

Aus dem Inhalt

Bd. 2,1 (2016)

Editorial:
Michaela Rücker / Sven-Philipp Brandt

Dariya Rafiyenko:
Gicht im Altgriechischen. Eine korpus-basierte Studie

Christine Roughan:
Digital Texts and Diagrams: Representing the Transmission of Euclid's Elements

Friedrich Meins:
**Digitale Editionen für die Altertumswissenschaften:
Eine Auseinandersetzung mit P. Sahles Kritik kritischer Texteditionen**

Leif Scheuermann:
Die Abgrenzung der digitalen Geisteswissenschaften

Charlotte Schubert:
Die Visualisierung von Quellennetzwerken am Beispiel Plutarchs

Projektankündigungen:
Reinhold Scholl:
Mehrsprachiges Online Wörterbuch zum Fachwortschatz der Verwaltungssprache des griechisch-römisch-byzantinischen Ägypten

Christopher Blackwell / Charlotte Schubert:
Annotating and Editing with Canonical Text Services (CTS) Project funded by the Andrew W. Mellon Foundation: 2016-2017



Zum Inhalt

Das erste Heft des Jahrgangs 2016 umfasst thematisch sehr unterschiedliche Beiträge, die Grundsatzfragen der digitalen Geisteswissenschaften, editorische, technische und methodische Bereiche der Digital Classics ansprechen.

So stellt zunächst Dariya Rafiyenko ihre Untersuchungen im Kontext der Krankheit Gicht vor. Bei dieser Krankheit lässt sich aufgrund der eindeutigen Symptomatik dieser chronischen Krankheit eine konstante Beschreibung seit der Antike problemlos nachvollziehen – im Gegensatz zu den vielen Pest-Bezeichnungen.¹ So gilt sie als eine der häufigsten Krankheiten der Antike² und ist bereits im Corpus Hippocraticum beschrieben.³ Rafiyenko untersucht in diesem Zusammenhang die verschiedenen Begriffe für Gicht im Altgriechischen sowie deren Kontext im Rahmen einer korpus-basierten Analyse, deren Ziel es ist einen Beitrag zur Erforschung der medizinischen Begrifflichkeit mit dem Hauptaugenmerk auf der byzantinischen Periode (4./6. Jh. - 15 Jh.) zu leisten. Zudem soll der Aufsatz die Methodik für eine korpus-basierte Untersuchung von Konzepten vorstellen.

Christine Roughan befasst sich in ihrem Beitrag „Digital Texts and Diagrams: Representing the Transmission of Euclid’s Elements“ mit dem Digital Euclid Project, das sämtliche Zeugnisse der Texte und Diagramme zu Euklids *Elementa* digital frei zugänglich machen will. Im ersten Teil ihres Beitrags setzt sie sich dabei mit der allgemeinen Überlieferungssituation zu Euklids *Elementa* auseinander und betrachtet auch die bereits digitalisierten und im Kontext der Digital Classics erzielten Ergebnisse. Im zweiten Teil des Beitrags diskutiert Roughan schließlich in einem größeren Kontext die Möglichkeiten, das vielschichtige Material zu Euklids *Elementa* methodisch und editorisch aufzubereiten und so für eine digitale Edition nachhaltig verwertbar zu machen.

Einen wissenschaftstheoretischen Beitrag liefern sowohl Friedrich Meins als auch Leif Scheuermann in ihren Aufsätzen. Meins setzt sich dabei mit der Rolle der digitalen Editionen im Bereich der Altertumswissenschaften auseinander und bezieht sich unter anderem auf Fragen der Autorintention, der Stemmologie und der Normalisierung. Ein besonderes Augenmerk legt er auf die Rolle der Konjekturen in kritischen Texteditionen, die unter anderem von Patrick Sahle im ersten Band seines dreiteiligen Werkes über digitale Editionsformen sehr kritisch behandelt worden sind.⁴ Meins plädiert hingegen dafür, den Wert der kritischen Edition als

1 Vgl. hierzu u.a. die sog. Pestbeschreibung in Thuk. 2,47-55, deren Symptomatik sich doch deutlich von späteren Pestausbrüchen, wie der großen Pandemie im 14. Jahrhundert in Europa, unterscheidet und heute in medizinhistorischen Untersuchungen nicht mehr als ‚Pest‘ klassifiziert wird, vgl. Tack (2012), S. 88 sowie ausführlich zur Forschungsdiskussion Hornblower (2003), S. 316f.

2 Vgl. Stamatu (2005), S. 356.

3 Hippokr. aff. 30f.

4 Vgl. Sahle (2013).

eigene Leistung gegenüber rein dokumentarischen Ansätzen nicht aus den Augen zu verlieren, sieht aber die dokumentarischen (digitalen) Ansätze als Möglichkeit, die Qualität der kritischen Editionen und die Diskussion über die Textbeschaffenheit zu verbessern.

Scheuermann setzt sich in seinem Beitrag wiederum mit der landläufigen Vorstellung auseinander, dass die Digital Humanities ein Sammelsurium all dessen seien, „was mit Geisteswissenschaften und den Computern zusammenhängt“. Hierzu versucht er zunächst eine genaue Bestimmung des Geltungsbereichs der Digital Humanities abzustecken, indem er eine Definition aus der Abgrenzung zu den Nachbardisziplinen bei gleichzeitiger Herausarbeitung der gemeinsamen Charakteristika der jeweiligen Einzeldisziplinen entwirft. Darauf aufbauend wendet Scheuermann diese Charakteristika auf die digitalen Geisteswissenschaften an und formuliert abschließend Forderungen an eine noch zu begründende Digitale Geisteswissenschaft auf der Basis eben jener herausgearbeiteten Charakteristika.

Den Abschluss bildet ein Aufsatz zur Visualisierung von Quellennetzwerken am Beispiel Plutarchs von Charlotte Schubert, in dem die Möglichkeiten des *distant reading* über die Visualisierung als eine wichtige Erweiterung der linearen Lektüre dargestellt werden. Die Multimodalität der Digitalität ermöglicht neue Repräsentations- und damit auch Rezeptionsformen, die gerade deshalb neue, textimmanente und relationale Beziehungen offenlegen und zu einer Erweiterung der Editions- und Interpretationsmethoden beitragen können. Schubert veranschaulicht dies anhand verschiedener Visualisierungsmöglichkeiten des Quellennetzwerks Plutarchs mit dem Netzwerkvisualisierungsprogramm Gephi, die aber gleichzeitig die traditionellen Ansätze Wests, Maas' und Diels' mit einbeziehen. Ein Vorschlag zu einer neuen Taxonomie der Quellen und Editionen soll es ermöglichen, die Voraussetzungen der Anordnung und Zuordnung verschiedener Schichten für digitale Repräsentationsformen zu systematisieren.

Der Band schließt mit zwei Projektbeschreibungen. Hierbei stellt zunächst Reinhold Scholl ein Projekt zu einem neuen, mehrsprachigen Online Wörterbuch zum Fachwortschatz der Verwaltungssprache des griechisch-römisch-byzantinischen Ägypten vor, das das nach über 100 Jahren immer noch maßgebende Wörterbuch Preisigkes nicht nur digitalisiert, sondern auch mit den Möglichkeiten der digitalen Repräsentation um wichtige Funktionen erweitert. So werden unter anderem zu jedem Lemma neben der Übersetzung Schreibvarianten, Synonyme, Antonyme, Hyperonyme und Hyponyme sowie verwandte Begriffe aufgenommen und damit die Lemmata um viele Aspekte erweitert. Das Projekt ist als Pilotprojekt für ein allgemeines Wörterbuch für papyrologische Texte angelegt und wird im Zuge dessen vom Sächsischen Ministerium für Wissenschaft und Kunst auf Vorschlag der Sächsischen Akademie der Wissenschaften zu Leipzig gefördert.

Ein weiteres Projekt im Bereich der Digital Humanities stellen Christopher Blackwell und Charlotte Schubert vor. Dazu werden in dem Projekt „Annotating and Editing with Canonical Text Services (CTS)“, das von der Andrew W. Mellon Foundation gefördert wird, zwei Projekte, nämlich CITE/CTS⁵ und eComparatio⁶ verbunden, um entsprechende Synergien zu schaffen. In dem Projekt wird anhand von Plutarchs Periklesvita ein use case ausgearbeitet, der die Vorteile des Einsatzes von CTS im Hinblick auf die technische, philologische und historische Ebene verdeutlichen soll. Hierzu sollen die von eAQUA aufgefundenen Parallelstellen und die mit eComparatio angezeigten Varianten mit Hilfe von CITE-URNs gespeichert und mit CTS-URNs verbunden werden, so dass sie Nutzern auch nachhaltig und mit den URNs eindeutig klassifiziert, zugänglich gemacht werden.

5 <http://www.homermultitext.org/hmt-doc/cite/>.

6 <http://ecomparatio.net/> (Login notwendig). Für eine genaue Beschreibung s. <http://www.eaqua.net/>.

Ausblick

Im weiteren Verlauf des Jahrgang 2016 sind Ausgaben mit verschiedenen thematischen Schwerpunkten vorgesehen. So wird zunächst unter der Herausgeberschaft von Roxana Kath ein Heft im Kontext des BMBF-Verbundprojekts „eXChange: Exploring Concept Change and Transfer in Antiquity“⁷ erscheinen. In diesem Projekt wurde in interdisziplinärer Zusammenarbeit von Altertumswissenschaftlern und Informatikern ein Recherchesystem aufgebaut, das Bedeutungsverschiebungen semantischer Räume nach Ort und Zeit visualisiert. Darüber hinaus können durch dieses System Begriffe, die sowohl in medizinischen Fachtexten als auch in politischen und alltäglichen Texten verwendet werden, auf ihren Kontext hin in argumentativen Strategien analysiert werden.

Ein weiteres Heft wird sich mit den Entwicklungen im Bereich der digitalen Papyrologie befassen. Hierzu werden unter der Herausgeberschaft von Reinhold Scholl neben dem neuesten Projekt „Mehrsprachiges Online Wörterbuch zum Fachwortschatz der Verwaltungssprache des griechisch-römisch-byzantinischen Ägypten“, das mit einer Projektbeschreibung bereits in diesem Band skizziert wurde, noch weitere Projekte und Ergebnisse aus dem Bereich der digitalen Papyrologie vorgestellt. Dabei handelt es sich unter anderem um das Papyrus-Projekt Halle-Jena-Leipzig,⁸ in dem die jeweiligen Papyrussammlungen die Papyrus- und Ostrakabestände nach gemeinsam entwickelten Kriterien katalogisiert, digitalisiert sowie sicherheitsverfilmt wurden.

Ein weiteres Papyrus-Projekt, das unter anderem auf den Erfahrungen des Papyrus-Projekts Halle-Jena-Leipzig aufbaut, ist das ‚Papyrus Portal‘.⁹ Hierbei handelt es sich um ein Projekt, das dem Nutzer eine effiziente und effektive Suche durch alle digitalisierten und elektronisch katalogisierten Papyrussammlungen Deutschlands ermöglicht und zudem eine einheitliche Präsentation der Suchergebnisse mit den wichtigsten Informationen zu einem Papyrus gewährleistet. Es ist, wie schon das Papyrus-Projekt Halle-Jena-Leipzig, mit dem Advanced Papyrological Information System (APIS) und anderen Metadatenbanken kompatibel. Dazu wird die Open Source Software MyCoRe verwendet, die aus der Zusammenarbeit einer Reihe deutscher Universitäten entstand und stetig weiterentwickelt wird.

Darüber hinaus ist ein weiteres Heft geplant, das verschiedene Aufsätze aus dem weiten Spektrum der Altertumswissenschaften vereinen soll. Hierbei wird ein Beitrag von Monica Berti und Marcel Mernitz dem Digital Hill Projekt gewidmet werden. Im Fokus stehen dabei die Arbeitsweise mit dem Treebanking-Tool Arethusa und dem Textaligning-Tool Alpheios, die von Perseids bereitgestellt werden, sowie die Darstellung der daran angegliederte Website. Darüber hinaus sind in diesem Heft noch weitere Beiträge aus der Archäologie sowie aus der Ägyptologie vorgesehen und auch weitere Beiträge aus dem Kontext der digitalen Altertumswissenschaften können bei Interesse eingereicht werden.

⁷ <http://www.exchange-projekt.de/index.html>.

⁸ <http://papyri.uni-leipzig.de/content/below/start.xml>.

⁹ <http://www.papyrusportal.de/content/start.xml>.

Literatur

Hornblower (2003), S. Hornblower, A Commentary on Thucydides (Vol. 1, Books I-III), Oxford 2003.

Moretti (2007): F. Moretti, Graphs, Maps, Trees, London / New York 2007.

Sahle (2013): P. Sahle, Digitale Editionsformen. Zum Umgang mit der Überlieferung unter den Bedingungen des Medienwandels (3 Bde.) Norderstedt 2013.

Stamatu (2005): M. Stamatu, Gicht, in: K.-H. Leven (Hg.), Antike Medizin. Ein Lexikon, München 2005, S. 356–358.

Tack (2012): R. Tack, Thukydides. Textauswahl und Kommentar, Hamburg 2012.

Quellen

Hippocrates et Corpus Hippocraticum, De affectionibus Bd. 5, with an English translation by P. Potter, London 1988, S. 1–91.

Thukydides, Historiae Bd. 1, ed. H.S. Jones / J.E. Powell, Oxford 1955.

Weitere Ressourcen (zuletzt abgerufen am: 18.02.2016):

eAQUA: Extraktion von strukturiertem Wissen aus Antiken Quellen für die Altertumswissenschaft, URL:

<http://www.eaqua.net/>

eComparatio: Visualisierungstool für den Vergleich von verschiedenen Textversionen, URL:

<http://ecomparatio.net/>

eXChange: Exploring Concept Change and Transfer in Antiquity, URL:

<http://www.exchange-projekt.de/index.html>

Papyrus und Ostraka Projekt Halle-Jena-Leipzig, URL:

<http://papyri.uni-leipzig.de/content/below/start.xml>

Papyrus Portal, URL

<http://www.papyrusportal.de/content/start.xml>

The CITE architecture, URL:

<http://www.homermultitext.org/hmt-doc/cite/>

Autorenkontakt¹⁰

Michaela Rücker

Universität Leipzig
Historisches Seminar
Lehrstuhl für Alte Geschichte

URL: <https://www.gko.uni-leipzig.de/historisches-seminar/seminar/alte-geschichte/mitarbeiter/michaela-ruecker.html>

Sven-Philipp Brandt

Universität Leipzig
Historisches Seminar
Lehrstuhl für Alte Geschichte
Email: sven-philipp.brandt@uni-leipzig.de

URL: <https://www.gko.uni-leipzig.de/historisches-seminar/seminar/alte-geschichte/mitarbeiter/sven-philipp-brandt.html>

¹⁰ Die Rechte für Inhalt, Texte, Graphiken und Abbildungen liegen, wenn nicht anders vermerkt, bei der Autorin.

Gicht im Altgriechischen. Eine korpus-basierte Studie

Dariya Rafiyenko

The article comprises a comprehensive analysis of the Greek lexemes referring to the concept of gout (a disease characterised by intensely painful, swelling joints) for the period from 6th c. AD to 15th c. BC. In the article, two major goals are pursued. The first one is to contribute to the study of medical terminology of the period under review. The second goal is to develop and refine the methodology which would allow to trace the concepts in historical corpora. For the purpose of this study a sample with approximately 1500 utterances was created. Each entry has been annotated along a set of metadata variables such as, for example, the specific lexeme used to denote this disease and its part of speech, the period of the attestation, type of a context in which the utterance is attested, etc. Using basic statistical methods (above all frequency analysis), the distribution of metadata in the sample is analysed. This way a historical overview of the concept of gout according to the corpus is given. In addition a number of regularities in the use of different parts of speech and lexemes emerge.

1 Einleitung

Unter *Gicht* versteht man im allgemeinen Sprachgebrauch eine Erkrankung, die sich besonders in schmerzhaften akuten und chronischen Entzündungen von Gelenken äußert. Die durch die Entzündung verursachten Schmerzen treten an den rötlichen, weichen, heißen und angeschwollenen Gelenken auf. Charakteristisch für Gicht sind auch krankheitsbedingte Gelenkdeformationen in Gliedmaßen.

Dank einer Reihe deutlicher Symptome, die nur für Gicht typisch sind (v.a. heftige Schmerzen und gelähmte Gelenke), ist diese Krankheit als solche bereits seit zweieinhalbtausend Jahren diagnostizierbar und gehört wohl „zu den häufigsten chronischen Erkrankungen der Antike“.¹ Die erste klinische Beschreibung ist in den Hippokratischen Schriften zu finden, welche in das 5./4. Jh. v. Chr. zu datieren sind.² Seit der griechischen und lateinischen Antike über das europäische Mittelalter bis in die Neuzeit beschäftigten sich verschiedene Autoren mit der Krankheit, auf die sich deshalb unzählige Verweise in den Quellen finden.³ Alles in allem kann das klinische Bild der „antiken“ und „mittelalterlichen“ Gicht mit der „modernen“ Gicht – etwa im alltäglichen Sprachgebrauch – mit ziemlicher Sicherheit identifiziert werden. Diese Tatsache macht es möglich, die Geschichte der Krankheit in der Zeitspanne von mehr als zweitausend Jahren aus kultur- und medizinhistorischen Gesichtspunkten zu untersuchen.⁴

1 Stamatu (2005), S. 356.

2 Hipp. et corp. hipp., *Aff.* 30-31 (VI 242–244 Littré); *Aph.* 6,28–30 (IV 570 Littré). Zur Datierung von *De affect.* ins 5./4. Jh. v. Chr. vgl. Potter (1988) ix–xi; Jouanna/DeBevoise (1995), S. 374. Vgl. die Geschichte der Textüberlieferung dieser Werke in Anastassiou/Irmer (2012), S. 39 und 47–161.

3 Vgl. *inter alia* Ebstein (1906), S. 6–11; Hassmann (1969); Mertz (1990), S. 6–17; Gourevitch (1984), S. 217–247; Porter/Rousseau (1998), S. 13–21, Stamatu (2005).

4 Klußmann (1998), Mertz (1990) und Porter/Rousseau (1998) mit weiterführender Literatur.

Dies gilt auch für die Beispiele, bei denen eine retrospektive Diagnose nicht gewährleistet werden kann,⁵ sondern von einer historischen Deutung der Textstellen, in denen die Krankheit erwähnt ist, auszugehen ist.

Der modernen Medizin sind mehrere Formen und Ursachen dieser Krankheit bekannt, weshalb auch mehrere Bezeichnungen existieren.⁶ Antike und mittelalterliche Medizin besaß ebenfalls eine große Zahl an Bezeichnungen für Gicht.

Für diese Untersuchung werden allein solche Gichtbezeichnungen herangezogen, die in den Texten in der altgriechischen Sprache aus der Zeit vom 7. Jh. v. Chr. bis zum 15. Jh. n. Chr. vorkommen und mit den oben dargestellten Symptomen beschrieben werden (d.h. Schmerzen und Schwellungen in Gelenken sowie Deformationen der Gelenke in den Gliedmaßen). Dabei werden alle Wortarten berücksichtigt: Verba, Adjektiva und Nomina. Insgesamt handelt es sich um etwa fünfundzwanzig unterschiedliche Lexeme, die alle unten in Tabelle 1 aufgelistet sind.

Unberücksichtigt dagegen bleiben zahlreiche Bezeichnungen der Krankheiten oder krankhaften Zustände, in denen ein Bezug zur Gicht vermutet werden kann, jedoch nicht eindeutig feststellbar ist. Es handelt sich um Krankheiten, zu deren Symptomen Entzündungen in Gelenken gehören (z.B. *ισχιάς*, *γονάγρα*, *ῥεῦμα*) sowie Bezeichnungen, die in ironischen Kontexten, v.a. in Komödien, vorkommen (z.B. *βασαναστραγάλα*, *περικονδυλοπωροφίλα*, *σφυροπρησιπύρα*, *χειραλγοποδάγρα* bei Lucianus, *Podagra* und Mazaris, *Peregr.*)

Im Folgenden werden so ausgewählte Gichtbezeichnungen philologisch analysiert und nach linguistischen, historischen, kulturgeschichtlichen und soziologischen Gesichtspunkten interpretiert. Da es mehrere Bezeichnungen für diese Krankheit gibt, stellt sich die Frage nach den Kriterien und Faktoren, die die Verteilung dieser Begriffe in den Texten erklären. Relevant sind hier Fragen, wie oft dann jeder einzelne davon benutzt wird, oder welche Bedingungen sich für die Wahl eines bestimmten Lexems feststellen lassen. Ferner geht es darum, ob es Bedeutungsunterschiede zwischen den auf den ersten Blick eher als synonym erscheinenden Bezeichnungen gibt, oder ob verschiedene Wortarten (Substantive, Adjektive und Verben) lediglich grammatischen bzw. syntaktischen Anforderungen dienen oder hier auch konzeptuelle Unterschiede vorliegen. Ich gehe ferner der Frage nach, ob fachspezifische Gichtbezeichnungen oder andere Konventionen im Gebrauch der Gichtbezeichnungen feststellbar sind, die für die medizinischen Texte charakteristisch sind.

Diese Fragestellung ist als Teil der breiteren Diskussion anzusehen, basierend auf der Fragestellung, ob es eine spezielle (Fach-)Sprache der Medizin in der griechischen Antike und in Byzanz gab, und in welchem Umfang man von wissenschaftlichen Fachsprachen in prämodernen Gesellschaften sprechen kann.⁷ Unter einer Fachsprache wird die Sprachform einer

⁵ Es ist hier zu beachten, dass eine retrospektive Diagnose nach heutigen Kategorien der Medizin wenig sinnvoll ist, da die Krankheiten nur in Abhängigkeit vom wissenschaftlichen Beobachter existieren, also kultur- und epochenspezifisch sind, vgl. Leven (1998), Leven (2004), Potter (2005). Nach dem Bruch der naturwissenschaftlich-technischen Methode mit der Nosologie der antiken Medizin ist die „Unzulänglichkeit einer direkten Interpretation des antiken Gegenstandes in modernen Begriffen“ klar (Potter (2005, S. 221); zur Antike vgl. Kollesch (1998), S. 2271–2272.

⁶ So werden unter *Arthritis* Gelenkschmerzen unabhängig von deren Ursache verstanden; als *Urikopathie* wird eine Stoffwechselerkrankung bezeichnet, die zu schmerzhaften Attacken an den Gelenken führt. Da das Großzehgelenk in mehreren Fällen betroffen wird, wird die Krankheit oft auch als *Podagra* bezeichnet; für gichtbedingte Schmerzen in den Handgelenken wird die Bezeichnung *Chiragra* verwendet, usw.

⁷ Zu Fachsprachen in der lateinischen und griechischen Antike und Spätantike allgemein s. Wenskus (1998), Fögen (2003), Schironi (2014). Fögen behauptet, dass griechische und römische Schriftsteller schon eine „recht genaue Vorstellung von den Charakteristika von Fachtexten und Fachsprachen haben, die sie zum Teil sehr präzise formulieren“ (Fögen (2003), S. 55).

fachlichen Kommunikation verstanden, welche spezifische Varietät einer Einzelsprache darstellt.⁸ In der Tat ist sich die bisherige Forschung darin einig, dass es eine relativ klar definierte Sprache der Medizin schon in der Antike gab.⁹ Das Phänomen wurde aber bis jetzt hauptsächlich lexikalisch untersucht, während andere linguistische Ebenen, wie etwa Syntax, Semantik, Pragmatik oder Stilistik zum großen Teil unberücksichtigt blieben.¹⁰

Bevor ich nun zur Analyse übergehe, seien hier noch kurz die Gründe für die Wahl speziell der Gichtbezeichnungen als Untersuchungsobjekt angeführt:

(1) Die Gesamtzahl der Belegstellen für die Untersuchung ist optimal. Es handelt sich um insgesamt ca. 1500 Belegstellen im Korpus des Altgriechischen. Diese Zahl ist einerseits ausreichend groß um statistisch signifikante Tendenzen feststellen zu können, andererseits erlaubt sie eine gewisse Übersichtlichkeit.

(2) Die Belegstellen decken einen großen Zeitraum (vom 5. Jh. v. Chr. – 15. Jh. n. Chr.) ab. Dieser erlaubt die Untersuchung von Bedeutungsveränderungen der gewählten Bezeichnungen.

(3) Die Texte, in welchen Gichtbezeichnungen belegt sind, gehören zu mehreren literarischen Gattungen, wobei etwa 70 % der Belegstellen den medizinischen Texten zu entnehmen sind. Während es sich in den medizinischen Texten um einen *terminus technicus* handeln kann, bietet eine beachtliche Zahl der Belegstellen aus nicht-medizinischen Texten eine gute Voraussetzung dafür, um den Zusammenhang und die Wechselwirkungen zwischen Medizin und anderen Textgattungen (wie etwa Historiographie, Astrologie, Philosophie u.a.) in der antiken und byzantinischen Gesellschaft zu untersuchen.

(4) Für alle Lexeme, die als Gichtbezeichnungen verwendet werden, liegt eine relativ klar definierbare und diachron konstante semantische Kernbedeutung vor: Ein für die Gicht typischer Krankheitszustand. Demgegenüber weisen die meisten anderen Lexeme viel mehr Komplexität und Polysemie auf. Die Übersichtlichkeit der Bedeutung der Gichtbezeichnungen kann teilweise in der Natur der Medizinsprache liegen, wo lexikalische Mehrdeutigkeit grundsätzlich vermieden wird. Eine feste Kernbedeutung ist in dieser Untersuchung, die eine manuelle Überprüfung aller Belegstellen unmöglich macht, erforderlich.

2 Das Untersuchungskorpus

Das Untersuchungskorpus wurde in drei Arbeitsphasen erstellt: In einer lexikographischen Recherche wurden *zuerst* Lexeme ausfindig gemacht, die die passende Bedeutung aufweisen. In der *zweiten* Phase wurden alle Belegstellen für die ausgewählten Lexeme aus dem *Thesaurus Linguae Graecae Online* (TLG) extrahiert. In der *dritten* Phase wurden die Belegstellen annotiert, d.h. mit zusätzlichen, für die Untersuchung relevanten Informationen versehen.

⁸ Definition einer Fachsprache in der modernen Linguistik s. bei Fluck (1996), S. 11–26, 47–59, 193–198; Roelcke (1999), S. 15–31.

⁹ Vgl. Langholf (1977), S. 9ff., Hellweg (1985), Kollesch (1999), Fögen (2009), Geschi (2014). Zur Fachsprache der Medizin im Lateinischen von der Antike bis ins Mittelalter vgl. Baader (1982), Langslow (1989) und Langslow (2000).

¹⁰ Vgl. Langslow (1989), S. 35–36.

Für den ersten Schritt wurden (a) Wörterbücher des antiken und byzantinischen Griechisch sowie (b) Online-Datenbanken, in erster Linie TLG und *papyri.info* abgefragt. Die Suche in Wörterbüchern erfolgte ausgehend von den Begriffen der modernen Sprachen. Alle erschlossenen Lexeme wurden in einer Gruppe zusammengeführt, die ferner als *Lexemgruppe Gicht* bezeichnet wurde (s. Tabelle 1).

Lexem	bezeugt in folgenden Wörterbüchern und/oder Online-Datenbanken	Bedeutung	Wortstippe	Zahl der Belegstellen im TLG
*άρθρητικός, 3	TLG	Schreibvariante, siehe άρθρικός	ARTHRIT	5
άρθρικός, 3	LSJ, Durling, TLG	gichtkrank	ARTHRIT	279
άρθρῖτις, f	LSJ, Durling, TLG	Gicht	ARTHRIT	208
ποδάγρα, f	LSJ, Durling, LBG, TLG	Fußgicht	PODAGR	387
ποδαγράω	LSJ, TLG	an Fußgicht leiden	PODAGR	32
*ποδάγρη, f	TLG	Schreibvariante, siehe ποδάγρα	PODAGR	5
ποδαγρία, f	LBG	wie ποδάγρα	PODAGR	0
ποδαγρίαω	LSJ, Durling, TLG	wie ποδαγράω	PODAGR	47
ποδαγρικός, 3	LSJ, Durling, TLG	fußgichtkrank; gegen Fußgicht; charakteristisch für Fußgicht	PODAGR	320
ποδαγρίτης, m	LBG	an Fußgicht Erkrankter	PODAGR	0
ποδαγρός, 2	LSJ, Durling, TLG	wie ποδαγρικός	PODAGR	58
*ποδαλγεία, f	papyri.info	Schreibvariante, siehe ποδαλγία	PODALG	0
ποδαλγέω	LSJ, LBG, TLG	an Fußgicht leiden, Fußgicht hervorrufen	PODALG	14
ποδάλγη, f	Lampe, TLG	wie ποδάγρα	PODALG	2
ποδαλγής, 2	LSJ, TLG	gichtkrank	PODALG	5
ποδαλγία, f	LSJ, TLG	Gicht (in Füßen)	PODALG	85
ποδαλγίαω	LBG, TLG	an Fußgicht leiden	PODALG	1
ποδαλγικός, 2/3	LSJ, Durling, LBG, TLG	gichtkrank; Fußgicht (νόσος, πάθημα)	PODALG	26
ποδαλγός	Lampe, TLG	gichtkrank	PODALG	32
χειράγρα, f	LSJ, Durling, TLG	Handgicht; die Lähmung der Hand durch die Gicht	CHEIRAGR	20
*χειραγράω	TLG	an Handgicht leiden	CHEIRAGR	2
χειραγρικός, 3	LSJ, TLG	an Handgicht leidend	CHEIRAGR	3
χειραγρός, 3	LSJ, TLG	an Handgicht leidend	CHEIRAGR	8
χειραλγία, f	Lampe, TLG	Handgicht	CHEIRALG	4
χειραλγός, 3	LSJ, Lampe, TLG	an Handgicht leidend	CHEIRALG	3

Tabelle 1: Lexemgruppe Gicht

(*): in den gedruckten Wörterbüchern nicht belegtes Lexem

Im Verlauf der *zweiten* Arbeitsphase wurden alle Belege für die oben angeführten Lexeme aus dem TLG extrahiert. Letzteres stellt heute die größte digitalisierte Sammlung altgriechischer und byzantinischer Texte dar. Es erfasst Texte aus dem Zeitraum vom 7. Jh. v. Chr. bis zum 16. Jh. n. Chr. Der Umfang des Korpus beträgt ca. 105 Mio. Worte mit über 4000 Autoren und insgesamt ca. 12 000 Werken.¹¹ Die Suchanfrage lautete wie folgt: „ποδα[λγ][γρ]|αρθρ[ιη]τ[χειρα][λγ][γρ]“. Hierfür wurde die Option „Wildcard Search“ zusammen mit „Textual Search“ unter <http://stephanus.tlg.uci.edu/inst/textsearch> abgerufen. Es wurden insgesamt 1574 Belegstellen zur Lexemgruppe Gicht im TLG gefunden.

¹¹ www.tlg.uci.edu/about/.

Schließlich wurden alle Belege mit zahlreichen Zusatzinformationen (im Folgenden *Metadaten*) versehen:

- Der Name des Autors (falls bekannt).
- Der Titel des Werkes (der Titel entspricht meistens dem konventionellen Titel der gedruckten Ausgabe).
- Die Lebenszeiten des Autors bzw. die Zeit der Entstehung des Werkes mit einer Genauigkeit von ca. 100 Jahren.
- Die Angaben zur Gattung oder mehreren Gattungen, mit welchen ein Autor zu identifizieren ist, sowie ähnliche Angaben zur Gattung und zum Inhalt des Werkes. So steht im TLG für Platon (tlg:0059) beispielsweise die Bezeichnung *Phil.* (,Philosoph‘), für Homer (tlg:0012) *Epic.* (,epischer Dichter‘). Für den Dialog *Phaidon* des Platons (tlg:0059.004) findet sich neben dem Eintrag *Phil.* (,Philosophie‘) zusätzlich noch der Eintrag *Dialog* (,Dialog‘ als literarische Form). Diese Angaben werden hier als *Textmarkierungen* bezeichnet (s. unten).
- Grammatische Informationen zu einzelnen Wortformen.

Bezüglich der Textmarkierung des TLG müssen folgende Punkte berücksichtigt werden:

(a) Die Gestaltung der Textmarkierung ist relativ frei. Die Textmarkierungen des Autors und seiner Werk(e) können gleich oder unterschiedlich sein. So findet man beispielsweise sowohl bei Aëtius von Amida (tlg:0718) selbst als auch bei seinem Werk *Iatricorum liber i–xvi* (tlg:0718.001 bis tlg:0718.016) dieselbe Markierung *Med.*, die beim Autor als ‚medizinischer Autor‘, bei dem Werk als ‚Werk zur Medizin‘ zu interpretieren ist. Andererseits wird bei Lucianus (tlg:0062) die Markierung *Soph.*, also ‚Sophista‘, angegeben, während bei seinem Werk *Podagra* (tlg:0062.071) die Markierung ‚Satura‘, also ‚satirisches Werk‘, erscheint. Insgesamt befinden sich etwa einhundert Textmarkierungen dieser Art sowohl für Autoren als auch für Werke im TLG. Es gibt aber auch Fälle, in denen keine Markierung des Autors bzw. des Werkes eingetragen ist, wie dies z.B. bei Zenobius Sophista (tlg:0098), Joannes VI Cantacuzenus (tlg:3169) oder dem anonymen Werk *Ecthesis chronica* (tlg:2738.006) der Fall ist.

(b) Die Textmarkierungen im TLG werden nach verschiedenen Prinzipien vergeben. So beziehen sich einige davon auf den Inhalt des Werkes, wie z.B. *Theol.*, *Med.*, *Hist.*, *Astrol.* (,Theologie‘, ‚Medizin‘, ‚Historiographie‘ und ‚Astrologie‘), während andere auf die literarische Form des Werkes, wie z.B. *Hexamet.*, *Dialog* (,Dichtung in Hexametern‘, ‚Dialog‘), Bezug nehmen.

Die Textmarkierungen aus dem TLG wurden bei der Annotation des Untersuchungskorpus – unter Berücksichtigung der oben angeführten Eigenschaften – ausgiebig verwendet. Folgende Tabelle (Tabelle 2) fasst alle annotierten Metadaten zusammen. Letztere werden in zwei Haupttypen aufgeteilt: (a) grammatische Merkmale und (b) textologische Merkmale:

(a) Grammatische Merkmale	
• Wortarten	<ul style="list-style-type: none"> • Substantiv • Adjektiv • Verb
• Wortsippen und Lexeme	<ul style="list-style-type: none"> • 5 Wortsippen (ARTHRIIT; PODAGR; PODALG; CHEIRAGR; CHEIRALG) • 22 Lexeme (wie in der Tabelle 1, jedoch ohne ποδαγρία, ποδαγρίτης und ποδαλγεία)
(b) Textologische Merkmale	
• Autoren und Werke	<ul style="list-style-type: none"> • 187 Autoren (z.B. Plato, Xenophon, Galenus, Michael Psellus u.a.) • 386 Werke (z.B. <i>Alcibiades</i> von Plato, <i>De anatomicis administrationibus</i> von Galenus u.a.)
• Datierung nach historischen Perioden, angelehnt an Horrocks (2010) (grobe Kategorisierung sowohl für gut datierte als auch für mehrere schlecht datierte Belegstellen)	<ul style="list-style-type: none"> • klassisch (450 - 323 v. Chr.) • hellenistisch (323 - 31 v. Chr.) • römisch (31 v. Chr. - 235 n. Chr.) • spätantik (235 - 337 n. Chr.) • frühbyzantinisch (337 - 650 n. Chr.) • mittelbyzantinisch (650 - 1261 n. Chr.) • spätbyzantinisch (1261 - ca. 1600 n. Chr.) • postbyzantinisch (nach 1600 n. Chr.)
• Kontexte (inhaltliche Zuordnung, die auf der Basis der im TLG vorhandenen Markierung angebracht wird)	<ul style="list-style-type: none"> • Medizin (Texte, die als <i>Med.</i> im TLG markiert sind) • Historiographie (hauptsächlich <i>Hist.</i> und <i>Chronogr.</i> im TLG) • Theologische Inhalte (<i>Theol.</i>, <i>Hagiogr.</i>, <i>Apocryph.</i>, <i>Eccl.</i>, <i>Caten.</i>, <i>Homilet.</i>, <i>Ser.Eccl.</i> und <i>Exeget.</i> im TLG) • <i>Comica</i> (<i>Comic.</i>, <i>Satyr.</i> und <i>Satura</i> im TLG) • Sprachwissenschaftliche Texte (<i>Gramm.</i>, <i>Lexicogr.</i>, <i>Philol.</i> u.a. im TLG) • Astrologie (<i>Astrol.</i>, <i>Astron.</i>, aber auch manchmal <i>Magica</i>, <i>Math.</i> im TLG) • Philosophie (<i>Phil.</i> im TLG) • Biographie (<i>Biogr.</i> im TLG) • <i>Historia Naturalia</i> (<i>Hist. Nat.</i> im TLG) • unmarkiert (manche weitere Kategorien der inhaltlichen Zuordnung)

Tabelle 2: Metadaten im Untersuchungskorpus

Um dem Leser die Möglichkeit zu geben, sich den Aufbau des Untersuchungskorpus und der Annotation genau vorzustellen, werden unten in der Tabelle 3 drei Einträge aus dem Untersuchungskorpus angeführt.

