

Aesthetics are (ir)relevant: Für eine Neue Visuelle Hermeneutik in den Geisteswissenschaften

Roxana Kath*

Schlüsselwörter: Digital Humanities, Textanalyse, Diskursanalyse, Kookkurrenzanalyse, New Visual Hermeneutics

Abstract: Der Aufsatz diskutiert die Chancen und Risiken der Nutzung von computergestützten Tools zur Textanalyse in den Altertumswissenschaften am Beispiel von Kookkurrenzanalysen. Er vertritt die These, dass solche Tools in einen neuen methodischen und erkenntnistheoretischen Rahmen verortet sein müssen und skizziert mit dem Ansatz einer New Visual Hermeneutics einen solchen Rahmen.

Abstract: This paper discusses the chances and risks of using computerized tools for text analysis in the Humanities exemplified by cooccurrence analyses. It will be argued that these tools have to be assessed in the context of a new methodological and epistemic framework which will be sketched by an approach here referred to as New Visual Hermeneutics.

0. Einleitung

Die Tools zur computergestützten Textanalyse aus den Digital Humanities (DH) werden die Forschungspraxis in den Geistes- und den Sozialwissenschaften in den nächsten Jahren und Jahrzehnten nachhaltig verändern.¹ In Deutschland werden innerhalb der Geisteswissenschaften computergestützte Textanalysetools insbesondere in den Altertumswissenschaften innovativ genutzt. Für die Nutzung von computergestützten Textanalysetools in der Politischen Theorie – oder genereller: für alle sozialwissenschaftlichen Forschungen, die text- oder diskursanalytisch vorgehen – ist es daher sinnvoll, die Herausforderungen und Probleme der DH, die bereits heute in der Forschungsarbeit der Altertumswissenschaften aufscheinen, zur Kenntnis zu nehmen.

Der vorliegende Aufsatz möchte den Brückenschlag von der Alten Geschichte zur Politischen Theorie leisten. Er vertritt auf Basis der bisherigen Forschungserfahrung in den Verbundprojekten *eAQUA* und *eXChange* die These, dass der Einsatz computergestützter Analysetools kein *add-on* ist, das sich bruchlos in bestehende geisteswissenschaftliche For-

* Dr. Roxana Kath, Universität Leipzig
Kontakt: rkath@uni-leipzig.de

1 Ich danke Michaela Rücker und Gary S. Schaal für ihre konstruktive Kritik an einer ersten Fassung dieses Aufsatzes.

schungspraxen integrieren lässt. Vielmehr wird ein neuer, methodisch und erkenntnistheoretisch fundierter Ansatz benötigt, um mit den Tools der DH zu *validen* wissenschaftlichen Erkenntnissen zu gelangen. Dieser Ansatz wird als *New Visual Hermeneutics* bezeichnet (vgl. Kath et al. 2014; Schaal/Kath 2014) und in seinen Grundzügen im Abschnitt 3 ausgeführt. Die Notwendigkeit einer Neuen Visuellen Hermeneutik wird anhand einer exemplarischen Diskussion der Chancen und Risiken von Kookkurrenzanalysen belegt (Abschnitt 2). Das Argument lautet, dass die primäre Form der Wissensgenerierung in den DH die Form von Visualisierungen annimmt (vgl. Keim et al. 2010). In Anlehnung an eine kanonische Differenzierung von Moretti (2013) wird erstens argumentiert, dass die strikte Opposition von *close* und *distant reading* zugunsten eines integrierten Lesemodus, dem *blended reading* (vgl. Lemke/Stulpe 2014) überwunden werden muss. Doch auch *blended reading* muss die Tatsache zur Kenntnis nehmen, dass jede Visualisierung das Ergebnis eines mehrstufigen Prozesses mit mehrfachen Datentransformationen ist, dessen Gesamtkenntnis für die Validität der Interpretation einer Visualisierung entscheidend ist (vgl. Kath et al. 2014).

Die skizzierte Diskussion über die Chancen und Risiken der Nutzung von DH Tools wird im Abschnitt 1 eingebettet in die Analyse eines quasisokratischen Dialoges, geführt von Seven of Nine und Naomi an Bord des Raumschiffes Voyager, denn in diesem Dialog spiegeln sich – *avant la lettre* – die Hoffnungen und Ängste, die mit den computergestützten Analysetools heute verbunden werden, bereits 1999 wider.

1. Hive-Mind & Big Data

Stardate: unknown

„Seven: Naomi Wildman, state your purpose here.

Naomi: I was waiting for you. Kadis-Kot, remember? Our weekly game?

Seven: Reschedule for tomorrow. I'm working now.

Naomi: Redecorating isn't work.

Seven: Aesthetics are irrelevant. I'm modifying the alcove to function as a cortical processing subunit.

Naomi: That was my next guess.

Seven: Then I won't need to explain its purpose.

Naomi: Okay, you win. What do you do with the cortical processing, uh Subunit?

Seven: On Borg vessels, there's one in each unimatrix. It downloads newly assimilated data to the drones.

Naomi: But Voyager isn't a Borg vessel, and you're not a drone anymore.

Seven: Voyager collects a great deal of information – sensor scans, navigational projections, engineering updates, away team reports, scientific analyses.

Naomi: The results of our Kadis-Kot tournament? That, too.

Seven: The crew must read and study this information – an inefficient procedure. These data nodes are downloading the information into the alcove.

Naomi: Which is going to download it into you.

Seven: Precisely. In a matter of hours, I will assimilate several months worth of data.

Naomi: Can I try it sometime?

Seven: Your physiology is different from mine.

Naomi: Guess I'll stick to reading my lessons and listening to Neelix tell stories.²

2 Dialog aus *The Voyager Conspiracy* (VOY 6, 09; Production Nr. 229, erste Ausstrahlung: 24.11.1999; http://www.springfieldspringfield.co.uk/view_episode_scripts.php?tv-show=star-trek-voyager&episode=s06e09, 25.05.2014.

Dieses ‚sokratische‘ Gespräch entstammt einer Science-Fiction-Serie. Die Gesprächspartner sind Seven of Nine, die als kleines Mädchen von den Borg, einer transhumanen Lebensform, assimiliert wurde, und Naomi Wildmann, die an Bord des Raumschiffes Voyager aufwächst. Ort der Handlung ist eben dieses Raumschiff, das durch einen Unfall in den Delta-Quadranten der Milchstraße gelangte und dessen Crew unter der Führung von Captain Janeway eine Odyssee nach Hause erlebt. Unterwegs trifft die Mannschaft auf eine Vielzahl fremder Kulturen, die die moralischen Prinzipien der Sternenflotte, das heißt der Menschheit, immer wieder neu auf die Probe stellen und dadurch den ZuschauerInnen einen Spiegel vorhalten. In diesem Sinne erinnert die Erzählstruktur der Serie sehr stark an antike und mittelalterliche Reiseberichte in der Tradition von Herodot. Der Reiz der Gespräche zwischen Seven und Naomi besteht darin, dass im Gegensatz zum platonischen Dialog Seven nur scheinbar ‚allwissend‘ ist und man nie ganz sicher sein kann, wer am Ende von wem etwas lernen wird. Tatsächlich ist es sehr oft so, dass es eigentlich Naomi ist, die Seven eine Lektion in ‚Menschlichkeit‘ erteilt.

Die Borg sind eine fiktive kybernetische Spezies,³ deren oberstes Ziel Perfektion ist. Das Streben danach ist geprägt von Effizienz. Das bedeutet beispielsweise, dass die Borg Wissen nicht durch eigene Forschung erwerben, sondern durch Assimilation anderer Zivilisationen, sofern deren (technische) Entwicklungsstufe für würdig befunden wird, die Borg dem Ziel der Perfektion näherzubringen. Das vollständige Wissen und alle Erfahrungen jedes einzelnen Individuums gehen in dem Prozess der Assimilation in das sogenannte *Hive-Bewusstsein* ein und werden Bestandteil eines kollektiven Gedächtnisses. Wenn eine Drohne ‚funktionsunfähig‘ wird, bleibt ihr Wissen dennoch im *Hive-Mind* erhalten. Dadurch sind die Borg potentiell unsterblich. Es gibt auch kein Vergessen. Im *Hive-Mind* befindet sich relevantes neben nicht relevantem Wissen. Damit dies nicht ein Chaos erzeugt, sorgt an Bord der Borg-Raumschiffe ein sogenanntes *Vinculum*⁴ dafür, dass nur relevantes Wissen über ein Kortikalimplantat an die einzelnen Drohnen weitergegeben wird.

Diese Technik versucht Seven an Bord der Voyager nachzubilden. Sie will die vorhandenen Datenbanken mit Schiffslogbüchern, Sensoraufzeichnungen und dergleichen der vergangenen fünf Jahre in ihr auch nach der Trennung vom Borg-Kollektiv noch vorhandenes Kortikalimplantat laden. Zunächst funktioniert diese Übertragung auch sehr gut und führt zu einer für die Crew erstaunlichen Kombinationsgabe und Problemlösungsfähigkeit. Doch je mehr Daten Seven herunterlädt, desto komplexer werden ihre Beweisketten. Dies kulminiert schließlich darin, dass Seven innerhalb kürzester Zeit drei verschiedene (Verschwörungs)theorien darüber vorlegt, wie die Voyager in den Delta-Quadranten gelangt sei, die alle falsch sind. Sie verbindet ausgehend von einem Schlüsselereignis – der Zerstörung der Phalanx des Fürsorgers (gezeigt in der Pilotfolge der Serie) – verschiedene Fakten zu Ereignisketten, die zum einen das Erscheinen der Voyager im Delta-Quadranten als einen Interventionsplan der Sternenflotte, zum anderen als eine Verschwörung des sogenannten Maquis erscheinen lassen. Die dritte Theorie fokussiert auf Seven of Nine selbst, deren Assimilation durch die Borg und Reintegration in die Mensch-

3 Vergleiche zu den Borg unter anderem Stoppe (2011); Balinisteanu (2007); Bostic (1988); Brkich/Barko (2012).

4 „The processing device at the core of every Borg vessel. It interconnects the minds of all the drones. It purges individual thoughts and disseminates information relevant to the Collective. Janeway: It brings order to chaos. Seven: Precisely.“ Dialog aus: *Infinite Regress* (VOY 5, 07; erste Ausstrahlung: 25.11.1998); <http://www.chakoteya.net/voyager/509.htm>, 25.05.2014.

heit ebenfalls ein Plan der Sternenflotte zum Erwerb von Wissen zur Vernichtung der Borg sei. Aufgrund des anfänglichen Erfolges der Technologie und der starken Suggestivkraft der ‚Beweise‘ verbreitet sich zeitweilig Misstrauen unter der Crew – insbesondere zwischen dem Captain und dem ersten Offizier. Die zentrale Botschaft der Episode *The Voyager Conspiracy* ist es zu zeigen, dass nur ein starkes, auf langjähriger persönlicher Erfahrung basierendes gegenseitiges Vertrauen die Suggestivkraft dieser algorithmisch erzeugten ‚Fakten‘ zu überwinden vermag.

