1 (2011), S. 107–124. DOI:10.1093/lc/fqq017. [\[Nachweis im GBV\] Nerbonne 2015](#).

Michael Wimmer / Claus Scheiblaue: »Instant points: Fast rendering of unprocessed point clouds«. In: Proceedings of the Eurographics Symposium on Point-Based Graphics. Hg. von Mario Botsch / Christine Chen. Zürich 2006, S. 129–136. [\[Nachweis im GBV\] Lange 2015](#).

Werner Wirth / Matthias Hofer: Präsenzerleben. Eine medienpsychologische Modellierung. In: Montage AV. Zeitschrift für Theorie und Geschichte audiovisueller Kommunikation 2008, S. 159–175. [\[Nachweis im GBV\] Jeschke et al 2015](#).

Ulla Wischerman: Frauenbewegungen und Öffentlichkeiten um 1900. Netzwerke – Gegenöffentlichkeiten – Protestinszenierungen. Königstein 2003. [\[Nachweis im GBV\] Vogeler 2015](#).

Bob G. Witmer / Michael J. Singer: Measuring Presence in Virtual Environments: A Presence Questionnaire. In: Presence: Teleoperators and Virtual Environments 7 (1998), H. 3, S. 225–240. [\[Nachweis im GBV\] Jeschke et al 2015](#).

Kerstin Wolff: Ein ungewöhnlicher Schreib-Ort? Frauenrechtlerinnen im deutschen Kaiserreich und ihr politisches Schreiben im Frauenverein – Eine Annäherung. In: Frauen in der literarischen Öffentlichkeit 1780–1918. Hg. von Caroline Bland / Elisa Müller-Adams. Bielefeld 2007, S. 121–142. [\[Nachweis im GBV\] Vogeler 2015](#).

Franz Josef Worstbrock: Wiedererzählen und Übersetzen. In: Mittelalter und frühe Neuzeit. Übergänge, Umbrüche und Neuansätze. Hg. von Walter Haug. Tübingen 1999, S. 128–142. [\[Nachweis im OPAC\] Viehhauser 2015](#).

X

XBRL Global Ledger Taxonomy, Specifications. 17. April 2007. [\[online\] Vogeler 2015](#).

XSL Transformations (XSLT) Version 2.0. W3C Recommendation. Hg. von Michael Kay. 23. Januar 2007. [\[online\] Vogeler 2015](#).

Y

Basil S. Yamey: Scientific Bookkeeping and the Rise of Capitalism. In: English History Review N.S. 1 (1949), S. 99–113. [\[Nachweis im GBV\] Vogeler 2015](#).

Mahir Yavuz: Mapping The Archive: 30 Years of Ars Electronica Visualized in Huge Scale. Information Aesthetics Weblog. 16.09.2009. [\[online\] Kwastek 2015](#).

Z

Raphael Zender / Ulrike Lucke: Synchronisationsmechanismen in Mixed Reality-Arrangements. In: Lernen und Lehren in virtuellen Welten. Zeitschrift für E-Learning – Lernkultur und Bildungstechnologie 7 (2012), H. 1, S. 47–60. [\[Nachweis im GBV\] Jeschke et al 2015](#).