Wortform	ἄρθριτιδα	ποδαγρικαῖς	ἐχειράγρασεν
Wortsippe	ARTHRIIT	PODAGR	CHEIRAGR
Wortart	Substantiv	Adjektiv	Verb
Lexem	ἄρθριτις	ποδαγρικός	χειραγρέω
Periode	Mittelbyz.	Röm.	Frühbyz.
Kontext	Medizin	Medizin	Theol. Inh.
Autor	Hippiatrica	Galenus	Antonius Hagiographus
Werk	<i>Hippiatrica Cantabrigiensia</i>	<i>In Hippocratis aphorismos commentarii vii</i>	<i>Vita Symeonis Stylitae senioris</i>
TLG-Nummer	tlg:0738.006	tlg:0057.092	tlg: 2807.001
Stelle	2,15,1	18a,84,3	27,12
Zitat	Διάρρινον πρὸς μᾶλιν ἄρθριτιδα Ἱπποκράτους.	τῆς ἐξῶθεν ὑγρότητος, ὄντα πυκνὰ καὶ σκληρὰ τὰ τοιαῦτα μόρια, τῷ τοιοῦτῳ λόγῳ καὶ μεθῆσι μόγισ. διὰ τοῦτο οὖν καὶ ταῖς ποδαγρικαῖς φλεγμοναῖς ὁ Ἱπποκράτης ὄρον ἔθετο λύσεως παντελοῦς οὐ τὴν τεσσαρεσκαδεκάτην ἡμέραν, ἀλλὰ τὴν τεσσαρακοστήν.	καὶ διαπονηθεὶς Ἀρδαβούριος ἔτεινεν τὸ τόξον αὐτοῦ βουλόμενος αὐτὸν τοξεῦσαι καὶ εὐθέως ἐνάρκησεν αὐτοῦ ἡ χεὶρ καὶ οὐκ ἴσχυσεν ἀπολῦσαι τὴν σαγίταν. οὐ μόνον δὲ ἐχειράγρασεν ἀλλὰ καὶ ἐποδάλγησεν ἕως τῆς τελευτῆς αὐτοῦ.

Tabelle 3: Beispiele aus dem Untersuchungskorpus

An grammatischen und textologischen Merkmalen sind folgende verwendet worden:

Wortstippe: Eine Wortstippe besteht aus mehreren Lexemen, die alle vom selben Stamm abgeleitet sind. Die Zahl der zugrundeliegenden Stämme ist dabei gering. Die Stämme selbst bestehen aus zwei Morphemen: Als erstes Glied erscheinen immer die Körperteile: ἄρθρον ‚Glied, Gelenk‘; πούς ‚Fuß, Bein‘ oder χεῖρ ‚Hand‘; das zweite Glied verweist darauf, das damit ein Leiden verbunden ist: -ιτ- wird oft bei der Bezeichnung einer entzündlichen Krankheit verwendet; -αλγ- von ἄλγος ‚Schmerz, Leid‘ und -αγρ- von ἄγρα ‚Jagd, Beute‘, etwa in der Bedeutung ‚Fessel‘. Somit werden fünf zusammengesetzte Stämme abgeleitet: ἀρθρ-ιτ (ARTHRIT), ποδ-αγρ (PODAGR), ποδ-αλγ (PODALG), χεῖρ-αγρ (CHEIRAGR) und χεῖρ-αλγ (CHEIRALG).

Wortart: Von den oben genannten Stämmen werden drei Wortarten abgeleitet: Substantive, Adjektive und Verben. Die Wortartenverteilung setzt sich wie folgt zusammen:

- Sieben Substantiva: ἀρθρῖτις, ποδάγρα (bzw. ποδάγρη), ποδαγρία, ποδάλγη, ποδαλγία (bzw. ποδαλγεία), χειράγρα und χειραλγία.

- Neun Adjektiva: ἀρθρικὸς (bzw. ἀρθρητικὸς), ποδαγρικός, ποδαγρός, ποδαλγικός, ποδαλγός, ποδαλγής, χειραγρικός, χειραγρός und χειραλγός.

- Fünf Verben: ποδαγράω, ποδαγριάω, ποδαλγέω, ποδαλγιάω und χειραγράω.

Autor/Werk: Hier wurden die Belegstellen dem Autor und Werk zugeordnet. Die Angaben wurden in den allermeisten Fällen ohne Änderungen aus dem TLG übernommen. Unberücksichtigt blieben Doppelseinträge, wie z.B. die identische Textpassage mit dem Adjektiv χειραλγός, die einmal als Fragment aus der Geschichte des Cassius Dio (*Hist.*, 2.–3. Jh. n. Chr., tlg:0385) und einmal als Teil der Exzerptensammlung des Kaisers Constantine VII. (*Hist., Scr. Eccl.*, 10. Jh. n. Chr., tlg:3023) angezeigt wird.¹² In Wirklichkeit handelt es sich um dieselbe Textpassage eines anonymen Geschichtsschreibers; Exzerpte aus seinem Werk wurden in der Exzerptensammlung *De sententiis* überliefert und in der Rekonstruktion der sonst verlorenen Teile des Geschichtswerks von Cassius Dio verwendet. Daher ist die fragliche Textpassage in zwei gedruckten Ausgaben vorhanden – einmal der Exzerptensammlung selbst, einmal in der Edition von Cassius Dio. Solche und ähnliche Fälle wurden berichtigt.

Datierung: Für die Datierung wurden die im TLG vorhandenen Lebenszeiten der Autoren angesetzt, selbst in den Fällen, wo der Name des Autors unbekannt bleibt, und von einem *Anonymus* die Rede ist. Auf diese Weise stellt die Zugehörigkeit zum Autor die Grundlage für die Datierung einer Belegstelle dar.

Bei gut datierten Belegstellen wird im TLG vom entsprechenden Jahrhundert ausgegangen. Dabei wird ein Schritt von 50 Jahren angenommen. So werden beispielsweise auf diese Weise die Texte, die zwischen den Jahren 400 und 499 n. Chr. entstanden sind (z.B. Werke von Priscus, Lebenszeiten 420–480, tlg:2946), von denen unterschieden, die zwischen den Jahren 450 und 549 n. Chr. geschrieben wurden (z.B. Werke von Johannes Malalas, Lebenszeiten 491–557, tlg:2871). Im ersteren Fall wird das 5. Jh. als Datierung angegeben, im letzteren Fall die Zeitspanne von der Mitte des 5. bis zur Mitte des 6. Jhs. Diese Vorgehensweise ermöglicht

¹² Vgl. Cass. Dio, *Hist.* S. 744,7 und Const. VII, *De sent.*, S. 266,17.

die Aufteilung der Gesamtperiode vom 5. Jh. v. Chr. bis zum 16. Jh. n. Chr. in mehr als vierzig kleine Zeitabschnitte. Dazu kommen noch fragliche Fälle, beispielsweise schlecht datierbare Texte wie Scholien, wofür im TLG eine Zeitspanne von bis zu mehreren Jahrhunderten angegeben ist.

Einfachheitshalber wurde im Untersuchungskorpus eine Datierung nach sieben historischen Perioden eingeführt (s. Tabelle 2 oben). Dabei wurden unter der jeweiligen historischen Periode die Belegstellen zusammengefasst, die innerhalb der entsprechenden zeitlichen Grenzen zu datieren waren, ob gut oder schlecht datierbar. Erstens reduziert die Datierung nach Perioden die Zahl der Zeitabschnitte von vierzig auf sieben. Die zeitliche Verteilung der Belegstellen lässt sich auf diese Weise besser analysieren. Zweitens erlaubt sie, auch mehrere schlecht datierbare Belegstellen in der Analyse zu berücksichtigen.

Kontext: Als wichtig hat sich im Rahmen der vorliegenden Untersuchung die Kategorie *Kontext* erwiesen. Unter dem *Kontext* allgemein wird der umgebende Text einer sprachlichen Einheit verstanden, oder auch – etwas breiter gefasst – der inhaltliche Gedanken- und Sinnzusammenhang, in dem eine sprachliche Einheit steht, und innerhalb dessen sie verstanden werden muss. In der vorliegenden Untersuchung wird unter Kontext die Gattung bzw. die inhaltliche Zuordnung des Quellentextes verstanden.

Die Kategorie „Kontext“ wurde auf der Grundlage der im TLG vorhandenen Textmarkierungen (s. oben) ausgearbeitet. Im Rahmen der vorliegenden Studie haben sich folgende Typen der Kontexte als relevant erwiesen (vgl. Tabelle 2): Medizinische Texte, Historiographie *sensu lato*, theologische Kontexte, *Comica*, sprachwissenschaftliche Abhandlungen, Astrologie, Philosophie, Biographie und *Historia Naturalia*. In Tabelle 2 werden die Textmarkierungen aus dem TLG aufgelistet, welche als Grundlage für die Zuweisung der entsprechenden Annotation dienen.

Die Zuordnung zu einem Kontexttyp erfolgte folgendermaßen. Unter dem Kontext „Medizin“ wurden sowohl klassische medizinische Traktate berücksichtigt, wie die von Hippocrates (tlg:0627), Galenus (tlg:0057), Leo Medicus (tlg:0723) u.a., als auch Rezeptsammlungen, wie etwa *Cyranides* (tlg:1482). Indessen wurden unter dem Kontext „Historiographie“ sowohl klassische historiographische Werke, die mit der Markierung *Hist.* („Geschichtswerk“) im TLG versehen sind (z.B. Werke von Xenophon (tlg:0032) oder Herodian (tlg:0015)), als auch chronographische Werke, die mit der Markierung *Chronogr.* („Chronographie“) versehen sind (z.B. Werke von Georgius Cedrenus (tlg:3018), Georgius Monachus (tlg:3043), Joel (tlg:3140)), berücksichtigt. Unter „theologischen Inhalten“ werden Texte verstanden, in denen der Bezug zur Religion stark ist. Es handelt sich beispielsweise um hagiographische, gleichzeitig aber historiographische Werke von Eusebius (tlg:2018) und Evagrius Scholasticus (tlg:2733), um Stellen aus Septuaginta (tlg:0527), Kommentare zu Bibelstellen und unterschiedliche exegetische Werke von Johannes Chrysostomus (tlg:2062), Theodoretus (tlg:4089) u.a. unter diese Kategorie fallen auch beispielsweise Briefe mit theologischen Inhalten, wie die von Photius (tlg:4040) oder Alexander von Nicaea (tlg:3341). Als „sprachwissenschaftliche Texte“ wurden Werke zur Lexikographie, Grammatiken, Onomastica u.ä. bezeichnet. Der Kontext „Biographie“ erfasst Lebensbeschreibungen in Briefen.

Alle Texte, die keinen Bezug zur Medizin haben, werden als nicht-medizinische Kontexte bezeichnet, um die Unterschiede zwischen der speziell medizinischen und jeder anderen Verwendung der Gichtbezeichnungen klarer aufzeigen zu können.

3 Methodologische Präliminarien

Bevor ich zur eigentlichen Analyse der in dem Untersuchungskorpus gesammelten Daten übergehe, werden hier kurz die Verfahren zur Auswertung der Daten besprochen. Methodologisch wird hier, neben den in der Klassischen und Byzantinischen Philologie üblichen Methoden der historischen und philologisch-literarischen Analyse, der Weg einer computer-basierten Analyse des Materials eingeschlagen, welcher eine Reihe eigener Methoden und Lösungen bietet. Grundsätzlich sollen hier die im Untersuchungskorpus vorhandenen Metadaten mit Hilfe von statistischen Auswertungsverfahren analysiert werden, um auf dieser Grundlage Schlüsse über die Verwendung von Gichtbezeichnungen im Altgriechischen zu ziehen. Dort, wo es angebracht erschien, wurden unterschiedliche Visualisierungsformen verwendet.

Ich stütze mich dabei methodologisch auf Franco Moretti (2005) und seine in der Computerphilologie als *Distant Reading* bezeichneten Methodologie. Diese erlaubt es, anhand von Metadaten die Texte sowie die dahinterstehenden kulturellen Prozesse zu analysieren, und wird als Gegenstück zum sogenannten *Close Reading* verstanden, d.h. einer detaillierten und qualitativen Betrachtung eines Textes, wie dies bei einer rein philologischen Untersuchung üblich ist.

Bei der Analyse (Sektion 4) wurden folgende Schritte abgearbeitet: Zuerst wurde die Verteilung der Belegstellen nach Zeit, Kontext und anderen Kategorien der Metadaten einzeln analysiert. Im zweiten Abschnitt wurden die Kategorien der Metadaten miteinander verglichen. Es wurde nach überproportional häufig auftretenden Kombinationen der Metadaten gesucht, welche auf Zusammenhänge zwischen den Kategorien hinweisen. Zuletzt wurden die entdeckten Assoziationen im Untersuchungskorpus analysiert und Schlussfolgerungen in Bezug auf die Lexemgruppe gezogen.

Kategorien der Metadaten im Untersuchungskorpus stellen die statistischen *Variablen* dar, welche nach einer Reihe der *Merkmalsausprägungen* kategorisiert werden (z.B. in der Variable „Wortsippe“ sind „Adjektiv“, „Substantiv“ und „Verb“ die Merkmalsausprägungen). Häufigkeitsverteilungen der Metadaten werden als Kreuztabellen dargestellt (vgl. Tabellen 4, 5 und 6).

Für die Analyse der Häufigkeitsverteilung wird der *Chi-Quadrat-Test* (χ^2) angesetzt (Bortz/Schuster 2010, 137–152), die Berechnungen werden in R ausgeführt.¹³ Entscheidend für die Wahl des Tests war die Natur der Daten. Der *Chi-Quadrat-Test* wird bei der Untersuchung *nominaler* Daten eingesetzt, welche entsprechend nach mehreren Merkmalen kategorisiert werden. Mit dem Test wird geprüft, ob ein Zusammenhang zwischen zwei nominalskalierten Variablen besteht. Im Rahmen des Tests wird eine erwartete Verteilung für jede Zeile der Tabelle geschätzt und mit der beobachteten Verteilung verglichen. Wenn die beobachteten und die erwarteten Werte gleich sind, wird vermutet, dass es keinen Zusammenhang zwischen diesen Variablen gibt. Demzufolge wird die Verteilung im Datensatz als Zufallsergebnis beschrieben. Wenn die beobachteten und erwarteten Werte stark voneinander abweichen, wird angenommen, dass die Verteilungen der zwei Variablen voneinander beeinflusst sind. So weist die beobachtete Häufigkeitsverteilung auf eine *Assoziation* zwischen zwei Variablen hin. Unter Assoziation wird dabei eine Beziehung zwischen zwei Variablen verstanden, sei es, dass diese einander auf irgendeine Weise beeinflussen oder vom Zufall beeinflusst sind.

¹³ <https://www.r-project.org>; Funktion `chisq.test()`.

Um beurteilen zu können, was eine „kleine“ und was eine „große“ Abweichung ist,¹⁴ wird in der vorliegenden Studie ein einheitliches Maß für die Abweichungen, die sog. *standardisierten Residuen* (im Folgenden als SR abgekürzt), eingesetzt. Bei SR ist der Mittelwert 0 und die Standardabweichung 1. Werte größer als 2 oder kleiner als -2 weisen auf erhebliche Abweichungen von einer zufälligen Verteilung und somit auf statistisch signifikante Abhängigkeiten unter den Merkmalen hin.

4 Der analytische Teil

Die Recherchen ergaben, dass es sich insgesamt um 25 Lexeme handelt, welche die Bedeutung „Gicht“, „gichtkrank“, „krank an Gicht sein“, „an Gicht leiden“ haben (s. Tabelle 1 oben): ἄρθρητικός, ἄρθριτικός, ἄρθριτις, ποδάγρα, ποδαγράω, ποδάγρη, ποδαγρία, ποδαγρίαω, ποδαγρικός, ποδαγρίτης, ποδαγρός, ποδαλγία, ποδαλγέω, ποδάλγη, ποδαλγής, ποδαλγία, ποδαλγίαω, ποδαλγικός, ποδαλγίτης, ποδαλγός, χειράγρα, χειραγράω, χειραγρικός, χειραγρός, χειραλγία, χειραλγός.

Alle nachgewiesenen Lexeme werden in den Übersetzungen und Wörterbüchern weitgehend ähnlich wiedergegeben. Bei den Substantiven handelt es sich um die Bezeichnung der Krankheit (außer ποδαγρίτης aus dem Lexikon zur Byzantinischen Gräzität,¹⁵ welches sich auf die kranke Person bezieht). Adjektive bezeichnen einen Verweis auf die Krankheit, sei es in der Beschreibung einer Person, die krank an Gicht ist, einem Mittel gegen Gicht, ein für Gicht charakteristisches Symptom o.ä. Verben verweisen auf einen Krankheitszustand und werden im Deutschen mit einer periphrastischen Konstruktion „an Gicht krank, an Gicht leiden“ wiedergegeben.

In den wichtigsten Wörterbüchern der klassischen und byzantinischen Periode wurden 21 Lexeme (von den 25 Lexemen in der Tabelle 1 oben) mit der relevanten Bedeutung auffindig gemacht.¹⁶ Die Sichtung weiterer Wörterbücher ergab keine weiteren Lexeme.¹⁷ In den Online-Datenbanken wurden 23 Lexeme (von den 25 Lexemen in der Tabelle 1 oben) gefunden. Dabei stieß ich auf das lexikographisch nicht dokumentierte, sondern nur im TLG nachgewiesene Lexem χειραγράω.¹⁸ Außerdem wurden orthographische Varianten ἄρθρητικός, ποδάγρη und ποδαλγία zu den in den Wörterbüchern bezeugten ἄρθριτικός, ποδάγρα und ποδαλγία nachgewiesen. Indessen waren in den Online-Datenbanken die Lexeme ποδαγρία und ποδαγρίτης aus LGB nicht auffindbar. Im weiteren Verlauf der Analyse werden nur die im TLG nachweisbaren Lexeme berücksichtigt, d.h. 22 von insgesamt 25 in Wörterbüchern und online-Datenbanken nachweisbaren Lexemen. Die Lexeme ποδαγρία, ποδαγρίτης und ποδαλγία, für welche es im TLG keine Belegstellen gibt, bleiben in der vorliegenden Untersuchung vorerst unberücksichtigt. Die orthographischen Varianten werden als ein Lexem betrachtet.

¹⁴ So beispielsweise gilt 10 als eine kleine Abweichung in der Kategorie mit 510 Belegen und als große in der Kategorie mit 15 Belegen.

¹⁵ Trapp (1994–2011).

¹⁶ Liddell/Scott (1968), Lampe (1961), Lexikon der Byzantinischen Gräzität von E. Trapp (1994–2011) sowie im Wörterbuch der medizinischen Termini von Galen von R. J. Durling (1993).

¹⁷ Wie z.B. Sophocles (1914), Pape (1866).

¹⁸ Georg. Choer., *De flexione nominum*: Τὰ δὲ ἔχοντα τὸ ρ ὑπὲρ μίαν συλλαβὴν μετὰ ἐτέρου συμφώνου, ἔτι δὲ καὶ τὰ ἔχοντα τὸ λ μετὰ ἐτέρου συμφώνου, διὰ τοῦ ἡ ἐκφέρονται κατὰ τὸν μέλλοντα, οἷον ὀμοκλῶ ὀμοκλήσω, ποδαγρῶ ποδαγρήσω, χειραγρῶ χειραγρήσω.

4.1 Die Verteilung der Belege nach einzelnen Metadatenkategorien

Im Folgenden wird die Verteilung der Belegstellen hinsichtlich einzelner Kategorien der Metadaten (Autoren und Werke, historische Perioden, Kontext, Wortsippen, Lexeme und Wortarten) beschrieben und analysiert.

Häufigkeit der Lexemgruppe pro Autor/Werk.¹⁹ Wie im Folgenden zu sehen sein wird, ist die Verteilung der Lexemgruppe *Gicht* bei verschiedenen Autoren und in verschiedenen Werken nicht gleichmäßig, sondern folgt gewissen Mustern.

Insgesamt stellt die Lexemgruppe *Gicht* mit 1570 Belegstellen im TLG eine eher unbedeutende Lexemgruppe dar. Der Gesamtumfang des TLG ist mit ca. 105 Mio. Wörtern anzusetzen und der Durchschnittsanteil unserer Lexemgruppe beträgt ca. 0.01495 pro 1000 Wörter, während beispielsweise der Anteil der Konjunktion *καί* im TLG ca. 53.5 pro 1000 Wörter beträgt.

Die 1570 Belege entstammen 187 Autoren des TLG. Die meisten Belege stammen aus den Werken zur Medizin und Astrologie, aus Komödien, historiographischen Werken und Werken theologischen Inhalts. Davon kommt die Lexemgruppe in den medizinischen Texten am häufigsten in der Sammlung *Cyranides* (mit 1.59 Wörtern pro 1000 Wörter, tlg:1482) sowie bei einigen anderen medizinischen Autoren wie Aëtius von Amida (0.6 pro 1000, tlg:0718), Dioscorides Pedanius (0.46 pro 1000, tlg:0656), Paulus von Nicaea (0.49 pro 1000, tlg:3219), Leo Medicus (0.4 pro 1000) u.a. vor. Von den Einzelwerken, in denen die Lexemgruppe am häufigsten vorkommt, können v.a. *Iatricorum liber xii* von Aëtius (6.37 pro 1000, tlg:0718.012) und *De lapidibus* von Pseudo-Dioscorides (3.65 pro 1000, tlg:1118.003) genannt werden.

Neben den rein medizinischen Autoren ist eine Reihe anderer Autoren zu erkennen, bei denen Gichtlexeme ebenfalls häufig auftreten. Von den Texten theologischen Inhalts können beispielsweise Briefe von Alexandros, Metropolit von Nicaea (0.72 pro 1000, tlg:3341) sowie das Werk *De fato et providentia (orationes 1–6)* von Johannes Chrysostomus (0.35 pro 1000, tlg:2062.060) genannt werden. Etwas unerwartet kommt die Lexemgruppe in mehreren Werken zur Astrologie auffällig oft vor, wie z. B. bei Teucer (0.48 pro 1000, tlg:1704) und Antiochus (0.43 pro 1000, tlg:1144). Schließlich ist sie in den satirischen Werken und Komödien von Lucianus (0.15 pro 1000, tlg: 0062) sowie Timarion (0.2 pro 1000, tlg:3132) relativ häufig belegt. Hier ist natürlich vor allem die Komödie *Podagra* von Lucianus zu nennen, in der die Lexemgruppe *Gicht* mit der Häufigkeit von 19.3 pro 1000 erscheint (tlg: 0062.071).

Verteilung der Lexemgruppe nach historischen Perioden. Wie oben dargestellt, wurde das Untersuchungskorpus mit acht historischen Perioden annotiert: Klassisch (450 – 323 v. Chr.), hellenistisch (323 – 31 v. Chr.), römisch (31 v. Chr. – 235 n. Chr.), spätantik (235 – 337 n. Chr.), frühbyzantinisch (337 – 650 n. Chr.), mittelbyzantinisch (650 – 1261 n. Chr.), spätbyzantinisch (1261 – ca. 1600 n. Chr.) und postbyzantinisch (nach 1600 n. Chr.). 1440 Belegstellen (91.7%) konnten einer dieser Perioden mit Sicherheit zugewiesen werden (1381 davon sind mit einer Genauigkeit von unter 100 Jahren datierbar). Die restlichen 130 Belegstellen (8.3%) sind nur schlecht datierbar und müssen deshalb hier unberücksichtigt bleiben. Abbildung 1 visualisiert die ungleichmäßige Verteilung der Lexemgruppe in den acht Perioden.

¹⁹ Bei der Berechnung der Zahl der Belege pro Autor/Werk wird die *relative Frequenz pro 1000 Wörter* angegeben, d.h. die Häufigkeit, mit welcher das Wort bzw. das entsprechende Lexem in einem Werk bzw. in allen Werken eines Autors pro 1000 Wörter vorkommt. Dieses Maß wird herangezogen, weil es sich zum Vergleich einzelner Werke/Autoren untereinander besser eignet. Wenn der Umfang einzelner Werke sehr unterschiedlich ist, sind die absoluten Häufigkeiten pro Werk/Autor nicht aussagekräftig.

Die Lexemgruppe ist zwar durch die gesamte Überlieferungszeit vom 5. Jh. v. Chr. bis zum 15. Jh. n. Chr. belegt, die meisten Belegstellen gehören aber erst in die römische (32%), früh- und mittelbyzantinische Zeit (35% und 16%). Deutlich seltener kommt die Lexemgruppe in der klassischen, hellenistischen und spätantiken Zeit vor. Auch in der postbyzantinischen Zeit sinkt die Zahl der Belege dramatisch.

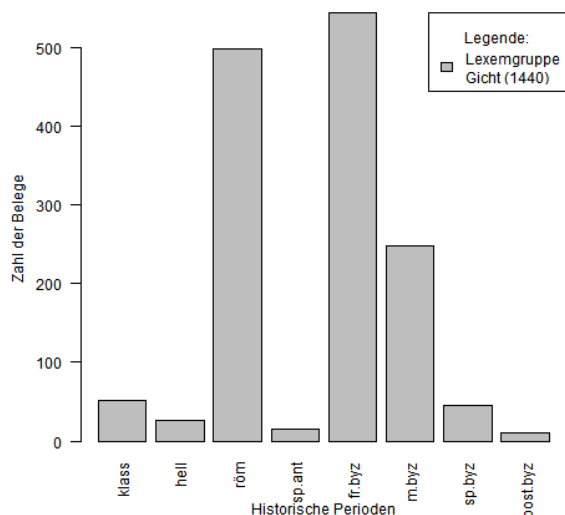


Abb. 1: Lexemgruppe Gicht nach Zeit (absolute Häufigkeit)

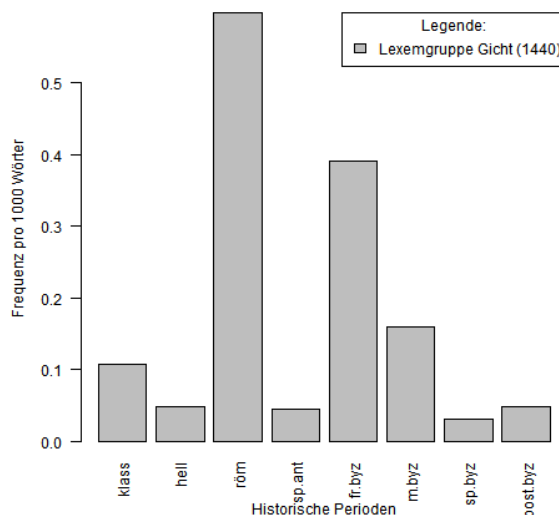


Abb. 2: Lexemgruppe Gicht nach Zeit (Frequenz pro 1000 Wörter)

In der Abbildung 2 wird die relative Frequenz der Lexemgruppe pro 1000 Wörter für einzelne historische Perioden illustriert.²⁰

Relative Frequenz und absolute Häufigkeit (s. Abbildung 1) der Lexemgruppe nach historischen Perioden ist weitgehend ähnlich. Drei Perioden, in welchen die Lexemgruppe am häufigsten relativ zum Wortumfang im Korpus auftritt, bleiben dieselben. Die römische Periode hat den größten Anteil an der Lexemgruppe (0.6 pro 1000); ihr folgen früh- und mittelbyzantinische Perioden (jeweils 0.4 und 0.17 pro 1000). Durch die Berechnung der relativen Frequenz wird ein höherer Anteil der Lexemgruppe in bezug auf den Wortumfang in der klassischen Periode sichtbar (0.1 pro 1000). In der hellenistischen, spätantiken, spät- und postbyzantinischen Periode weist die Lexemgruppe dagegen nicht nur weniger Belegstellen (vgl. Abbildung 2), sondern auch einen deutlich kleineren Anteil im Vergleich zum Gesamtwortumfang des Korpus auf: Weniger als 0.05 pro 1000 in allen Fällen.

Kontext. Mehr als die Hälfte der Belegstellen für die Lexemgruppe (1052 oder 67%) kommen in den medizinischen Texten vor, worunter medizinische Traktate, Rezeptsammlungen, Beschreibungen der Wirkstoffe gegen unterschiedliche Krankheiten und einige andere Textgattungen fallen. Fast 400 (29.2%) andere Belegstellen verteilen sich inhaltlich wie folgt: Historiographie (7.5%), theologische Inhalte (6%), *Comica* (4.5%), Sprache (3.6%), Astrologie (3.2%), Philosophie (1.8%), Briefe (1.6%) und *Historia Naturalia* (1%). Schließlich konnten die restlichen 3.9% der Belegstellen nicht sicher zugeordnet werden.

²⁰ Der Textumfang des TLG ist natürlich nicht in allen Perioden gleich. Der TLG bietet aber die Möglichkeit, den Wortumfang relativ zur Periode zu ermitteln, und zwar ist es möglich, die Zahl der Wortformen im gesamten Korpus pro Jahrhundert unter <http://stephanus.tlg.uci.edu/lris/demo/stat.jsp> (Login notwendig) abzurufen. Die ermittelten Zahlen werden hier nach historischen Perioden geordnet.

Wortsippen. Die Verteilung der Belegstellen nach den fünf Wortsippen ist wie folgt: Am häufigsten ist die Wortsippe um den Stamm PODAGR mit 866 Belegen (55.2%), dann folgt der Stamm ARTHRIT mit 492 Belegen (31.3%) und der Stamm PODALG mit 170 Belegen (10.8%). Seltene Stämme sind die beiden Stämme um die Wurzel ‘Hand’ nämlich CHEIRAGR und CHEIRALG mit jeweils 34 (2.2%) und 8 (0.5%) Belegen.

Lexeme. Im Hinblick auf die einzelnen Lexeme ist aus der Tabelle 1 oben ersichtlich, dass einzelne Lexeme im Korpus unterschiedlich oft vorkommen. Die häufigsten sind ποδάγρα (24.6%), ποδαγρικός (20.4%), ἀρθριτικός (17.8%) und ἀρθρίτις (13.2%), die insgesamt 76% aller Belegstellen ausmachen. Weitere 19% der Fälle zeigen ποδαγράω (2%), ποδαγριάω (3%), ποδαγρός (3.7%), ποδαλγός (2%), ποδαλγικός (1.7%), ποδαλγία (5.4%) und χειράγρα (1.3%). Die restlichen 11 Lexeme kommen alle zusammen nur in 5% der Fälle vor.

Wortarten. Die Verteilung der Wortarten im Untersuchungskorpus setzt sich wie folgt zusammen: Adjektive und Substantive werden etwa gleich oft mit dem Anteil von jeweils 47.1% (739 Fälle) und 45.3% (711 Fälle) verwendet; die Häufigkeit der Verwendung des Verbs steht deutlich nach und beträgt lediglich 6.1% (96 Fälle).

4.2 Beschreibung und Feststellung statistischer Assoziationen zwischen einzelnen Metadatenkategorien

Um die komplexen Zusammenhänge zwischen den Belegen und Metadatenkategorien festzustellen, werden im Folgenden einzelne Kategorien miteinander paarweise verglichen. Dabei wird hier zunächst nur eine deskriptive Beschreibung geboten, welche anhand von Visualisierungen (Abbildungen) und Tabellen veranschaulicht wird. Die Verteilung der Wortarten (Abbildung 2), Wortsippen (Abbildung 4 und 5) und Kontexte (Abbildung 6) wird in Abhängigkeit zur Datierung behandelt. Die Verhältnisse zwischen den Wortsippen, Wortarten und dem Kontext untereinander werden mit Tabellen 4, 5 und 6 veranschaulicht. Die eigentliche Analyse der Zusammenhänge folgt erst in der nächsten Sektion (4.3).

Verteilung der Wortarten nach Zeit. Adjektive kommen häufiger als alle anderen Wortarten in der römischen und frühbyzantinischen Zeit vor. Substantive sind eher für die byzantinische Zeit typisch. Zu bemerken ist ebenfalls, dass die Zahl der Substantive die der Adjektive ab der mittelbyzantinischen Zeit übersteigt.

Verteilung der Wortsippen nach Zeit. Zwei der häufigsten Wortsippen – PODAGR und ARTHRIT – sind ab der klassischen bis in die spätbyzantinische Zeit belegt. Drei weitere Wortsippen – PODALG, CHEIRALG und CHEIRAGR – sind erst später belegt: PODALG ab der hellenistischen, CHEIRAGR ab der römischen und CHEIRALG ab der frühbyzantinischen Periode.

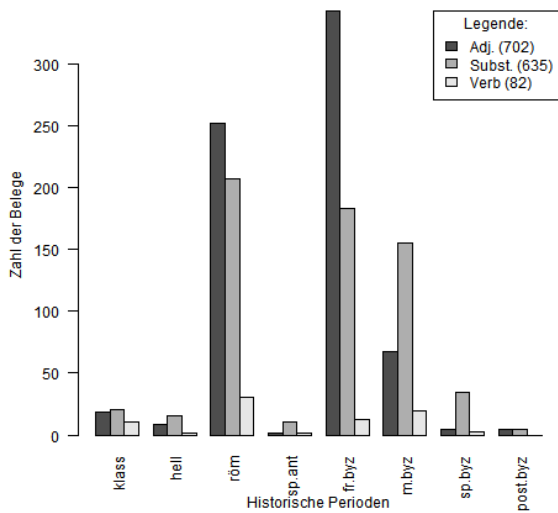


Abb. 3: Anzahl der Belege nach Wortarten

Bei den Wortsippen mit dem Bestandteil AGR (d.h. PODAGR und CHEIRAGR) fällt der Höhepunkt der Verwendung in die römische Zeit; bei denen mit ALG (d.h. PODALG und CHEIRALG) – in die byzantinische Epoche, und zwar in die früh- und mittelbyzantinische Periode. Auffällig ist dabei, dass eine gleich hohe Zahl der Belegstellen sowohl in der früh- als auch in der mittelbyzantinischen Periode bezeugt ist (vgl. ähnliches Muster der Verteilung bei Substantiven in Abb. 3). Bei allen anderen Wortsippen geht die Zahl der Belegstellen in der mittelbyzantinischen Periode stark zurück.

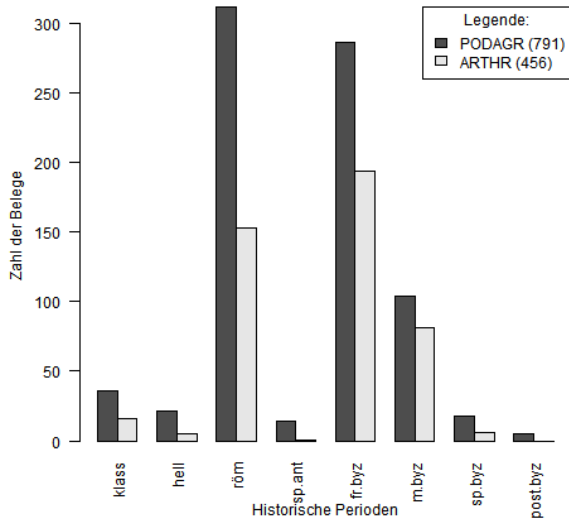


Abb. 4: Anzahl der Belege nach Wortsippen podagr- und arthrit-

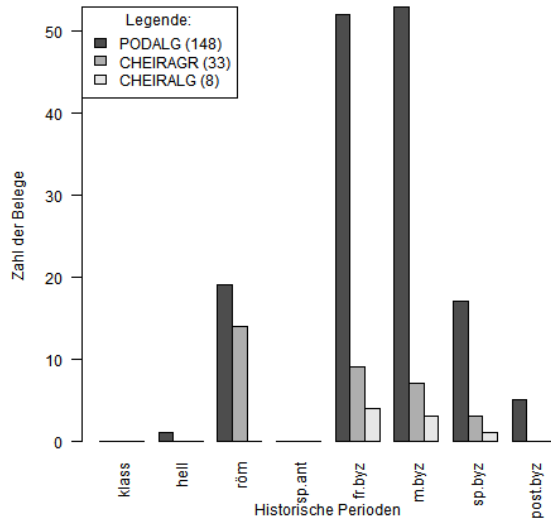


Abb. 5: Anzahl der Belege nach Wortsippen podalg-, cheiralg- und cheirarg-

Verteilung der Kontexte nach Zeit.²¹ Die medizinischen und nicht-medizinischen Kontexte sind verhältnismäßig ähnlich hinsichtlich der Zeit verteilt. Auffällig ist jedoch die Art, wie die Lexemgruppe in der byzantinischen Zeit verteilt ist: Der Höhepunkt der Verwendung in den medizinischen Texten fällt auf die frühbyzantinische, in den nicht-medizinischen Texten jedoch auf die mittelbyzantinische Periode. Typisch für die Zeit nach der mittelbyzantinischen Periode ist auch die Tendenz, dass die Zahl der nicht-medizinischen Texte die der medizinischen übersteigt.

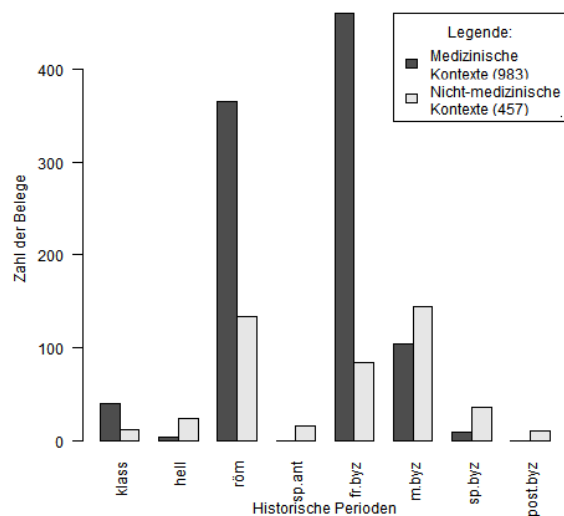


Abb. 6: Anzahl der Belege nach Kontexttypen

²¹ Hier werden alle nicht-medizinischen Belegstellen einfachheitshalber in einer Kategorie zusammengefasst.