Letztlich gelangt der Captain zu der Überzeugung, dass mit Seven beziehungsweise dem von ihr modifizierten Alkoven etwas nicht stimmen kann. Die Diagnose des Doctors bestätigt dies: „Seven downloaded too much data into her cortical implant. She’s trying to make sense of more information than she can process.“ Danach gelingt es dem Captain, Seven selbst davon zu überzeugen, dass sie der Informationsfülle und deren Suggestivkraft erlegen ist: „Your modified alcove threw your synaptic patterns into chaos, and your mind can’t make sense of all the information, so you’re generating theory after theory in an attempt to bring order to that chaos.“ Einmal mehr fügt sich Seven daraufhin in ihre (Re)transformation in einen Menschen und wendet sich von der Borg-Technologie ab.

„Naomi: Guess who assimilated three books and ten reports in two days?

Seven: Naomi Wildman.

Naomi: That’s correct. Are you taking your new alcove apart?

Seven: Yes.

Naomi: Why?

Seven: My attempt to download Voyager’s database failed.

Naomi: Maybe you can fix it.

Seven: The alcove functioned within expected parameters. Unfortunately, I did not. Three books and ten reports is impressive, but quantity is less relevant than quality. You must be able to interpret the data and enjoy the process.“⁵

Zum Zeitpunkt der Ausstrahlung der Episode im November 1999 war Big Data in der heutigen Form noch Science-Fiction. Dennoch nahm mit dem Auftreten von Google am 27. September 1998⁶ eine Entwicklung ihren Anfang, die in Form von Facebook, Twitter, Amazon & Co. heute sowohl unseren Alltag als auch unter dem Etikett *Digital Humanities*⁷ zunehmend die geisteswissenschaftliche Forschung bestimmt.⁸ Entstanden ist ein Internet ohne Vergessen, in dem Algorithmen wie das Borg-Vinculum versuchen, „Ordnung ins Chaos zu bringen“ und wichtiges und weniger wichtiges Wissen allzeit abrufbar zu machen, nach Relevanzkriterien, die der NutzerIn häufig entzogen sind (vgl. Becker 2012). „Resistance is futile“ könnte man angesichts der Omnipräsenz und beständigen Verbreitung der algorithmenbasierten Technologien denken und Sevens Sinneswandel wirkt aus heutiger Perspektive geradezu antiquiert. Im Sinne Naomis („Maybe you can fix it“) scheint der allgemeine Trend eher in die Richtung immer besserer Algorithmen zu verlaufen.

Das sich beständig vergrößernde Angebot neuer digitaler Recherchertools sowie neuer digitaler Editionen und Forschungsliteratur ist nicht nur eine Chance für Geistes- und SozialwissenschaftlerInnen, sondern auch eine Last. Was darf man als ForscherIn ignorieren

5 Vergleiche Anmerkung 3.

6 Von diesem Zeitpunkt an war die Suchmaschine, deren Vorläufer *BackRub* (Start 1996) hieß, unter dem Namen *Google* online.

7 Zur Begrifflichkeit vergleiche Manovich (2012) und Schubert (2013).

8 So haben wahrscheinlich die meisten ForscherInnen schon einmal den *Google Ngram Viewer* benutzt.

beziehungsweise weglassen? Was muss man unbedingt benutzen? Im besten Fall unterstützen die in den DH entwickelten Tools die WissenschaftlerInnen darin, das Material schneller nach relevanten Informationen zu durchsuchen und (vor)zustrukturieren.

Doch fehlt es uns noch an Erfahrungen, an einer *best practice*, nicht nur hinsichtlich der Nutzung der neuen Tools, sondern auch bezüglich der Interpretation ihrer Ergebnisse. Am problematischsten ist jedoch, dass die DH *community* – vor allem in den USA – eine grundlagentheoretische und epistemische Reflektion für überflüssig hält. Ihr Mantra lautet „doing is theory“. Die nicht unsympathische Hinwendung zur faktischen Analyse und zur Programmierung immer neuer Tools führt jedoch dazu, dass das grundlegend Neue der DH aus dem Blick gerät. Das Neue, Innovative soll im Folgenden näher betrachtet werden.

2. „You can’t always predict how Borg technology will affect you“

Die Erstellung digitaler Bibliotheken in großem Umfang hat in den Altertumswissenschaften (Alte Geschichte, Klassische Philologie, Epigraphik, Papyrologie) eine lange Tradition und die mit Texten arbeitenden Disziplinen können heute auf digitalisierte Textcorpora in unterschiedlichsten Formen (vor allem Online-Bibliotheken, Datenbanken auf CD-ROM) und Formaten (Beta-Code, UTF-8, ASCII) zurückgreifen.

Datenbank	Bestände
TLG: Thesaurus Linguae Graecae www.tlg.uci.edu	Sprache: Altgriechisch Umfang: 76.000.000 Worte auf der TLG E CD ROM (2000); TLG-online 105.000.000 Worte. Digitalisierung der antiken griechischen Texte von 800 v. Chr. bis 600 n. Chr., seit einigen Jahren auch Erfassung der byzantinischen Texte bis 1453 n. Chr. als Projekt der University of California, Irvine unter der Leitung von Maria Pantelia.
Perseus Digital Library: www.perseus.tufts.edu	Sprachen: Altgriechisch; Latein; Englisch; Arabisch; Deutsch; Italienisch Umfang - Klassische Texte: 68.137.849 Worte Im Rahmen seiner Humboldt Professur für Digital Humanities an der Universität Leipzig initiiert Gregory R. Crane, der Editor-in-Chief der Perseus Digital Library gerade ein Open Philology Project, mit dem jeder antike Texte in einer Edition frei verfügbar gemacht werden soll.
PHI 5.3: Packard Humanities Institute: www.packhum.org	Sprache: Latein Die CD ROM des Packard Humanities Institute (Los Altos in Kalifornien) enthält im Wesentlichen die gesamte lateinische Literatur bis zum Jahr 200 n. Chr., einige spätere Autoren (z. B. Servius, Porphyrius, Zenon, Justinian) und mehrere Fassungen des Bibeltextes. PHI # 5.3 steht mittlerweile frei verfügbar im Netz.
BTL: Bibliotheca Teubneriana Latina	Sprache: Latein Umfang - online-Edition von BTL: ca. 13 Millionen Worte Die CD Roms BTL 1–4 (erschienen 1999–2006, 2009) sowie die Online-Version enthalten die lateinischen Autoren vom 3. Jh. v. Chr. bis zur Spätantike; in BTL 4 erfolgte die systematische Integration von mittel- und neulateinischen Teubnerausgaben und die Neuaufnahme der lateinischen Texte, die von 2003 bis 2004 in der BT neu erschienen sind.

Tab. 1: Übersicht über die wichtigsten und umfangreichsten digitalen Textcorpora

Damit haben die Altertumswissenschaften einen Stand erreicht, der sie vor vielen anderen Geistes- und Sozialwissenschaften für die Anwendung von Text-Mining-Verfahren⁹ und Visualisierungen prädestiniert. Diese Tatsache hat entscheidend dazu beigetragen, dass das BMBF die Weiterentwicklung und Anwendung von Methoden aus dem *Information*

9 Definition von Gerhard Heyer (2006: 3): „Mit dem Terminus Text Mining werden computergestützte Verfahren für die semantische Analyse von Texten bezeichnet, welche die automatische bzw. semi-automatische Strukturierung von Texten, insbesondere sehr großen Mengen von Texten, unterstützen.“ Vergleiche zur Einführung auch Mehler (2004); Gorunescu (2011).

Retrieval im Rahmen der Verbundprojekte eAQUA¹⁰ und eXChange¹¹ am Lehrstuhl für Alte Geschichte der Universität Leipzig seit 2008 fördert.

Die in eAQUA entwickelten Werkzeuge ermöglichen der NutzerIn die Bearbeitung sehr großer Textmengen in kurzer Zeit. Gerade dadurch neigen sie aber dazu, den gewohnten Forschungsprozess auf den Kopf zu stellen. Diese These soll entlang der bereits eingangs angesprochenen kanonischen Differenzierung von *close* und *distant reading* von Moretti (2013) ausgeführt werden. Traditionell wurde und wird ein großer Teil der geisteswissenschaftlichen Arbeit ‚händisch‘ in der Bibliothek geleistet. Bücher, Aufsätze, Fragmente und Ähnliches werden *einzel*n und *seriell* von einer WissenschaftlerIn gelesen. Dieses genaue Lesen realer Textquellen, *one at a time*, bezeichnet Moretti als *close reading*. Dem *close reading* sind jedoch Grenzen gesetzt, die in der Endlichkeit der menschlichen Existenz und der mit ihr verbundenen Ressourcen (unter anderem Zeit, Geld, Aufmerksamkeit) begründet liegen. Dem stellt Moretti die Idee des *distant reading* gegenüber. Hier wird quasi eine Vogel- oder Makroperspektive eingenommen; computergestützt können fast unendlich viele Texte gleichzeitig aufgearbeitet werden und in synchroner oder asynchroner und komparativer Perspektive analysiert werden. Der Blick richtet sich auf die Makroebene und verlässt notwendigerweise die Ebene eines konkreten Textes. Es ist unmittelbar ersichtlich, dass beide Methoden des Lesens in einem trade-off stehen. Das Detailniveau der Analyse sinkt vom *close* zum *distant reading* dramatisch; dafür steigt der potentielle abstrakt-generalisierte Erkenntnisgewinn et vice versa. Das *distant reading* ist somit ein Antidot gegen die Informationsüberflutung¹² und insbesondere Visualisierungen sind für *distant reading* geeignet. *Visual Analytics* ist jener Bereich innerhalb der DH, in der die visuelle Aufarbeitung von Daten adressiert wird: „On a grand scale, visual analytics provides technology that combines the strengths of human and electronic data processing. Visualization becomes the medium of a semi-automated analytical process, where humans and machines cooperate using their respective, distinct capabilities for the most effective results.“ (Keim et al. 2010: 2) Die Frage lautet nur: Unter welchen Bedingungen kommen wir beim *distant reading* zu *validen* wissenschaftlichen Einsichten? Sind die Methoden hierfür bereits ausgearbeitet und die *traps* und *pitfalls* bekannt?

Eine große Herausforderung sind die Visualisierungen von *distant readings*, die den direkten ‚Zugriff‘ auf alle Quellenstellen auf einmal in einer bisher *unbekannten* Form ermöglichen und zwar im Extremfall noch vor der Konsultation von Sekundärliteratur. Positiv gewendet ordnet sich dieser Zugang in die lange Tradition der Wissenschaftsgeschichte ein, denn bereits:

„Bernhard von Chartres sagte, dass wir nur Zwerge seien, die auf den Schultern von Riesen sitzen, damit wir mehr und weiter als diese sehen können, nicht etwa dank eigener größerer Sehschärfe oder Körpergestalt, sondern weil uns die Größe der Riesen in die Höhe erhebt.“¹³

10 www.eaqua.net (BMBF-Förderung 2008–2011/12).

11 Das Verbundprojekt „eXChange: Exploring Concept Change and Transfer in Antiquity“ untersucht das Verhältnis zwischen Wissenschaftssprache und alltäglichem Handeln von der Antike bis in die frühe Neuzeit (www.exchange-projekt.de; BMBF-Förderung 2012–2015).

12 Vergleiche Moretti (2009; 2013) sowie zur Kritik an Moretti unter anderem Goodwin et al. (2011).

13 *Dicebat Bernardus Carnotensis nos esse quasi nanos, gigantium humeris incidentes, ut possimus plura eis et remotiora videre, non utique proprii visus acumine, aut eminentia corporis, sed quia in altum subvehimur et extollimur magnitudine gigantea.* Joannes Saresberiensis; *Metalogicon*, Migne, *Patrologia Latina* Volumen 199, Sp. 900; (= Johannes of Salisbury: *Metalogicon* 3,4,46–50; hrsg. John B. Hall: *Ioannis*

Die Visualisierungen versetzen uns gleichsam in die Lage, aus der Vogelperspektive auf die Masse der überlieferten Texte zu schauen. Die Algorithmen hinter den Visualisierungen sorgen dafür, dass *bestimmte* Phänomene visuell besonders herausgehoben werden. Doch wie interpretiert man visuelle Zusammenhänge *inhaltlich*? Und wie berücksichtigt man angemessen die Tatsache, dass keine Form der Visualisierung erkenntnistheoretisch neutral ist?¹⁴

2.1 „Seven has some compelling evidence“

Eine wichtige Form der Visualisierung ist die Visualisierung von Kookkurrenzen. Als Kookkurrenzen werden Wörter bezeichnet, die innerhalb eines zuvor definierten Rahmens zusammen auftreten. Dieser Rahmen kann ein Satz, ein Absatz, ein Artikel et cetera sein. Man differenziert zwischen linken und rechten Kookkurrenzen: In dem Satz „Die Kinder fahren nach Hause.“ ist „Kinder“ zum Beispiel die linke Kookkurrenz von „Hause“. Visualisierungen von Kookkurrenzen sind in der Lage, semantische Zusammenhänge aufzuzeigen, die durch herkömmliche Suchstrategien (Verwendung von Lexika, Nachschlagewerken, Konkordanzen, Indizes, Bool'sche Wortsuche) nicht erkannt werden. Insbesondere seltene Kookkurrenzen, die im gesamten Korpus – das heißt in unseren Forschungsprojekten in der gesamten griechischen oder lateinischen Literatur der Antike – lediglich ein- bis dreimal auftreten, können zu fachlich neuen Erkenntnissen führen. Da der entsprechende Zusammenhang durch Lesen lediglich durch einen ‚glücklichen Zufall‘ zu finden wäre, hat Charlotte Schubert (2013) in diesem Kontext in Anlehnung an Robert K. Merton und Elinor Barber (2006) auch von *serendipity* gesprochen:¹⁵

„Diese Kombination aus Überraschung, nicht-beabsichtigtem Ergebnis und strategischer Bedeutung erweist ‚serendipity‘ als einen speziellen Zufall, der sich vor allem durch seine Wirkung und Anschlussfähigkeit auszeichnet, aber auch grundsätzlich auf einer völlig anderen Prämisse ruht. Das informationswissenschaftliche Verständnis der explorativen Suche geht demgegenüber von einer analytisch definierten Suchstrategie aus, die mit einer facettierten Suche kombiniert wird.“ (Schubert 2013: 183; vgl. Heyer et al. 2006)

Entdeckungen im Sinne von *serendipity* sind in der Arbeit mit Kookkurrenzgraphen meines Erachtens nicht die Ausnahme, sondern die Regel. Dies liegt vor allem an der Eigenart der interaktiven Visualisierungen, die immer neue Blickwinkel und Fokussierungen (Ausschnitte, Zoom und dergleichen) auf dasselbe (bekannte) Material erlauben.

Saresberiensis metalogicon, Turnhout 1991: 116); zur Traditionslinie dieses häufig Isaac Newton zugeschriebenen Aphorismus vergleiche Robert K. Merton (1983).

14 Die konkrete Gestaltung einer Visualisierung – zum Beispiel die Position, Schrift- und Hintergrundfarbe oder die Größe von Wörtern in Kookkurrenzgraphen – hat immensen Einfluss auf das interpretative Ergebnis. Zudem erfassen WissenschaftlerInnen abhängig von ihrem Kulturkreis Inhalte von Visualisierungen in unterschiedlicher Reihenfolge – also von links nach rechts, von oben nach unten oder umgekehrt. Unbewusst kann entweder Begriffen in der Bildmitte oder am Rand größere Aufmerksamkeit zuteil- oder gar Bedeutung zugeschrieben werden (vgl. Schaal/Kath 2014). So wie Seven sind Menschen Sinnsucher und neigen dazu, eher Struktur und Kausalität zu erkennen als Zufall und Kontingenz. Daher tendieren Visualisierungen grundsätzlich dazu, überinterpretiert zu werden. Dies ist besonders dann problematisch, wenn die ForscherIn das zugrundeliegende Textkorpus nicht gut kennt (vgl. Schmidt 2012; Schaal/Kath 2014; Kath et al. 2014).

15 Zur Diskussion von *serendipity* in der Wissenschaft vergleiche unter anderem Campa (2008; 2013).

skythischen Nomaden. Die Athener selbst verglichen die Evakuierung ihrer Stadt während der Perserkriege mit der Taktik der Skythen, dem Feind auszuweichen (Herodot 7,10,2), die (jeweils) zum Sieg über die Perser geführt habe (Herodot 8,61; 7,140,1). Nicht zuletzt bezeichnete Aristoteles (Politik 1319a 21–26) die Nomaden aufgrund von Charakteristika ihrer Lebensweise (Freiheit, Gleichheit, Gerechtigkeit) nach den Ackerbauern als am besten geeignet für die Entwicklung der Demokratie (vgl. Schubert 2010c: 18, 183 ff.). Insofern haben die Historiographen des vierten Jahrhunderts vor Christus anscheinend ein philosophisches Modell auf ihre Konzeption der athenischen Geschichte übertragen (Schubert 2011: 44).

An dieser Stelle wird die Frage relevant, ob hier ein *systematischer* Vorteil der computergestützten Textanalyse vorliegt oder ein kontingenter Effekt der beteiligten Forscherin, denn möglicherweise hätte Charlotte Schubert diese Relation zwischen athenischer Frühgeschichte und dem Nomadismus ohne ihre Vorerfahrung aus dem sogenannten Nomaden-SFB 586 „Differenz und Integration“ (www.nomadsed.de) nicht hergestellt. Wie bereits Goethe (1948: 52) in einem Brief an Friedrich von Müller schrieb: „Man erblickt nur, was man schon weiß und versteht.“¹⁷ Die Abhängigkeit vom Vorverständnis ist jedoch bei der Analyse von Visualisierungen besonders groß.¹⁸ Dass das Aufscheinen von zufälligen Zusammenhängen in Kookkurrenzgraphen (zunächst) auch in die Irre weisen kann, zeigt das folgende Beispiel.

2.2 „Your reasoning is flawed“

Die folgende Abbildung zeigt einen Kookkurrenzgraphen zu dem griechischen Wort βακτηρία (Bakteria = Stab, Stütze). Die Analyse des Wortfeldes ist Bestandteil einer Diskursanalyse zur vielfältigen symbolischen Verwendung des Stab-Begriffes in der antiken Literatur¹⁹ – insbesondere als Attribut von Wanderern oder Bettlern (zum Beispiel auch Odysseus als Bettler), Hirten, Ärzten und Philosophen (zum Beispiel die Kyniker; vgl. Kath 2012: 137).