Kontext vs. Wortart. Tabelle 4 zeigt, wie sich die Belege in den unterschiedlichen Texttypen relativ zu den drei Wortarten (Adjektiv, Substantiv und Verb) zahlenmäßig verteilen:

	Adjektiv			Substantiv			Verb			Summe
	beob.	erw.	(SR)	beob.	erw.	(SR)	beob.	erw.	(SR)	
Medizin	615	510.27	(4.64)	383	471.67	(-4.08)	45	61.06	(-2.06)	1043
Historiographie	28	56.75	(-3.82)	85	52.46	(4.49)	3	6.79	(-1.45)	116
theologische Inhalte	25	45.50	(-3.04)	62	42.06	(3.08)	6	5.44	(0.24)	93
Comica	19	31.80	(-2.27)	34	29.39	(0.85)	12	3.81	(4.20)	65
Sprache	5	26.91	(-4.22)	38	24.87	(2.63)	12	3.22	(4.89)	55
Astrologie	22	23.48	(-0.31)	26	21.71	(0.92)	0	2.81	(-1.68)	48
Philosophie	6	13.21	(-1.98)	15	12.21	(0.80)	6	1.58	(3.51)	27
Biographie	6	11.74	(-1.68)	17	10.85	(1.87)	1	1.41	(-0.34)	24
Historia Naturalia	1	7.34	(-2.34)	12	6.78	(2.00)	2	0.88	(1.20)	15
Summe	727	727		672	672		87	87		1486

Tabelle 4: Kontext vs. Wortart

Lemma (vgl. Sektion 3): (beob.): beobachtete Häufigkeit; (erw.): statistisch modellierte erwartete Häufigkeit für den Vergleich; (SR): standardisierte Residuen (Mittelwert 0, Standardabweichung 1; erhebliche Abweichung >2 oder <-2, s. Sektion 3)

Zwei wichtige Schlüsse lassen sich daraus ziehen. Zum Einen hebt sich die Medizin von allen anderen Texttypen durch häufigeres Verwenden von Adjektiven (SR 4.64) und gleichzeitiges Vermeiden der Substantive (SR -4.08) und – jedoch in kleinerem Maße – der Verben (SR -2.06) ab. Zum Anderen sind bei der Historiographie und den theologischen Inhalten ähnliche Charakteristiken zu beobachten: Beide Kategorien verhalten sich dahingehend gleich, dass sie beide die Neigung zur Verwendung der Substantive (SR jeweils 4.49 und 3.08) und das Vermeiden der Adjektive (SR jeweils -3.82 und -3.04) aufweisen. Somit können hier zwei Textsortencluster angesetzt werden: Zum Einen der Cluster „Medizin“ und zum Anderen der Cluster „Historiographie und theologische Inhalte“.

Ferner können einige weitere Beobachtungen gemacht werden. So zeichnen sich die Inhalte „Comica“, „Sprache“ und „Philosophie“ durch ein überzufällig häufiges Vorkommen an Verben (SR 4.20, 4.89 und 3.51) aus, während die Inhalte „Sprache“ durch ein überzufällig seltenes Vorkommen der Adjektive (SR -4.22) geprägt sind.

Kontext vs. Wortsippe. Tabelle 5 unten liefert eine weitere Evidenz für die beiden oben festgestellten Cluster „Medizin“ und „Historiographie und theologische Inhalte“: Diese werden auch in der Wortsippenverteilung bestätigt.

	PODAGR			ARTHRIT			PODALG		
	beob.	erw.	SR	beob.	erw.	SR	beob.	erw.	SR
Medizin	567	572.74	(0.24)	432	337.64	(5.13)	36	112.32	(-7.20)
Historiographie	46	63.7	(2.22)	23	37.55	(-2.37)	42	12.49	(8.35)
theologische Inhalte	31	51.18	(2.82)	6	30.17	(-4.40)	51	10.04	(12.93)
Comica	57	38.11	(0.06)	4	22.47	(-3.90)	8	7.47	(0.19)
Sprache	40	30.49	(0.72)	6	17.97	(-2.82)	7	5.98	(0.42)
Astrologie	27	27.77	(0.15)	7	16.37	(-2.32)	8	5.44	(1.09)
Philosophie	23	15.24	(0.99)	4	8.99	(-1.66)	1	2.99	(-1.15)
Biographie	15	13.61	(0.38)	2	8.02	(-2.13)	8	2.67	(3.26)
Historia Naturalia	15	8.17	(0.39)	0	4.81	(-2.19)	0	1.6	(-1.27)
Summe	821	821		484	484		161	161	

Tabelle 5: Kontext vs. Wortsippe

	CHEIRAGR			CHEIRALG			Summe
	beob.	erw.	SR	beob.	erw.	SR	
Medizin	17	23.72	(-1.38)	0	5.58	(-2.36)	1052
Historiographie	3	2.64	(0.22)	3	0.62	(3.02)	117
theologische Inhalte	2	2.12	(-0.08)	4	0.5	(4.96)	94
Comica	0	1.58	(-1.26)	1	0.37	(1.03)	70
Sprache	3	1.26	(1.55)	0	0.3	(-0.55)	56
Astrologie	9	1.15	(7.32)	0	0.27	(-0.52)	51
Philosophie	0	0.63	(-0.79)	0	0.15	(-0.39)	28
Biographie	0	0.56	(-0.75)	0	0.13	(-0.36)	25
Historia Naturalia	0	0.34	(-0.58)	0	0.08	(-0.28)	15
Summe	34	34		8	8		1508

Tabelle 5: Kontext vs. Wortsippe (Fortsetzung)

Lemma (vgl. Sektion 3): (beob.): beobachtete Häufigkeit; (erw.): statistisch modellierte erwartete Häufigkeit für den Vergleich; (SR): standardisierte Residuen (Mittelwert 0, Standardabweichung 1; erhebliche Abweichung >2 oder <-2, s. Sektion 3)

Der Kontext „Medizin“ unterscheidet sich von allen anderen Kontexten durch überzufällig²² häufiges Verwenden von ARTHRIT (SR 5.13) sowie das Vermeiden von PODALG (SR -7.20). Für den Cluster „Historiographie und theologische Inhalte“ ist die Vorliebe zu PODALG (SR 8.35 und 12.93) und CHEIRALG (SR 3.02 und 4.96) ausgeprägt, während die Wortsippen PODAGR (SR 2.22 und 2.82) und ARTHRIT (SR -2.37 und -4.40) untypisch sind. Des Weiteren wird in allen Kontexten außer „Medizin“ das Vermeiden von ARTHRIT beobachtet. In „Astrologie“ ist eine starke Vorliebe für CHEIRAGR festzustellen (SR 7.32).

Wortart vs. Wortsippe. Es folgt nun die zahlenmäßige Verteilung der Belegstellen nach Wortart und -art:

	Adjektiv			Substantiv			Verb			Summe
	beob.	erw.	SR	beob.	erw.	SR	beob.	erw.	SR	
PODAGR	378	405.83	(-1.38)	392	390.45	(0.08)	79	52.72	(3.62)	849
ARTHTRIT	284	235.18	(3.18)	208	226.27	(-1.21)	0	30.55	(-5.53)	492
PODALG	63	78.87	(-1.79)	87	75.88	(1.28)	15	10.25	(1.49)	165
CHEIRAGR	11	15.77	(-1.20)	20	15.18	(1.24)	2	2.05	(-0.03)	33
CHEIRALG	3	3.35	(-0.19)	4	3.22	(0.44)	0	0.43	(-0.66)	7
Summe	739	739		711	711		96	96		1546

Tabelle 6: Wortsippe vs. Wortart

Lemma (vgl. Sektion 3): (beob.): beobachtete Häufigkeit; (erw.): statistisch modellierte erwartete Häufigkeit für den Vergleich; (SR): standardisierte Residuen (Mittelwert 0, Standardabweichung 1; erhebliche Abweichung >2 oder <-2, s. Sektion 3).

Aus den Daten in Tabelle 6 folgt, dass es die Wortsippe PODAGR ist, die in breiterem Maße (SR 3.62) als die anderen Stämme mit der Möglichkeit kompatibel ist, die Krankheit verbal zum Ausdruck zu bringen. In anderen Worten, möchte der Autor die Krankheit mit einem Verb

²² Der Begriff der *Überzufälligkeit* wird in der Inferenzstatistik verwendet. Als *überzufällig* gelten solche Zusammenhänge, die sich aus überdurchschnittlich häufigen und statistisch signifikanten Korrelationen erschließen lassen.

ausdrücken, muss er in der Regel zu den Verben genau dieser Sippe und nicht zu denen einer der anderen Sippen greifen. Zwar ist diese Wortsippe auch zahlenmäßig die stärkste, dieses Ergebnis ist jedoch nicht davon abhängig.

Ferner ist von einem Zusammenhang zwischen Adjektiven und ARTHRIT (SR 3.18) auszugehen.

4.3 Auswertung der festgestellten Zusammenhänge

Die in Sektion 4.2 aufgezeigten Assoziationen sollen im Folgenden ausgewertet werden, indem zuerst die im Untersuchungskorpus beobachteten Tendenzen in den einzelnen Wortsippen beschrieben werden und dann die Geschichte der Lexemgruppe zu verschiedenen historischen Perioden analysiert wird. Zuletzt werden die empirisch festgestellten Cluster „Medizin“ und „Historiographie und theologische Inhalte“ näher charakterisiert.

Übersicht über die Wortsippen (Abbildung 7). Zwar ist die Lexemgruppe seit der klassischen bis in die spätbyzantinische Periode durchgehend bezeugt, doch sind es nur PODAGR und ARTHRIT (zugleich auch die zwei häufigsten Wortsippen), die in allen Perioden belegt sind. Es handelt sich dabei hauptsächlich um vier Lexeme: ποδάγρα, ποδαγρικός, ἄρθριτις und ἄρθριτικός (respektive 25%, 20%, 13% und 18% im Untersuchungskorpus).

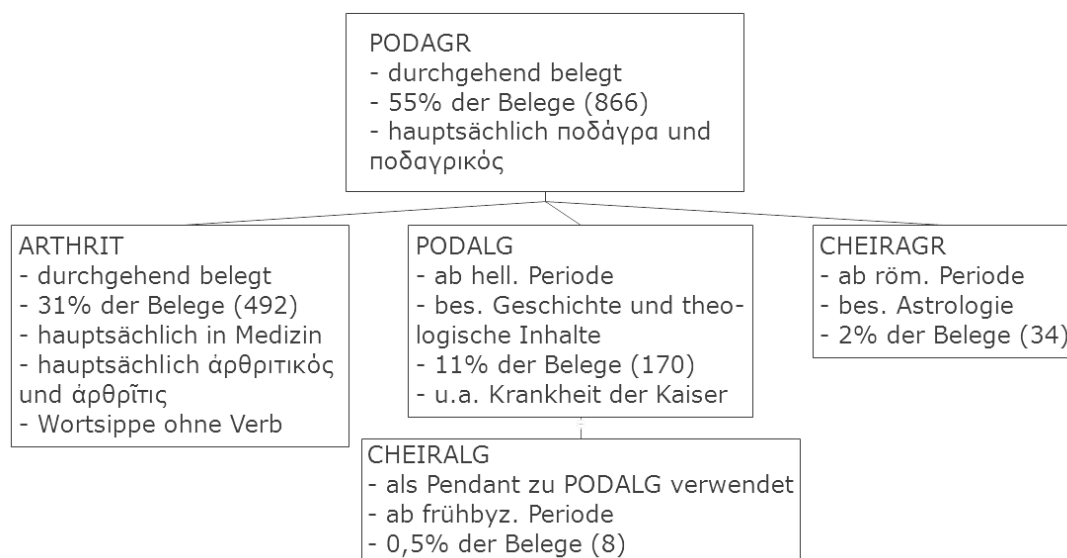


Abb. 7: Übersicht über die Wortsippen

Das Auftreten der Wortsippen in den Texten ist nicht zufallsbedingt, sondern folgt klaren semantischen Vorgaben (vgl. Tabelle 7). Diese sehen folgendermaßen aus: Die Wortsippe PODAGR, die in allen literarischen Kontexten verwendet wird und 55% des Untersuchungskorpus ausmacht, kann als stilistisch und kontextneutrale Kernwortsippe angesehen werden. Die Wortsippe ARTHRIT hingegen, die mit 31% der Belegstellen im Untersuchungskorpus repräsentiert ist, ist hauptsächlich auf die medizinischen Kontexte beschränkt und kann daher als *terminus technicus* der Medizin bezeichnet werden. Diese Wortsippe unterscheidet sich ferner von den anderen zwei durch die Wortartenverteilung. Während Adjektive bei ARTHRIT überzufällig häufig erscheinen, stellt ARTHRIT gleichzeitig die einzige Wortsippe dar, von der kein Verb abgeleitet wurde. Diese Eigenschaften lassen sich vermutlich auf die Eigenarten der medizinischen Fachsprache zurückführen.

	PODAGR	ARTHRIT	PODALG	CHEIRAGR	CHEIRALG
Medizin		+	-		-
Historiographie	+	-	+		+
theologische Inhalte	+	-	+		+
Biographie		-	+		
Astrologie		-		+	
<i>Comica</i>		-			
Sprache		-			
<i>Historia Naturalia</i>		-			
Philosophie					
	PODAGR	ARTHRIT	PODALG	CHEIRAGR	CHEIRALG
Adjektiv		+			
Substantiv					
Verb	+	-			

Tabelle 7: Übersicht über die Wortsippen
 Lemma (vgl. Sektion 3): (+): SR >2; (-): SR <-2

Die Wortsippe PODALG, die 11% aller Belegstellen des Untersuchungskorpus ausmacht und ab der römischen Zeit belegt ist, kommt zwar fast in allen Kontexten vor, ist aber besonders in historiographischen und theologischen Werken üblich. Die Wortsippe kann als nicht-medizinisches Äquivalent der Gichtbezeichnung charakterisiert werden.

Bei den Wortsippen mit dem Bestandteil CHEIR- handelt es sich um äußerst seltene und relativ späte Wortsippen, die erst seit römischer Zeit im Falle von CHEIRAGR bzw. seit frühbyzantinischer Zeit im Falle von CHEIRALG belegt sind. CHEIRAGR kommt überzufällig oft in astrologischen Schriften, CHEIRALG in Historiographie und theologischen Inhalten vor. Die Überprüfung der äußerst selten belegten Wortsippe CHEIRALG zeigte, dass es sich dabei um Fälle handelt, in denen die Wortsippe kontrastiv zu einem unmittelbar vorher verwendeten PODALG steht und allein auf die Erkrankung der Hände verweist, wie z.B.:

Ephraem Syr., *De virt.* 4,65: Εἰ δὲ καὶ περὶ τοῦ κόπου τοῦ ἔργου δυσφορεῖ τις, πόσοι εἰσὶν κατὰ τὸν κόσμον χειραλγοὶ καὶ ποδαλγοί, ἄνδρες πλούσιοι [...].

Übersicht der historischen Perioden. Die Abbildungen (oben, Sektion 4.2), die die Verteilung der Kontexte (Abbildung 3), Wortsippen (Abbildung 4, 5) und Wortarten (Abbildung 6) in Abhängigkeit zur Datierung visualisieren, erlauben einen Gesamtüberblick über die Geschichte der Lexemgruppe. In klassischer und hellenistischer Periode sind nur einige dutzend Belegstellen sowie lediglich zwei Wortsippen bezeugt: PODAGR und ARTHRIT. Die Belegstellen aus der Medizin sind verhältnismäßig zahlreich. Es handelt sich um *Hippocrates et Corpus Hippocraticum* (33 Belegstellen, tlg:0627) und Diocles (7 Belegstellen, tlg:0664).

Die römische Periode, aus der wir eine extrem hohe Zahl der Belegstellen haben, ist für die Untersuchung der Lexemgruppe von besonderer Wichtigkeit. PODAGR und ARTHRIT, die aus der klassischen und hellenistischen Periode ererbt sind, genießen in dieser Zeit eine zahlenmäßig starke Ausbreitung. Außerdem kommen zwei weitere Wortsippen – PODALG und CHEIRALG – zum ersten Mal hinzu. Damit sind für die römische Periode vier Wortsippen anzusetzen. Zahlreich sind die Belegstellen aus sowohl medizinischen als auch nicht-medizinischen Kontexten. Zu dieser Zeit ist die Lexemgruppe in allen im Untersuchungskorpus vorhandenen Kontexten bezeugt. Bei den medizinischen Kontexten handelt es

sich um folgende Autoren: Galenus (223 Belegstellen, tlg:0057), *Cyranides* (58, tlg:1482), Dioscorides Pedanius (65, tlg:0656), Aretaeus (11, tlg:0719) u.a. Bei den nicht-medizinischen Kontexten ist die Lexemgruppe am häufigsten in folgenden Kontexten bezeugt: *Comica* (43 Belegstellen, alle aus Lucianus, tlg: 0062), Philosophie (17), Sprache (18) und Astrologie (19). Die Substantive und Adjektive werden gleich häufig verwendet; verhältnismäßig zahlreich sind auch Verben.

Aus der Spätantike sind nur einige wenige Belegstellen erhalten, welche jedoch weitgehend ihre Charakteristiken mit denen aus der römischen Zeit teilen.

In byzantinischer Zeit, d.h. in der Zeit ab der zweiten Hälfte des 4. Jh. n. Chr. bis zum 15. Jh. kommen wieder sehr viele Belege vor, die die Zahl der Belege bis zum 5. Jh. n. Chr. insgesamt übersteigen. Charakteristisch für die gesamte byzantinische Zeit ist die Vorliebe für die Wortsippen mit dem Bestandteil –ALG. Die Wortsippe PODALG ist zwar schon seit der römischen Zeit bekannt, nimmt aber ab der frühbyzantinischen Zeit zahlenmäßig wesentlich zu. Die Wortsippe CHEIRALG ist nur in byzantinischer Zeit bezeugt.

Die frühbyzantinische Periode hebt sich durch eine sehr hohe Zahl der Belegstellen aus der Medizin hervor. Diese Belegstellen stammen von folgenden Autoren: Aëtius von Amida (247, tlg:0178), Oribasius (110, tlg:0722), Alexander (43, tlg:0744), Paulus (59, tlg:0715) u.a. Verhältnismäßig häufig (insgesamt 51 Belegstellen) kommen Gichtbezeichnungen in den Abhandlungen theologischen Inhalts vor: Bei Johannes Chrysostomus (26, tlg:2062), Sophronius (4, tlg:4042), Theodoretus (2, tlg:4089), Sozomenus (2, tlg:2048), Antonius Hagiographus (2, tlg:2807) u.a.

Die mittelbyzantinische Zeit ist durch ein häufiges Verwenden der Gichtbezeichnungen in den nicht-medizinischen Kontexten geprägt, die die eigentlich medizinischen Kontexte hinsichtlich der Belegstellenzahl weit übertreffen. Die medizinischen Belege stammen aus *Hippiatrica* (49, tlg:0738), Leo Medicus (7, tlg:0723), Symeon Seth (15, tlg:3113) u.a.

Insgesamt 57 Mal ist die Lexemgruppe in den historiographischen Werken belegt: In Georgius Monachus (5, tlg:3043), Georgius Cedrenus (6, tlg:3018), Symeon Logothetes (5, tlg:3070), Theophanes Confessor (5, tlg:4046), Joel (4, tlg:3140) sowie manchen anderen. Verhältnismäßig oft ist die Lexemgruppe in der mittelbyzantinischen Zeit in den Abhandlungen zur Sprache anzutreffen (26 Mal insgesamt): Bei Photius (7, tlg:4040), in der Suda (9, tlg:9010), bei Georgius Choeroboscus (5, tlg:4093), Eustathius Thessalonicensis (4, tlg:4083) sowie bei manchen anderen Autoren. Charakteristisches Merkmal der mittelbyzantinischen Zeit ist die häufige Verwendung der Wortsippe PODALG. Dies erscheint umso auffälliger, da alle anderen Wortsippen zahlenmäßig zurückgehen. Ein weiteres besonderes Merkmal dieser Periode ist die häufige Verwendung der Substantive – zum ersten Mal übersteigen diese die Zahl der Adjektive.

In der spätbyzantinischen Periode kommen einige Tendenzen deutlicher zum Vorschein, die sich schon in der mittelbyzantinischen Periode abzeichnen: So findet sich ein häufiges Vorkommen der Gichtbezeichnungen in nicht-medizinischen Kontexten, zum großen Teil in der Historiographie, eine Vorliebe zu Substantiven sowie zur Wortsippe PODALG. Die Letztere ist verhältnismäßig häufig. Sie ist mit insgesamt 24 Belegstellen in historiographischen Werken bezeugt: Bei Georgius Pachymeres (5, tlg:3142), Nicephorus Callistus Xanthopulus (5, tlg:3236), Laonicus Chalcocondyles (4, tlg:3139) u.a.

Soweit zu den historischen Perioden und deren Entwicklungen. Im Hinblick auf die Textinhalte wurden oben aufgrund der Verwendung von Wortarten und Wortsippen zwei Kontextcluster angesetzt: „Medizin“ und „Historiographie und theologische Inhalte“. Im Folgenden wurde jeder der beiden Cluster für sich betrachtet (vgl. Tabelle 8).

Cluster „Medizin“. Die medizinischen Texte stellen einen Cluster dar, d.h. sie verhalten sich in Bezug auf einige Metadatenkategorien (vor allem Wortarten > Adjektive) einheitlich und gleichzeitig mit Abstand anders als alle anderen Texttypen. In der Verwendung von Wortarten unterscheidet sich dieser Cluster durch das häufigere Verwenden von Adjektiven und das Vermeiden von Substantiven (vgl. oben Tabelle 4); Verben sind hingegen in den medizinischen Texten eher selten. In der Verwendung von Wortsippen ist für diesen Cluster überzufällig häufiges Verwenden von ARTHRIT und das Vermeiden von PODALG charakteristisch (s. oben Tabelle 5 und 6). Die häufigsten Lexeme in der Medizin sind die zwei Adjektive ἄρθριτικός und ποδαγρικός. Die nicht-medizinischen Kontexte von ἄρθριτικός betragen indessen lediglich 5% und von ποδαγρικός 7%.

	Adjektiv	Substantiv	Verb	PODAGR	ARTHRIT	PODALG	CHEIRAGR	CHEIRALG
Medizin	+	-	-		+	-		-
Historiographie	-	+		+	-	+		+
theologische Inhalte	-	+		+	-	+		+
<i>Historia Naturalia</i>	-	+			-			
Sprache	-	+	+		-			
<i>Comica</i>	-		+		-			
Philosophie			+					
Astrologie					-		+	
Biographie					-	+		

Tabelle 8: Übersicht über die Kontexte
Lemma (vgl. Sektion 3): (+): SR >2; (-): SR <-2

Cluster „Historiographie und theologische Inhalte“. Zwei Kontexttypen, welche in der vorliegenden Untersuchung als „Historiographie“ und „theologische Inhalte“ bezeichnet werden, unterscheiden sich von allen anderen Kontexten durch eine beachtenswert ähnliche Verwendung von Wortsippen und Wortarten. Von den Wortsippen werden hier PODALG und CHEIRALG weitgehend bevorzugt, während PODAGR und ARTHRIT vermieden werden. Man könnte somit annehmen, dass dieser Cluster den Bestandteil -ALG bevorzugt. Wie dem auch sei, es sind die beiden einzigen Kontexte, in denen alle fünf Wortsippen belegt sind. In der Verwendung von Wortarten kann die Neigung zu Substantiven festgestellt werden (vgl. oben Tabelle 4). Adjektive sind für den Cluster eher untypisch.

Hier schließt sich die Frage an, warum die Kontexte „Historiographie“ und „theologische Inhalte“ weitgehend ähnliche Züge aufweisen. Es scheint, dass diese beiden Kontexte einen wichtigen Berührungspunkt haben, der letztendlich ähnliche Prinzipien für die Textgestaltung mit sich bringt. Dieser Berührungspunkt sind die Lebensbeschreibungen, die in den beiden Fällen besonders häufig vorkommen. In der Historiographie geht es dabei um Lebensbeschreibungen von historisch wichtigen Personen, in den Texten theologischen Inhalts hingegen handelt es sich um Lebensbeschreibungen von Märtyrern oder Heiligen. Die Verwendung von allen fünf Wortsippen findet dabei ihre Erklärung in der Neigung zur Vielfältigkeit in der Narration, da es sich vermutlich um literarisch anspruchsvollere Texte handelt, als bei anderen Kontextkategorien dieser Untersuchung.

Demzufolge kann im Falle von Gichtbezeichnungen das zusammengefügte Cluster „Historiographie + theologische Inhalte“ als Cluster „*Vitae*“ bezeichnet werden. Denn wenn es um Lebensbeschreibungen geht, bedienen sich die Autoren der profanen als auch kirchlichen Texte im Falle von Gichtbezeichnungen derselben Konventionen. Dabei unterscheiden sich diese Konventionen erheblich von denen, die in anderen, in dieser Untersuchung definierten Kontexten wie z.B. „Medizin“, „*Comica*“, „Sprachwissenschaftliche Texte“ u.a., auftreten.

Kaiser und Podalgia. Ein weiterer Befund, der im Laufe der vorliegenden Untersuchung gemacht wurde, betrifft die Bezeichnung dieser Krankheit bei den Kaisern. In mehr als ein Dutzend Quellen wurde für folgende Kaiser die Krankheit Gicht mit der Wortsippe PODALG beschrieben: Commodus²³ (Regierungsjahre 180 – 192); Constantinus I.²⁴ (Regierungsjahre 306 – 337); Mauricius²⁵ (Regierungsjahre 582 – 602); Michael VI.²⁶ (Regierungsjahre 1056 – 1057); Alexius I. Comnenus²⁷ (Regierungsjahre 1081 – 1118); Isaak II. Angelus²⁸ (Regierungsjahre 1185 – 1192 und 1203 – 1204); Alexius III. Angelus²⁹ (Regierungsjahre 1195 – 1203). Bis dato konnten keine weiteren Textstellen festgestellt werden, in denen die Gicht bei Kaisern mit einer anderen Wortsippe beschrieben wird.

5 Schlussfolgerungen und Ausblicke

In der vorliegenden Untersuchung wurden alle altgriechischen Begriffe für die Bezeichnung der Krankheit *Gicht* vergleichend untersucht. Die Analyse erfolgte auf der Grundlage eines Untersuchungskorpus, das sämtliche Belegstellen der Lexeme mit der entsprechenden Bedeutung im gesamten Thesaurus Linguae Graecae Online enthält. Diese Belegstellen wurden zusätzlich mit den relevanten Metadaten annotiert (s. Sektion 2). Vier Metadatenkategorien ergaben sich als wichtig für die Untersuchung: Zugehörigkeit zu einer Wortart, Wortsippe, historischen Periode und zu einem Kontexttyp. Insgesamt beinhaltet die Lexemgruppe 22 Lexeme, die mit insgesamt 1570 Belegstellen im Korpus repräsentiert sind.

Anhand der hier vorgestellten Methode konnten neue Erkenntnisse für die klassische und byzantinische Philologie gewonnen werden. Zum Einen wurden Konventionen im Bereich Medizin festgestellt: Die Neigung zur Verwendung von Adjektiven sowie die Vorliebe für die Wortsippe ARTHRIT. Damit lässt sich die These von einer schon in der Antike vorhandenen Fachsprache der Medizin im Bereich Morphologie und Syntax bestätigen. Zum Anderen leistet diese Studie einen Beitrag zum Verständnis der Gattungen, und zwar dergestalt, dass sich die Historiographie und die Werke theologischen Inhalts ähnlicher Konventionen in der Verwendung der Lexemgruppe Gicht bedienen und somit einen stärkeren Bezug zueinander als zu den anderen Gattungen haben.

23 Georg. Cedr., *Comp. hist.*, vol. 1 p. 441,6; Georg. Mon., *Chron. breve*, vol. 110 p. 532,10; Sym. Log., *Chron.*, p. 94,4.

24 Georg. Mon., *Chron.*, p. 729,6; Georg. Mon., *Chron. breve*, vol. 110 p. 896,7; Sym. Log., *Chron.*, p. 168,43; Niceph. I., *Brev. hist.*, p. 35,7; J. Zonaras, *Epitome hist.*, p. 227,10; Theoph. Conf., *Chronogr.*, p. 358,28.

25 Const. VII Porph., *De ins.*, p. 148,24 (= J. Antioch., *Fragmenta*, fr. 318,11); Georg. Mon., *Chron.*, p. 662,8; Joel, *Chronogr.*, p. 46,2; Theoph. Conf., *Chronogr.*, p. 288,29.

26 Joel, *Chronogr.*, p. 63,11.

27 Mich. Glycas, *Annales*, 622,15; *Anonyma Metaphrasis ‚Alexiadis‘*, p. 159,4.

28 Ephraem Aeniensis, *Hist. Chron.* 6994.

29 Ephraem Aeniensis, *Hist. Chron.* 6453; 6630.

Schließlich kann die hier gebotene Übersicht der Lexemgruppe *Gicht* sowohl nach Wortsippen als auch nach der Zeit eine Grundlage für die weitere Beschäftigung mit den besprochenen Lexemen sein. Sie ist als erster Schritt in der Untersuchung der Bedeutung und Geschichte der *Gicht*-Bezeichnungen anzusehen (vgl. Beschreibung nach historischen Perioden in 4.3).

Die weitere Forschung kann in unterschiedliche Richtungen gehen: Besonders aussichtsreich scheint die Untersuchung der Gattungskonventionen zu sein. Im besten Fall könnte sie sogar zur Klassifizierung der Texte beitragen. Sinnvoll ist auch die Untersuchung anderer Lexeme und Lexemgruppen, nicht zuletzt für die Untersuchungen der gesellschaftlichen Prozesse oder auch für das Wörterbuchwesen.

Abkürzungsverzeichnis

- LBG Lexikon der Byzantinischen Gräzität, s. Trapp (1994–2011)
SR standardisiertes Residuum
TLG Thesaurus Linguae Graecae online, s. Pantelia (Hg.)

Literatur

- Anastassiou / Irmer (2012): A. Anastassiou / D. Irmer, Testimonien zum Corpus Hippocraticum. Teil III: Nachleben der hippocratischen Schriften in der Zeit vom 4. bis zum 10. Jahrhundert n. Chr., Göttingen 2012 [u.a.].
- Baader (1982): G. Baader, Die Entwicklung der medizinischen Fachsprache in der Antike und im frühen Mittelalter, in: G. Baader / G. Keil (Hrsg.) Medizin im mittelalterlichen Abendland, Darmstadt 1982, S. 417–442.
- Bortz/Schuster (2010): J. Bortz / Chr. Schuster, Statistik für Human- und Sozialwissenschaftler. 7. Aufl. Berlin [u.a.] 2010.
- Durling (1993): R.J. Durling, A Dictionary of Medical Terms in Galen. Leiden [u.a.] 1993.
- Ebstein (1906): W. Ebstein, Die Natur und Behandlung der Gicht. 2. Aufl. Wiesbaden 1906.
- Fluck (1996): H.-R. Fluck, Fachsprachen. Einführung und Bibliographie, 5. Aufl., Tübingen 1996.
- Fögen (2003): Th. Fögen, Metasprachliche Reflexionen antiker Autoren zu den Charakteristika von Fachtexten und Fachsprachen, in: Horster, M./Reitz, Chr. (Hrsg.), Antike Fachschriftsteller: Literarischer Diskurs und sozialer Kontext, Wiesbaden 2003, S. 31–60.
- Fögen (2009): Th. Fögen, Wissen, Kommunikation und Selbstdarstellung. Zur Struktur und Charakteristik römischer Fachtexte der frühen Kaiserzeit, München 2009.
- Geschi (2014): G. Geschi, Medical Vocabulary, in: Giannakis u.a. (2014) (Hrsg.), 2, S. 404–407.
- Giannakis u.a. (2014) (Hrsg.): G.K. Giannakis, Encyclopedia of Ancient Greek language and linguistics, 3 Bde. Leiden [u.a.] 2014.
- Gourevitch (1984): D. Gourevitch, Le triangle hippocratique dans le monde gréco-romain, Paris 1984.
- Hassmann (1969): P. Hassmann, Zur Medizingeschichte und Sozialpathologie der Gicht, Med. Diss., Düsseldorf 1969.
- Hellweg (1985): R. Hellweg, Stilistische Untersuchungen zu den Krankengeschichten der Epidemienbücher I und III des Corpus Hippocraticum, Bonn 1985.
- Horrocks (2010): G. Horrocks, Greek: A History of the Language and its Speakers, Chichester [u.a.] 2010.
- Jouanna / DeBevoise (1995): J. Jouannam / M.B. DeBevoise, Hippocrates. Translated by M. B. DeBevoise, Baltimore [u.a.] 1995.
- Klußmann (1998): R. Klußmann, Gicht – Gier – Größe – Macht. Herrscher im Spannungsfeld von Lust und Leid, Gießen 1998.

- Kollesch (1999): J. Kollesch, Medizin und ihre Fachsprache im Altertum, in: Hoffmann u.a. (Hrsg.) Lothar, Fachsprachen. Languages for Special Purposes. Ein internationales Handbuch zur Fachsprachenforschung und Terminologiewissenschaft. Hoffmann, L./Kalverkämper, H./Wiegand, H. E. (Hgg.) In Verbindung mit Christian Galinski, Werner Hüllen. Bd. 2. Berlin [u.a.] 1999, S. 2270–2277.
- Lampe (1961): G.W.H. Lampe, A patristic Greek lexicon, Oxford 1961.
- Langholf (1977): V. Langholf, Syntaktische Untersuchungen zu Hippokrates-Texten. Brachylogische Syntagmen in den individuellen Krankheits-Fallbeschreibungen der hippokratischen Schriftensammlung, Wiesbaden 1977.
- Langslow (2000): D. Langelow, Medical Latin in the Roman Empire, Oxford 2000.
- Leven (1998): K.–H. Leven, Krankheiten: Historische Deutung versus retrospektive Diagnose, in: N. Paul / Th. Schlich (Hgg.), Medizingeschichte: Aufgaben, Probleme, Perspektiven. Unter Mitarbeit von Stefanie Kuhne, Frankfurt [u.a.] 1998, S. 153–185.
- Leven (2004): K.–H. Leven, „At times these ancient facts seem to lie before me like a patient on a hospital bed“ – Retrospektive Diagnosis and Ancient Medical History, in: H.F.J. Horstmanshoff / M. Stol, M. (Hgg.) Magic and Rationality in Ancient Near Eastern and Graeco-Roman Medicine, Leiden 2004, S. 369–386.
- Leven (2005): Leven, K.–H.(Hrsg.), Antike Medizin. Ein Lexikon, München 2005.
- Liddell / Scott (1968): H.G. Liddell / R. Scott (Hgg.), A Greek-English Lexicon. Revised and augmented throughout by Sir H. S. Jones with the assistance of R. McKenzie. With a Supplement. (Nachdr. der 9. Aufl. von 1940). Oxford 1968.
- Mertz (1990): D.P. Mertz, Geschichte der Gicht. Kultur- und medizinhistorische Betrachtungen, Stuttgart [u.a.] 1990.
- Moretti (2005): F. Moretti, Graphs, Maps, Trees: Abstract Models for a Literary History. London and New York 2005.
- Pantelia (Hrsg.): Thesaurus Linguae Graecae Online. URL: <http://www.tlg.uci.edu/>.
- Pape (1866): W. Pape, Deutsch-griechisches Handwörterbuch. Bearbeitet von Maximilian Sengebusch. 3. Aufl., Braunschweig 1866.
- Porter / Rousseau (1998): R. Porter / G.S. Rousseau, Gout : the Patrician malady, New Haven [u.a.] 1998.
- Potter (1988): P. Potter, Hippocrates, Bd. V. With an English translation by P. Potter, London 1988.
- Potter (2005): P. Potter, Diagnose, retrospektive, in: Leven (2005) (Hrsg.), S. 220–222.
- Roelcke (1999): T. Roelcke, Fachsprachen, Berlin 1999.
- Schironi (2014): F. Schironi, Scientific Vocabulary, in: Giannakis u.a. (2014) (Hrsg.), 3, S. 262–265.
- Sophocles (1914): E.A. Sophocles, Greek lexicon of the Roman and Byzantine periods. (From B. C. 146 to A. D. 1100). Cambridge 1914
- Stamatu (2005): M. Stamatu, Gicht, in: Leven (2005) (Hrsg.), 356–358.
- Trapp (1994–2011): E. Trapp (Hrsg.) Lexikon zur byzantinischen Gräzität, besonders des 9.–12. Jahrhunderts. Fasz. 1–7. Wien 1994–2011.

Wenskus (1998): O. Wenskus, 26. Reflexionen zu fachsprachlichen Phänomenen in der Antike und Spätantike, in: L. Hoffmann u.a. (1998) (Hrsg.) Fachsprachen. Languages for Special Purposes. Ein internationales Handbuch zur Fachsprachenforschung und Terminologiewissenschaft. L. Hoffmann / H. Kalverkämper / H. E. Wiegand (Hgg.) In Verbindung mit Chr. Galinski, W. Hüllen. Bd. 1. Berlin [u.a.] 1998, S. 295–301.