17 Vergleiche auch Kurt (2008: 371).

18 Auch die eindrucksvollen Analysen von Moretti (2009; 2013) wären ohne seine immensen Vorkenntnisse nicht möglich.

19 Die Ergebnisse werden an anderer Stelle zusammen mit Michaela Rücker publiziert, der ich an dieser Stelle ganz herzlich für wertvolle Hinweise danke.

Richterbestellung für die Dikasterien, wonach die Richter einen Stab in der Farbe ihres Gerichtssaals erhalten. Im *Plutos* sagt Karion in Anspielung darauf zum Chor, sie hätten ihr *Symbolon* – das heißt den Stab als die farblich codierte Eintrittskarte für den Gerichtssaal – von Charon erhalten. Aristophanes hat zuvor ein Wortspiel eingefügt, indem er den Gerichtssaal mit einem Grab vergleicht – die Nennung von Charon als Überbringer des Zeichens nimmt diese mythologische Deutung wieder auf (vgl. Droysen 1871: 431, Anmerkung 277). Es handelt sich folglich – anders als zunächst erwartet – um die witzige Vermischung eines institutionellen mit einem mythologischen Kontext. Das alles geht weder aus der Visualisierung noch aus der zu dieser Suchanfrage gehörenden Belegstelle hervor, da diese aufgrund der Satzgrenze den Kontext nicht wiedergibt (vgl. zu dieser Problematik Abschnitt 3.1).

Einen ähnlichen Fall von Überraschung oder Irreführung beschreibt Kirschenbaum (2011: 34). In einer frühen Phase des *nora*-Projektes haben Steve Ramsay und Bei Yu versucht, Shakespeares Werke den traditionellen Kategorien entsprechend automatisch zu klassifizieren. Der Computer habe die meisten Werke erwartungsgemäß richtig zugeordnet, allerdings habe er dabei *Othello* als Komödie klassifiziert. Auf der Suche nach der Ursache für diese ‚Fehlklassifizierung‘ und der daran anschließenden Frage, was *Othello* von den anderen Tragödien unterscheidet, sei Ramsay auf einen Forschungszweig gestoßen, der gerade diese Frage adressiert.

In beiden Fällen konnten Text-Mining und Data-Mining durch das Moment der Überraschung möglicherweise interessante Forschungsfragen aufzeigen. Die eigentliche/reich-tige Bedeutung beziehungsweise der Hintergrund der Beobachtung war allerdings erst durch ‚klassische‘ Recherchen zu ermitteln. In diesem Sinne sind auch die Grenzen eines *distant reading*, das auf die Rückbindung an die Quellen verzichten will, offensichtlich. Erst die Verbindung aus *distant* und *close reading* konnte die interpretatorischen Kurzschlüsse und Fehler, die sich aus dem *distant reading* alleine ergaben, überwinden. Diese Verbindung, die von Lemke und Stulpe (2014) als *blended reading* bezeichnet wird, erscheint erfolgversprechend. Doch selbst *blended reading* reicht nicht aus, wenn die Fenster des *close reading* zu eng gewählt werden. Das nächste Beispiel verdeutlicht daher die Relevanz der Kenntnis des größeren (historischen) Kontextes für die valide Interpretation von Visualisierungen von *distant readings*.

2.3 „Speculation is not evidence“

Mit Hilfe des sogenannten Zitationsgraphen von eAQUA lassen sich Zitationsprofile einzelner Autoren und Werke visualisieren. Das Programm vergleicht hierzu alle vorhandenen Texte der antiken griechischen und byzantinischen Literatur innerhalb der Einheit eines Satzes miteinander (etwa 76 Millionen Sätze) und ermittelt Zitate, die auf der Übereinstimmung identischer Wortketten von mindestens fünf Wörtern beruhen (vgl. Geßner 2010: 29 f.; ausführlich: Büchler 2013).

dem Eingang eine Farbe aufgetragen. Der Richter nimmt also seinen Stab und geht in das Gericht, das dieselbe Farbe wie sein Stab und denselben Buchstaben wie seine Eichel hat. Wenn er eintritt, erhält er von Staats wegen eine Kennmarke [*Symbolon* – gr. *σύμβολον*; Anmerkung d. A.] von dem, der für dieses Amt erlost worden ist.“ (Übersetzung Dreher).

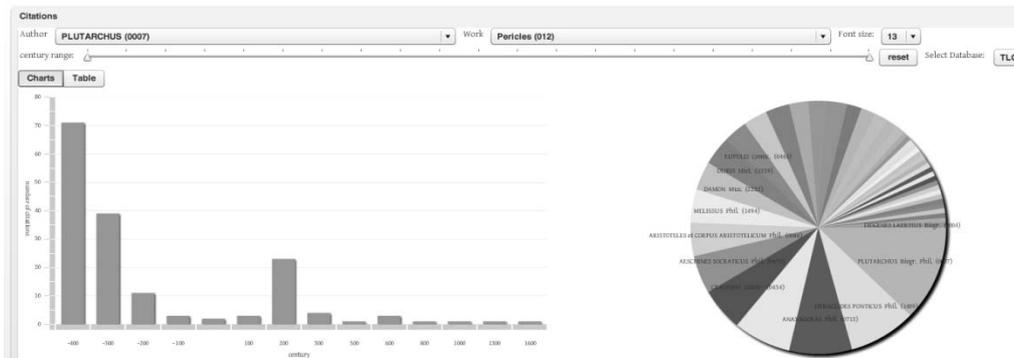


Abb. 3: Visualisierung des Zitationsprofils von Plutarch, *Perikles*

Im Rahmen einer Analyse des Zitationsprofils der Periklesvita gelangte Charlotte Schubert (2010a) zu dem Ergebnis, dass im Gegensatz zur Annahme der Forschung Plutarch (um 45–ca. 125 n. Chr.) durchaus Quellenarbeit betrieb und vor allem zeitgenössische Autoren des fünften Jahrhunderts (zum Beispiel Komödienzitate der perikleischen Zeit) sowie solche des vierten Jahrhunderts als Grundlage seiner Viten ‚großer Griechen‘ benutzte.²⁴ Legt man diese Überlegung zugrunde, wäre in einem weiteren Schritt ein Blick auf die römischen Viten interessant.

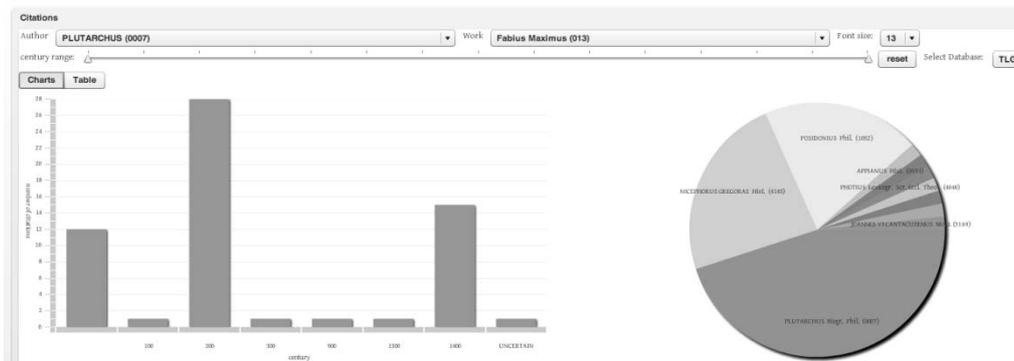


Abb. 4: Visualisierung des Zitationsprofils von Plutarch, *Fabius Maximus*

Ein Blick auf das Zitationsprofil von Plutarchs Biographie von Quintus Fabius Maximus (um 275–203 v. Chr.) zeigt neben einer Fülle von Selbstzitat, dass Plutarch hier anscheinend keine zeitgenössischen Quellen anführt.²⁵ Eine erste Erklärung fällt leicht: Da sich die Zitationssuche in eAQUA nur auf griechische Texte erstreckt, werden lateinische Autoren natürlich nicht erfasst. Die naheliegende Vermutung, Plutarch habe hierfür die lateinische Annalistik des zweiten und ersten Jahrhunderts vor Christus benutzt, ließe sich im Rahmen einer (zu entwickelnden) sprachübergreifenden Recherche möglicherweise belegen. Die Frage, ob dies der Mühe wert sei, ist zu bejahen, denn letztlich hängt davon

24 Eine detaillierte Untersuchung weiterer Viten (Themistokles, Aristeides, Kimon und Alkibiades) durch Studierende des Masterstudiengangs Classical Studies der Universität Leipzig hat diese Beobachtung bestätigt (vgl. Schubert/Klank 2012: 36, 39, 40, 47).

25 Die gleiche Beobachtung lässt sich auch für die anderen römischen Viten machen.

die Einschätzung des Quellenwerts dieses antiken Autors ab. Konkret geht es um die Frage, ob man ihn für die Analyse der Römischen Republik (3. Jh. v. Chr.) heranziehen kann oder nicht und ob die historischen und gesellschaftlichen Modelle, die darauf gegebenenfalls aufbauen, anwendbar sind oder nicht. Da Plutarch unter AlthistorikerInnen bisher keinen sonderlich guten Ruf genießt, wäre nicht viel verloren, es ließe sich aber viel gewinnen, wenn analog zu den griechischen Viten der Beweis gelänge, dass Plutarch auch für die römischen Biographien eigene Quellenarbeit betrieben hat und zeitgenössische oder zeitnahe²⁶ Autoren für sein Werk benutzte. Dennoch besteht die Möglichkeit, dass das Ergebnis negativ oder uneindeutig ausfällt, entweder weil uns die benutzten Autoren nicht erhalten sind und deren Zitate folglich nicht gefunden werden können oder weil Plutarch aus unbestimmten Gründen deren Werke ignoriert hat. Doch im Moment ist das alles (noch) Spekulation.

Ohne Zweifel hat die Neuartigkeit der Technologie(n) Auswirkungen auf unsere Arbeitsweise und unsere Fragestellungen. Die Arbeit mit Visualisierungen tendiert zu einer Analyse von Mustern (*pattern*), das heißt von Linien und verbindenden Elementen (vgl. Moretti 2009; 2013). Sie eignen sich besonders für explorative Verfahren und für das Testen von Hypothesen. Die (Re)präsentation des Materials erfordert dabei einen eigenen methodischen Schritt, der das Geworden-Sein des jeweiligen Befundes reflektiert und dessen Falsifizierbarkeit gewährleistet. Die Notwendigkeit eines solchen Ansatzes wurde schon mehrfach betont.²⁷ Haber (2011) hat hierfür den Begriff „Quellenkritik des Digitalen“ vorgeschlagen, Rawson (2011) sprach mit Blick auf den gesamten Interpretationsprozess von „Algorithmic Criticism“. Kath et al. (2014) haben aus einer interdisziplinären Perspektive eine Neue Visuelle Hermeneutik zur Diskussion gestellt. Die Grundzüge einer *New Visual Hermeneutics* als Forschungsansatz werden ebenda ausgeführt.

3. New Visual Hermeneutics

Der Ansatz einer Neuen Visuellen Hermeneutik recurriert epistemisch auf die empirische Phänomenologie von Don Ihde (1998; 2012), inkludiert methodische Ansätze aus den Sozialwissenschaften und soll den besonderen epistemischen Herausforderungen der DH Rechnung tragen. Der besondere Fokus der Neuen Visuellen Hermeneutik liegt auf einer integrativen Betrachtung des gesamten Prozesses der Visualisierung und differenziert unter Erweiterung einer in der *Visual Analytics* allgemein akzeptierten Grundstruktur vier distinkte, aber in konstitutiver Verbindung stehende Arbeitsschritte:

26 Die römische Geschichtsschreibung setzte erst sehr spät am Ausgang des dritten Jahrhunderts vor Christus ein und ist bis ins erste Jahrhundert vor Christus hinein auch nur fragmentarisch, das heißt durch Zitate, in anderen Werken überliefert. Zu den sich daraus ergebenden technisch-methodischen Problemen vergleiche 3.1.

27 Bereits Lücknerath (1968: 272) hat für die Geschichtswissenschaften eine ebensolche „vorgeschaltete Theorie“ gefordert. Vergleiche dazu auch Haber (2011: 11–23).

Aufhebung der Grenze zwischen Text und Bild die „Textualität von Geschichte“ (vgl. Montrose 1995: 67; Baßler 1995).²⁹

Die folgenden Überlegungen sollen veranschaulichen, welche Informationen im Rahmen der einzelnen Schritte ermittelt werden könnten, die wiederum helfen sollen, die Ergebnisse zu falsifizieren. Der Fokus liegt auf den Arbeitsschritten eins bis drei – vor allem auf der Phase der Visualisierung selbst.

3.1 Analyseschritt 1: Acquisition / Repository / Semantic Enrichment

Bevor Texte mit Hilfe (semi-)automatischer Verfahren analysiert und Visualisierungen erzeugt werden können, müssen sie zunächst digital erschlossen und aufbereitet werden. Bei jedem der folgenden vier Schritte – a) Print-Edition; b) Digitalisierung; c) Aufbereitung der Daten für die Datenbank (Preprocessing); d) Anreicherung von/mit Meta-Daten – werden von FachwissenschaftlerInnen, InformatikerInnen und/oder BibliothekarInnen Entscheidungen getroffen, die die Textgestalt verändern und damit auch die nachfolgenden Analysen beeinflussen.

a) Textkritische Editionen antiker Werke sind als wissenschaftliche Ergebnisse philologischer Arbeit eigentlich (Re)konstruktionen einer *möglichen* Textgestalt unter Kennzeichnung unsicherer Buchstaben oder Worte und unter Angabe alternativer Lesarten im Textkritischen Apparat. Dabei sind aus der Zeit zwischen dem achten Jahrhundert vor Christus und dem dritten Jahrhundert nach Christus 59 Prozent der Autoren nur fragmentarisch erhalten (Romanello et al. 2009: 158). Das heißt, wir kennen sie in der Regel nur durch Zitationen bei anderen Autoren.³⁰ Seit dem 18. Jahrhundert sind eine Reihe wissenschaftlicher Fragmentsammlungen entstanden, die versuchen, den Aufbau dieser Werke durch die Einteilung in Bücher und Kapitel unter Zuordnung der Fragmente zu rekonstruieren.

b) Im Zuge der Digitalisierung sind alle diese Texte in die entsprechenden Datenbanken aufgenommen worden. Daraus ergibt sich, dass Texte einzelner (fragmentarisch erhaltener) Autoren doppelt oder sogar mehrfach in der Datenbank vorhanden sind, da zum Beispiel der TLG (Thesaurus Linguae Graecae) mehrere Fragmentsammlungen (Müller und Jacoby) aufgenommen hat. Gleichzeitig sind während der Digitalisierung Informationen aus den Printausgaben wieder verloren gegangen, da eine wissenschaftlich befriedigende Nachbildung eines Textkritischen Apparates bisher nicht gelungen ist. Das heißt, in der Datenbank befindet sich eine Textgestalt, deren Qualität von der verwendeten Edition abhängig ist. Da die Digitalisierung der antiken Texte bereits in den 1960er Jahren begann – in einer Zeit, in der auf dem Computer noch nicht zwischen Groß- und Kleinbuchstaben unterschieden werden konnte, weicht der Text in diesem Punkt gegebenenfalls auch von der Edition ab.³¹

c) Eine weitere Veränderung erfahren die digitalisierten Texte, wenn sie aus verschiedenen Quellen und Formaten in einer Datenbank vereinigt werden sollen, da im Zuge nöti-

29 Zur ‚Zirkulation sozialer Energie‘ in und zwischen den verschiedenen Medien vergleiche Greenblatt (1993). Meines Erachtens erweitert Moretti den New Historicism um eine digitale Komponente.

30 Neufunde auf Papyri sind die Ausnahme.

31 Für diesen Hinweis danke ich Eddy Gouder (CTLO) und Bart Janssens (Brepols); zum Informationsverlust durch Digitalisierung vergleiche auch Bárány (2004).

ger Konvertierungen (zum Beispiel Betacode in UTF) unter Umständen bestimmte Zeichen (zum Beispiel Unterpunkte zur Kennzeichnung unsicherer Buchstaben) nicht mehr richtig dargestellt werden können. Anschließend durchlaufen die Texte in der Regel eine Normalisierung, Lemmatisierung oder Grundformenreduktion, die das spätere Suchergebnis beeinflussen (vgl. Abb. 6).

Word *βακτηρία* (24630)

Number of occurrences 209

Class of frequency 14

Words with same normalized form: *βακτηρία* (209); *βακτηρία* (158); *Βακτηρία* (24); *ΒΑΚΤΗΡΙΑ* (1); *βακτηρία* (1);

Words with same base

form: *βακτηρίαν* (246); *βακτηρία* (209); *βακτηρία* (158); *βακτηρίας* (149); *βακτηρία* (34); *βακτηρίας* (28); *Βακτηρία* (24); *Βακτηρίαν* (8); *βακτηρίη* (1); *Βακτηρίας* (1); *βακτηρίαν* (1); *Βακτηρία* (1);

Abb. 6: Suchanfrage zu *βακτηρία* in eAQUA³²

Ebenso entscheidend ist die Frage, welche Einheit als Grundlage für weitere Analysen verwendet werden soll. Häufig ist dies der Satz. Alternativ könnten auch ein Absatz oder bei Papyri und Inschriften der gesamte Text verwendet werden. Allerdings variiert dann die Länge der Texte sehr stark und wirkt negativ auf die Statistik ein. Dennoch gibt es zahlreiche fachwissenschaftliche Anwendungsbeispiele, in denen die Begrenzung der Analyse auf den Satz unter Ausblendung des Kontextes nicht zum ‚richtigen‘ Ergebnis führt (siehe oben; vgl. Rücker 2010: 98; Schubert/Klank 2012: 33; 46; 48).

d) Die Erhebung von Metadaten antiker/historischer Autoren ist wie die Erstellung einer Textedition Ergebnis eines wissenschaftlichen Forschungsprozesses (vgl. Kath et al. 2014). Die getroffenen Vorentscheidungen wirken sich sehr direkt auf die Visualisierungen aus, sobald die Metadaten weiterverarbeitet werden. Häufig gehen beispielsweise Angaben zu unsicher zu datierenden Autoren verloren, da sie nicht oder nur schlecht darstellbar sind (vgl. 3.3). Aus Unsicherheiten werden Fakten.

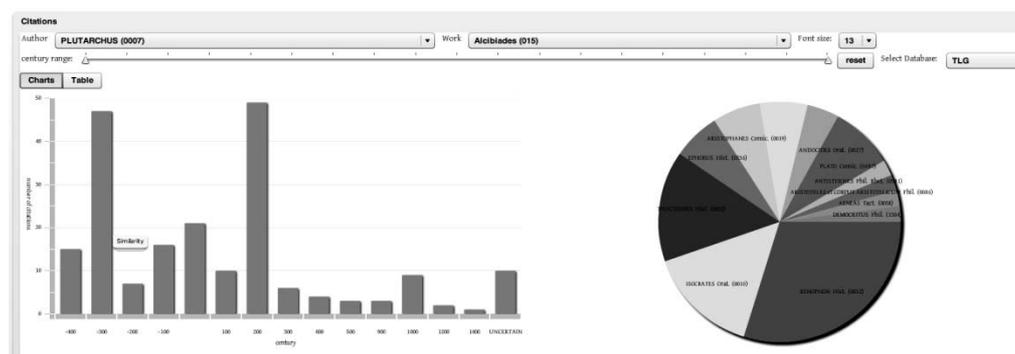


Abb. 7: Visualisierung des Zitationsprofils von Plutarch, *Alcibiades* (Fokus im Kreisdiagramm auf Autoren des vierten Jahrhunderts vor Christus)

32 Aufgrund der fehlenden Grundformenreduktion in eAQUA werden alle Wortformen einzeln aufgeführt (Anzahl der Belegstellen in Klammern). Die Berechnung und Visualisierung der Kookkurrenzen erfolgt auch für jede Wortform einzeln.