Quellen

(Anmerkung zur Zitationskonventionen: Die gedruckte Edition wird nur bei denjenigen Autoren und Werken angeführt, aus denen Beispiele im Text zitiert werden. Andere erwähnte Autoren und Werke werden mit ihren Nummern im TLG im folgenden Format zitiert: „tlg:1234“ für den Autor und „tlg:1234.123“ für das Werk.)

Anonyma Metaphrasis ‚Alexiadis‘, lib. XI–XIII. Edition: H. Hunger, Anonyme Metaphrase zu Anna Komnene, Alexias XI–XIII [Wiener Byzantinistische Studien 15. Wien 1981], S. 31–151.

Cassius Dio, *Historiae Romanae* (Petri Patricii excerpta Vaticana sive Maiana). Edition: U.P. Boissevain, Cassii Dionis Cocceiani historiarum Romanarum quae supersunt, vol. 3. Berlin 1901 (repr. 1955), S. 731–749.

Constantinus VII Porphyrogenitus Imperator, *De insidiis*. Edition: C. de Boor, *Excerpta historica iussu imp. Constantini Porphyrogeniti confecta*, vol. 3: *excerpta de insidiis*, Berlin 1905.

– *De sententiis*. Edition: U.P. Boissevain, *Excerpta historica iussu imp. Constantini Porphyrogeniti confecta*, vol. 4: *excerpta de sententiis*, Berlin 1906.

Ephraem Aeniensis, *Historia Chronica*. Edition: O. Lampsides, *Ephraem Aenii Historia Chronica* [Corpus Fontium Historiae Byzantinae. Series Atheniensis 27. Athens 1990].

Ephraem Syrus, *De virtute, ad novitium monachum* (capita quattuor). Edition: K.G. Phrantzoles, Ὁσίου Ἐφραίμου τοῦ Σύρου ἔργα, vol. 2. Thessalonica 1989, S. 77–110.

Georgius Cedrenus, *Compendium historiarum*. Edition: I. Bekker, *Georgius Cedrenus Ioannis Scylitzae ope*, 2 vols. [Corpus scriptorum historiae Byzantinae. Bonn 1:1838; 2:1839], 1: S. 3–802; 2: S. 3–638.

Georgius Choeroboscus, *Prolegomena et scholia in Theodosii Alexandrini canones isagogicos de flexione nominum*. Edition: A. Hilgard, *Grammatici Graeci*, vol. 4.1. Leipzig 1894 (Repr. Hildesheim 1965), S. 103–417.

Georgius Metochites, *Historiae dogmaticae liber III*. Edition: J. Cozza-Luzi, *Novae Patrum Bibliothecae*, Tomus 10/1. Rom 1905, S. 319–370.

Georgius Monachus, *Chronicon* (lib. 1–4). Edition: C. de Boor, *Georgii monachi chronicon*, 2 vols. Leipzig 1904, 1: S. 1–382; 2: S. 383–804.

– *Chronicon breve* (lib. 1–6) (redactio recentior). Edition: J.-P. Migne, *Patrologiae cursus completus* (series Graeca) (MPG) 110, Paris 1857–1866, S. 41–1260.

Georgius Pachymeres, *Συγγραφικὰ ἱστορία* (libri vi de Michaelae Palaeologo). Edition: A. Failler and V. Laurent, *Georges Pachymérés. Relations historiques*, 2 vols. [Corpus Fontium Historiae Byzantinae. Series Parisiensis 24.1–2. Paris 1984], 1: S. 3–325; 2: S. 329–667.

Hippocrates et *Corpus Hippocraticum*, *Aphorismi*. Edition: Hippocrates, Bd. IV. *Heracleitus on the Universe*. With an English translation by W. H. S. Jones. London 1931 (repr. 1992), S. 97–221.

– *De affectionibus*. Edition: Hippocrates, Bd. V. With an English translation by P. Potter. London 1988, S. 1–91.

Ioannes Antiochenus, *Fragmenta ex historia chronica*. Edition: U. Roberto, *Ioannis Antiocheni fragmenta ex historia chronica* [Texte und Untersuchungen zur Geschichte der Altchristlichen Literatur 154. Berlin - New York 2005], S. 2–574.

Ioannes Zonaras, *Epitome historiarum* (lib. 13–18). Edition: T. Büttner-Wobst, *Ioannis Zonarae epitomae historiarum libri xviii*, vol. 3 [Corpus scriptorum historiae Byzantinae. Bonn 1897], S. 1–768.

Joel, *Chronographia compendiaria*. Edition: I. Bekker, *Ioelis chronographia compendiaria* [Corpus scriptorum historiae Byzantinae. Bonn 1836], S. 3–66.

Lucianus, *Podagra*. Edition: M.D. Macleod, *Lucian*, vol. 8. Cambridge 1967, S. 324–354.

Mazaris, *Peregrinatio Mazaris ad inferos*. Edition: Mazaris' *Journey to Hades: or, Interviews with dead men about certain officials of the imperial court* [Arethusa Monographs 5. Buffalo: Department of Classics, State University of New York, 1975], S. 2–99.

Michael Glycas, *Annales*. Edition: I. Bekker, *Michaelis Glycae annales* [Corpus scriptorum historiae Byzantinae. Bonn 1836], S. 3–457, S. 460–625.

Nicephorus I, *Breviarium historicum de rebus gestis post imperium Mauricii* (e cod. Vat. gr. 977). Edition: C. de Boor, *Nicephori archiepiscopi Constantinopolitani opuscula historica*. Leipzig 1880, S. 3–77.

Symeon Logothetes, *Chronicon*. Edition: St. Wahlgren, *Symeonis Magistri et Logothetae Chronicon*, Berlin [u.a.] 2006.

Theophanes Confessor, *Chronographia*. Edition: C. de Boor, *Theophanis chronographia*, vol. 1. Leipzig 1883 (repr. Hildesheim 1963), S. 3–503.

Weitere Ressourcen (zuletzt abgerufen am: 11.11.2015):

The R Project for Statistical Computing:

<https://www.r-project.org>

Thesaurus Linguae Graecae, A Digital Library of Greek Literature:

www.tlg.uci.edu/about/

<http://stephanus.tlg.uci.edu/lris/demo/stat.jsp>

Autorenkontakt³⁰

Dariya Rafiyenko

Universität Leipzig

Institut für Klassische Philologie und Komparatistik

Lehrstuhl Byzantinistik und griechische Philologie

Email: dariya.rafiyenko@uni-leipzig.de

URL: <http://rafiyenko.info/>

³⁰ Die Rechte für Inhalt, Texte, Graphiken und Abbildungen liegen, wenn nicht anders vermerkt, bei der Autorin.

Digital Texts and Diagrams: Representing the Transmission of Euclid's Elements

Christine Roughan

Abstract: The Digital Euclid project aims to publish an open, digital edition of every extant witness to the text and diagrams of Euclid's *Elements*. This paper discusses the required groundwork and is divided in two parts. It first covers a survey of the surviving manuscript and print sources for the *Elements* that intends to identify the extent of these materials, how many of these works have already been digitally imaged, and what challenges they pose to current data extraction methods. The latter part of the paper discusses the methods used to produce machine-actionable texts and diagrams and focuses especially on the development of tools for the identification and extraction of diagrammatic data.

Introduction

The *Elements* of Euclid is a text which has received continuous publication and use since it was first authored in the third century B.C.E. This geometrical text appears in hundreds of manuscripts; combined its manuscript and print editions number well over one thousand and span languages from across the globe.

The many and evolving forms that the *Elements* has taken throughout its lengthy transmission history have been a challenge to detail in their entirety. This is a text which has been well-studied: I. L. Heiberg for instance provides the current critical edition of the Greek text,¹ and editions even exist for specific translations of the *Elements* as well (consider H. L. L. Busard's critical editions of various medieval Latin translations).² The original Greek, after all, is only one form that the text has taken, and it cannot answer any questions about how Euclid was read and understood in — for example — the medieval Latin West. The sheer volume of material leaves the transmission of this text difficult to navigate and grasp in its entirety. In the case of just one translation within this transmission, Busard himself states that „131 manuscripts of Campanus' version of Euclid's *Elements* are known. Thus it was impossible to collate all of them.”³

When dealing with a large set of varied texts, print is not the ideal medium to convey it in its entirety. Navigating and analyzing such a dataset of the *Elements* in print would be slow and unwieldy, requiring thousands of pages. A second limitation of print appears in the case of the diagrams. Each proposition of the *Elements*, after all, is accompanied by a mathematical diagram, which itself contains essential information that cannot always be gleaned from the

¹ Heiberg (1883–6).

² Busard (1968), (1983), (1984), (1987), (1992), (1996), (2001), (2005).

³ Busard (2005), S. 46.

text alone.⁴ Variation appears during the transmission of these figures just as it does during the transmission of the text; however, they have received little scholarly attention. Until recently, in the many critical editions of Greek mathematical texts, there have been no *apparati critici* detailing the diagrams — Reviel Netz is perhaps the first to take steps towards such an apparatus by providing written explanations of variation between Archimedes manuscripts and occasional thumbnails showing alternate diagrams.⁵ This is an important start, but this approach does consume space rapidly, especially when attempting to provide alternate diagrams from a transmission as wide as that of the *Elements*.

The Digital Euclid project takes a digital approach to the transmission of the *Elements* in order to present for scholarly access and reuse the text and diagrams from these many witnesses. What would take thousands of pages in print and either lengthy or incomplete *apparati critici* can be published much more efficiently in an electronic format.

The project takes its cues from University of Leipzig’s Open Greek and Latin Project, which will publish at least one version of all extant Greek and Latin sources and which ultimately aims to represent every surviving version of these texts.⁶ The Digital Euclid project keeps in mind the latter goal and strives to represent every edition of the *Elements* in a form that is open, machine-actionable, and annotated. This paper discusses the initial work that is necessary for a project of this kind. It has two parts: firstly, a survey of the manuscript and print transmission of Euclid’s *Elements*; secondly, the testing of extant tools and the development of new ones to aid in digitization.

A survey on the transmission of the elements

The survey comprises various editions, translations, revisions, and recensions of the *Elements*, as well as its adaptations into school texts. Commentaries have not yet been included. For the most part, separate works which only quote the *Elements* are not included – the current exceptions to this are texts which preserve fragments of Boethius’s Latin translation.

When considering the extent of the survey, one must acknowledge that only a portion of material transmitted as Euclid’s *Elements* can be attributed to him (the apocryphal Books XIV and XV were once considered Euclidean, for example). However, the Digital Euclid project does include this material in its entirety, not solely that which scholars presently attribute to Euclid. The goal of doing so is to provide a fuller picture of how the *Elements* was read and understood throughout history. A side result of this is that the survey does include manuscripts and texts that contain only apocryphal material.

Multiple authors have produced bibliographies of Euclid’s works, including Pietro Riccardi (1887), Georges J. Kayas (1977), and Max Steck (1981). These bibliographies are of varying comprehensiveness; the Digital Euclid survey takes Riccardi’s bibliography as its starting point, which offers a very thorough listing of print editions of the *Elements* up through 1887. This date is also convenient for the purposes of the project, as material published before 1887 is public domain and is therefore available for scholarly reuse and republication. The list totals

⁴ Saito (2009), S. 817.

⁵ Netz (2004).

⁶ <http://www.dh.uni-leipzig.de/wo/projects/open-greek-and-latin-project/>.

over a thousand editions. Riccardi also describes more than 180 Euclidean manuscripts, but notes that this is an incomplete list. The project therefore turned to manuscript lists provided by other scholars (such as the formerly mentioned Heiberg and Busard, as well as Folkerts and Lo Bello, etc.)⁷ as well as manuscript catalogs.

The goals of the survey were threefold. For each extant version, the survey would:

- 1) record the relevant bibliographic data and assign identifiers both to the physical codex or book and the abstract work
- 2) note the current status of digitization, current copyright on the book itself, and – where applicable – current copyright on digital images of the version in question
- 3) record factors that could impact automated text and diagram extraction workflows (page layouts, languages or fonts, and quality of the page images)

The survey covered 415 scanned printed editions and 477 manuscripts, spanning fourteen languages and the first through the nineteenth century C.E.

Manuscript editions in the survey

The survey presently contains 477 codices, papyri, and fragments, which appear in ten languages and date from the first through the eighteenth century C.E. These materials preserve 522 distinct versions and translations of the *Elements* text, 84 of which are Greek, 51 of which are Arabic, and 344 of which are Latin. The Latin portion of the survey is currently the most complete, followed by the Greek. Other languages include French and Middle French, Hebrew, Italian, Modern Greek, Persian, and Turkish.

	Full Coverage (Absolute)	Full Coverage (Percentage)	Partial Coverage	Enunciations Only
Book I	226	63.8%	51	13
Book II	226	63.8%	19	13
Book III	221	62.4%	15	11
Book IV	212	59.9%	12	8
Book V	213	60.2%	15	8
Book VI	200	56.5%	15	8
Book VII	186	52.5%	20	7
Book VIII	182	51.4%	14	6
Book IX	174	49.1%	14	6
Book X	167	47.2%	29	8
Book XI	169	47.7%	15	8
Book XII	166	46.9%	7	8
Book XIII	161	45.5%	7	8
Book XIV	142	40.1%	6	6
Book XV	136	38.4%	12	6

The coverage of the *Elements* is known for 354 of these texts; perhaps unsurprisingly, the books that receive the most coverage are Books I and II (the tendency of geometrical manuscripts to provide Book I's definitions results in Book I also receiving the most partial coverage). Figure 1 provides an illustration of coverage across books. With only two exceptions, later books appear less often than the ones preceding them. Nevertheless, between the ninth and the eighteenth centuries C.E. even Book XV appears in five languages and over 130 manuscripts — more than enough examples of a lengthy and far-ranging transmission.

Abb. 1: Coverage of the *Elements* in 354 manuscript versions of the text.

⁷ Folkerts (1989), Lo Bello (2003).

As of July 2015, 112 of these manuscript texts have been at least partially imaged and made available online; 82 have been imaged in full. Of the texts with complete digital publication, 59 are available either as public domain images or under Creative Commons licenses that would allow for scholarly reuse. These 59 manuscript texts still provide a wide selection of *Elements* editions and translations: they represent examples from the first to the seventeenth century C.E. and span six languages (although the overwhelming majority are Latin). Most of these digital manuscripts are provided by the Münchener Digitalisierungszentrum and Gallica.⁸ A little over half of these are presented in greyscale, but the remainder are full-color images. Nearly all of the manuscripts provided by Gallica are digitized microfilm copies.

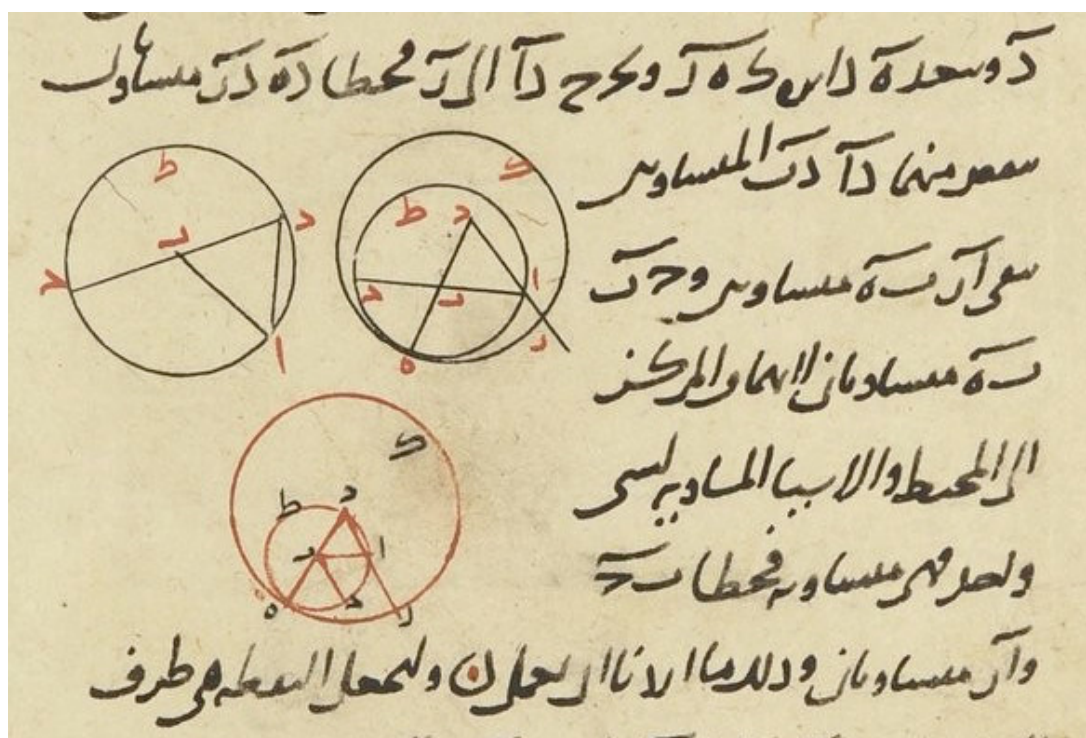


Abb. 2: A selection of folio 145v of BnF Arabe 2484, showing diagrams relevant to an Arabic version of Book I, proposition 2. Source: gallica.bnf.fr.

In surveying these manuscripts, the Digital Euclid project determined some categories in which to organize the mathematical diagrams. The project encountered three types of layouts: inline, where the diagrams appear within the margins of the main text; marginal, where they remain outside these borders; and a combination, where they appear in both spaces throughout the work. In this initial survey, the majority of the manuscripts viewed contain diagrams in both locations. Inline diagrams appear second most often. Where full-color texts were available, the Digital Euclid project also recorded when differently colored inks were used between texts and diagrams with the hope that this distinction might aid in automated attempts to locate the diagrams. The overwhelming majority uses ink that is the same color, but not all: figure 2 offers one example. The diagrams are drawn with both red and brown ink, while text is in a brown ink. The diagram layout illustrated here is inline.

⁸ MDZ: <http://www.digitale-sammlungen.de/>. Gallica: <http://gallica.bnf.fr/>.

Print editions in the survey

Riccardi lists over one thousand publications of the *Elements* between 1482 and 1887. As of July 2015, the project’s survey has covered 614 unique examples of the various editions, or 415 unique editions. Several of these editions either were multivolume works or contained multiple translations, so the total number of abstract works surveyed comes to 429. All of these have been imaged and made available online: the project’s initial focus was on works that can be located in one of four databases: Google Books (contains 45.4% of the 614 versions), the Internet Archive (38.7%), HathiTrust (30.0%), and SLUB Dresden (13.4%).⁹ The works contained in these databases provide a wide selection already: they date from 1482 to 1908 and span sixteen languages.

Figure 3 outlines the distribution of works across languages currently in the survey. The counts consider not physical books, but abstract versions of a text (and so treat a text published across multiple volumes as one). While Euclid has certainly been published far more often in English than in Modern Greek, for example, the lack of representation in the survey for certain languages is partially caused by the focus so far on scanned copies present in the four databases mentioned above.

	Number of works		Number of works
English	175	Arabic	2
Latin	112	Sanskrit	2
Italian	38	Chinese	1
French	29	Russian	1
Ancient Greek	28	Polish	1
German	25	Modern Greek	1
Spanish	6	Danish	1
Dutch	6	Hungarian	1

Abb. 3: Languages represented in the Digital Euclid survey of print editions as of July 2015.

	Full Coverage (Absolute)	Full Coverage (Percentage)	Partial Coverage
Book I	362	85.2%	21
Book II	350	82.4%	21
Book III	343	80.7%	2
Book IV	336	79.1%	2
Book V	329	77.4%	19
Book VI	323	76.0%	19
Book VII	85	20.0%	15
Book VIII	88	20.7%	15
Book IX	84	19.8%	15
Book X	85	20.0%	16
Book XI	197	46.4%	33
Book XII	189	44.5%	29
Book XIII	87	20.5%	15
Book XIV	65	15.3%	12
Book XV	62	14.6%	12

Abb. 4: Coverage of the Elements in 425 print versions of the text.

In comparison to the manuscript situation, a far more dramatic preference for certain books crystallizes during the print transmission. Again, Book I takes the lead, appearing in 85.2% of the surveyed texts. More notably, Books I–VI all appear in more than 75% of these texts, Books XI and XII in around 45%, and the rest in approximately 20% or less. The first six books were published frequently, often with the eleventh and twelfth attached. The survey reveals that the prevalence of this combination varies across languages: about two thirds of the surveyed English texts comprise Books I–VI, sometimes with Books XI and/or XII added. The same is true for less than half of the Latin texts and less than a fifth of the Greek ones.

⁹ Google Books: <https://books.google.com/>; Internet Archive: <https://archive.org/>; HathiTrust: <https://www.hathitrust.org/>; SLUB Dresden: <http://www.slub-dresden.de/startseite/>.

As noted, all of the print editions included thus far in the survey have already been made available as digital scans, and all surveyed materials from the Internet Archive, Google Books, HathiTrust, and Dresden SLUB are public domain. The majority of the scans are provided already binarized: 73.9% in Google Books, 87.9% in HathiTrust, and 77.3% in the Internet Archive (Dresden SLUB provides full-color scans).

During the survey, four potential manuscript layouts were recorded. Like manuscript diagrams, print diagrams might have inline or marginal layouts. A third option is provided by those texts that consistently locate the diagram at the top of the page. Lastly, many print editions did not locate the diagrams beside their proposition, but rather placed diagrams together on pages that folded out, usually from the back of the book. The layout represented most often in the survey is inline, appearing over 70% of the time. Foldout layouts are the second most common and are used in about 15% of the editions. Very few editions use marginal or top-of-page layouts.

The Digital Euclid project found one major issue in the scanning process for texts with foldout diagrams: across 138 scans of books with that layout, only 6 of them actually imaged the diagrams. This is a serious omission. While the overwhelming majority of printed diagrams are imaged successfully because they appear inline with the text, the failure to image diagram foldout pages means that digital methods currently cannot be used to analyze an entire tradition of print diagrams.

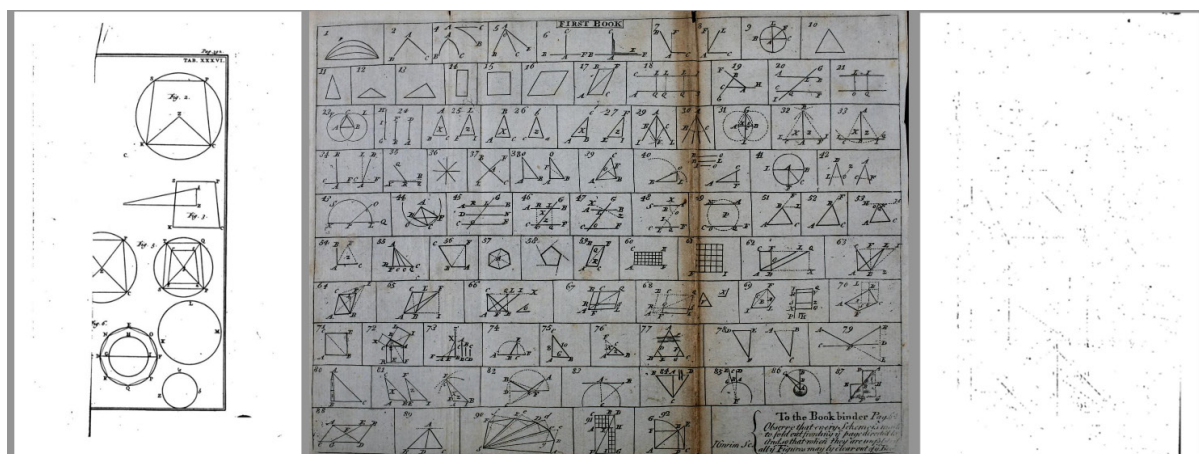


Abb. 5: Three examples of scanned diagram foldouts. The leftmost page was partially imaged, but not unfolded. The middle is an example of an unfolded page. The rightmost page remained folded, and nothing of the diagrams was captured. Source for all three images: the Internet Archive.

Final remarks on the survey

Work with the survey, even in its initial form, indicates that it is a useful dataset: Digital Euclid has made great use of it during development of appropriate workflows for data extraction across the *Elements*. As this survey approaches completion, it is hoped that it can be of continuing use. With this goal in mind, a portion of the survey will be published as part of the Perseus Catalog, a digital catalog that unites classical bibliographies and metadata.¹⁰ The full dataset of the survey is located in the Github repository for the Digital Euclid project.¹¹

¹⁰ <http://catalog.perseus.org/>.

¹¹ <https://github.com/cmroughan/digital-euclid>.

The survey already shows comparative potential, as demonstrated by the quick analyses on each book above. As more texts are transcribed and the details of their content are described by CTS URNs (which are discussed below), it will become possible to investigate questions like these with far deeper levels of granularity.

This year's usage of the survey has also suggested several possibilities for expansion. Continuing along the coverage example, it would be beneficial to note where lacunae are caused by a physical absence of material (and therefore it is unknown whether or not the absent portion of the text was originally included or not) and where the absent text is recognized to have never been included in the first place. Company is also a feature that would be illuminating to trace through the history of a text — how often does the *Elements* appear alongside another particular text? How is this affected across time, languages, and the transition from manuscript to print? In a similar vein, when do scholia or other scholarly commentaries accompany the text? Another dimension could eventually be added to the survey by noting scholarship on the editions; for example, to whom is a manuscript attributed? Has this answer changed over time? In the future, as the survey is completed for versions of the *Elements*, it will come to comprise further works. The Digital Euclid project will include commentaries on the *Elements* in later versions, and other Euclidean texts are likely eventual candidates as well.

Data Extraction Workflow in the Digital Euclid project

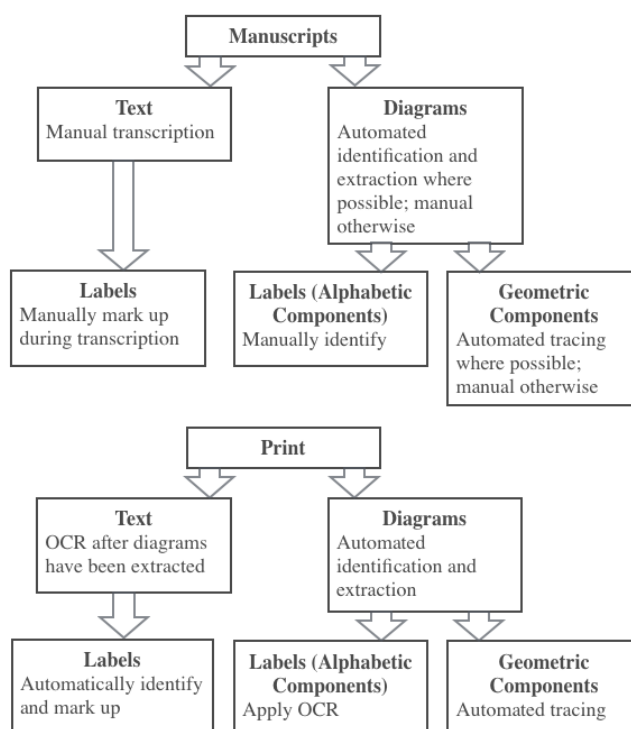


Abb. 6: Workflows for extraction of textual and diagrammatic data from manuscript and print editions.

Like the Open Greek and Latin Project, Digital Euclid has as its end goal the publication of digital, annotated EpiDoc- and CTS-compliant XML texts. The Digital Euclid project additionally notes the importance of the mathematical diagrams, and so aims to produce digital, annotated SVG traces of the mathematical figure.

With a multitude of print editions and manuscripts revealed by the survey to have been digitally imaged already, Digital Euclid began to develop workflows and to test tools for the extraction of both textual and diagrammatic information. The second part of this paper will discuss these efforts.

The different challenges that manuscripts and print editions pose necessitate two separate approaches. This is already accepted in the case of texts: optical character recognition (OCR) for manuscripts still has a significant way to go. Manuscript diagrams also pose slightly different challenges than print ones, as will be discussed below. Figure 6 outlines the two workflows.

Like the survey of *Elements* editions, the resultant transcriptions and traces are being published in the project's Github repository.

Regarding the Text of the *Elements*

The Digital Euclid project will produce texts that are machine-actionable and CTS- and EpiDoc-compliant, following the lead of Open Greek and Latin. CTS URNs serve as a means of identifying texts and selections of those texts unambiguously; one advantage this provides is that CTS texts are automatically aligned.¹² EpiDoc is a subset of the Text Encoding Initiative's guidelines that is especially suitable for primary sources.¹³

Two tags used by the Digital Euclid project in the markup of the text are worthy of note, since they highlight information particular to mathematical texts. Firstly, because the labels function as identifiers linking text and diagram, these receive referencing string tags. A label or cluster of labels is tagged as `<rs type="labels"></rs>`, making it easy to point a computer either towards a structure that is identified with multiple labels or towards the individual point or shape that corresponds to a single label (once these are automatically generated in the diagram portion of the workflow, that is.). By looking between a list of labels or label groupings in the text and labelled points or objects in the diagram, a computer can automatically create explicit links between the text and the diagram. It must be recognized that this automated method cannot work where there are disagreements between text and diagram (which might occur due to an error within one or the other). The second element used in the Digital Euclid project is the `<figure/>` tag. This marks the existence of a diagram in the text that accompanies the section in question, even if the diagram is not physically near that section, as would be the case for foldout diagrams. Figure tags reference the unique identifier assigned by the project for the diagram in question.

The question of how to extract text from page images has already received a great deal of work, and the current answer — optical character recognition (OCR) — is generally sufficient for the many printed texts of the *Elements*. Certain editions do contain mathematical notation (and these are noted in the survey) and so require OCR that can handle those characters. Such software does exist — a math detection module has been developed for Tesseract OCR, for example.¹⁴ Manuscript OCR is an area of ongoing research, but for now, manual transcription (usually accomplished through large citizen science projects) has proved to be successful, if slower.

¹² <http://cite-architecture.github.io/ctsrn/overview/>.

¹³ <https://sourceforge.net/p/epidoc/wiki/Home/>.

¹⁴ <https://github.com/tesseract-ocr>.

Markup can also be accomplished automatically or manually, depending on the text in question. For manuscript editions, the labels would be tagged as such during the manual transcription process. For printed texts, labels are usually distinct enough to be automatically tagged, although this process does require review. In most cases, a group of capitalized letters that are separated from the others by periods or spaces can be identified as a group of labels, especially when they do not spell out a word in the language of the text. Figure elements can be added to the text's XML once diagrams and text have been aligned.

Regarding the Diagrams of the *Elements*

The major challenge to automated data extraction in the case of Euclid's *Elements* comes in the form of the mathematical diagrams. These are worthy of study in their own right: they might contain information not in the text, reveal insights into their function in ancient and medieval times, or offer their own clues regarding the process of manuscript creation and transmission.

The data, the diagrams themselves contain thus makes it desirable to produce digital and annotated traces of diagrams similar to the digital and annotated transcriptions of text. Since OCR is able to rapidly obtain textual data from page images, the lack of similar tools to handle diagrams is a major speed bump in the digitization and extraction process for a complete edition. Furthermore, the diagrams are best located before any OCR of the text takes place: the alphabetic characters¹⁵ — or simply geometric components that a computer might mistake for a character — can be picked up during OCR and introduce errors into the resultant output, as seen in Figure 7. It is therefore worthwhile to seek automated means of identifying diagrammatic regions on a page or folio, both to remove non-textual data before later OCR workflows and to link images of the diagram with the appropriate proposition.

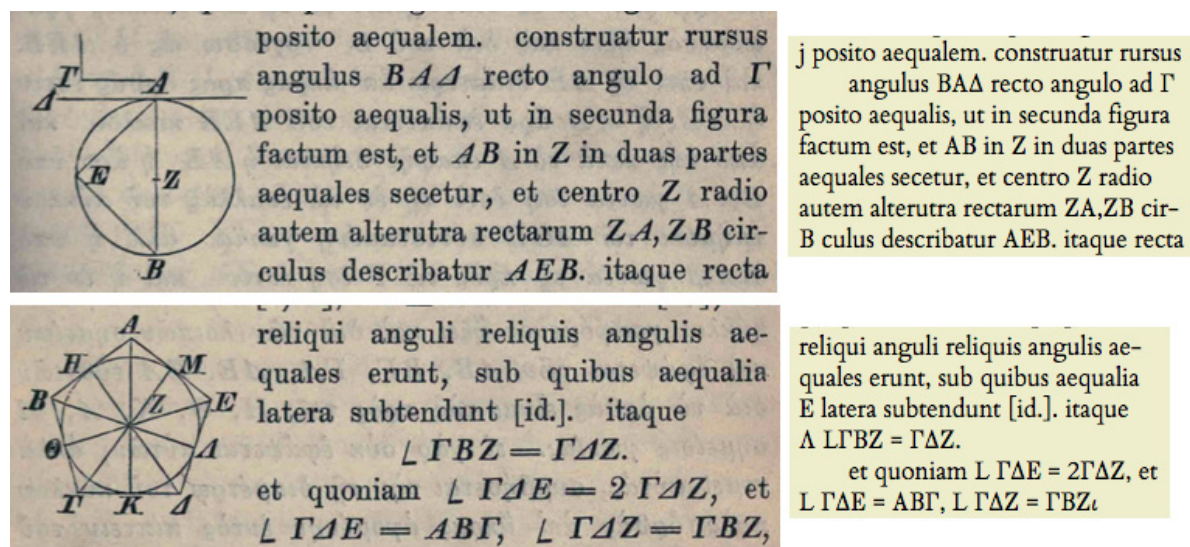


Abb. 7: Screenshots of OCR output from the Lace Greek OCR project website. In the output for the first example, the 'j' at the start of the first line and 'B' in the last were introduced from the diagram. The second includes 'A' at the start of the fourth line. Source: heml.mta.ca/lace.

¹⁵ The mathematical diagrams in the *Elements*, like many Ancient Greek mathematical figures, contain both alphabetic and geometric content. Most, though not all, consist of an arrangement of shapes and the alphabetic labels that identify them.

Several layout analysis tools were tested to determine whether extant software would be capable of distinguishing diagrams from text. Digital Euclid first investigated the open source document analysis system OCRopus and the layout analysis component of the Tesseract OCR system.¹⁶ Neither were successful in recognizing diagrams.

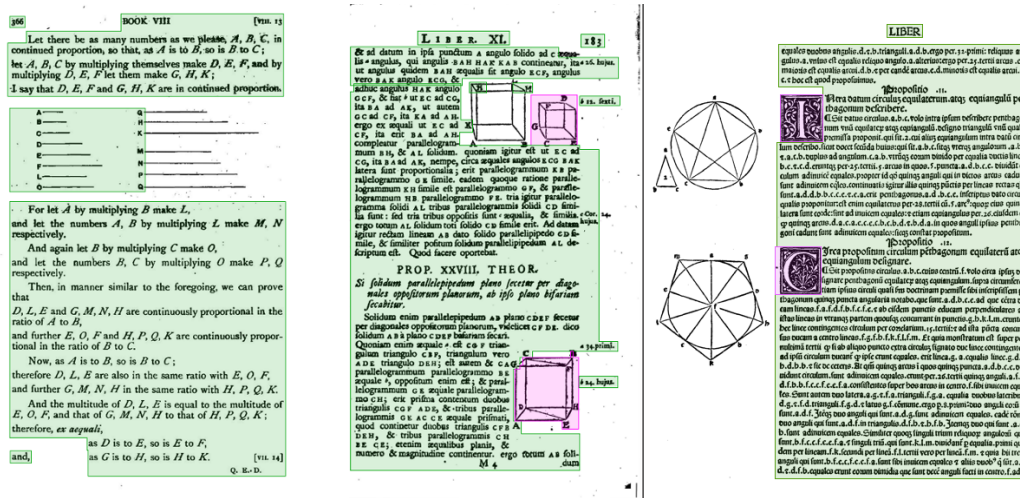


Abb. 8: Three sample pages with incorrect layout analysis by ABBYY FineReader. Source for all three original images: the Internet Archive.

The project then tested out ABBYY FineReader¹⁷ and found that while the program did have some success, there was considerable room for improvement. The software was tested on 93 diagrams, in which were represented all varieties of diagram layout except for foldout, and it properly identified 18 of them (a further 11 were poorly or incompletely captured).

While this initial sample was small, the output ABBYY FineReader provided from it was illuminating. As the rightmost page in figure 8 illustrates, the program often returned false positives when drop caps were present, and this is to be expected: the program was not trained to distinguish between different kinds of ‘illustrations’. Additionally, in the case of line diagrams like the ones seen on the leftmost page of figure 8, only 7.4% were recognized. ABBYY FineReader was inconsistent in how it handled the diagrams’ labels and showed no preference towards handling them as text, handling them as part of an illustration, or skipping them entirely.

Identification and Extraction of Diagram Data

Since extant tools proved inapt to handle the mathematical diagrams, new ones were necessary. The Digital Euclid project therefore used the Gamera framework to build tools that would recognize diagrams in the *Elements*.¹⁸ Automated diagram recognition posed a different challenge than automated character recognition. Texts are composed of a limited set of characters, in many languages separated from each other by whitespace — it therefore makes sense to train a computer to locate connected components and recognize the individual characters. Diagrams can contain both geometric and alphabetic content and can comprise multiple elements

16 OCRopus: <https://github.com/tmbdev/ocropy>, Tesseract: <https://github.com/tesseract-ocr>.

17 <http://www.abbyy.com/finereader/>.

18 <http://gamera.informatik.hsnr.de/>.

separated by whitespace; they are more akin to a larger structure of text such as a word or a line rather than a character. Furthermore, the connected components that form the geometric portion of the diagram vary dramatically, with many being unique.

The constructed tools currently used by the project work best when input is provided regarding the page layout of the diagrams. Diagrams that are consistently located in the margins, for example, or at the top of the page can be identified easily. Some editions have diagrams consistently located on the outer edge of the page, which again simplifies the search. Editions with inline or mixed-layout diagrams are more complicated to handle.

The Digital Euclid project has also found success by distinguishing between two types of connected components, here termed 'line' and 'quad'. The project considers line components to be those portions of a diagram whose bounding box fits the profile of a line, exceeding certain aspect ratio thresholds. Quad components are bounded by boxes with less extreme aspect ratios and which exceed certain area thresholds. In figure 8, the diagram on the leftmost page is composed of line components, while the rest consist of quad components.

The first challenge is how to identify lines reliably when 1) line length varies, so there are no consistent dimensions to look for and 2) line cross-sections (since the printed line is not truly one-dimensional), are not consistent across different editions. Digital Euclid first identifies 'definite lines'¹⁹ as connected components that surpass one of two extreme aspect ratios corresponding to horizontal or vertical orientations. Meanwhile, 'possible lines' are identified as connected components that surpass more moderate aspect ratios. This first pass records all possibilities, as well as data on the definite lines. The second pass learns from the first and considers various factors (heights of definite horizontal lines, widths of definite vertical ones, and nearest neighbors for each possibility) to remove false positives from the initial results.

There are some remaining issues with lines. Accidental intersection between lines and labels occurs often, causing the bounding boxes to not match the profile for a line. Future work will test projection analysis as a means of identifying these problematic components. Dotted and other non-continuous lines are also an issue. The segments of broken lines still usually result in matches, but dotted lines must be handled manually.

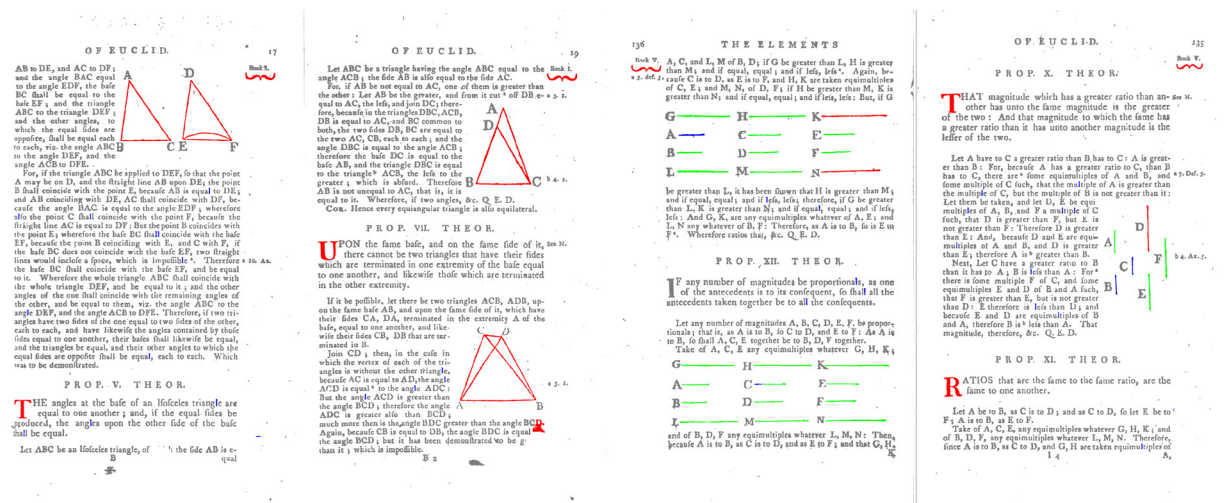


Abb. 9: Four pages from Simson's edition put through the first phase of the identification process. Potential diagram components are colored for manual review. Red: quad component; green: definite line component; blue: possible line component. Source for all four images: the Internet Archive.

¹⁹ Definite lines are not necessarily components of diagrams: the three non-diagram examples that are often picked up are 1) page edges, 2) characters such as the letter 'l', the number '1', or the symbol '=', and 3) lines used elsewhere as part of the page layout (such as separation of main text and critical apparatus).

Meanwhile, the added dimension causes quad components to have much more variety in comparison to line components. Where lines can generally be distinguished through the aspect ratios of their bounding boxes, quad components can be distinguished primarily through the area of theirs. With the exception of dots in dotted lines, these components are always larger than the regular characters that make up the text. Digital Euclid's initial approach is therefore as follows: remove noise from the page, then analyze the connected components. Assuming the page contains text, the majority of these will be textual characters. Any connected components that have areas significantly larger than the text are potentially full or partial diagrams.

This method results in its own share of false positives, for example large drop caps or character clusters that are not properly separated by whitespace. These can be cut down through knowledge of the layout, analysis of nearest neighbors, or determining the percentage of black pixels within the bounding box. In the future, the Digital Euclid project plans to test the use of distance transforms to improve this automated correction step. While manual review is still necessary to handle the false positives, these methods do shorten the time it takes to identify diagrams across works.

The method works best in print editions. Within manuscripts, ligatures, abbreviations, and scripts where the characters are not separated by whitespace result in connected components that can be larger than geometric components. While this alone isn't necessarily an insurmountable barrier, intersection between the diagrams and either the *Elements* text or scholia text is also frequent. Further work is needed to develop tools specifically for manuscript diagrams. While more time-consuming, identifying manuscript diagrams manually is still a reasonable option today.

Once the diagrams have been identified, they can be separated from the text, allowing for diagram images to continue down their path in the workflow and for OCR to be applied to the remaining page. The approach discussed above finds the geometric components, but the separate alphabetic labels must be captured as well, both to create complete traces and to prevent them from interfering with the OCR of the text. Digital Euclid locates these by considering all material within the bounding boxes of the geometric components part of the diagram, and by expanding the bounding boxes by a certain input amount in order to locate outer labels.

Producing Machine-actionable Diagram Data

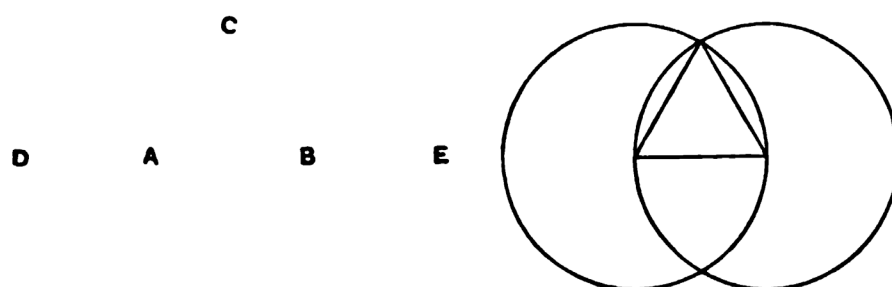


Abb. 10: Diagram for Elements Book I proposition 1, separated into two images containing either alphabetic or geometric components. Source: the Internet Archive.

Once diagram selections have been extracted, they can be split into two images that will go through separate workflows – figure 10 provides an example. The prior identification process already found most of the geometric material, so it is not complicated to separate this from the labels.

Currently, this step cannot be completed solely through automated means because of the high tendency of labels to intersect with the geometric portions. How often this occurs varies across editions: in Heath’s translation, only 25 of the 494 diagrams have this issue. In some cases this intersection is accidental, the result of poor binarization for instance, but in others it is simply how the diagram was produced (an example can be seen in figure 11).

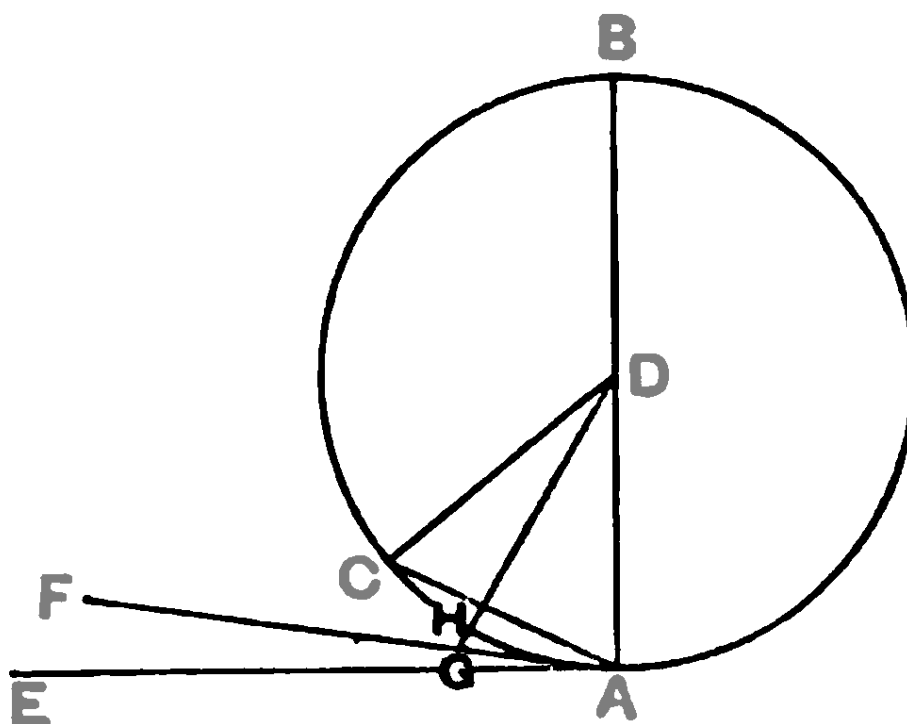


Abb. 11: Diagram from Heath’s English edition. Connected components that are identified as labels appear in grey. This diagram contains two labels, H and G, that overlap with the figure and that must be removed manually. Source: the Internet Archive.

OCR can be used to transcribe the labels. Although most out-of-the-box systems will return errors on these images because they expect lines, words, or at least a significant number of characters, Gamera for instance can be used to train a basic OCR system that will function on images with scattered characters like figure 10. The OCR process also provides a bounding box defining the region of interest for each label, which will be useful in later steps when traced diagrams are annotated according to their labels.

As with the main text, OCR has difficulty dealing with manuscript labels, and these are currently handled manually.

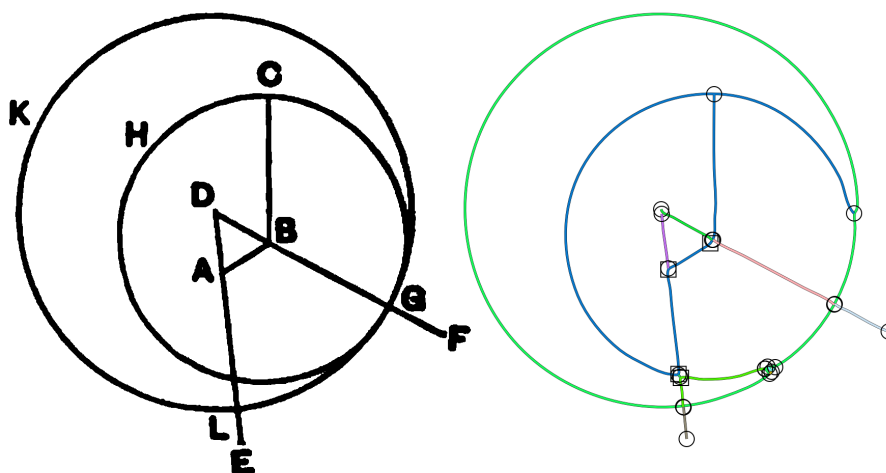


Abb. 12: Left: diagram for Elements Book I proposition 2 from Heath's edition. Right: a segmented autotraced SVG version with endpoints and corners located and marked for manual review. Source for original image: the Internet Archive.

Tracing the geometric portion requires a different approach. Autotrace, a command line utility, can produce a rough SVG version of a raster file.²⁰ The resultant output is not precise enough to serve as the final digital trace, but does contain enough data to more easily locate certain points of interest, namely corners, endpoints, and intersections. By using these points and referring back to the original raster image, the Digital Euclid project hopes to produce cleaner SVG traces. Figure 12 shows a diagram that has been analyzed in this way, and which has also been segmented into the smallest elements of the overall shape.

Additionally, traces can be produced manually using vector-editing software. For this the Digital Euclid project uses Inkscape.²¹

Recombining the data

Lastly, the workflow turns to the recombination of the now-extracted data. Within diagrams, the transcribed labels are reunited with the SVG traces. This is accomplished by analyzing the distance between each label and each point of interest to find pairs that are in closest proximity to each other. In the Elements of Euclid, labels are most often located at these points. For labels that are not within the expected distance to one of these points, the method checks for the nearest arbitrary point on the diagram and assigns the label to that point.

Assigning labels to points allows for the automatic creation of identifiers for each segment of the diagram. Sections that contain labelled points can be identified as 'line.AB', for example, or 'arc.ΓΔ'. These identifiers are added to the diagram's SVG file.

It is also necessary to unite the diagram with the proposition it accompanies. Except in the case of fold-out diagrams, this can be accomplished by comparing the regions of interest for

²⁰ <http://autotrace.sourceforge.net/>.

²¹ <https://inkscape.org/en/>.

the diagrams and propositions. Whether appearing at the beginning, middle, or end of a proposition, the diagram is generally located in close vicinity. The Digital Euclid project preserves bounding box information obtained during the OCR workflow for the main text, so once the text is structured in CTS, the region of interest for the propositions can be determined. Even in a manual transcription or tracing process, the region of interest for that text or diagram is recorded and can be used at this stage.

Conclusion

This initial phase of the Digital Euclid project investigated the potential for a digital representation of the transmission of Euclid's *Elements*, and this goal is certainly feasible. Institutional and large-scale digitization efforts have produced high quality images of hundreds of editions. This is important: further steps along the digitization process are dependent upon the first step of imaging the texts. While print editions currently dominate the digitized set, more and more manuscripts are being made available for scholarly use.

The project used a combination of extant tools and experimental new ones to test out identifying and extracting the textual and diagrammatic data. The text of the *Elements* can for the most part be approached through typical OCR workflows. Diagrams were the major challenge in the groundwork for the Digital Euclid project: especially in comparison to the text, they were the speed bump to the extraction process. However, this past year with Digital Euclid has demonstrated that this was the case primarily due to lack of work on the problem. Ancient mathematical diagrams are not unapproachable by automated means.

When looking forward, plenty of work remains to be done. The survey is not yet comprehensive: the project will continue to add to this and publish updates to the Digital Euclid Github repository. Similarly, the identification and extraction tools and methods that were newly developed this year will be improved before the project begins to use them in earnest. Ultimately these approaches will allow the Digital Euclid project to publish numerous machine-actionable texts, diagrams, and datasets from and regarding the *Elements* that can serve as a flexible resource for further scholarship.

Literatur:

- Busard (1968): H. L. L. Busard, *The Translation of the Elements of Euclid from Arabic into Latin by Hermann of Carinthia(?)*, Leiden 1968.
- Busard (1983): H. L. L. Busard, *The First Latin Translation of Euclid's Elements commonly ascribed to Adelard of Bath. Books I–VIII and Books X.36–XV.2*, Toronto 1983.
- Busard (1984): H. L. L. Busard, *The Latin Translation of the Arabic Version of Euclid's Elements commonly ascribed to Gerard of Cremona*, Leiden 1984.
- Busard (1987): H. L. L. Busard, *The Mediaeval Latin translation of Euclid's Elements Made Directly from the Greek*, Stuttgart 1987.
- Busard (1992): H. L. L. Busard, *Robert of Chester's Redaction of Euclid's Elements, the so-called Adelard II Version*, Birkhäuser 1992.
- Busard (1996): H. L. L. Busard, *A Thirteenth-Century Adaption of Robert of Chester's Version of Euclid's Elements*, München 1996.
- Busard (2001): H. L. L. Busard, *Johannes de Tinemue's redaction of Euclid's „Elements“, the so-called Adelard III version*, Stuttgart 2001.
- Busard (2005): H. L. L. Busard, *Campanus of Novara and Euclid's Elements*, Stuttgart 2005.
- Folkerts (1989): M. Folkerts, *Euclid in Medieval Europe*, The Benjamin Catalogue 1989.
- Heiberg (1883–6): I. L. Heiberg, *„Euclidis Elementa“, Euclidis Opera Omnia Vol.1–4*, Leipzig 1883–6.
- Kayas (1977): G. J. Kayas, *Vingt-trois siecles de tradition euclidienne: essai bibliographique*, Palaiseau 1977.
- Lo Bello (2003): A. Lo Bello, *Gerard of Cremona's Translation of the Commentary of Al-Nayrizi on Book I of Euclid's Elements of Geometry*, Leiden 2003.
- Netz (2004): R. Netz, *The Works of Archimedes: Translation and Commentary*, Cambridge 2004.
- Riccardi (1887): P. Riccardi, *Saggio di una bibliografia euclidea*, Tipografia Gamberini e Parmeggiani 1887.
- Saito (2009): K. Saito, *„Reading ancient Greek mathematics“, The Oxford Handbook of the History of Mathematics*, Oxford 2009.
- Steck (1981): M. Steck, *Bibliographia Euclidean: Die Geisteslinien der Tradition in den Editionen der 'Elemente' des Euklid um 365–300*, Arbor scientiarum, Hildesheim 1981.

Weitere Ressourcen (Zuletzt aufgerufen am 29.12.2015):

ABBYY FineReader:

<http://www.abbyy.com/finereader/>

Autotrace:

<http://autotrace.sourceforge.net/>

Inkscape:

<https://inkscape.org/en/>

Gallica:

<http://gallica.bnf.fr/>

Google Books:

<https://books.google.com/>

HathiTrust:

<https://www.hathitrust.org/>

Internet Archive:

<https://archive.org/>

Münchener Digitalisierungszentrum (MDZ):

<http://www.digitale-sammlungen.de/>

OCROPUS:

<https://github.com/tmbdev/ocropy>

Open Greek and Latin Project of the Open Philology Project:

<http://www.dh.uni-leipzig.de/wo/projects/open-greek-and-latin-project/>

Perseus:

<http://catalog.perseus.org/>

SLUB Dresden:

<http://www.slub-dresden.de/startseite/>

Sourceforge (Find, Create, and Publish Open Source Software):

epidoc.sourceforge.net/

Tesseract:

<https://github.com/tesseract-ocr>

<https://code.google.com/p/tesseract-ocr/>

The Gamera Project:

<http://gamera.informatik.hsr.de/>

The CITE Architecture:

<http://cite-architecture.github.io/cturn/overview/>

Autorenkontakt²²

Christine Roughan

Universität Leipzig

Lehrstuhl für Digital Humanities

Email: cmroughan@gmail.com

²² Die Rechte für Inhalt, Texte, Graphiken und Abbildungen liegen, wenn nicht anders vermerkt, bei den Autoren.

Digitale Editionen für die Altertumswissenschaften: Eine Auseinandersetzung mit P. Sahles Kritik kritischer Texteditionen

Friedrich Meins

Abstract: In his extensive study „Digitale Editionsformen“, Patrick Sahle gives a broad survey of the history of text-editions in general and the new possibilities as provided by digital technology. In the course of this, he passes harsh criticism on the „Critical edition“. In his eyes the „Critical edition“ is an outdated means, whose ontological and hence scientific limitations are mainly a result of the technological limitations and ideological preconditions of the times when it came into being.

This short paper argues that at least in the case of the „Critical edition“ as it is still the basis of Classical studies both historical and linguistical, its own significance as an outcome of research cannot simply be depreciated against a general claim for „editions“ in a mere documentary meaning.

Die Möglichkeiten digitaler Medien sind auch an den Diskussionen über das Wesen klassisch-philologischer Texteditionen nicht spurlos vorübergegangen. Die Option, scheinbar beliebige Datenmengen speichern und verfügbar machen zu können, hat dazu geführt, dass Textkritik, verstanden als die Etablierung eines verbindlichen Textes nach unterschiedlichen Kriterien, mitunter als ein Relikt vergangener Zeiten, ja als das Resultat der Notwendigkeit, mit beschränkten materiellen Ressourcen auskommen zu müssen, dargestellt wird.

Im Zuge dessen sind zahlreiche Kritikpunkte an der klassisch-philologischen Edition, aber auch an anderen Formen der Textedition geäußert worden. Diese Kritikpunkte sind im Einzelnen oftmals keineswegs neu. Allerdings ist in einigen Publikationen der Anspruch vertreten worden, dass die neuen Möglichkeiten des digitalen Edierens in der Lage seien, den angemahnten Problemen Abhilfe zu schaffen. Insbesondere in Patrick Sahles umfangreicher Dissertation über „Digitale Editionsformen“ sind zahlreiche, mitunter recht verschiedenartige und auch auf verschiedene Formen kritischer Textausgaben gemünzte Kritikpunkte zusammengetragen worden, mit dem Ziel, für neue Formen der Textrepräsentation zu werben.¹

Dieser kurze Aufsatz soll sich konkret mit der Frage beschäftigen, inwiefern sich den insbesondere in der Monographie Sahles gesammelten Kritikpunkten, die sich als Hinweise auf allgemeine erkenntnistheoretische Probleme der klassisch-philologischen Edition betrachten lassen, durch digitale Editionsformen tatsächlich wirkungsvoll begegnen lässt, und inwiefern solche Editionsformen als eine Ergänzung oder gar als ein Ersatz der klassisch-philologischen Edition angesehen werden können.

Dazu ist einleitend festzustellen, dass die Kritik an der klassischen Edition bei Sahle nur einen geringen Teil seines enorm umfangreichen Werkes einnimmt, die hier vorgenommene Kritik der Kritik also notwendigerweise ein Zerrbild entstehen lässt. Allerdings fordert Sahle

¹ Sahle (2013): P. Sahle, Digitale Editionsformen. Zum Umgang mit der Überlieferung unter den Bedingungen des Medienwandels, Norderstedt 2013, 3 Bde. (Schriften des Instituts für Dokumentologie und Editorik Bd. 7).

selbst zu einer solchen selektiven Nutzung seines Werkes auf. Meine Perspektive, die hier zu Wort kommt, ist dezidiert diejenige eines Althistorikers, für den die kritischen klassisch-philologischen Editionen bislang ein unverzichtbares Hilfsmittel darstellen.²

Patrick Sahle widmet im ersten Band seines umfangreichen Werkes über digitale Editionen ein Kapitel der Kritik an der „historisch-kritischen Ausgabe“, die seines Erachtens in einem „unkritischen Verständnis noch heute gültiges Leitbild“ sei.³

Sahle bedient sich hierbei mit dem Begriff der „historisch-kritischen Ausgabe“ – von ihm abgekürzt als „HKA“ – eines Sammelbegriffes, der auf die unterschiedlichen praktischen Anforderungen an moderne und klassisch-philologische Editionen zunächst ebenso wenig eingeht wie auf den Unterschied zu im engeren Sinne dokumentarischen, also etwa Editionen mittelalterlicher Urkunden, die sich in der Regel als Repräsentation exakt eines materiellen Textzeugnisses verstehen. Das allein weist schon darauf hin, dass für Sahle das zentrale Problem die kritische Auswahl bzw. Entscheidung für eine Lesart an sich ist, der er die hierarchiefreie Verfügbarmachung sämtlicher Zeugnisse entgegenstellt.

Zunächst soll nun versucht werden, seiner Argumentation zu folgen, indem sowohl die explizit der klassisch-philologischen Edition gewidmeten Kritikpunkte, als auch solche, die in der Praxis aus den Anforderungen und Problematiken kritischer Editionen anderer Disziplinen erwachsen, betrachtet werden, letztere immer im Hinblick auf die Frage, ob ihr abstrakter Gehalt auch auf die klassisch-philologische Edition sinnvoll übertragen werden kann.

Dass Sahle die unterschiedlichen Problemstellungen verschiedener kritischer Editionstypen durchaus bewusst sind, wird aus dem ersten von ihm pauschal gegen die „HKA“ geführten Vorwurf deutlich: Deren „Indifferenz“ beruhe darauf, dass auch im Falle anderer, „nicht altphilologische[r] Gegenständ[e] und Bedingungen [...] die Methode einfach beibehalten“ worden sei.⁴ Dass dieser Kritikpunkt schwerlich gegen die klassisch-philologische Edition an sich ins Feld zu führen ist, versteht sich von selbst, dennoch verfolgt Sahle weiter seine unter verschiedenen Schlagwörtern geführte Pauschalkritik gegen die „HKA“ als Sammelbegriff.

Der für Sahle offensichtlich zentrale Kritikpunkt an der explizit klassisch-philologischen Edition selbst versammelt unter dem zunächst etwas uneindeutigen Lemma des „Idealismus“ vor allem verschiedene Beispiele dafür, dass gewisse „gesteckte Ziele“ – das zeige der „empirische Befund“ – nicht erreicht worden seien: So gebe es weder „*finale*“ bzw. „«definitive»“, noch „vollständige“ Editionen.⁵

Im Folgenden geht Sahle explizit auf die Problematiken der Stematologie ein, also eines zentralen Bestandteiles der klassisch-philologischen Editionstechnik.

So sei die Stematologie keinesfalls „mathematisch-regelhaft“, d.h. kein Verfahren, das *ceteris paribus* immer zu den gleichen Ergebnissen führen würde. Vielmehr gebe es auch ohne veränderte Umstände oftmals abweichende Editionen, mitunter sogar mit völlig widersprüchlichen Stemmata.⁶

2 Es sei ausdrücklich darauf hingewiesen, dass ich hinsichtlich vieler von Sahle erwähnter Quellengattungen die Auffassungen und Verbesserungsvorschläge durchaus teile. In gewisser Hinsicht handelt es sich hier also um eine Art Teilrezension, die ihrerseits zur Diskussion anregen will, und das über den Teilgegenstand der klassisch-philologischen Edition.

3 Vgl. Sahle (2013a), S. 133.

4 Vgl. ebd., S. 137; vgl. auch ebd., S. 34–5.

5 Vgl. ebd., S. 133–4. Kursivierung und Anführungszeichen i.O.

6 Vgl. ebd., S. 134–5.

Daran zeige sich – hier offenbart Sahle seine Auffassung des Begriffes „Idealismus“ –,⁷ dass, „[w]enn die Methode nicht in der Lage ist, die Herausgeber unter den gleichen Bedingungen der jeweiligen Überlieferung zu den gleichen Schlüssen zu zwingen, [...] ihre Objektivität offensichtlich nur eine idealistische Wunschvorstellung ohne praktische Relevanz“ sei.⁸

Auch eine weitere Gruppe von Kritikpunkten wird mit einem großen Schlagwort versehen, das von Sahle als mittelbare Folge der Tatsache dargestellt wird, dass die „positivistische“ Methode ihrem Ideal der Objektivität nicht gerecht werden könne, nämlich dem des „Konstruktivismus“.

Dabei richtet Sahles Kritik sich zunächst gegen sprachliche „Normalisierungen“. „[D]as klassische Latein als historisch-geographische Partikularsprache“ werde „auf alle Zeiten, Räume und Genres übertragen.“⁹ Auch hier wird keine deutliche Unterscheidung zwischen verschiedenen Formen der historisch-kritischen Ausgabe vorgenommen. Dass Sahle allerdings gerade hier die klassisch-philologische Methode als Ursprung des Problems ansieht, wird an anderer Stelle dennoch deutlich.¹⁰ Neben dieser Kritik der Normalisierungen, als deren Folge auch die schließlich noch gesondert behandelte „Ahistorizität“ gesehen werden müsse,¹¹ kommt er noch auf eine andere Problematik zu sprechen:

„Konstruktion – um nicht zu sagen: Mythos – ist ebenfalls die angebliche Intention des Autors, sein «Wille», aber auch die Einheit des «Werkes» und die aus diesen Idealvorstellungen abzuleitende höhere Autorität jenes Textes, den der Editor bestimmt.“¹²

Das hier angesprochene Problem der Autorenintention ist zumeist explizit im Zusammenhang mit den Problemen des Edierens moderner Autoren diskutiert worden, und auch die in Sahles umfangreicher Sekundärliteraturschau zusammengetragenen Belege dafür, dass es sich bei der Intention um einen „Mythos“ handle, befassen sich mit den Problematiken moderner Editionen. Das liegt vor allem daran, dass die Diskussion ursprünglich im Zusammenhang mit der Problematik, den richtigen Copy-Text auszusuchen, entstanden ist, sprich etwa, wenn im Falle mehrerer vorhandener Autorenmanuskripte, Korrekturfahnen, lektorierte Fassungen usw. dasjenige Textexemplar auszuwählen ist, das der Intention des Autors angeblich am nächsten kommt. Bei Sahle allerdings findet sich keinerlei Hinweis darauf, dass er das Problem der Autorität des einen Textes nicht auch auf den Archetyp der klassischen Philologie übertragen wissen will, der in seiner Diktion ebenfalls als „rein subjektiv konstruiert“, und „pseudo-objektiv“ gelten muss.¹³

Aber inwiefern lassen sich nun die recht allgemeinen Erkenntnisse, dass die Unterstellung einer Autorenintention, Konjekturen, „Vermischung“ von Textzeugnissen, Emendationen etc. durchaus problembehaftet sind und stark von den subjektiven Vorlieben und Anschauungen

7 Deutlicher wird Sahles Verbindung des Begriffes „Idealismus“ mit der hier geschilderten Problematik eines positivistischen Objektivitätsanspruches, vgl. Sahle (2013a), S. 36: „Die Idee der klassischen Idealität steht hinter der kritischen Edition wie sie auch die Wissenschaftstradition des 19. Jahrhunderts bestimmt. Zu der historischen Sichtweise, dass alles Gewordene in der Zeit verortet ist, kommt die Vorstellung von der zeitlosen Wahrheit, von einer finalen Richtigkeit der Erkenntnis, die mit der Etablierung der positivistischen Naturwissenschaften konvergiert.“

8 Ebd., S. 134.

9 Ebd., S. 137–8.

10 Vgl. ebd., S. 30.

11 Vgl. ebd., S. 141–3.

12 Ebd., S. 138.

13 Vgl. ebd., S. 139. Es erfolgt vielmehr sogar eine direkte Identifikation auch der klassischen Philologie mit dem Problem der Autorenintention: „Aber selbst wenn man auf die belegfreien inhaltlich-stilistischen Konjekturen verzichtete, musste der Vorrang der Autorenintention und die Vermischung der Überlieferungsträger nicht nur bei den alphilologischen, sondern ebenso bei den mittelalterlichen und den modernen Werken dazu führen, dass die Editionen Texte boten, die nie so existiert hatten, zugleich aber vorgaben, die ganze Zeit so existiert zu haben.“

eines Editors geprägt sind, als tatsächlich valide Ansatzpunkte einer Fundamentalkritik ansehen, der als einzige Alternative die Bevorzugung einer unkritischen Dokumentation vorzuschweben scheint? Allgemein muss angemerkt werden, dass der Unterscheidung, die Sahle zwischen den Schlagwörtern Konstruktivismus, Idealismus und Positivismus zieht, keine scharfe terminologische Trennung der von ihm unter diesen Lemmata behandelten Kategorien zu Grunde liegt; eher handelt es sich um verschiedene Ausdrucksformen des stetig wiederholten Vorwurfs, dass eine Kritik notwendigerweise einen subjektiven Eingriff in einen Text darstelle, der in aller Regel eher schade als nütze. Sämtliche der von Sahle angeführten Fehler lassen sich prinzipiell sinnvoller als Ausdruck – in seiner Diktion – ‚positivistischer‘ Verfahren ansehen, die auf dem trügerischen ‚Ideal‘ einer absoluten Wahrheit beruhen würden und daher notwendigerweise rein ‚subjektive‘ Konstrukte zur Folge hätten. Diesen Vorwurf versucht er anhand jedes einzelnen Punktes erneut zu belegen. Da im Folgenden nachvollzogen werden soll, inwiefern dies im Einzelnen gelingt, erfolgt die weitere Diskussion anhand der konkret formulierten Kritikpunkte und nicht anhand der allgemeinen Lemmata Sahles.

i) Autorenintention als Mythos?

Zunächst scheint es notwendig zu sein, sich eine in recht kurzer Zeit selbst zum modernen Klassiker gewordene Kritik der Textkritik, McGanns „Critique of Modern Textual Criticism“,¹⁴ und die hierin vorgenommene Kritik der „Ideologie der finalen Intentionen“ noch einmal zu vergegenwärtigen.

Im Vorwort zur 1992 erschienenen Neuauflage von McGanns Kritik bemerkt D. C. Greetham:

„It has become common for McGann to have to insist, in conference and published article, that he did not intend in the *Critique* to unseat the author as source of authority and meaning, but rather to place the author in the historical continuum, and to overcome the ideology of the solitary originator“.¹⁵

Diese Feststellung verdeutlicht, dass McGann keineswegs zum Ziel hatte, die Intention des Autors als einen reinen „Mythos“ zu disqualifizieren. McGann sieht die idealisierte Figur des Autors vor allem kritisch, weil er die Mechanismen der modernen Literaturproduktion in die Textkritik mit einbezogen wissen möchte. Neben der Person des Autors selbst seien zahlreiche weitere Personen an der Produktion eines Werkes beteiligt, etwa Lektoren, Verleger, eventuell staatliche Zensurbehörden und dergleichen mehr – der Autor ist Bestandteil einer Kulturindustrie. Schließlich ist auch die Auswirkung zahlreicher weiterer sozialer Faktoren auf die Person des Autors selbst zu berücksichtigen. In McGanns Kritik zeigt sich somit ebenfalls Kritik an einem Idealismus – McGann selbst spricht von einer „Romantic ideology“ – dessen zentrales Kriterium aber weniger die Vorstellung einer objektiven Wahrheit ist, die Gegenstand positivistischer Verfahren des Erkenntnisgewinns sein kann, sondern die Vorstellung des Autors als einer „autonomous authority“, die von den ihn umgebenden Verhältnissen unabhängig agiere.¹⁶ Diese Betonung der sozialen Umstände der Entstehung literarischer Werke allerdings führt gerade nicht dazu, die Rolle, die der Wille oder die Intentionen des Autors spielen, als Mythos abzutun, sondern vielmehr dazu, diese Faktoren anderen beizuordnen: McGann wehrt sich lediglich gegen die Annahme, es gebe „a single *fons et origo*“.¹⁷ Dass auch ein an-

14 McGann (1992).

15 Vgl. McGann (1992), S. xiii.

16 Vgl. ebd., S. 42.

17 Vgl. McGann (1992), S. 49.

tiker Autor kein autarkes Wesen, sondern eine soziale Person war, und somit ebenso wie sein Werk als Produkt eines sozialen Seins gesehen werden sollte, kann mittlerweile sicherlich als *opinio communis* gelten und hat in Einzelfällen auch zu kontroversen

Diskussionen geführt, wie man beispielsweise an den augusteischen Autoren sehen kann, deren Verhältnis zum *princeps* schon seit jeher die Diskussionen bestimmt. Auch gibt es kaum einen antiken Großautor, dem man nicht verschiedene Stufen der Überarbeitung nachsagt, die sich mal an verschiedenen Arbeitsschichten, mal an großangelegten Umsortierungen, oftmals auch an klassischen Autoritätskonflikten mit späteren Überarbeitern zeigen.

Das alles ändert allerdings nichts daran, dass McGann das Problem der Autorenintention deutlich als eines definiert, das in der Praxis dort begegnet, wo es ein Editor mit Texten zu tun hat, die in einem unmittelbaren Verhältnis zum Autor stehen, und es explizit von den Problemen der klassischen Philologie mit ihren Archetypen abgegrenzt wissen will,¹⁸ während der Archetyp seinerseits *per definitionem* als (Re-)Konstruktion und größtmögliche Annäherung an einen stets hypothetischen und meist Jahrhunderte von den ersten Manuskripten entfernten Ursprungstext angenommen wird. Dadurch kann das von McGann beschriebene und kritisierte Autoritätsargument, es mit einem unmittelbaren und daher unanfechtbaren Zeugnis aus erster Hand zu tun zu haben, hier sowieso keinerlei unmittelbare Relevanz beanspruchen.¹⁹

ii) Stematologie als pseudoobjektiver ‚Idealismus‘?

Das führt auch direkt zur weiteren Feststellung, dass nämlich die von Sahle zusammengetragene Sekundärliteraturschau des 19. Jahrhunderts das ablehnende Urteil vor allem hinsichtlich eines zentralen Punktes nicht zu untermauern vermag: Das betrifft den Vorwurf, die Stematologie sehe sich bis heute als eine quasi-naturwissenschaftliche Angelegenheit, die finale, objektive, allgemeingültige und stets wiederholbare Ergebnisse liefere. Die historische Parallelisierung mit den – angeblich – rein positivistischen und reflexionsarmen Naturwissenschaften macht es allzu leicht, so zu tun, als habe seit den Zeiten Lachmanns keine kritische Reflexion innerhalb der klassischen Philologie stattgefunden. Dass das nicht haltbar ist, verdeutlicht schon Wests Definition des Archetyps.