In eAQUA mussten zudem für die Berechnung der Visualisierungen im Zitationsgraphen und in Mental Maps die Lebensdaten der Autoren auf ein Datum verkürzt werden. Dieses Datum ist derzeit das Sterbedatum (soweit bekannt), unter der Annahme, dass die meisten Autoren ihre Werke erst am Lebensende verfasst haben. Problematisch wird dieser Kompromiss aber bei Autoren, deren Lebenszeit sich über die Jahrhundertgrenze erstreckt. Aus historischer Perspektive ist es dabei wahrscheinlich weniger problematisch, Platon (428–347 v. Chr.) als Autor des vierten Jahrhunderts zu verorten, als Aristophanes (480–380 v. Chr.) oder Thukydides (460–397 v. Chr.), deren Werke man gern im Kontext des Perikleischen Zeitalters oder des Peloponnesischen Krieges betrachten würde.³³

Die genannten Beispiele verdeutlichen, dass die Datengrundlage, auf der die folgenden Schritte aufbauen, Ergebnis vieler Einzelentscheidungen ist. Die NutzerIn der entsprechenden Tools muss sich dabei immer vergegenwärtigen, dass das „Archiv“ selbst somit ein Konstrukt ist (Foucault 1981; vgl. Haber 2011: 58).

3.2 Analyseschritt 2: Text-Mining & Information Retrieval

Die Phase des Text-Mining und Information Retrieval steht grundsätzlich in der Gefahr, sogenannte *Black Boxes* zu erzeugen (vgl. unter anderem Rieder/Röhle 2012: 76), da sie in der Regel von der Informatik allein verantwortet wird. Die FachwissenschaftlerIn kann und muss aber den Prozess ihrer jeweiligen Suchanfragen dokumentieren. Das heißt, sie sollte die jeweils gewählten Einstellungen in den interaktiven Tools festhalten (zum Beispiel den gewählten (Sub-)korpus, den Analysezeitraum, die Metadatengrundlage, Signifikanzmaß und dergleichen). Die Abbildungen 8 und 9 zeigen den Einfluss unterschiedlicher Einstellungen bei gleicher Suchanfrage im oben skizzierten Beispiel zu Charon.

33 Eine Alternative wäre die Datierung einzelner Werke. Dies scheitert jedoch vorerst an zwei Dingen. Zum einen wäre dieses Vorhaben extrem aufwändig (das heißt teuer), zum anderen fehlen in den allermeisten Fällen die notwendigen Informationen. Bereits jetzt verfügen wir in etwa 50 Prozent der Fälle über keinerlei Angaben zu Datierung und/oder Ort. Im Rahmen des eXChange-Projektes soll der NutzerIn daher eine flexible Gestaltung der Metadatenauswahl ermöglicht werden. Die Auswirkungen der Veränderungen werden in der Rechercheumgebung live zu verfolgen sein.

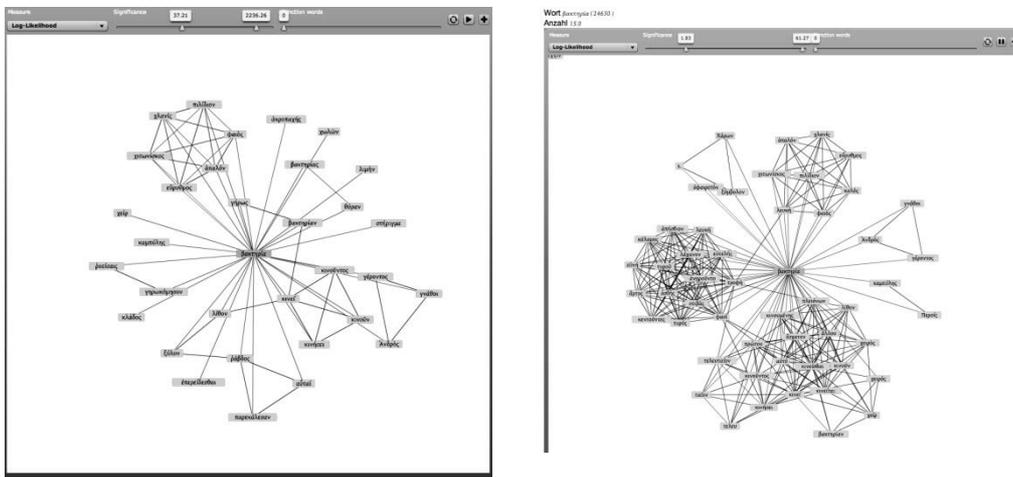


Abb. 8 und 9: Kookkurrenzgraphen zu *βακτηρία*

Abbildung 8 zeigt den Kookkurrenzgraphen zu *βακτηρία* in der Voreinstellung der Suchanfrage von eAQUA. Der oben beschriebene Zusammenhang zu Charon wird hier nicht angezeigt. Für Abbildung 9 wurde das Suchwort zunächst in die Suchmaske von Mental Maps (vgl. Abb. 10; vgl. zur Funktionsweise Kath 2011) eingetragen. Mit Hilfe der interaktiven Einstellung des Tools wurde der Zeitraum auf ‚vor Christus‘ eingeschränkt und anschließend ein Kookkurrenzgraph erzeugt. Für Abbildung 9 wurden einige der Einstellungen an den Schieberegler verändert (zum Beispiel wurden die sogenannten Stoppwörter entfernt). Auffällig ist, dass die Position des Zusammenhangs Charon-Symbolon innerhalb des Wortnetzes variiert.³⁴

Hier sollte zudem über die Auswirkung von Besonderheiten der Datengrundlage reflektiert werden wie zum Beispiel die Dopplung von Textstellen durch Fragmentsammlungen im Korpus oder die Auswirkung der Satzgrenze auf die Anfrage (vgl. 3.1). Auch die Entscheidung über die Bedingung einer Übereinstimmung von fünf identischen griechischen Worten für die Suche nach Zitationen in eAQUA ist solch ein Fall.

3.3 Analyseschritt 3: Visualisierung der Daten

Unabhängig vom technischen Perfektionsgrad muss es sich auch und gerade bei den Visualisierungen selbst immer um Konstruktionen handeln. Dies muss insofern fortwährend betont werden, da ihnen eine außergewöhnliche Suggestivkraft innewohnt, die Unsicherheiten und Wahrscheinlichkeiten für die BetrachterIn in Fakten überführt (vgl. Drucker 2012; Rieder/Röhle 2012: 69, 73).

³⁴ Derzeit ist es mit Hilfe der Suchanfrage von eAQUA auch bei identischen Einstellungen technisch bedingt grundsätzlich nicht möglich, Kookkurrenzgraphen zu erzeugen, die identisch aussehen.

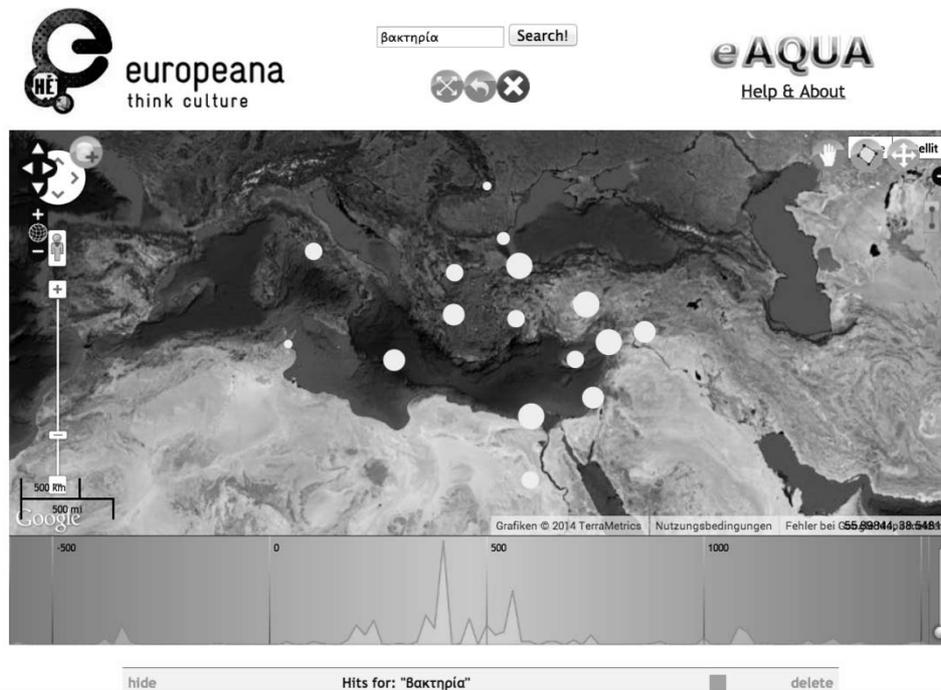


Abb. 10: Mental Maps-Suchanfrage zu *βακτηρία*

Die Zuordnung von Geokoordinaten zu antiken Autoren (Aufenthalts- und Wirkungsorte) erscheint beispielsweise mindestens so problematisch wie deren Datierung. Selbst der beliebte Rückgriff auf *Google Maps* als Grundlage geotemporaler Visualisierungen ist nicht frei von Problemen (vgl. Drucker 2012). Bereits die Projektion eines Satellitenbildes auf eine zweidimensional gestreckte (Karten)darstellung ist eine Konstruktion. Zudem sind Räume und Grenzen sozial konstruiert (Kath/Rieger 2009) und unterliegen historischen Wandlungsprozessen – das heißt, die moderne Stadt muss nicht mit der gegebenenfalls gleichnamigen antiken Polis identisch sein. Die Abbildung des Wirkungsortes eines antiken Autors auf eine historische Kartendarstellung oder auch nur die Lokalisierung auf einer physischen Karte kann lediglich eine Annäherung sein, die unter Umständen mehr verwirrt als erklärt.