Die Tatsache allein, dass der Anspruch einer möglichst exakten Annäherung an ein Ideal und dessen Wahrheitsanspruch in den Anfangszeiten der klassischen Philologie im 19. Jahrhundert in nicht wenigen Fällen wohl lediglich behauptet wurde, lässt solche qualifizierenden Attribute wie die von Sahle gewählten im Zusammenhang eines Pauschalurteils also zumindest unangemessen erscheinen. Der Editionstext, so sehr er auch Ergebnis von Konstruktionen und teilweise willkürlichen Entscheidungen sein mag, bildet als wissenschaftliche Hypothese, die sich dieser Eigenschaft selbstredend stets bewusst sein muss, eine unumstößliche Grundlage für altertumswissenschaftliches Arbeiten.²⁰ Eine befriedigende Alternative dazu bieten rein dokumentarische Ansätze nicht. Umgekehrt können ein verbesserter Zugang zu Manuskripten und eine dokumentarische Aufarbeitung aber eventuell sogar dazu beitragen, durch verbesserte Nachvollziehbarkeit die Autorität der klassischen Edition zu stärken.

18 Vgl. Greetham im Vorwort zu Mc Gann (1992), S. xiv–xv, der betont, dass McGann die Unanwendbarkeit der klassischen Methode auf moderne Editionen kritisiert und nicht die Methode selbst.

19 Vgl. West (1973), S. 32: „[T]he lowest common ancestor of the known manuscripts.“

20 Was hier ausdrücklich aus der Perspektive des Historikers gesagt sei.

iii) Normalisierung

Besser geeignet, die Vorteile dokumentarischer Editionsformen hervorzuheben, scheinen vor allem Sahles Hinweise auf das Problem sprachlicher Normalisierungen. Sahle behandelt dieses Problem unter anderem im Kontext einer Diskussion des Konzeptes einer sogenannten Textologie,²¹ die ihre Stärken vor allem dort entfalten kann, wo man es mit Quellengattungen zu tun hat, bei denen der materielle Charakter des Objektes notwendigerweise stark im Vordergrund steht, was auch für die in den Altertumswissenschaften relevanten Quellengattungen mit materiellem Charakter, also etwa Inschriften, Münzen, Ostraka und Papyri einen großen Nutzen bringen kann und in vielen Fällen auch schon gebracht hat.²²

Gewichtiger scheint im hier behandelten Fall der Textedition aber vor allem die von Sahle ebenfalls herausgehobene Möglichkeit, problematische Idealisierungen etwa der Orthographie, Metrik usw. zu vermeiden, d.h., die historische oder linguistische Bedeutung etwa regionaler Besonderheiten der einzelnen Textzeugnisse beizubehalten. Hier kann eine stärkere dokumentarische Ausrichtung dezidiert sprachwissenschaftlichen Fragestellungen, für die der Kontext in der Textgeschichte weniger relevant sein mag als der konkrete empirische Befund, sicher entgegenkommen.

iv) Konjekturen

Sahles Einwände gegen Konjekturen scheinen, ebenso wie die Kritik der Stematologie, über den in Grenzen sicherlich berechtigten Skeptizismus nicht hinauszukommen. Auch hier ließe sich die Frage stellen, was seines Erachtens an die Stelle begründeter und gekennzeichnete Hypothesen – denn darum sollte es sich im Idealfall ja handeln –, die einen Text im Zweifel erst lesbar machen,²³ treten sollte. Natürlich können auch hier digitale Formen genutzt werden, um einen leichteren und eventuell besseren Zugriff auf verschiedene Varianten etwa verschiedener Ausgaben zu ermöglichen.²⁴

v) Die einzelnen Punkte als Ausdruck einer Fundamentalkritik?

Die hier angesprochenen Kritikpunkte dienen Sahle, wie bereits erwähnt, vor allem dazu, die erkenntnistheoretischen Grundlagen der kritisch-philologischen Edition als hoffnungslos veraltet darzustellen. Dass er hierbei vor allem die historischen Anfänge der Debatte nachverfolgt und die explizite Diskussion über die klassische Philologie dort verlässt, wo sich eine Diskussion über die Voraussetzungen der neuen Philologien entspinnt, wurde ebenfalls bereits angedeutet.

21 Auch in der Gegenüberstellung einer auf die Materialität ausgerichteten „Textologie“, die um eine dokumentarische Abbildung des Textträgers bemüht ist, und einer hermeneutisch-interpretativen „Editionsphilologie“, deren Ergebnis von vornherein durch Synthese und Interpretation bestimmt ist und dadurch verfälscht zu werden drohe, deutet sich Sahles genereller Skeptizismus an, vgl. Sahle (2013a), S. 241–3.

22 Verwiesen sei hier, freilich stellvertretend für viele, auf das Papyrusportal: <http://www.papyrusportal.de>.

23 Zur Problematik der Konjektur im Spannungsfeld zwischen „bewusst und unbewusst ablaufenden Formen der Wissensverarbeitung“, vgl. Wirth (2008), S. 269–294 (vorangegangenes Zitat ebd., S. 269), der auf die Bedeutung konjekturalen Wissens in unterschiedlichen Theorien der Hypothesenbildung, etwa bei Pierce und Popper verweist, und dabei auch auf die von Schleiermacher implizierte Annahme einer „konjekturalen Kompetenz“ als Resultat eines „durch Übung gebildeten Talents“ (ebd. S. 286, Schleiermacherzitat aus Hermeneutik und Kritik, Frankfurt a.M. 1977, S. 283, zitiert nach ebd.) und Schnittstelle zwischen hermeneutischer und philologischer Kritik eingeht.

24 Das Projekt eComparatio hat neben einer plaintextbasierten Textvergleichssoftware unter anderem die Anbindung an Faksimiles zu Grunde liegender Handschriften zum Ziel. Darüber hinaus soll die Möglichkeit geschaffen werden, neben der automatischen Erstellung von Apparaten auch umgekehrt aus Apparaten Textvarianten zu erstellen. Vgl. www.eaqua.net.

Hier soll nun noch kurz – und sicher ebenfalls verkürzt – versucht werden zu erläutern, weshalb darin tatsächlich die Absicht einer Fundamentalkritik gesehen werden kann.²⁵

Das hängt erneut mit der vor allem an der Medientheorie orientierten Materialismusdefinition Sahles zusammen. Neben der schon angesprochenen grundsätzlichen Materialität von Textträgern, die es aus reinem Erkenntnisinteresse an ihnen als Objekten der Überlieferung zu bewahren gilt, betont Sahle deutlich die Auswirkung jener Objekte auf die Auffassungen von Text selbst.²⁶ Problematisch scheint dabei weniger diese Erkenntnis, sondern seine wiederholte Qualifizierung dieser Entwicklung als „evolutionär“.²⁷ Die von Sahle an anderer Stelle angesprochene Entscheidung des Editors²⁸ zwischen den von ihm dargelegten „Textbegriffen“²⁹ erscheint so letztlich als determiniert, die Entscheidung für die „idealistische“ Variante des „einen“, kritisch etablierten Textes als schlichtweg überholt im Sinne einer noch dazu als natürlich und damit unaufhaltsam voranschreitend dargestellten Entwicklung. Bestenfalls kann hier die kritische Edition noch den Platz neben anderen Textformen beanspruchen, da ihr materielles Vorhandensein immerhin dies noch rechtfertigt.

Sahles allgemeine Haltung zu kritischen Editionen ist also offenkundig nicht bereit anzuerkennen, dass es sich, zumindest im Falle dezidiert philologischer Arbeiten, bei ihnen bereits um wissenschaftliche *Auseinandersetzungen mit der Überlieferung* handelt, die damit prinzipiell auch eine andere ontologische Qualität beanspruchen können als die ihnen zu Grunde liegenden materiellen *Überlieferungsträger*. Die Kritik ist kein notwendiges Übel, sondern eine wissenschaftliche Leistung, die natürlich fehlerbehaftet ist. Aber: Erst dezidiert philologische und linguistische Forschungen machen die antike Literatur zu einer nutzbaren und damit nützlichen Quellengattung für die Geschichts- und Literaturwissenschaft.

Wollte man Sahles mitunter polemischen Ton³⁰ gegen ihn selbst wenden, könnte man die Aufforderung zum unterscheidungslosen und unkritischen Sammeln von Informationen aller Art und die Suche nach einer digitalen Umsetzungsmöglichkeit, die „wahrer“ sei als „der Text selbst“, und den Anspruch, die „überkommenen Grenzen zwischen Edition und Überlieferung“ aufzulösen, vermutlich mit mehr Berechtigung als „Positivismus“ bezeichnen, als ein kritisch-konstruktivistisches Verfahren, wie es bei der klassischen Edition zur Anwendung kommt.

Ein weiteres Problem ist die von Sahle verwendete Terminologie, die die technische Entwicklung als eine natürliche bezeichnet. In Verbindung mit der Feststellung, dass Vorstellungen von Text sich technischen Bedingungen ihrer materiellen Träger verdanken, führt das zur Vorstellung einer wissenschaftlichen Methode, die sich den unabwendbaren technischen Entwicklungen unterzuordnen und anzupassen habe, anstatt in diesem Zusammenhang selbst eine aktive Rolle zu übernehmen.

Diese Vision, durch Markup-Verfahren eines Tages dazu in der Lage zu sein, Text „wahrer“ als er selbst sei, nicht mehr allein zu repräsentieren, sondern zu reproduzieren und ihm hierbei eine Stellung zwischen Materialität und Idealität zusprechen zu können, führt in letzter Konsequenz erneut dazu, dass diese Editionen die Quellen genauso unstrukturiert abbilden, wie die Wirklichkeit ist.³¹

²⁵ Man betrachte dazu lediglich das Schlusswort in Sahle (2013b), o. S., das seinerseits zu einer polemischen Auseinandersetzung aufruft, wenn es eine Auflösung der „überkommenen Trennung von Edition und Überlieferung“ fordert.

²⁶ Vgl. ebd., S. 81–98.

²⁷ Ebd. passim. Vgl. programmatisch in S. 111ff. sowie S. 391: „Die Evolution der Techniken ist eine Evolution der Textbegriffe“.

²⁸ Vgl. ebd., S. 19–20.

²⁹ Ebd., S. 1–98.

³⁰ Man beachte auch das Schlusszitat in Bd. 3.

³¹ Dass das auf einen anderen Adressaten hinausläuft als die klassische Edition, nämlich nicht mehr auf einen menschlichen Leser, ist naheliegend.

Dass Sahle Textkritik in seinem abschließenden Plädoyer noch als einen Bestandteil digitaler Editionen ansieht, ist vor allem der Tatsache geschuldet, dass die „technische Evolution“ noch nicht dort angelangt ist, wo er sie sich wünscht.

Es bleibt zu sagen, dass auch im Falle der klassisch-philologischen Edition eine Verbindung dokumentarischer und kritischer Editionsformen – mit Primat der letzteren! –, ebenso das Nebeneinanderstellen und Vergleichen verschiedener Editionen dazu führen könnte, dass die klassischen Editionen und ihre Textrekonstruktionen in noch stärkerem Maße das werden, was sie dem ursprünglichen Verständnis nach sein sollten: begründete Hypothesen, die Grundlage, aber auch Gegenstand einer wissenschaftlichen Diskussion sind.

Der hier im Zusammenhang mit der tatsächlichen Intention der Kritik McGanns an der Autorenintention angesprochene Aspekt, nämlich die stärkere Betonung historischer Bedingungen für die Entstehung eines antiken Textes, kann so über den Bereich hinaus, den ihr die klassische Textkritik zuweist, verstärkt auch von denjenigen geführt werden, die bisher vor allem die Rolle reiner Adressaten der kritischen Ausgaben gespielt haben, eben Historikern, Archäologen und anderen Altertums-, Geistes- und Kulturwissenschaftlern. Der mitunter in der Diskussion über das Wesen und die Zukunft der Digital Humanities recht trivial rezipierte Begriff der Hermeneutik, der meist gegen eine ebenso trivialisierte Vorstellung von naturwissenschaftlicher Methodik als bloßer Appell an die Bewusstmachung eigener Erkenntnisvoraussetzungen angeführt wird und nicht selten in eben die Skepsis mündet, die in Sahles Kritik an der kritischen Edition anklingt, könnte so tatsächlich durch eine Perspektive bereichert werden, in der sich die historische (Quellen-)kritik wieder stärker auf die Diskussion über die Beschaffenheit und den Entstehungs- und Überlieferungskontext der Texte auswirkt, was durch die Spezialisierung der einzelnen Disziplinen zunehmend aufgegeben worden ist.

Literatur:

McGann (1992): J. J. McGann, A Critique of Modern Textual Criticism, Chicago / Charlottesville / London 1992.

Sahle (2013a): P. Sahle, Digitale Editionsformen. Zum Umgang mit der Überlieferung unter den Bedingungen des Medienwandels Bd. 1, Norderstedt 2013 (Schriften des Instituts für Dokumentologie und Editorik Bd. 7).

Sahle (2013b): P. Sahle, Digitale Editionsformen. Zum Umgang mit der Überlieferung unter den Bedingungen des Medienwandels Bd. 3, Norderstedt 2013 (Schriften des Instituts für Dokumentologie und Editorik Bd. 7).

West (1973): M. L. West, Textual Criticism and Editorial Technique, Stuttgart 1973.

Wirth (2008): U. Wirth, Die Konjektur als blinder Fleck einer Geschichte des bedingten Wissens, in: C. Welsh / St. Willer (Hg.), Interesse für bedingtes Wissen. Wechselbeziehungen zwischen den Wissenskulturen, München 2008, S. 269–294.

Weitere Ressourcen (Zuletzt aufgerufen am 18.02.2016):

Das ‚Papyrus Portal‘: digitalisierte und elektronisch katalogisierte Papyrussammlung,
URL: <http://www.papyrusportal.de>

eAQUA: Extraktion von strukturiertem Wissen aus Antiken Quellen für die Altertumswissenschaft,
URL: www.eaqua.net

Autorenkontakt³²

Friedrich Meins

Universität Erfurt / Universität Leipzig
Max-Weber-Kolleg / Lehrstuhl für Alte Geschichte

Email: friedrich_meins@uni-leipzig.de

³² Die Rechte für Inhalt, Texte, Graphiken und Abbildungen liegen, wenn nicht anders vermerkt, bei den Autoren.

Die Abgrenzung der digitalen Geisteswissenschaften

Leif Scheuermann

Abstract: Subject of this paper is a demarcation of Digital Humanities as human sciences in the tradition of Wilhelm Dilthey. On this foundation the DH have to focus on the relation between life, expression, and understanding with the methodology of genuine digital hermeneutics. In a next step requirements on this methodology are verbalized, an approach for the implementing in a web based digital research environment are shown and risk as well as chances discussed. A new form of multidimensional and multimedial argumentation has to be developed, without neglecting the suggestive power of images.

Einführung

Sprechen wir heute von Geisteswissenschaften, so meinen wir gemeinhin eine Agglomeration verschiedener Fächer, der die Geschichts-, Literatur-, Kunst-, Religions-, und Sprachwissenschaften angehören, die sich universitär administrativ in geisteswissenschaftlichen Fakultäten organisieren. Diese oft pragmatisch getroffene Einteilung bestimmt zwar weitestgehend das Bild der Geisteswissenschaften nach außen, sie kann aber keine Grundlage für eine im Entstehen begriffene Disziplin der „Digital Humanities“¹ bilden. Die folgenden Ausführungen stellen eine Reihe von Gedanken zu einer Methode der genuin digitalen Geisteswissenschaften dar, die auf keinen Fall Vollständigkeit beanspruchen und bewusst in einigen Bereichen provokant formuliert sind. Ziel ist es dabei von der Vorstellung der DH als Sammelbecken für alles, was mit den Geisteswissenschaften und Computern zusammenhängt, – von der Frage des Einsatzes neuer Hardware über die Massendigitalisierung und Auszeichnung von Texten, Bildern und Objekten, die Programmierung von Spezialanwendungen bis hin zur Langzeitarchivierung – abzukommen.²

Für eine stringente Methodologie einer solchen Fachrichtung, ist in einem ersten Schritt eine genaue Bestimmung ihres Geltungsbereiches von Nöten, will man sich nicht der Beliebigkeit hingeben. Um die Digitalen Geisteswissenschaften als Konzept zu formulieren, bedarf es einer starken Definition in zweierlei Hinsicht: in Abgrenzung zu den benachbarten Disziplinen und in der Bestimmung der gemeinsamen Charakteristika der Einzeldisziplinen.³ In einem zweiten Schritt sollen die herausgearbeiteten Charakteristika auf den Bereich der Digitalen Geistes-

1 Die Bedeutungsfelder der Begriffe „humanities“ und „Geisteswissenschaften“ sind weitgehend deckungsgleich, so dass beide Begriffe synonym verwendet werden können.

2 Auf eine Ausführung des zum Teil sehr disparaten Stands der internationalen Debatte über DH muss an dieser Stelle verzichtet werden. Es empfehlen sich zum Einstieg folgende beiden Rezensionen bei H-Soz-Kult:

[Hobohm \(2015\): Sammelrezension Digital Humanities;](#)

[Herrmann \(2015\): Sammelrezension Understanding / Debates in the Digital Humanities.](#)

3 Für die Einzeldisziplinen mag eine Umwandlung hin zu „Kulturwissenschaften“, wie sie u.a. Reinhard Koselleck 1991 implizit in seinem Aufsatz „Wie sozial ist der Geist der Wissenschaft?“ [Koselleck (1991)] gefordert hat, durchaus sinnvoll sein, doch führt dies zwangsläufig zur Aufgabe eines gemeinsamen Apparates, welcher hingegen durch die Vertreter der Digital Humanities vorausgesetzt wird. Möchte man den Begriff der digitalen Geisteswissenschaften erhalten und stark machen, so kann dies eben gerade nicht über eine Wendung hin zu „digital cultural studies“ geschehen.

wissenschaften angewandt werden. Dies soll in Form der Formulierung von Forderungen an eine zu begründende Digitale Geisteswissenschaft geschehen. In einem letzten Schritt werden ferner Ansätze zur Umsetzung dieser aufzuzeigen sein, wobei speziell in diesem Abschnitt eine Vision appellativen Charakters im Zentrum steht.

Abgrenzung der Geisteswissenschaften

Für eine notwendige Bestimmung des Begriffs der Geisteswissenschaften soll im Folgenden kurz auf Wilhelm Diltheys programmatischen Aufsatz „Abgrenzungen der Geisteswissenschaften“⁴ eingegangen werden. Die folgenden Ausführungen beziehen sich dabei maßgeblich auf die zweite Fassung der Studie vom 7. Januar 1909. Hierin formuliert Dilthey als Ziel „die Sonderung der Wissenschaften in solche der Natur und des Geistes.“⁵ In die Gruppe der Wissenschaften des Geistes ordnet Dilthey die Geschichtswissenschaften, die Ökonomie, die Kunst, die Rechtswissenschaften, die Religionswissenschaften die Philosophie sowie die Sprachwissenschaften, da all diese „das Erlebnis, das Verstehen fremder Erlebnisse und Urteile und Begriffe, welche die erlebten und verstandenen Sachverhalte ausdrücken“⁶ thematisieren. „Alle systematischen Wissenschaften des Geistes beruhen auf der Beziehung, die zwischen dem Erlebten und Verstandenen und den Begriffen, die es ausdrücken besteht“⁷, wobei der Betrachter selbst immer Teil dieses Geflechtes ist. Damit wird auch klar, dass Geisteswissenschaften in ihrer Definition bereits eine Wertneutralität ausschließen, ohne damit die Gerichtetheit auf die „objektive Erkenntnis ihres Gegenstandes“⁸ zu verlieren. Dilthey definiert als Ziel der Geisteswissenschaften die „objektive Erkenntnis der Gesellschaft, der Geschichte, des Menschen. Die Möglichkeit einer solchen ist überall ihre Voraussetzung.“⁹

Doch liegt diese objektive Erkenntnis im „Verhalten des Geistes, welches im Unterschiede von dem naturwissenschaftlichen Erkennen der Gegenstand Geisteswissenschaft gebildet wird. Die Menschheit wäre, aufgefasst in Wahrnehmung und Erkennen, für uns eine physische Tatsache, und sie wäre als solche nur den Naturwissenschaften zugänglich. Als Gegenstand der Geisteswissenschaften entsteht sie aber nur, sofern menschliche Zustände **erlebt** werden, sofern sie in Lebensäußerungen zum **Ausdruck gelangen** und sofern diese Ausdrücke **verstanden** werden.“¹⁰ Dilthey zieht die Grenze zwischen den Wissenschaftskategorien also im Erleben und Verstehen. Naturwissenschaften thematisieren nach dieser Definition nicht die Perspektive des Betrachters.¹¹ Hieraus ergibt sich nun auch seine methodologische Forderung. Einer naturwissenschaftlichen Empirie stellt er die geisteswissenschaftliche Hermeneutik gegenüber, welche über den Dreischritt des Erlebens, Verstehens und Ausdruckfindens sich einem Thema nähert, wobei eine Hauptaufgabe darin besteht, das Erlebte und Verstandene „als in Urteilen und Begriffen adäquat darstellbar aufzufassen“¹² und umzusetzen. Ziel ist also nicht

4 Siehe Dilthey (1970): Der Aufbau der geschichtlichen Welt in den Geisteswissenschaften.

5 Ebd. S. 376.

6 Ebd.

7 Ebd. S. 377.

8 Ebd. S. 379.

9 Ebd. S. 381.

10 Ebd. S. 98.

11 Mag diese Ansicht im Detail spätestens seit Einstein so nicht mehr haltbar sein, so gibt sie uns dennoch eine grobe Richtschnur für die Trennung der Wissenschaftsbereiche.

12 Ebd. S. 383.

die Nacherzählung oder „wirklichkeitsgetreue“ Rekonstruktion des Vergangenen, sondern das Verstehen und das Fassen in allgemeine Begriffe und Urteile, die „Ausbildung der >analytischen Wissenschaft der einzelnen Zweckzusammenhänge, die als Kultursysteme durch die Geschichte hindurchgehen...“¹³ Den Boden für diesen hermeneutischen Prozess bildet das empirische Bewusstsein, in welchem der Betrachter sich der Fakten jenseits seiner selbst sicher ist. Darüber hinaus schaffen die Geisteswissenschaften jedoch Ordnungen, welche zu Erkenntnissen führen. Sie bahnen Pfade des Verständnisses durch die Masse der empirischen Zeugnisse, welche wiederum zu einem neuen Verständnis und so einem hermeneutischen Fortschreiten führen.

Forderungen an eine Methodologie der „Digital Humanities“

Was ergibt sich nun aus dieser kurzen Nachskizzierung der „Abgrenzung der Geisteswissenschaften“ für die Frage einer Methodologie der „Digital Humanities“? Zum einen muss nochmals herausgestellt werden, dass – obgleich Dilthey als der Begründer der Geisteswissenschaften gilt – hier nicht der Eindruck erweckt werden darf, die Entwicklungen einer Gruppe von Fächern, die sich in den Geisteswissenschaften selbst definieren, seien in irgendeiner Art und Weise Anfang des 20. Jh. stehen geblieben. Der Text soll vielmehr als eine der Gründungs-urkunden einer methodologisch verbundenen Einheit verstanden werden, auf welche sich die DH in ihrer Genese gezielt beziehen sollten. Auf drei Bereiche soll hierbei im Folgenden näher eingegangen werden:

- a) Geisteswissenschaften als Erfahrungswissenschaften basieren auf einer empirischen Grundlage, erschöpfen sich jedoch nicht darin.
- b) Die zentrale Methode der Geisteswissenschaften ist die Hermeneutik, nicht ein Aufzählen von Fakten oder eine „möglichst originalgetreue“ Rekonstruktion bzw. Reproduktion.
- c) Geisteswissenschaftliches Arbeiten kann nicht ohne den Erfahrungshorizont des Forschenden betrachtet werden.

Mit Aufkommen der Nutzung von Computern in den Geisteswissenschaften spätestens ab den 70er Jahren des 20. Jh. war eines der zentralen Ziele, möglichst viele Quellen in das neue Medium zu übertragen, so dass bis heute eine überbordende Fülle an Informationen digital zur Verfügung stehen. Zudem finden sich, betrachtet man die beantragten und bewilligten Projekte, welche unter dem „Label“ DH firmieren, immer noch in Mehrheit Digitalisierungsprojekte verschiedenster Korpora, seien es Texte, Bilder oder Objekte. Doch hat dies mit einer Digitalen Geisteswissenschaft im strengen Sinne nichts zu tun und ist nicht als Teil der Methode zu betrachten. Dasselbe gilt für die Programmierung von Anwendungen und Werkzeugen für eine Analyse per se. Auch diese sind nicht Teil der DH, sondern als Hilfswissenschaft anzusehen, die eine weitere Arbeit ermöglicht.

Anders jedoch verhält es sich mit der Aufarbeitung der Informationen, sei es durch die Auszeichnung in einer der Standard-Ontologien oder der Integration in eine Datenbankstruktur. Bei dieser Tätigkeit handelt es sich nicht mehr um die reine Aufnahme von Daten, sondern um die Auswahl, Ordnung und Interpretation derselben. Es muss klargestellt werden, dass eine

¹³ Ebd. S. 384.

solche Ordnung der Daten bereits einen ersten hermeneutischen Akt darstellt¹⁴ und die Ebene des empirischen Bewusstseins durch die strukturierte Eingabe in ein neues Medium verlassen wurde. Dem entsprechend ist es für eine Weiternutzung der Daten von zentraler Bedeutung, die Fragestellung der Datenaufnahme und die Kriterien unter denen diese stattfand klar zu dokumentieren. Der Forschende muss sich selbst als Teil des Forschungsprozesses verstehen, selbstreferenziell arbeiten und so die Möglichkeit zur Kollaboration mit anderen Forschenden schaffen. Auch muss die Vorstellung einer automatisierten Auszeichnung bzw. Strukturierung in einer Datenbank verworfen werden, da die Ordnung letzten Endes immer in der Verantwortung des Forschenden liegen muss. Allenfalls können Assistenzsysteme erstellt werden, die die großen Datenmengen bewältigen können. Die Resultate jedoch unhinterfragt als „wahr“ gelten zu lassen widerspricht den aufgezeigten methodologischen Grundlagen Diltheys.

Kommen wir nach der Datenaufnahme zur Frage der Analyse. Hier muss es das Ziel der DH sein, Methoden zur Umsetzung einer digitalen Hermeneutik zu schaffen. Es genügt nicht, dass der Forschende Abfragen in einer Datenbank macht oder Sachverhalte auf einer digitalen Karte visualisiert und diese dann in herkömmlicher Form für seine Arbeit an einem Text nutzt. Dies wäre allenfalls unter dem Begriff „Computer aided Humanities“ zu fassen. Das zweite zentrale Thema einer zu begründenden digitalen Geisteswissenschaft ist also in der Art der Abfrage der Daten und in der Wahl und Kombination der genutzten Werkzeuge zu sehen. Es muss Anspruch der Disziplin sein,

1. alle möglichen geisteswissenschaftlichen Forschungsdaten in einem System zusammenzuführen und mit einer beliebigen Anzahl unterschiedlichster Anwendungen zu kombinieren und weiterzuverarbeiten;
2. darüber hinaus den digitalen hermeneutischen Prozess im genutzten Medium zu dokumentieren und so nachvollziehbar zu machen.¹⁵

Eine Forderung an eine digitale Geisteswissenschaft ist es eine umfassende Formalisierung zu schaffen, und dies gilt sowohl für den ersten hermeneutischen Schritt der Datenstrukturierung, als auch für die Verarbeitung der Daten und deren Ausgabe in unterschiedlichsten miteinander gekoppelten Medien. Es bedarf einer Metaebene,¹⁶ welche die den hermeneutischen Prozess begleitenden Analyseschritte dokumentiert, um ihn verstehbar, nachvollziehbar und vermittelbar zu machen sowie neue, genuin digitale Visualisierungsformen für die Vermittlung der hermeneutischen Prozesse¹⁷ – darin besteht die zentrale Forderung an eine Digitale Geisteswissenschaft. Zur Begründung einer solchen Wissenschaft muss der Blick vom empirischen Bewusstsein der Objekte weg, hin zu einer Beschreibung des Prozesses der Begriffs- und Urteilsbildung im neuen Medium sowie deren formaler Mittel erfolgen.

14 Dass nicht das wahl- und ziellose Sammeln, sondern das Ordnen und Systematisieren des Gesammelten als integraler Bestandteil des hermeneutischer Prozesses von zentraler Bedeutung für die Philologie, wie auch der gesamten Geisteswissenschaften ist, übersieht Lauer, wenn er schreibt: „Wissenschaftshistorikerinnen wie Lorraine Daston haben gezeigt, wie in den aufkommenden modernen Wissenschaften an die Stelle der Exotika und Rarissima der Kunstkammern und Jahrmärkte die kontrollierte Genauigkeit der Beobachtung beim Sammeln der Daten trat. Das, und nicht die Hermeneutik, macht Wissenschaften modern, auch die Philologien. Und diese Traditionslinie schreiben nun auch die Digital Humanities fort.“ [Lauer (2013), S. 102].

15 Es muss nochmals betont werden, dass die technische Erstellung eines solchen Systems oder einzelner Anwendungen nicht in den engeren Methodenkanon der DH zu rechnen ist, ebenso wenig wie die Frage der Publikation oder Zitierweisen in den traditionellen Geisteswissenschaften.

16 Hierfür wäre eine Ontologie der hermeneutischen Prozesse unabdingbar.

17 Maßgeblich zur Frage einer neuen visuellen Hermeneutik: Kath / Schaal / Dumm (2015).

Als letzten zu betrachtenden Schritt im Prozess der Datenverarbeitung¹⁸ wie auch der Hermeneutik ist die Ausgabe der Daten bzw. Vermittlung des neu geschaffenen Wissens zu nennen, welche einen unverbrüchlichen Bestandteil des hermeneutischen Arbeitens darstellt. Die Neuen Medien schaffen neue Möglichkeiten des Ausdrucks, weg von der zweidimensionalen Form des Narrativs hin zu einer multimedialen, vernetzten und somit mehrdimensionalen Form, wobei allerdings nicht auf die Form des Textes verzichtet werden kann. Dabei ist eine Verzahnung mit Bild, Animation etc. jedoch ein wichtiger Bestandteil einer neuen digitalen Form der Wissensrepräsentation. Eine weitere zentrale Forderung an die DH muss es in diesem Rahmen sein, die Grenzen der Visualisierungsmöglichkeiten auszuloten und die Auswirkungen der Neuen Medien auf den Vorgang der Rezeption kritisch zu begleiten. Besonders in der Suggestivkraft des Bildes bzw. der Animation, welche zwangsläufig einen Wahrheitsanspruch vermittelt, besteht eine zentrale Herausforderung, möchte man sich nicht einer Romantik des 19. Jh. ergeben. Es sind also wiederum nicht die technischen Möglichkeiten, die die Digitalen Geisteswissenschaften betreffen, sondern der Umgang mit ihnen. Eine Vision der „Venice time machine“ wie sie 2013 von Frederic Kaplan propagiert wurde, ist ein erschreckendes Beispiel für ein fehlendes Problembewusstsein.¹⁹ Wenn der hermeneutische Prozess nicht ohne den Betrachter möglich ist, so muss klar sein, dass eine Visualisierung der Vergangenheit, nie ohne die lebensweltlichen Erfahrungen des Forschers existieren kann, oder, um es nochmals mit Dilthey auszudrücken, ist die Aufgabe der digitalen Geisteswissenschaften den Untersuchungsgegenstand „als in Urteilen und Begriffen adäquat darstellbar aufzufassen“²⁰ und nicht zu versuchen etwas nachzuerzählen, was nicht nacherzählbar ist.

Einen letzten Aspekt gilt es in diesem Zusammenhang wenigstens noch kurz anzusprechen, wenn er an dieser Stelle jedoch nicht in extenso ausgeführt werden kann, nämlich das durch Jürgen Mittelstraß formulierte „orientierende Wesen“²¹ der Geisteswissenschaften.

„Nicht der der viel weiß, ist der, der Orientierungsfragen beantwortet, sondern der, der im Rahmen einer Lebensform, auch einer wissenschaftlichen Lebensform, die geheimnisvolle Grenze zwischen Wissen und Können überschritten hat. Das machte übrigens schon den idealistischen Begriff der Bildung aus, ein Umstand, der in ihrerseits meist kurzfristiger Identifikation der Geisteswissenschaften als Bildungswissenschaften meist unberücksichtigt bleibt.“²²

Übertragen auf die digitalen Geisteswissenschaften ergibt sich hieraus die Forderung des „orientierenden Wesens“ für den digitalen Raum und die digitale Gesellschaft. Digitale Bildung darf sich nicht in einer lexikalischen Aufreihung von Faktenwissen nach dem Vorbild der Wikipedia erschöpfen, sondern muss den gezielten und verantwortungsvollen Umgang mit großen Datenmengen ebenso umfassen, wie die Auswirkungen der Erweiterung der medialen Vielfalt. Digitale Humanisten, um dieses Schlagwort wenigstens einmal kurz genannt zu haben, müssen ein Können im Verständnis der digitalen Welt besitzen, eine Bildung, und damit in die sich immer weiter digitalisierende Gesellschaft hinein wirken.

18 Nach dem IPO – Model.

19 Zu einer kritischen Auseinandersetzung mit den Möglichkeiten der raum-zeitlichen Visualisierung siehe Scheuermann (2014).

20 Dilthey (1970), S. 383.

21 Mittelstraß (1991). Hier lehnt Mittelstraß den für die Geisteswissenschaft durch Marquard geprägten Begriff der Kompenationswissenschaft ebenso ab wie den der Orientierungswissenschaften, da er auf der einen Seite eine Unter- und auf der anderen eine Überforderung sieht. Stattdessen findet er die Formulierung des „orientierenden Wesens“ der Geisteswissenschaften.

22 Ebd. S. 39.

Ansätze zur Verwirklichung der „Digital Humanities“

Nach einer skizzenhaften Definition der Geisteswissenschaften auf Basis von Wilhelm Diltheys Aufsatz „Abgrenzungen der Geisteswissenschaften“ und den sich daraus ergebenden Forderungen für eine Begründung der Digitalen Geisteswissenschaften sollen nun in einem Letzten Abschnitt die bereits bestehenden technischen Ansätze thematisiert werden.

Vier Thesen wurden bis dato herausgearbeitet:

- 1) Die Aufgabe der digitalen Geisteswissenschaft besteht im Strukturieren und Ordnen digitaler Daten sowie in der Begriffsfindung.
- 2) Hierfür ist eine Nutzung unterschiedlichster Datenquellen und Analyseverfahren sowie die freie Kombination dieser im digitalen Medium erforderlich.
- 3) Der Forschende steht dabei nicht außerhalb des digitalen hermeneutischen Prozesses, sondern ist Teil davon, weshalb immer der eigene Standpunkt wie auch der Verlauf der Forschungen genau zu dokumentieren sind.
- 4) Die Präsentation der Ergebnisse der Forschungen, welche den Prozess selbst beinhalten, muss in adäquater digitaler Weise digital erfolgen, wobei das Primat des Textes und der Erzählung zu brechen und durch eine vernetzte multimediale Form zu ersetzen ist. Wichtig ist dabei jedoch nicht in eine neue Romantik zu verfallen oder einen absoluten „Wahrheitsanspruch“ zu postulieren.

Wie ist nun eine solche digitale Geisteswissenschaft technisch zu verwirklichen? Zuerst einmal bedarf es einer digitalen Forschungsumgebung, welche die freie Kombination unterschiedlichster verteilter Daten und Anwendungen in einem kollaborativen System ermöglicht. Sie muss themenunabhängig sein, auf internationalen Standards beruhen, für jeden transparent, in der Anwendung möglichst einfach, ohne größere Programmierkenntnisse nutzbar und so einem weiten Anwenderkreis zugänglich sein. Die Plattform selbst nimmt also die Rolle einer Vermittlungsbasis ein, in die einzelne Anwendungen integriert werden können. Ihre zentrale Aufgabe ist die Verwaltung generischer Schnittstellen, in der die freie Kombination aller Anwendungen möglich ist.²³

Erste Ansätze zum Aufbau einer solchen Plattform bestehen bereits seit den 80er Jahren des 20. Jh. im Bereich der Meme Media Technologie. Die Grundidee dieser Technologie ist, dass zwischen Anwendungen und Dokumenten aus Sicht ihrer Architektur keine Unterschiede bestehen, so dass komplexe Abläufe, wie sie für den hermeneutischen Prozess benötigt werden, gleich einem Textdokument, welches aus einzelnen Buchstaben besteht, aus einfachen Modulen zusammengesetzt werden können.²⁴ Für die Interaktion dieser primitiven Einheiten (pads) wurden In- und Output-ports ohne weitere Spezifizierung (slots) definiert. Kuwahara beschreibt diese als: „A hole or a plug where one may connect a „contact“ in order to create a unilateral or a bilateral stream channel or path between two slots in two separate pads.“²⁵ In der synthetic media architecture besitzen diese Slots nur zwei Parameter: einen Namen und eine „standard message“, welche die Richtung des Datenaustausches festlegt.²⁶ Über diese einfache Struktur können nun beliebige Anwendungen miteinander kombiniert und zu Komplexen zusammen-

23 Zur barrierefreien Nutzung auch durch einfache Anwender ist es sinnvoll dies in einer GUI umzusetzen.

24 Tanaka (2013) S. 163: „From an achitectural view there is no difference between multi media documents and tools“.

25 Kuwahara (2013).

26 „set“ = von parent zu child; „gimme“ = von child zu parent; „update“ von einem parent zu children.

geführt werden. Mit dem Aufkommen des Internets und später des Web 2.0 wurden diese Ansätze in der Ausformung „Webbles“ in Form einer web basierten Plattform verwirklicht.²⁷ Auf dieser können aus einer großen Menge von einfachen, sogenannten „primitive webbles“, beliebig komplexe Strukturen zur Datenverarbeitung generiert werden. Gleichzeitig können bestehende webbasierte Anwendungen integriert werden, indem sie „gewrapped“, also die In- und Output Schnittstellen definiert werden. Die genutzte Anwendung selbst wird dabei als Blackbox betrachtet.²⁸ Die Arbeit eines Programmierers besteht also lediglich darin, Schnittstellen zu definieren und so den Datenaustausch zu ermöglichen. Dabei kommen Metadatensprachen bzw. Ontologien, wie die TEI²⁹ oder CIDOC CRM³⁰, eine besondere Bedeutung zu, da sie die Bedeutungen der jeweiligen Daten eindeutig definieren. Über sogenannte generische „wrapper“ können hierdurch alle referenzierten Daten automatisch in das System eingegliedert werden.

Noch ist diese Technologie in ihren Ansätzen begriffen, doch würde sie die Möglichkeit einer übergreifenden themenunabhängigen Plattform als Grundlage einer genuin digitalen Geisteswissenschaft bieten. Hierbei ist allerdings nicht gemeint, dass es nur eine solche Plattform geben kann – diese Vorstellung entbehrt jeglicher Realität und Pragmatik –, sondern vielmehr, dass es Sammlungen von Anwendungen, gleichsam Datenmarktplätzen, bedarf, in denen Benutzer ihre Daten und Anwendungen bis hin zu sehr komplexen Systemen und dokumentierten Arbeitsprozessen zur Verfügung stellen, so dass der Forschende aus einem reichen Sortiment schöpfen kann. Hierfür bedarf es jedoch klarer Metadatenstandards, die die Anwendungen hinreichend dokumentieren. Ferner ist ein Zusammenschluss von Forschenden zu ähnlichen Themenstellungen durchaus sinnvoll. In welchem institutionellen Rahmen dies nun geschehen mag, ist bis dato allerdings noch nicht abzusehen.

Neben der Nutzung einer Plattform zur beliebigen Kombination unterschiedlichster Anwendungen und Daten muss eine zentrale Funktion der Plattform in der Dokumentation und Visualisierung des Forschungsprozesses liegen. Wie kann eine genuin Digitale Argumentation, welche den Rahmen eines klassischen Textes verlässt, aussehen? Auch hier besteht noch ein weiterer Forschungsbedarf. Ein theoretischer Ansatz mag in der Konzeption des Deleuzschen Rhizoms³¹ liegen, eines mehrdimensionalen Netzwerkes aus Strängen, die neben den genutzten Daten auch Analysen und Visualisierungsformen beinhalten. Diese unterschiedlichen Argumentationsstränge stehen miteinander in einem räumlichen Verhältnis, sie kreuzen sich an manchen Stellen, führen wieder auseinander und bilden dadurch ein Gesamtsystem, in dem es verschiedene Bereiche unterschiedlichster Funktionen geben mag, „Segmentierungslinien, nach denen es geschichtet ist, territorialisiert, organisiert, bezeichnet, zugeordnet; aber auch Deterretorialisierungslinien, an denen es unaufhaltsam flieht“³², die aber dennoch ein Ganzes bilden. Der Rezipient ist dabei immer Teil des Systems und hat nun die Möglichkeit durch das Rhizom zu navigieren, die Stränge nachzuvollziehen, sie sich dadurch anzueignen und für sich zu manipulieren, Urteile zu fällen, neue Verknüpfungen zu bilden und so den hermeneutischen Prozess voranzutreiben. Welche Art der Visualisierung für dieses System adäquat erscheint, oder ob es wiederum einer Vielzahl von Visualisierungsmöglichkeiten bedarf, ist noch zu ermitteln. Die Grundlage für ein solches System und seine Visualisierung muss jedoch in

27 Eine Umsetzung findet sich unter: <http://www.meme.hokudai.ac.jp/WebbleWorldPortal/>.

28 Das gleiche gilt auch für Datenquellen, die nichts anderes als Anwendungen sind.

29 <http://www.tei-c.org/index.xml>.

30 <http://www.cidoc-crm.org/>.

31 Deleuze / Guattari (1977).

32 Ebd. S. 16.

jedem Falle in der Erstellung einer Ordnung zur vollständigen Beschreibung aller digitalen geisteswissenschaftlichen Forschungsprozesse sein, eine Ontologie des digitalen hermeneutischen Prozesses, ein weiteres Desiderat der Forschung.

Abschließend gilt als Quintessenz des Gesagten festzuhalten: Eine digitale Geisteswissenschaft muss das Augenmerk von den Daten und deren Erhebung weg und hin zum hermeneutischen Prozess des Erlebens, Verstehens und Vermittelns führen. Dabei gilt es von der Eindimensionalität des Textes hin zu einer Form des mehrdimensionalen multimedialen Argumentierens zu gelangen, ohne die Suggestivkraft von Bild und Animation dabei außer Acht zu lassen. Romantische Vorstellungen des in die historischen, literarischen, kunsthistorischen etc. Räume Eintauchens sind als solche zu benennen, aus einer wissenschaftlichen Perspektive abzulehnen, und durch die Prozesse des Urteilens, Bewertens und Begrifflich Fassens zu ersetzen, welche adäquat dokumentiert und präsentiert werden müssen. In welcher Art dies geschehen kann und welche technischen Umsetzungen dabei entstehen, das sind die zukünftigen Forschungsbereiche der Digital Humanities.

Literatur:

Arnold et. al. (2013): O. Arnold / W. Spickermann / N. Spyratos / Y. Tanaka (Hrsg.), Webble Technology: First Webble Summit, WWS 2013, Erfurt, Germany, June 3–5, 2013, Proceedings, Berlin / Heidelberg 2013.

Deleuze (1977): G. Deleuze / F. Guattari, Rhizom, Berlin 1977.

Dilthey (1970): W. Dilthey, Der Aufbau der geschichtlichen Welt in den Geisteswissenschaften, Frankfurt a.M. 1970.

Frühwald et. al. (1991): W. Frühwald / H. R. Jauß / R. Koselleck / J. Mittelstraß / B. Steinwachs (Hrsg.), Geisteswissenschaften heute, Frankfurt a.M. 1991.

Geiselberger / Moorstedt (2013): H. Geiselberger / T. Moorstedt (Hrsg.) (2013), Big Data - Das neue Versprechen der Allwissenheit, Frankfurt a.M. 2013.

Hobohm (2015): H.–Chr. Hobohm, Sammelrezension Digital Humanities: URL: <http://hsozkult.geschichte.hu-berlin.de/rezensionen/2015-1-001> [zuletzt gesehen am: 16.07.2015].

Herrmann (2015): F. Herrmann(2015), Sammelrez: Understanding / Debates in the Digital Humanities URL: <http://hsozkult.geschichte.hu-berlin.de/rezensionen/2015-1-012> [zuletzt gesehen am: 16.07.2015].

Kath /Schaal / Dumm (2015): R. Kath / G. Schaal / S. Dumm, New Visual Hermeneutics, in: Zeitschrift für germanistische Linguistik 43/1 (2015), S. 27–45.

Koselleck (1991): R. Koselleck, Wie sozial ist der Geist der Wissenschaft?, in: W. Frühwald / H. R. Jauß / R. Koselleck / J. Mittelstraß / B. Steinwachs (Hrsg.), Geisteswissenschaften heute, Frankfurt a.M. 1991, S. 112–241.

Koselleck (1991): R. Koselleck, The Power of Webble World and How to Utilize It, in: O. Arnold / W. Spickermann / N. Spyratos / Y. Tanaka (Hrsg.), Webble Technology: First Webble Summit, WWS 2013, Erfurt, Germany, June 3-5, 2013, Proceedings, Berlin / Heidelberg 2013, S. 31–55.

Lauer (2013): G. Lauer, Die Vermessung der Kultur. Geisteswissenschaften als Digital Humanities, in: H. Geiselberger / T. Moorstedt (Hrsg.) (2013), Big Data - Das neue Versprechen der Allwissenheit, Frankfurt a.M. 2013, S. 99–116.

Mittelstraß (1991): J. Mittelstraß, Die Geisteswissenschaften im System der Wissenschaft, in: W. Frühwald / H. R. Jauß / R. Koselleck / J. Mittelstraß / B. Steinwachs (Hrsg.), Geisteswissenschaften heute, Frankfurt a.M. 1991, S. 15–44.

Scheuermann (2014): L. Scheuermann, Lived time and space, in: Espace Temps.net. URL: <http://www.espacetemps.net/en/articles/lived-time-and-space/> [Zuletzt gesehen am: 16.07.2015].

Tanaka (2013): Y. Tanaka, Meme Media and Knowledge Federation: Past, Present, and Future, in: O. Arnold / W. Spickermann / N. Spyrtos / Y. Tanaka (Hrsg.), *Webble Technology: First Webble Summit, WWS 2013, Erfurt, Germany, June 3–5, 2013, Proceedings, Berlin / Heidelberg 2013*, S. 166–183.

Weitere Ressourcen (Zuletzt aufgerufen am 28.12.2015):

International Council of Museums Conceptual Reference Model (CIDOC CRM), URL: <http://www.cidoc-crm.org/>.

Meme Media World, URL: <http://www.meme.hokudai.ac.jp/WebbleWorldPortal/>.

Text Encoding Initiative (TEI), URL: <http://www.tei-c.org/index.xml>.

Autorenkontakt³³

Dr. phil. Leif Scheuermann

Karl-Franzens-Universität Graz
Lehrstuhl für Alte Geschichte und Altertumskunde
Email: leif.scheuermann@uni-graz.at

URL: <http://altegeschichte.uni-graz.at/de/mitarbeiten/>

³³ Die Rechte für Inhalt, Texte, Graphiken und Abbildungen liegen, wenn nicht anders vermerkt, bei den Autoren.

Die Visualisierung von Quellennetzwerken am Beispiel Plutarchs

Charlotte Schubert

Abstract: This paper focuses on the research on visualizations that support close and distant reading of textual data in the digital classics. The visualization and analysis of textual relationships will be thematised suggesting a taxonomy of classical textual sources as well as the surplus of potentially beneficial visualization approaches for research using graphs to analyze the textual relationship. Based upon the inherent textual content and using the texts' associated metadata network graphs (created with Gephi) are presented which allow an analysis of relationships among the textual contents of Plutarch's work.

Digitalität und Visualität¹

Die Digitalität als neue Repräsentationsform des Wissens fordert alle Wissenschaften heraus: Die von Franco Moretti so eindrücklich als ‚distant reading‘ beschriebenen quantitativen Analysemodelle haben dies deutlich gezeigt. Die Ergebnisse, dargestellt in der Form von Kurven, Karten und Stammbäumen, demonstrieren, wie graphische Visualisierungen eine schier unglaubliche Menge von Daten zusammenfassend darstellen können.² Dies ist aber auch schon als „Abschied von der Textologie des interpretativen ‚close reading‘“³ bezeichnet worden und zeigt, welche Skepsis diesen neuen Analysemethoden entgegengebracht wird. Kadens Kritik unterstellt, dass so keine Tiefenlektüre mehr möglich sei:

„Was die Auseinandersetzung dort aber zeigt, ist, dass die technisch quantitativen Erschließungsmethoden bislang ausgerechnet auf der Ebene der Bedeutung scheitern.“⁴

Diese massive Kritik ist nicht gerechtfertigt, denn auch bei Moretti geht es um textimmanente, relationale Beziehungen, denen selbstverständlich Bedeutungen inhärent sind. Aus der geographischen Visualisierung werden z.B. Muster erkennbar, die nicht offensichtlich und schon gar nicht über die lineare Lektüre zu erkennen sind, die aber Aufschluss über inhaltliche Bezüge geben können.

Gleichwohl bleibt die – nun schon öfter diskutierte – Frage, wie sich dieser Ansatz zu der traditionellen, historisch-philologischen Textanalyse verhält, die sich z.B. als Stilentwicklungsforschung oder Befundanalyse etabliert hat? Hier soll die These aufgestellt und auch anhand von Beispielen untermauert werden, dass sich aus der Multimodalität der Digitalität

1 An dieser Stelle sei sowohl den vielen Diskutanten aus den verschiedenen Seminaren gedankt, in denen ich die Gelegenheit hatte, über dieses Thema zu sprechen wie auch den Gutachtern, die mich vor einigen zu waghalsigen Schlüssen bewahrt haben. Möglicherweise noch verbliebene Ab- und Irrwege sind selbstverständlich allein mir zuzurechnen.

2 Moretti (2007); Manovich (2001).

3 Kaden (2009), S. 8.

4 A.a.O., S. 86.

neue Rezeptionsformen ergeben. Multimodalität umfasst heute nicht nur herkömmliche Text – Bild – Kombinationen, sondern diverse Verbindungen.

Visualisierung ist heute fast ubiquitär, sie wird spielerisch eingesetzt, explorativ angewendet und auch ihr ästhetischer Wert wird einbezogen. Mittlerweile werden zunehmend Überlegungen präsentiert, die zeigen, inwiefern bisher zu wenig berücksichtigt wurde, dass Visualisierung eine sehr abstrakte Form der Repräsentation darstellt und als eigene, semiotische Ebene betrachtet werden kann.⁵

Gemeinsam ist den unterschiedlichen Überlegungen, dass sie die Digitalität betonen, d.h. die andere Qualität von Daten als Medien gegenüber z.B. Texten und Bildern in gedruckter Form und vor allem auch auf die veränderte Größenordnung hinweisen, die in Massendaten vorliegt. Daraus ergeben sich andere Anforderungen der Verarbeitung und Analyse:

„The rationale is that data is too large for purely visual methods, requiring the use of data processing and mining; yet, the desired tasks are too exploratory for purely analytical methods, requiring the involvement of human analysts, using visualization as a medium for human interaction with the data.“⁶

Von der Seite der Informatik wird dies als *information overload* bezeichnet und Visualisierung wird als Mittel und neue Methode zur Auflösung desselben angeboten. Beschrieben wird diese als ‚Visual Analytics‘:

„On a grand scale, visual analytics provides technology that combines the strengths of human and electronic data processing. Visualisation becomes the medium of a semi-automated analytical process, where humans and machines cooperate using their respective, distinct capabilities for the most effective results.“⁷

Ein ebenfalls in der Diskussion um die Visualisierung verwendeter Begriff, der die Andersartigkeit der Vorgehensweise beschreiben soll, ist die Exploration:

„Visualisierungen transformieren, gewichten und filtern komplexe Daten und bringen sie dadurch in eine Form, die sie als Informationen erfassbar und interpretierbar machen.“⁸

Manovich hat diese Überlegungen weiterverfolgt und daraus einen Ansatz entwickelt, den er mit ‚Cultural Analytics‘ (als Analogiebildung zu Visual Analytics) umschreibt. Er bezieht sich auf die quantitative Auswertung und Analyse von visuellen Medien (Filme, Animationen, Videospiele, Comics), aber auch Printmedien (Bücher, Fotos), Kunst und weitere Medien. Sein Fokus liegt ganz auf der Dynamik der visuellen Repräsentationsformen, deren interaktive und animierte Formen er als das eigentlich Charakteristische dieser Medientypen ansieht.⁹ Visualisierung ist für Manovich eine darauf aufsetzende Repräsentationsform, die er als Umsetzung aus selbst nicht in visueller Repräsentation vorliegenden Daten definiert. Das bisher nicht Darstellbare wird nun darstellbar, sichtbar gemacht und für die weitere Analyse und Interpretation offengelegt. In dieser visuellen Repräsentation wird das nicht Darstellbare geordnet und angeordnet, neue Muster und Strukturen entstehen.

⁵ Vgl. Klawitter et al. (2012), S. 10 und insbesondere Manovich (2001) und (2008).

⁶ Endert et al. (2014), S. 2.

⁷ Keim et al. (2010) S. 2.

⁸ Bubenhofer (2014).

⁹ Manovich (2002).

Für Manovich liegt darin der Gegensatz zu dem Sublimen – dem Erhabenen als dem der direkten Sichtbarkeit nicht Zugänglichen – und daher beschreibt er die Visualisierung von Daten als ‚anti-sublime‘.¹⁰ Andererseits betont auch er, dass diese Art der Visualisierung eine neue Ebene der Abstraktion schafft.

Neben dieser auf Repräsentation fokussierten Sicht lassen sich mit guten Gründen weitere Positionen im Hinblick auf Visualisierungen einnehmen, die etwa von einer hermeneutischen Annäherung ausgehen:¹¹ Mit dem Terminus der ‚New Visual Hermeneutics‘ wird gegenüber dem Blick auf die Technologie oder die Dynamik der Repräsentation der gesamte Prozess vom Sampling und Aufbereiten der Daten (1.) über die algorithmenbasierte Analyse (2.) und die Visualisierung (3.) bis hin zur hermeneutischen Interpretation der Visualisierung (4.) beschrieben. Kath/Schaal/Dumm (2015) stellen hier die berechnete Forderung nach einem iterativen Kreislauf auf, der sehr viel mehr erfassen soll als die Technologie, die Exploration oder die Analyse der Repräsentation.

Zwei Aspekte dieser gegenwärtigen Diskussion sind jedoch bisher wenig diskutiert worden und sollen hier versuchsweise anhand eines konkreten Beispiels thematisiert werden:

1. Die unterschiedlichen Fachdisziplinen und Medien stehen in der Regel in jahrhundertalten Traditionen, in denen hochdifferenzierte Ordnungssysteme auf der Basis gedruckter Texte entwickelt wurden. Insofern ist schon Schritt 1 nach Kath/Schaal/Dumm (2015), das Sampling und Aufbereiten der Daten, nicht nur allgemein gesprochen komplex, sondern gerade die Rohdaten müssen nach fachspezifischen Kriterien der Auswahl, Kontrolle und ggf. Kritik angelegt werden – und dies je nach Text, Medien und Objekttyp mit dementsprechend unterschiedlichen Voraussetzungen. So scheint heutzutage die Datenaufbereitung im Bereich der digitalen Geisteswissenschaften eine Art Black Box zu sein: Fast nie wird in einschlägigen Veröffentlichungen Auskunft darüber gegeben, wie die einem präsentierten Ergebnis zugrundeliegenden Daten aufbereitet wurden – etwa in der Art wie es in naturwissenschaftlichen Veröffentlichungen mit dem Kapitel „Material und Methoden“ oder in SOPs (Standard Operation Protocol) üblich ist. So werden die einzelnen Schritte des Preprocessing i.d.R. nicht offengelegt und vor allem finden sich praktisch nie Überlegungen, die die der Datenaufbereitung zugrundeliegende Systematik analysieren und sie in ein Verhältnis zur jeweiligen fachspezifischen Wissensordnung setzen.¹²

2. Das Verhältnis von Visualisierung, Daten und Interpretation ist immer in Abhängigkeit von der Fragestellung und den daraus jeweils fachspezifisch geprägten Daten zu sehen. Es steht derzeit zur Diskussion, ob die epistemischen Prozesse, die sich in den geisteswissenschaftlichen Disziplinen zu stabilen Praktiken verdichtet haben, im Zeitalter der Digitalität unverändert weiterbestehen oder in das beschriebene, iterative Verfahren eingehen können – aber ebenso in der Diskussion ist (noch), inwiefern die algorithmenbasierten Analyseverfahren überhaupt mit genuin hermeneutischer Interpretation zusammengebracht werden können. Vor dem Hintergrund dieser sehr weitreichenden Fragen, die in einem einzelnen Beitrag weder erschöpfend noch wirklich angemessen behandelt werden können, sollen die folgenden Überlegungen eher als Prolegomena oder Anstoß für die weiterhin zu führende Diskussion zu verstehen sein. Entsprechend der oben genannten Forderung nach einem fachspezifischen Ansatz liegt der Fokus dieser Prolegomena auf dem Bereich der alttumswissenschaftlichen Textquellenanalyse.

10 Manovich definiert in dem Aufsatz vor allem seinen Begriff des ‚Anti-Sublimen‘ als einen Gegensatz zu dem ‚Sublimen‘, ohne allerdings auf den Begriff des Sublimen selbst näher einzugehen.

11 Kath/Schaal/Dumm (2015), S. 7.

12 Zur Kritik eines solchen Vorgehens in den Sozialwissenschaften: Koplenig (2014). Eine der seltenen Ausnahmen: Scharloth et al. (2013), S. 345–380 und Bubenhofer et al. (2015). Ein Beispiel für die zwar ansprechende, aber eher opake Visualisierung, insbesondere auch antiker Quellen: Andurand et al. (2015), S. 54–94.

Editionen und Texttraditionen: Die Taxonomie der Quellen

Die ‚Visuelle Textanalyse‘ erhebt den Anspruch,¹³ sowohl für die Analyse großer Dokumentensammlungen als ‚distant reading‘, als auch für ein sich der traditionellen Textinterpretation annäherndes Vorgehen einsetzbar zu sein.¹⁴ Sind größere Dokumentensammlungen vorhanden oder handelt es sich um einen Autor, von dem zahlreiche Werke zu einem Korpus zusammengestellt werden können, so dürften die o.g. Anforderungen nach dem iterativen Kreislauf in der Verbindung von algorithmenbasierter Auswertung und hermeneutischer Reflexion erfüllbar sein, wenn den Ansprüchen an Datenkritik Rechnung getragen wird: d.h. wenn eine Darstellung die Kriterien einer gesicherten Qualität der Daten, die Nachvollziehbarkeit der Auswertungswege und die Übertragbarkeit der Ergebnisse erfüllt. Besonders wichtig für die Akzeptanz der Visualisierung durch die Fachcommunities außerhalb der Informatik ist, dass die Ergebnisse mit mindestens einer anderen Methode, in den Textwissenschaften vorzugsweise aus dem historisch-philologischen Bereich, nachvollzogen werden können.

Für die Altertumswissenschaften mit ihrer 2500 Jahre alten Textgeschichte und -tradition ist dies allerdings für viele Bereiche wesentlich komplexer, da sehr spezifische Editionstechniken und -modelle entwickelt worden sind, die ihrerseits außerordentlich voraussetzungsreich sind. Von diesen Voraussetzungen sollen hier nur umrisshaft einige genannt werden, insofern sie für die weiteren Überlegungen zur Visualisierung relevant sind. Für die Repräsentation von Texten – soweit es sich um Quellentexte handelt –, haben sich bestimmte Rekonstruktionsmodelle für die Editionen etabliert. Die beiden wichtigsten Modelle sind das von K. Lachmann (v.a. in seinem *Commentarius* zur Lukrezäusgabe, Berlin 1850), das auf die Erstellung eines kritischen Textes mit dem Ziel der Annäherung an einen ‚Ur-Text‘ abzielt und das von Bédier (1928, S. 356), der die Auswahl einer Handschrift unter Angabe möglichst aller Varianten favorisierte. Die einzelnen Arbeitsschritte sind heute als feststehend in der Editionsarbeit definiert (*constitutio textus, recensio, examinatio, divinatio*).¹⁵

In den Altertumswissenschaften hat sich ein Fokus auf sog. Autorentexte herausgebildet und daher wird auch dort, wo kein Autor vorhanden ist, die Rekonstruktion eines Autors oder sogar auch eines Werkes angestrebt. Hinzu kommt die Vorstellung, dass der Text des Autors oder des Werkes als eine Konstante betrachtet werden kann,¹⁶ und auch dies hat dazu geführt, dass die Traditionsbildung hierarchisch gesehen und in der Folge auch so repräsentiert wurde. Deutlich wird dies in den Printeditionen in der hierarchischen Anordnung von Text und untergeordnetem kritischem Apparat: Das herkömmliche Layout-Muster der gedruckten Edition ist ein schwarz-weißer, typographisch gefasster Text im Hochformat, der zentral den Text und darunter den *Apparatus criticus* sowie – eher selten – den *Apparatus fontium* wiedergibt oder weitere Apparate wie einen Testimonien-, Similien- oder Zitatensystem.¹⁷ Die gedruckte Seite kann hier als Container für den Editionstext beschrieben werden: Der Container hat einen Rahmen mit Referenzpunkten (Seitenzahlen) für den Text und quasi einen Paratext mit dem dem Text untergeordneten, stark kodierten Apparat.¹⁸

Wenn die Rekonstruktion eines Autorentextes im Vordergrund steht, dann definieren die Bezüge zwischen dem einen Text, seinen Vorgängern, Nachbarn und Nachfolgern keine gleichrangige, sondern eine unterordnende und hierarchisierende Beziehung. Insbesondere die Stemmata, die die Beziehungen zwischen den Handschriften darstellen, visualisieren die Abhängigkeit von und Ausrichtung auf einen ursprünglichen Autorentext, der für diese Konzeption steht:

¹³ Von Seiten der Informatik dazu: Rohrdantz et al (2010).

¹⁴ Allg. Übersicht bei: Jänicke et al. (2015); Als konkretes Beispiel: Eder (2014); Schubert / Weiß (2015).

¹⁵ Maas (1960), S. 5ff.; West (1973), insb. zu Lachmann, S. 54; vgl. dazu Timpanaro (1963), S. 21.

¹⁶ Kramer (2011), S. 371ff.

¹⁷ Sahle (2013), S. 273.

¹⁸ Sahle a.a.O.

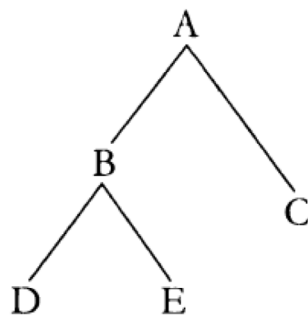


Abb. 1a: Stemma aus West (1973), S. 13;

Schema (über die kleinen Querstriche s. § 8i)

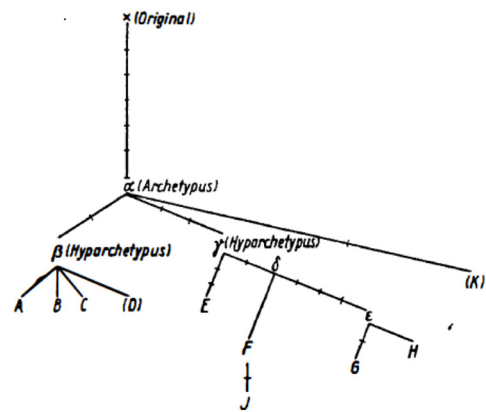


Abb. 1b: Stemma aus Maas (1960), S. 7

Auch die Beziehungen der Texte untereinander werden nach diesem Muster konzipiert:

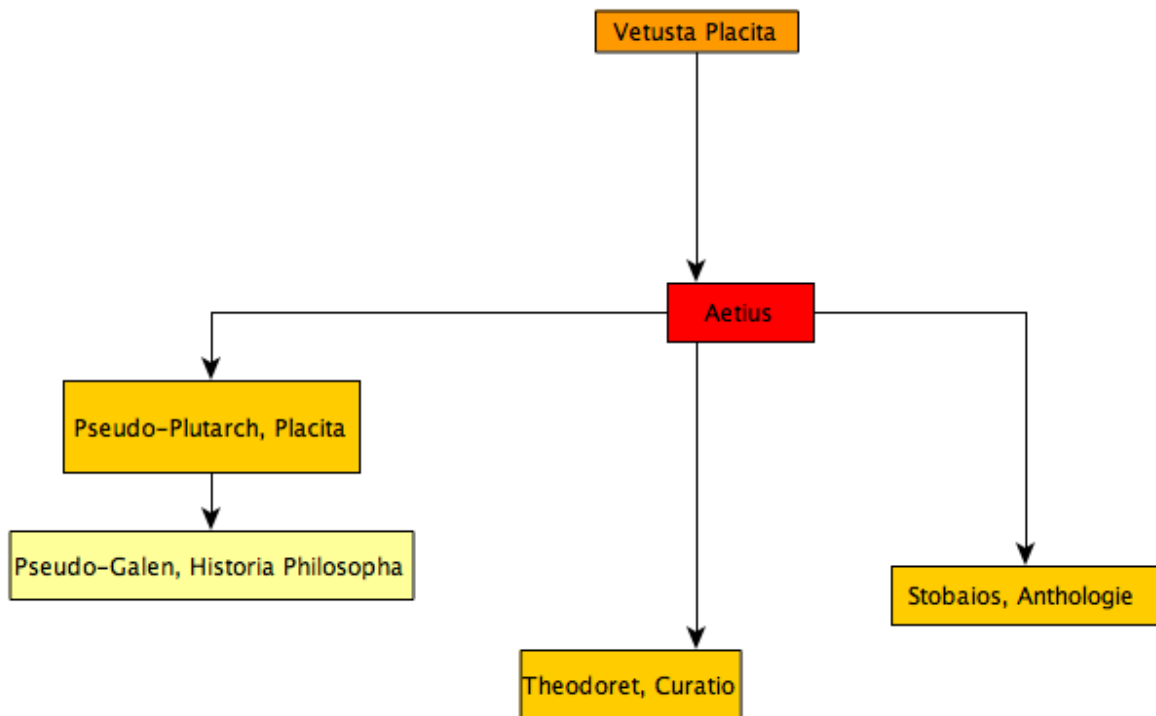


Abb. 2: Hermann Diels' Modell der antiken Doxographie (nach Mansfeld / Runia [1997], S. 4)

Für die antike Literatur gilt, dass fast alle Texte aus der Antike – mit der Ausnahme der Inschriften und einiger Texte und Fragmente auf Papyrus – nur aus der handschriftlichen Überlieferung der Spätantike und des Mittelalters erhalten sind, also in einer Zeit geschrieben wurden, die mehrere hundert oder sogar 1000 Jahre nach der Schreibzeit der jeweiligen Autoren liegt. Natürlich wird auch diese handschriftliche Tradition in den Abhängigkeitsdarstellungen berücksichtigt:

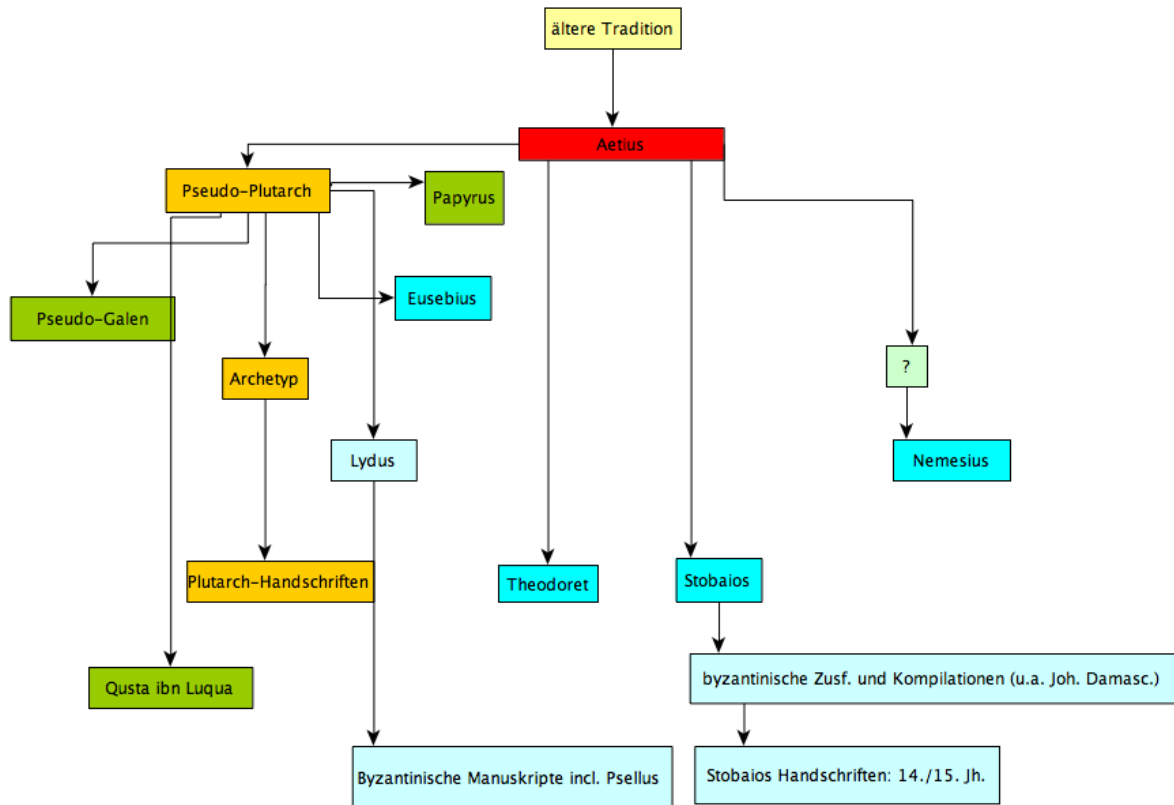


Abb. 3: Schema der Traditionsentwicklung in der antiken Doxographie nach Mansfeld / Runia (2010), S. 518

Angesichts der Digitalisierung und ganz anderer Repräsentations- und Auswertungsmöglichkeiten scheint alles daraufhin zu deuten, dass es durch die digitalen Medien zu einer Erweiterung des Blicks kommen kann, da die Präsentationsformen heute mehr Variationsreichtum haben und auch das Textverständnis pluralistischer wird.¹⁹

Hier soll der Versuch unternommen werden, diese Traditionsbildung sowohl begrifflich als auch visuell zu fassen.²⁰ Dazu wird die Vorstellung von einem Netzwerk der Quellen verwendet und darauf aufbauend werden entsprechende Visualisierungen entworfen. Um den Begriff des Quellennetzwerkes für eine Visualisierung nutzbar zu machen, soll hier in aller Kürze eine Taxonomie der Quellen skizziert werden, die Zeitphasen und Rezeptionscharakter nach den unterschiedlichen Perspektiven gliedert, ohne eine Hierarchie zu implizieren. In den Texten, d.h. den gedruckten Editionen, mit denen heute in der Regel gearbeitet wird, treten sehr verschiedene Beziehungen auf: Solche Beziehungen bestehen zwischen dem Text des eigentlichen Autors und darin enthaltenen Textpassagen, die für die Stimmen anderer Autoren stehen (Zitate, Paraphrasen, Anspielungen), sie bestehen aber auch im Verhältnis zu späteren Texten (Kommentaren, rezipierenden Texten) und modernen Editionen, die ihrerseits eine spezifische Sicht vertreten.

Diese Beziehungen sollen im Folgenden abstrakt als Schichten aufgegliedert werden. Als Ausgangspunkt gilt ein bestimmter Text, der im Schnittpunkt vorangegangener und nachfolgender Autoren steht. Konkret lassen sich die Beziehungen im Text zwischen Autor, anderen und früheren Autoren sowie Rezipienten späterer und heutiger Zeit nach den folgenden drei Schichten systematisieren:

¹⁹ Sahle a.a.O.

²⁰ Anders als bei McGann (1983) und McGann (1991) wird hier nicht nach Textproduktion, Textüberlieferung und Textrezeption unterschieden, denn die Überlieferungssituation ist für die antiken Texte sehr viel komplizierter und erfordert daher ein genaueres Einteilungsschema.

Eine erste Schicht ist der Text des Autors A: Hier liegt dann eine längere oder kürzere Abhandlung zu einem Thema als kompletter Text vor. Der Autor A ist namentlich bekannt und gibt in dem Text seine eigene Meinung wieder, benutzt aber auch verschiedene Quellen (B, C...) aus den Werken anderer Autoren, manche (doch sicher nicht immer alle) zitiert er namentlich, andere paraphrasiert oder kommentiert er.

Eine zweite Schicht ist die Rezeption des Textes von Autor A in späteren Texten, deren Autoren (D, E, ...) ebenfalls namentlich bekannt sind: In diesen späteren Texten der Autoren D, E, ... wird der Text von Autor A zitiert oder paraphrasiert, kommentiert (z.B. in reinen Kommentarwerken oder in Scholia), dies allerdings auch nicht immer unter Nennung des Namens von Autor A bzw. des dem zitierten Werk zugehörigen Titels.

Aufgrund der vielen verlorengegangenen Werke antiker Autoren ist man seit dem 19. Jahrhundert mit großer Intensität dazu übergegangen, sog. Fragmentsammlungen dieser verlorenen Autoren zusammenzustellen: Dabei werden Zitate, Paraphrasen, Kommentierungen und z.T. auch Textpassagen ohne Namensnennungen in reinen Rekonstruktionseditionen (d.h. Textpassagen B, C, ... aus Autor A werden zu Fragmenten der Autoren F, G, ...) versammelt. Diese Sammlungen repräsentieren eine dritte Schicht, indem durch Zuweisung von Textpassagen ganz unterschiedlichster Provenienz ganze Werkkontexte und Autoren konstruiert werden. In diesen Fragmenteditionen liegt ein fast unermessliches Verdienst im Hinblick auf die Sichtbarmachung und Rekonstruktion von Verlorenem – allerdings liegt darin auch im Hinblick auf die Analyse von Traditionswegen in der antiken Überlieferung die Gefahr der mangelnden Transparenz. Der autoritative Charakter der Textedition, der sich bei einer Fragmentedition besonders stark auswirkt, verdeckt die komplexe Schichtung der Texte. Dass diese verschiedenen Schichten der Überlieferung in Editionen, Kommentaren, insb. in den verschiedenen Formen der Apparate selten vollständig dargestellt sind, hat ganz unterschiedliche Gründe, die auch in dem Verständnis von Autorschaft und Konventionen der Editionsarbeit begründet sind, wodurch die Analyse der Schichtung aber noch zusätzlich erschwert wird.