Visual Analytics versucht auf diese Herausforderung zu reagieren und beschäftigt sich zunehmend mit der „Visualisierung von Unsicherheit“ (Brodie et al. 2012; Jänicke 2012; Jänicke/Wrisley 2013; Rees 2012). Die Forschung in diesem Bereich ist sehr jung, so dass hierzu noch keine (gültige) Einschätzung getroffen werden kann. Allerdings weisen die aktuellen Ansätze ein Maß an visueller Komplexität auf, so dass sich die Frage stellt, ob diese Informationsfülle von einer menschlichen NutzerIn noch verarbeitet werden kann.

Ohne speziell Berger/Luckmann (1995) bemühen zu müssen, verfügen die geisteswissenschaftlichen Disziplinen alle über ein Methodenspektrum, das ForscherInnen in die Lage versetzt, mit Unsicherheiten der Materialbasis und der Konstruiertheit von sozialen und virtuellen Räumen umzugehen. Wir müssen diese Erfahrung nur in die Arbeit mit den neuen Tools der DH integrieren!

Die Einsicht in die Konstruiertheit der Datenbasis und der Visualisierungen wirft zudem die Frage auf, ob wir es nicht mit einer Art Metakonstruktion zu tun haben, deren Verwendung wohlüberlegt sein will. Sie ist abhängig von einer gründlichen Vorkenntnis der Eigenarten der ursprünglichen Materialbasis (Texte) sowie einem Grundverständnis der technischen Prozesse³⁵ und den dahinterliegenden Epistemologien (Dixon 2012: 191 ff.; Ramsay/Rockwell 2012). Dass sich der nicht geringe Aufwand³⁶ – die semantischen Analysen sind schließlich nicht das Ende, sondern nur der Anfang historischer Untersuchungen (vgl. Lückerath 1968: 276) – auch für die Sozialwissenschaften lohnen kann, zeigen die produktiven Ergebnisse in den Altertumswissenschaften. Außerdem werden immer mehr der für SozialwissenschaftlerInnen interessanten Quellen nur noch digital vorliegen beziehungsweise entstehen. Bereits heute werden immer häufiger Analysen von Twiternachrichten durchgeführt, wobei nicht transparent ist, wie repräsentativ die analysierten Daten in Relation zur Gesamtheit aller Twiternachrichten sind (Boyd/Crawford 2013). Lev Manovich (2001) hat zudem auf die permanente Veränderlichkeit als Eigenart digitaler Texte, Fotos und Filme im Gegensatz zu ihren analogen Vorfahren hingewiesen.³⁷

Die in Entwicklung befindlichen explorativen Recherchertools der DH, die imstande sind, live auf Veränderungen von Metadaten und/oder Konzeptdesigns reagieren zu können, werden Analysen komplexerer Fragestellungen ermöglichen. Ihre Herausforderungen sind gleichzeitig ihre Stärke: Sie ermöglichen die Betrachtung desselben Materials aus wechselnden Blickwinkeln und Simulationen historischer und sozialer Phänomene in Echtzeit.

Die DH üben folglich Einfluss auf die weitere Entwicklung der Geisteswissenschaften aus, da sie die Fragen an das vorhandene Material verändern und nicht nur dessen Repräsentation. Indem wir uns der Aufgabe der Integration der Methoden der Digital Humanities in unsere Fächer stellen, können wir die Bereitschaft zur Nutzung dieser Werkzeuge erhöhen und die neuen Zugangswege zu Texten und Bildern gewinnbringend für die eigene fachliche Arbeit einsetzen.

Literatur

- Bárány, Balázs, 2004: Informationsverlust durch die Digitalisierung. Diplomarbeit, Wien; tud.at/uni/diplomarbeit/diplomarbeit.pdf, 08.06.2014.
- Baßler, Moritz, 1995: Einleitung. In: Moritz Baßler (Hg.), *New Historicism. Literaturgeschichte als Poetik der Kultur*, Frankfurt (Main), 7–28.
- Becker, Matthias, 2012: Der Algorithmus, bei dem man mit muss? Über die Kehrseite der digitalen Automatisierung, Feature. In: Deutschlandfunk vom 2. März 2012; <http://www.deutschlandfunk.de/der-algorithmus-bei-dem-man-mit-muss.1170.de.html?dram:articleid=184103>, 07.06.2014.
- Berger, Peter L. / Luckman, Thomas: *Die gesellschaftliche Konstruktion der Wirklichkeit. Eine Theorie der Wissenssoziologie*, Frankfurt (Main).
- Berry, David M., 2012: Introduction: Understanding the Digital Humanities. In: David M. Berry (Hg.), *Understanding Digital Humanities*, London, 1–20.

35 Lin (2012: 302–306) skizziert die typischen Schritte der Entwicklung von Text-Mining-Tools (1. scoping and data preparation, 2. data analysis and training the algorithms, 3. software development, 4. iterative development), die die in der New Visual Hermeneutics zugrunde gelegten Schritte auch spiegeln.

36 Moretti (2011: 106) selbst reflektiert dies: „Grad students are in general wary of projects like Graphs, Maps, Trees. They realize they are very time-consuming, as one may work weeks gathering data that reveal nothing.“

37 Vergleiche auch in diesem Zusammenhang Haber (2011: 102).

- Bostic, Adam I., 1988: Seeing Cyborg through the Eyes of Popular Culture, Computer-Generated Image-ry and Contemporary Theory. In: Leonardo 31, 357–361.
- Boyd, Danah / Crawford, Kate, 2013: Big Data als kulturelles, technologisches und wissenschaftliches Phänomen. Sechs Provokationen. In: Heinrich Geiselberger / Tobias Moorstedt (Hg.), Big Data. Das neue Versprechen der Allwissenheit, Berlin, 187–218.
- Brkich, Christopher A. / Barko, Tim, 2012: Our Most Lethal Enemy?: Star Trek, the Borg, and Methodological Simplicity. In: Qualitative Inquiry 18, 787–797.
- Brodlie, Ken / Osorio, Rodolfo A. / Lopes, Adriano, 2012: A Review of Uncertainty in Data Visualization. In: John Dill / Rae Earnshaw / David Kasik / John Vince / Pak Chung Wong (Hg.), Expanding the Frontiers of Visual Analytics and Visualization, London, 81–109; http://www.comp.leeds.ac.uk/kwb/publication_repository/2012/uncert.pdf, 04.06.2014.
- Büchler, Marco, 2013: Informationstechnische Aspekte des Historical Text Re-Use. Dissertation, Leipzig (Manuskript); <http://www.qucosa.de/fileadmin/data/qucosa/documents/10851/Dissertation.pdf>, 08.06.2014.
- Campa, Riccardo, 2013 [2008]: Making Science by Serendipity. A review of Robert K. Merton and Elinor Barber's The Travels and Adventures of Serendipity. In: Journal of Evolution and Technology 17, 75–83; <http://jetpress.org/v17/campa.htm>, 27.05.2014.
- Dixon, Dan, 2012: Analysis Tool or Research Methodology: Is there an Epistemology for Patterns. In: David M. Berry (Hg.), Understanding Digital Humanities, London, 191–208.
- Droysen, Johann G., 1868: Grundriss der Historik, Leipzig.
- Droysen, Johann G., 1871: Aristophanes Werke. 1. Theil, 2. Auflage, Leipzig.
- Drucker, Johanna, 2012: Humanistic Theory and Digital Scholarship. In: Matthew K. Gold (Hg.), Debates in the Digital Humanities, Minneapolis / London, 85–95.
- Foucault, Michel, 1981: Die Archäologie des Wissens, Frankfurt (Main).
- Geßner, Annette, 2010: Das automatische Auffinden der indirekten Überlieferung des Platonischen Timaios und die Bedeutung des Tools „CitationGraph“ für die Forschung. In: Charlotte Schubert / Gerhard Heyer (Hg.), Das Portal eAQUA – Neue Methoden in der geisteswissenschaftlichen Forschung I, Working Papers Contested Order 1, Leipzig, 26–41.
- Goodwin, Jonathan / Holbo, John / Moretti, Franco, 2011: Reading Graphs, Maps, and Trees: Responses to Franco Moretti, Anderson.
- Gorunescu, Florin, 2011: Data Mining. Concepts, Models and Techniques, Berlin / Heidelberg.
- Greenblatt, Stephen, 1993: Einleitung. Die Zirkulation sozialer Energie. In: Stephen Greenblatt, Verhandlungen mit Shakespeare. Innenansichten der englischen Renaissance, Frankfurt (Main), 9–33.
- Haber, Peter, 2011: Digital Past. Geschichtswissenschaften im digitalen Zeitalter, München.
- Heyer, Gerhard / Quasthoff, Uwe / Wittig, Thomas, 2006: Text Mining: Wissensrohstoff Text – Konzepte, Algorithmen, Ergebnisse, Bochum.
- Ihde, Don, 1998: Expanding Hermeneutics. Visualism in Science, Evanston (IL.).
- Ihde, Don, 2012: Experimental Phenomenology. Second Edition: Multistabilities, New York.
- Jänicke, Stefan, 2012: GeoTemCo: Comparative Visualization of Geospatial-Temporal Data; <http://www.informatik.uni-leipzig.de:8080/geotemco/>, 04.06.2014.
- Jänicke, Stefan / Wisley, David J., 2013: Visualizing Uncertainty: How to Use the Fuzzy Data of 550 Medieval Texts? In: Proceedings of the Digital Humanities.
- Kahneman, Daniel, 2012: Schnelles Denken, langsames Denken, München.
- Kath, Roxana, 2011: Das Mental Maps-Interface: Erforschung von Konzepten in Raum und Zeit. In: Charlotte Schubert (Hg.), Neue Methoden der geisteswissenschaftlichen Forschung. Eine Einführung in das Portal eAQUA II, Working Papers Contested Orders No. 3, Leipzig, 64–86.
- Kath, Roxana, 2012: Konsumverzicht im Nomadenbild und in der Philosophie der Antike. In: Roxana Kath / Michaela Rücker (Hg.), Die Geburt der griechischen Weisheit oder Anacharsis, Skythe und Griechen. Mitteilungen des SFB „Differenz und Integration“ 13, Halle, 127–145.
- Kath, Roxana / Rieger, Anna-Katharina, 2009: Einleitung. In: Roxana Kath / Anna-Katharina Rieger (Hg.), Raum – Landschaft – Territorium. Zur Konstruktion physischer Räume als nomadischer und sesshafter Lebensraum, Wiesbaden, 1–16.

- Kath, Roxana / Schaal, Gary S. / Dumm, Sebastian, 2014: New Visual Hermeneutics. In: Zeitschrift für germanistische Linguistik 42, Themenheft „Automatisierte Textanalyse“ (im Erscheinen).
- Keim, Daniel / Kohlhammer, Jörn / Ellis, Geoffrey / Mansmann, Florian, 2010 (Hg.): Mastering the information age. Solving problems with visual analytics; <http://www.diglib.org>, 14.05.2014.
- Kirschenbaum, Matthew, 2011: Poetry, Patterns and Provocation: The nora Project. In: Jonathan Goodwin / John Holbo / Franco Moretti (Hg.), Reading Graphs, Maps, and Trees: Responses to Franco Moretti, Anderson, 31–40.
- Kurt, Ronald, 2008: Vom Sinn des Sehens: Phänomenologie und Hermeneutik als Methoden visueller Erkenntnis. In: Jochen Dreher / Michaela Pfadenhauer (Hg.), Phänomenologie und Soziologie. Wiesbaden, 369–378.
- Lemke, Matthias / Stulpe, Alexander, 2014: Text und soziale Wirklichkeit. Theoretische Grundlagen und empirische Anwendung von Text-Mining-Verfahren in sozialwissenschaftlicher Perspektive. In: Zeitschrift für germanistische Linguistik 42, Themenheft „Automatisierte Textanalyse“ (im Erscheinen).
- Lin, Yu-wei, 2012: Transdisciplinarity and Digital Humanities: Lessons Learned from Developing Text-Mining Tools for Textual Analysis. In: David M. Berry (Hg.), Understanding Digital Humanities, London, 295–314.
- Lückerath, Carl A., 1968: Prolegomena zur elektronischen Datenverarbeitung im Bereich der Geschichtswissenschaft. In: Historische Zeitschrift 207, 265–296.
- Manovich, Lev, 2001: The Language of New Media, Cambridge / London.
- Manovich, Lev, 2010: What is Visualization?; <http://manovich.net/2010/10/25/new-article-what-is-visualization>, 06.05.2014.
- Mehler, Alexander, 2004: Textmining. In: Henning Lobin (Hg.), Texttechnologie. Perspektiven und Anwendungen, Tübingen, 329–352.
- Merton, Robert K., 1983: Auf den Schultern von Riesen. Ein Leitfaden durch das Labyrinth der Gelehrsamkeit, Frankfurt (Main).
- Merton, Robert K. / Barber, Elinor, 2006: The Travels and Adventures of Serendipity: A Study in Sociological Semantics and the Sociology of Science, Princeton.
- Montrose, Louis, 1995: Die Renaissance behaupten. Die Poetik und Politik der Kultur. In: Moritz Baßler (Hg.), New Historicism. Literaturgeschichte als Poetik der Kultur, Frankfurt (Main), 60–93.
- Moretti, Franco, 2009: Kurven, Karten, Stammbäume. Abstrakte Modelle für die Literaturgeschichte, Frankfurt (Main).
- Moretti, Franco, 2013: Distant Reading, London / New York.
- Plewe, Brandon, 2002: The Nature of Uncertainty in Historical Geographic Information. Transactions in GIS 6: 431–456; http://dusk.geo.orst.edu/buffgis/TGIS_uncertainty.pdf, 04.06.2014.
- Ramsay, Stephen / Rockwell, George, 2012: Developing Things: Notes toward an Epistemology of Building in the Digital Humanities. In: Matthew K. Gold (Hg.), Debates in Digital Humanities, Minneapolis, 75–84.
- Ramsay, Stephen, 2011: Reading Machines: Toward an Algorithmic Criticism (Topics in the Digital Humanities), Urbana / Chicago / Springfield.
- Rees, Gethin P., 2012: Uncertain Date, Uncertain Place: Interpreting the History of Jewish Communities in the Byzantine Empire using GIS; <http://www.dh2012.uni-hamburg.de/conference/programme/abstracts>, 08.06.2014.
- Rieder, Bernhard / Röhle, Theo, 2012: Digital Methods: Five Challenges. In: David M. Berry (Hg.), Understanding Digital Humanities, London.
- Romanello, Matteo / Berti, Monica / Boschetti, Frederico / Babeu, Alison / Crane, Gregory, 2009: Rethinking Critical Editions of Fragmentary Texts by Ontologies, 13th International Conference on Electronic Publishing: Rethinking Electronic Publishing: Innovation in Communication Paradigms and Technologies, Mailand, 155–174; http://elpub.architexturez.net/system/files/pdf/158_elpub2009.content.pdf, 08.06.2014.
- Rosenberger, Robert, 2009: Quick-freezing philosophy: An analysis of imaging technologies in neurobiology. In: Jan-Kyrre Berg Olsen / Evan Selinger / Søren Riis (Hg.), New waves in philosophy of technology, New York, 65–82.

- Rosenberger, Robert, 2011: A Case Study in the Applied Philosophy of Imaging: The Synaptic Vesicle Debate. In: *Science, Technology & Human Values* 36, 6–32.
- Rücker, Michaela, 2010: Die Möglichkeiten der automatischen Textergänzung auf Papyri. In: Charlotte Schubert / Gerhard Heyer (Hg.), *Das Portal eAQUA – Neue Methoden in der geisteswissenschaftlichen Forschung I*, Working Papers Contested Order 1, Leipzig, 91–100.
- Schaal, Gary S., 2013: Auf den Spuren der Deutungsmacht. Ein konzeptioneller Vorschlag zur empirischen Analyse mit Text Mining. In: André Brodocz / Stefanie Hammer (Hg.), *Variationen der Macht*, Baden-Baden, 199–221.
- Schaal, Gary S. / Kath, Roxana, 2014: Zeit für einen Paradigmenwechsel in der Politischen Theorie? Der Ansatz der neuen visuellen Hermeneutik. In: André Brodocz / Daniel Schulz / Julia Schulze Wessel (Hg.), *Die Verfassung des Politischen. Festschrift für Hans Vorländer (im Erscheinen)*.
- Schmidt, Benjamin, 2012: When You Have a Mallet, Everything Looks Like a Nail; <http://sappingattention.blogspot.de/2012/11/when-you-have-mallet-everything-looks.html>, 12.01.2014.
- Schubert, Charlotte, 2010a: Zitationsprofile, Suchstrategien und Forschungsrichtungen. In: Charlotte Schubert / Gerhard Heyer (Hg.), *Das Portal eAQUA – Neue Methoden in der geisteswissenschaftlichen Forschung V*, Working Papers Contested Order 1, Leipzig, 42–55.
- Schubert, Charlotte, 2010b: Formen der griechischen Historiographie: Die Atthidographen als Historiker Athens. In: *Hermes* 138, 259–275.
- Schubert, Charlotte, 2010c: Anacharsis der Weise. Nomade, Skythe Grieche, Tübingen.
- Schubert, Charlotte, 2011: Detailed description of eAQUA search portal. In: Charlotte Schubert (Hg.), *Das Portal eAQUA – Neue Methoden in der geisteswissenschaftlichen Forschung V*, Working Papers Contested Order 3, Leipzig, 33–53.
- Schubert, Charlotte, 2013: Zauberlehrling und Meister: Digital Humanities zwischen Informatik und Geisteswissenschaften? In: Jörn Kobes / Kai Ruffing / Wolfgang Spickermann (Hg.), *20 Jahre Arbeitsgemeinschaft Geschichte und EDV*, Gutenberg, 167–186.
- Schubert, Charlotte / Bräckel, Oliver / Willkommen, Corina, 2013: Visualisierung und Wissensrepräsentation: Die kleisthenische Phylenreform. In: Charlotte Schubert (Hg.), *Das Portal eAQUA – Neue Methoden in der geisteswissenschaftlichen Forschung I*, Working Papers Contested Order 10, Leipzig, 1–8.
- Schubert, Charlotte / Klank, Markus, 2012 (Hg.): *Das Portal eAQUA – Neue Methoden in der geisteswissenschaftlichen Forschung III*, Working Papers Contested Order 7, Leipzig.
- Stoppe, Sebastian, 2011: Ein transhumanistischer Leviathan? Die Borg als emotionslose Dystopie in Star Trek. In: *Arbeitstitel – Forum Für Leipziger Promovierende* 3, 69–82; www.wissens-werk.de/index.php/arbeitstitel/article/download/85/106, 08.06.2014.
- Wiedemann, Gregor / Lemke, Matthias / Niekler, Andreas, 2013: Postdemokratie und Neoliberalismus – Zur Nutzung neoliberaler Argumentation in der Bundesrepublik Deutschland 1949–2011. In: *Zeitschrift für Politische Theorie* 4, 99–115.