Die Visualisierung hat hier nun das Ziel, die Texte der Autoren A – G als Netzwerk darzustellen, so dass die unterschiedlichen Beziehungen und Abhängigkeiten abgebildet werden können. Dies reduziert nicht die Komplexität der hier abstrakt beschriebenen Beziehungen, bietet aber eine Form der Übersicht, die sonst nicht zu erreichen ist. Eine solche Netzwerkvisualisierung soll demnach ein Quellennetzwerk mit folgenden Schichten darstellen:

- der Text von A wird in seiner Beziehung zu B, C, ... dargestellt, d.h. die in einem Text vom Autor selbst verwendeten Quellentexte werden als Relation abgebildet (erste Schicht)
- der Text von A wird außerdem in seinem Verhältnis zu D, E, ... dargestellt, d.h. im Verhältnis zu den Texten späterer Zeit, in denen Passagen aus dem Text des Autors A erhalten sind (zweite Schicht)
- der Text von F, G, ..., d.h. die modernen Fragmentsammlungen mit ihren rekonstruierten Werkkontexten werden ebenfalls abgebildet (dritte Schicht).

Visualisierung eines Quellennetzwerkes am Beispiel von Plutarch

Aus der hier aufgestellten Taxonomie der Beziehungen von Texten und Autoren ergeben sich folgende Anforderungen an die Visualisierung, die die Quellen Plutarchs mit den beschriebenen drei Schichten als Netzwerk darstellen sollen:

1. Die Quellen, die Plutarch in einem bestimmten Werk benutzt hat, müssen identifiziert werden und aufgrund der von der Editionspraxis vorgegebenen Autorenezuordnung einem Autor zugewiesen werden.
2. Die Rezeption Plutarchs ist darzustellen, d.h. wer zitiert ihn später oder zitiert wie er Texte aus der Plutarch vorangegangenen Zeit.
3. Die Quellensammlungen, die Fragmente aus Plutarchs Texten oder anderen Quellen zu Werkkontexten oder Autorenwerken rekonstruieren, müssen sichtbar werden.

Als Visualisierungsmethode ist hierfür die graphische Netzwerkanalyse als die heute gängigste Möglichkeit, Beziehungen zwischen Texten zu visualisieren, gewählt worden.²¹ Graphische Layouts auf der Grundlage von force based Algorithmen (wie etwa Circular Layout und ForceAtlas2 Layout in dem Netzwerkvisualisierungsprogramm Gephi²²) zeigen die Autoren und Texte selbst als Knoten, wobei die Größe der Knoten die Häufigkeit einer Nennung repräsentiert und die Nähe der Knoten sowie die Dicke der Kanten die Stärke der Beziehung anzeigen. Die Quellentexte sind hier auf der Grundlage von String-Matching-Algorithmen in Textpassagen zerlegt worden, so dass alle Parallelpassagen innerhalb der zu vergleichenden Texte identifiziert werden können.²³ Auf dieser Grundlage werden den Textpassagen die Autorennamen, Werknamen, Datierungen und Genreklassifizierungen als Metadaten zugeordnet. Diese werden in Form von CSV-Tabellen extrahiert und können so in Gephi als Netzwerk visualisiert werden.

Für das hier vorgestellte Beispiel werden die Viten des Themistokles, Kimon und Perikles aus den Parallelbiographien des Plutarch ausgewählt,²⁴ die nach dem beschriebenen Verfahren mit Gephi sowohl in dem Circular Layout und in dem Layout ForceAtlas2 visualisiert werden (Abb. 4–6).

In Abbildung 4 (Circular Layout) ist die Auswertung auf Plutarchs Lebenszeit begrenzt, d.h. es werden nur Textpassagen ausgewertet, die Autoren zugeordnet werden können, die vor oder während seiner eigenen Lebenszeit geschrieben haben. Für die drei zugrunde gelegten Viten zeigt sich deutlich, dass die Hauptquellen, die Plutarch in den drei Viten verwendet hat, einerseits Werke der Komiker waren (Telekleides, Kratinos, Eupolis, Platon Comicus), die als Zeitgenossen von Themistokles, Kimon und Perikles im 5. Jahrhundert v. Chr. schrieben oder zumindest noch nah an der Zeit dieser Protagonisten waren. Viel stammt auch aus den Biographien des Hermippos (3. Jh. v. Chr.), der selbst sehr quellenreich schrieb. Entsprechend der ersten Schicht in der hier zugrunde gelegten Taxonomie lässt sich dies visualisieren:

²¹ Jänicke et al. (2015).

²² Bastian et al. (2009); Jacomy et al (2014).

²³ Ausführliche Beschreibung des Verfahrens bei: Büchler et al. (2010) und [Büchler et al. \(2011\)](#); zur fachwissenschaftlichen Anwendung: [Schubert, 2010](#) und [dies., 2012](#) sowie dies. / Weiß (2015). Zu den in eAQUA verwendeten Algorithmen, Methoden der sprachspezifischen Verarbeitung: s. <http://www.eaqua.net/> die Beschreibungen unter Dokumentation/Wissensdatenbank/; zu der Zitations-Analyse: Dokumentation/Wissensdatenbank/Neue Methoden in den Geisteswissenschaften/Zitations-Analyse.

²⁴ Textgrundlage sind die Ausgaben und Texte des TLG-E.

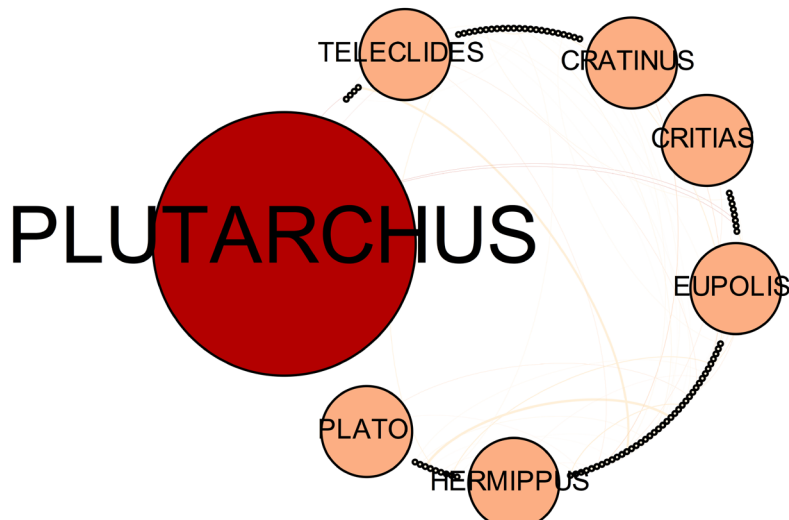


Abb. 4: Plutarchs Quellen in den Viten des *Themistokles*, *Kimon* und *Perikles* visualisiert mit Gephi nach Autor und Werk im Circular Layout, eingegrenzt auf die Zeit bis Plutarch (bis 125 n. Chr.)

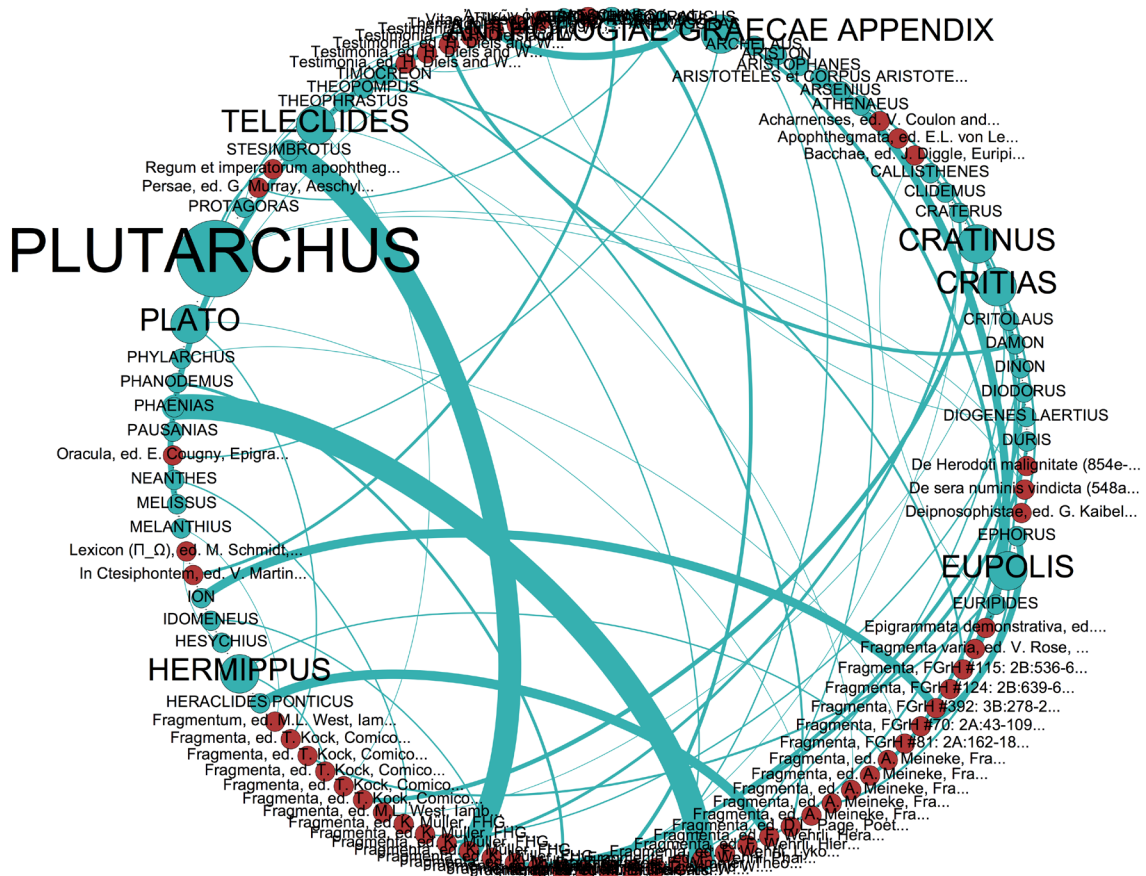


Abb. 5: Plutarchs Quellen in den Viten des *Themistokles*, *Kimon* und *Perikles* visualisiert mit Gephi nach Autor und Werk im Circular Layout, ohne zeitliche Eingrenzung auf die Zeit bis zu Plutarch (bis 125 n. Chr.)

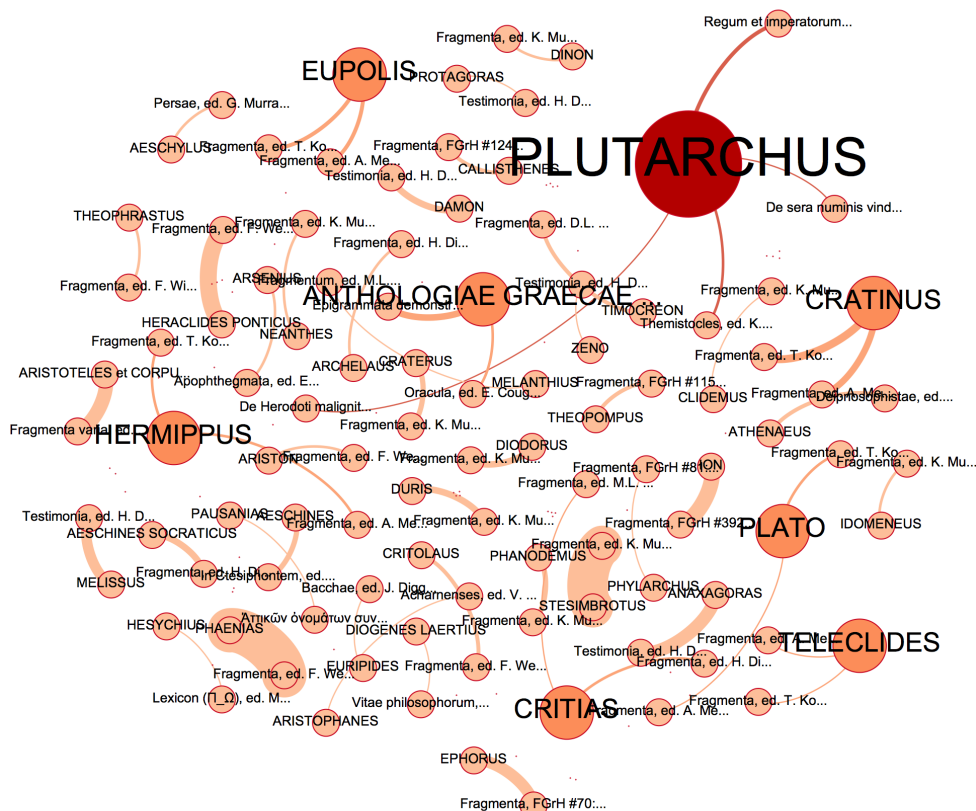


Abb. 6: Plutarchs Quellen in den Viten des *Themistokles*, *Kimon* und *Perikles* visualisiert mit Gephi im Layout ForceAtlas2, ohne eine zeitliche Eingrenzung: Die Knoten zeigen die Quellenautoren, die Plutarch verwendet hat und die späteren Texte, in denen Passagen aus seinem Werk enthalten sind; die Kanten zeigen die Verbindungen zu den Werken selbst.

Erweitert man die Auswertung, indem man auf die Eingrenzung der Texte aus der Zeit bis zu Plutarchs eigener Lebenszeit verzichtet, so sieht man noch deutlicher, wie hoch die Zahl der Fragmente bzw. der nur fragmentarisch erhaltenen Autoren ist, die hier als Quellen identifiziert werden können (entsprechend der zweiten und dritten Schicht der oben beschriebenen Taxonomie).

Gleichzeitig ist gut zu erkennen, wie groß die Zahl dieser Textpassagen ist, die Plutarch in seinen Werken verwendet und die sich auch in hoher Zahl in der *Anthologia Graeca* finden.²⁵

In der folgenden Abbildung (Abb. 7) ist durch unterschiedliche Einfärbung der Kanten visualisiert, welche Texte, die sowohl bei Plutarch wie auch in Texten, die ihm als Quellen vorlagen als auch in späteren Werken – nach der Lebenszeit des Plutarch datiert – Textparallelen zu seinem Werk aufweisen: Durch die Markierung nach Autoren (grün) und Werken (rot) ist sofort die überwältigende Zahl von Autoren erkennbar, deren Texte als Fragmente – d.h. aus anderen Werken entnommene Textpassagen – in einem rekonstruierten Werk ediert wurden. Aber auch spätere Autoren, in denen parallele Textpassagen zu den drei hier ausgewählten Viten des Plutarch enthalten sind – ohne dass dies Zitate aus Plutarch sein müssen – lassen sich sofort identifizieren wie etwa Athenaeus (grün) mit der Kante zu seinem Werk *Deipnosophistai* (rot) und Diogenes Laertius (grün) mit der Kante zu seinem Werk *Vitae Philosophorum* (rot) [beide mit roten Hinweisfeilen markiert].

²⁵ Die *Anthologia Graeca* ist eine Sammlung von Gedichten, Versen und Epigrammen von der hellenistischen bis zur byzantinischen Zeit, die ihrerseits aus zwei großen Sammlungen besteht (*Anthologia Palatina* und *Anthologia Planudea*), die im Mittelalter zusammengestellt wurden.

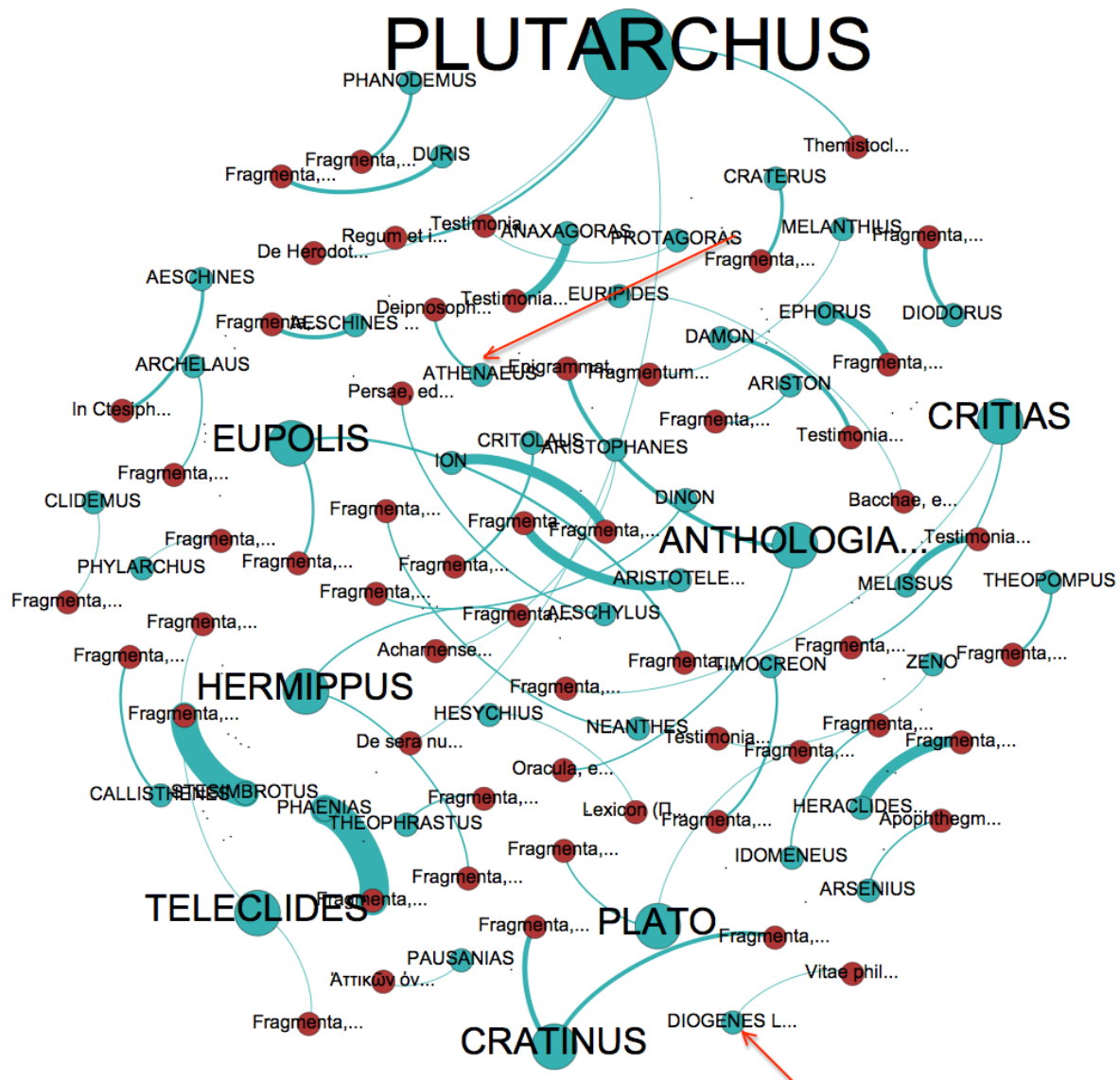


Abb. 7: Visualisierung der von Plutarch in den Viten des Themistokles, Kimon und Perikles häufig verwendeten Textpassagen aus seinen Quellen und deren Auftreten bei späteren Autoren (grüne Knoten) mit den Kanten zu den jeweiligen Werken (rote Knoten)

Auch für den gesamten Werkkontext eines Autors lassen sich diese Visualisierungen nach der hier verwendeten Taxonomie darstellen. Für das gesamte erhaltene Werk Plutarchs ist ein Subkorpus seiner Werke angelegt worden, das mit dem gesamten Bestand des digitalisierten Korpus griechischer Literatur verglichen wurde.²⁶ Die – allgemein formulierte – Ausgangsfrage ist, welche Textpassagen in seinem Werk zu finden sind, die auch in den Texten anderer Autoren auftreten (Schicht 1 und 2). Daran knüpft sich die Frage an, wo diese Textpassagen in der späteren Rezeption sowie den heutigen Editionen auftreten (entsprechend der dritten Schicht der oben beschriebenen Taxonomie).

Zum Vergleich ist in der nächsten Abbildung (Abb. 8) zuerst die zeitliche Eingrenzung auf die Lebenszeit Plutarchs gewählt worden, so dass erkennbar wird, in welchen Werken Plutarchs – jetzt bezogen nur auf das Werk von Plutarch selbst – sich die am häufigsten anderen Autoren zuzuweisenden Textpassagen finden.

²⁶ S.o. Anm. 19 und 20.

Mit der zeitlichen Eingrenzung und der Filterfunktion von Gephi, die die Anzeige auf die wichtigsten Beziehungen beschränkt,²⁷ zeigen sich im Ergebnis die von Plutarch hauptsächlich – bezogen auf sein eigenes Gesamtwerk – zitierten Autoren (Homer, Euripides, Aeschylus, Theophrast, Empedokles, entsprechend Schicht 1) und es wird gleichermaßen sichtbar, in welchen seiner Werke sich diese Zitate hauptsächlich finden: In den *Placita philosophorum* (auch einem Pseudo-Plutarch zugeschrieben, s. dazu unten), den Fragmenten, den *Quaestiones convivales*, der *Consolatio ad uxorem* und v.a. in den Biographien *Solon*, *Theseus* und *Perikles*!

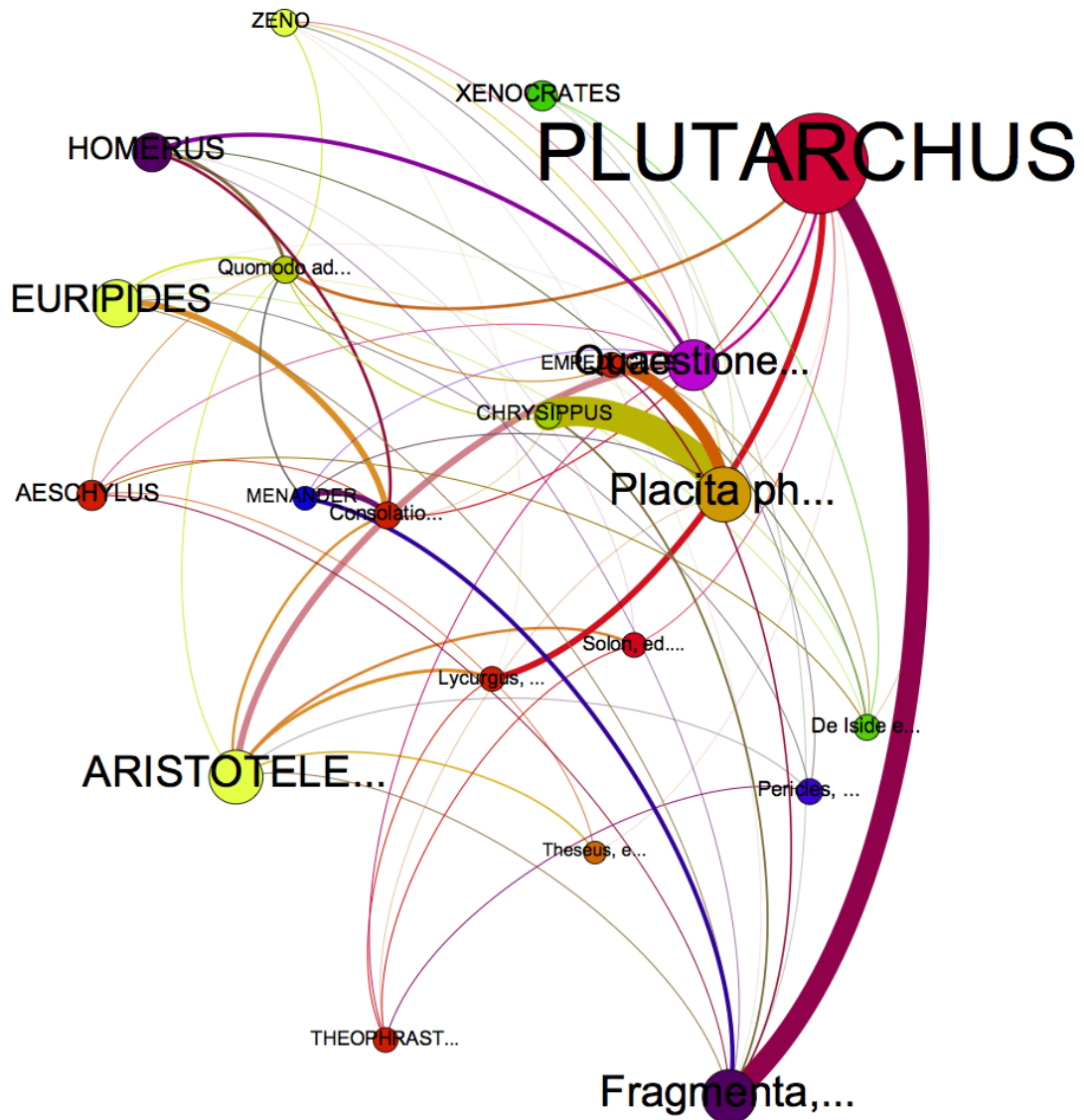


Abb. 8: Quellennetzwerk Plutarchs mit zeitlicher Eingrenzung auf die Zeit bis zu Plutarchs Lebenszeit (125 n. Chr.)

In Plutarchs erhaltenem Gesamtwerk (ca. 128 Werke) zeigt sich das Quellennetzwerk natürlich deutlich komplexer. In der Abb. 9 ist eine Einstellung ohne zeitliche Eingrenzung gewählt, d.h. alle Texte und Autoren werden angezeigt, sowohl diejenigen, die vor als auch diejenigen, die nach ihm zu datieren sind. Auch hier ist wie in Abb. 8 ein Filter gewählt worden, so dass von den Autoren, die entweder von Plutarch verwendet wurden, oder aus Plutarch zitieren oder dieselben Quellen wie er benutzt haben, nur diejenigen mit den häufigsten Nennungen dargestellt

²⁷ Anwendung der Filter (Topology/ Giant Component/ Degree Range 19–91), um die weniger relevanten Knoten auszublenden, die für die ausgewählten Beziehungen nur eine untergeordnete Rolle spielen. Die Titel der Werke sind hier nur gekürzt verwendet, um die Übersichtlichkeit und Lesbarkeit der Abbildung zu ermöglichen.

sind: Stobaios, die Suda, Eustathius, die Scholien und die – nicht einem Autor zuzuweisende - Sammlung der *Anthologia Graeca*; gleichermaßen verweisen die Bezüge auf Parallelen in den Zitaten aus Euripides, Homer und Aristoteles hin, ein wenig erstaunlicher Befund, da es sich hierbei um die am häufigsten in der Antike zitierten Autoren handelt.

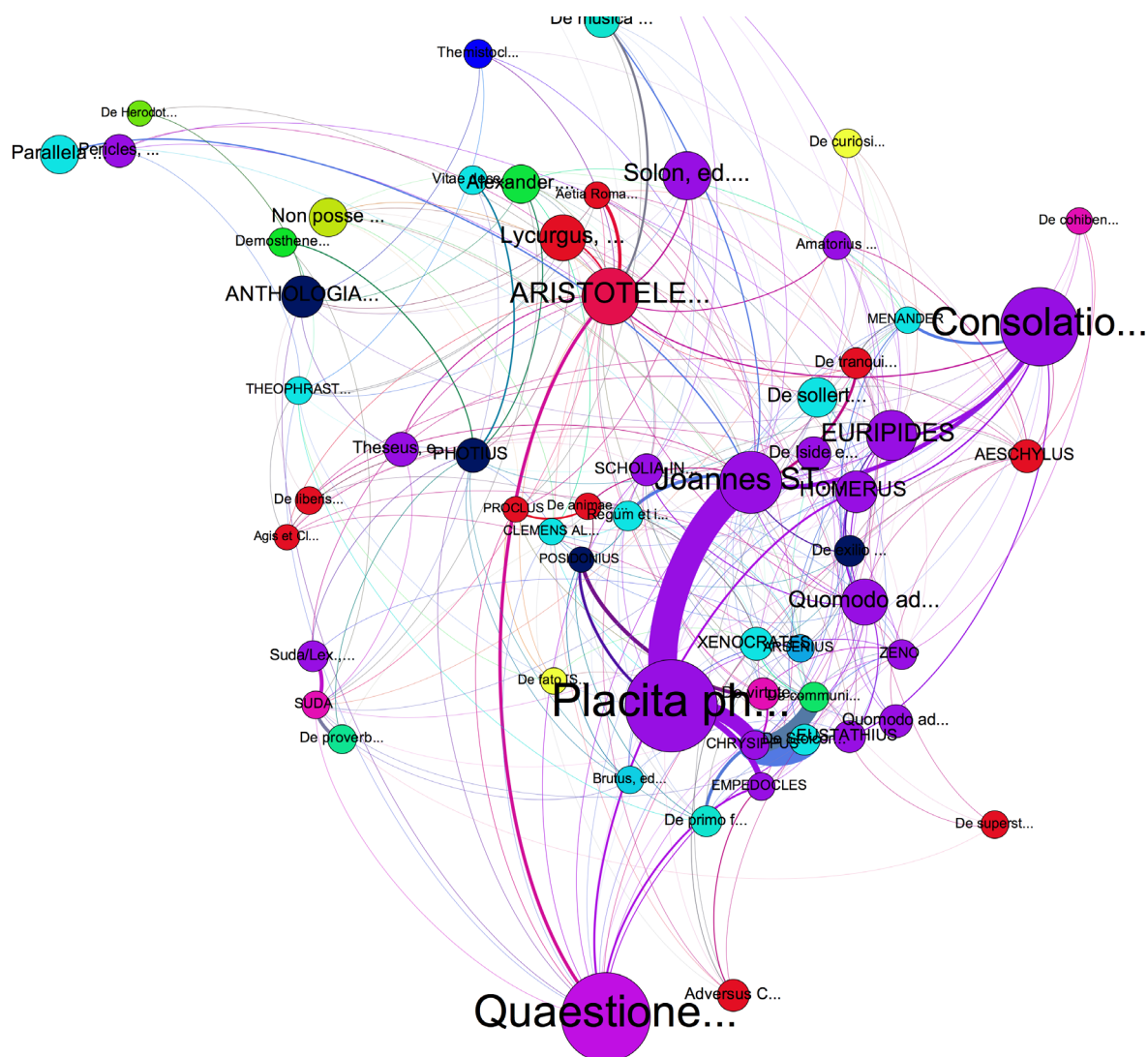


Abb. 9: Visualisierung der von Plutarch in seinem gesamten Werk häufig verwendeten Textpassagen sowie deren Auftreten in den Werken späterer Autoren

In der Datenaufbereitung sind für das Werk Plutarchs die *Placita Philosophorum* miteinbezogen worden. Diese *Placita Philosophorum* gelten seit Hermann Diels nur als Epitome aus einem größeren Werk, das Diels einem gewissen „Aëtius“ zugeschrieben hat und daher wurden die *Placita Philosophorum* von ihm Plutarch ab- und einem Pseudo-Plutarch zugesprochen.²⁸ Diese ebenso geniale wie spekulative Konstruktion von „Aëtius“ durch Diels beherrscht seit her alle Überlegungen zu den *Placita Philosophorum*: Der Autor wird als Pseudo-Plutarch bezeichnet, obwohl die Schrift Bestandteil des Konvoluts an Schriften der *Moralia* Plutarchs gewesen ist.²⁹ Die Visualisierung zeigt hier die Parallelen zwischen der Anthologie des Johan-

²⁸ Zu der heutigen Diskussion um die Identifizierung des Namen, der nur aus Theodoret's *Curatio* (Curatio II,95; IV,31; V,16) erhalten ist: Lebedev (1988 / 2013); Bremmer (1998), S. 154–160; Bottler (2014), S. 15ff.

²⁹ Die neuere Diskussion bei: Mansfeld / Runia (1997); dies. (2009) und (2010).

nes Stobaios und den *Placita Philosophorum*.³⁰ Dies ist an sich nicht überraschend und war bekanntlich einer der Gründe, der zu den Untersuchungen mit dem beschriebenen Ergebnis bei Diels geführt hat.

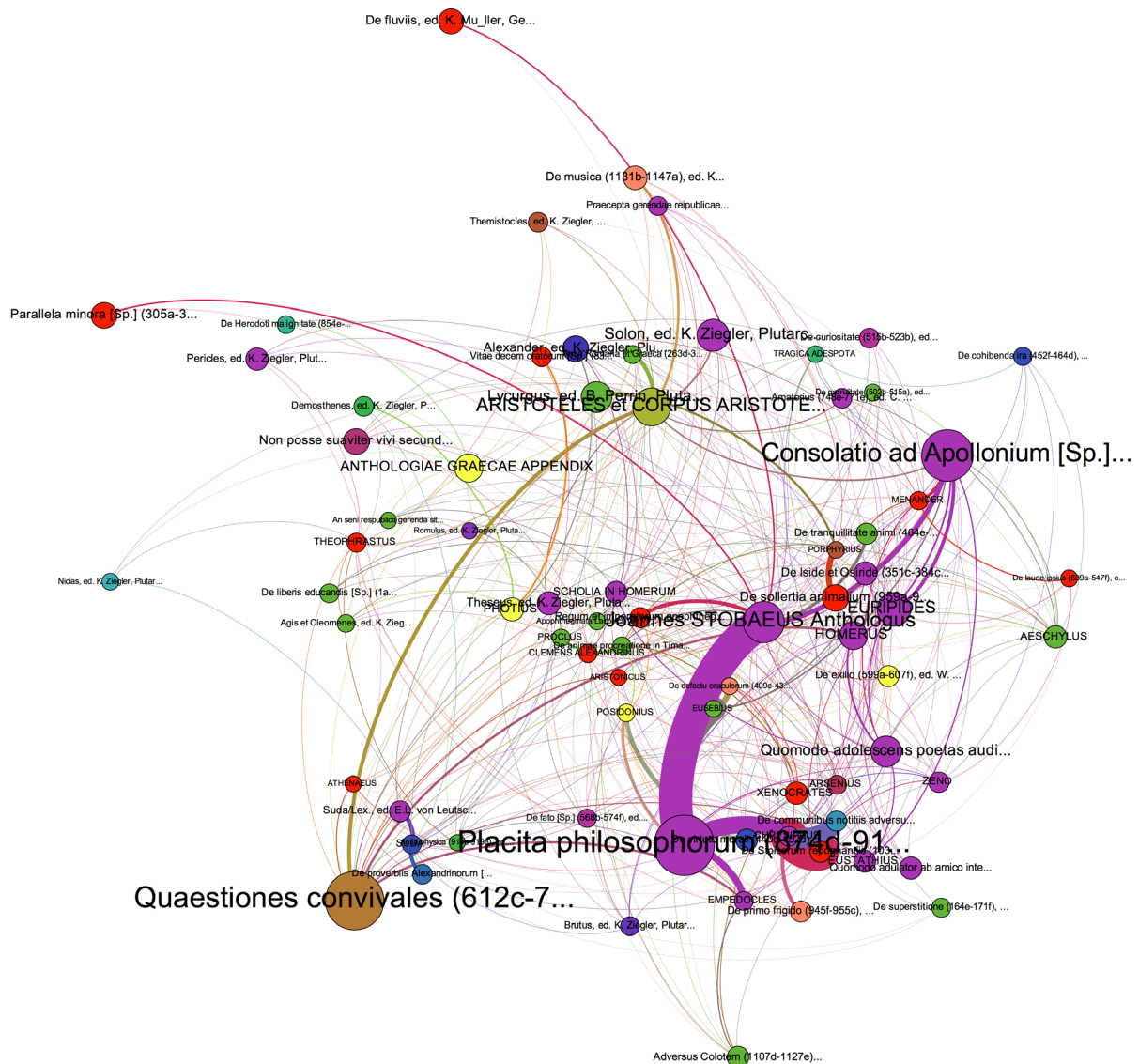


Abb. 10: Visualisierung der von Plutarch in seinem gesamten Werk häufig verwendeten Textpassagen sowie deren Auftreten in den Werken späterer Autoren (ohne Zugrundelegung eines Subkorpus für Plutarchs Gesamtwerk)

Noch deutlicher wird diese Beziehung, wenn man für den Vergleich, wie in Abb. 10 gezeigt, weitere statistische Berechnungen in Gephi durchführt (Modularity, Closeness Centrality).³¹ Im Detail zeigen sich dann sehr aufschlussreiche Parallelen, d.h. eine Beziehung zwischen den *Placita Philosophorum*, Empedokles und dem Werk *Adversus Colotem* des Plutarch:

³⁰ Diels (1879). Hier für die Datenaufbereitung verwendete Ausgabe der *Placita Philosophorum*: J. Mau, *Plutarchi Moralia*. Bd. V, Fasc. 2, Pars 1, Leipzig 1971 und für Stobaios die Ausgabe von Wachsmuth/Hense, Berlin 1884–1912 (Repr. 1958).

³¹ Zu den Berechnungsverfahren: Gewählt wurden insb. Modularity (Modularität), weil sie die Teile des Netzwerks hervorhebt, und Closeness Centrality (Nahzentralität), da sie die Stellung eines Knotens im Verhältnis zum gesamten Netzwerk zeigt. Anschließend sind die gleichen Filter angewendet worden wie für die Visualisierung in Abb. 9.

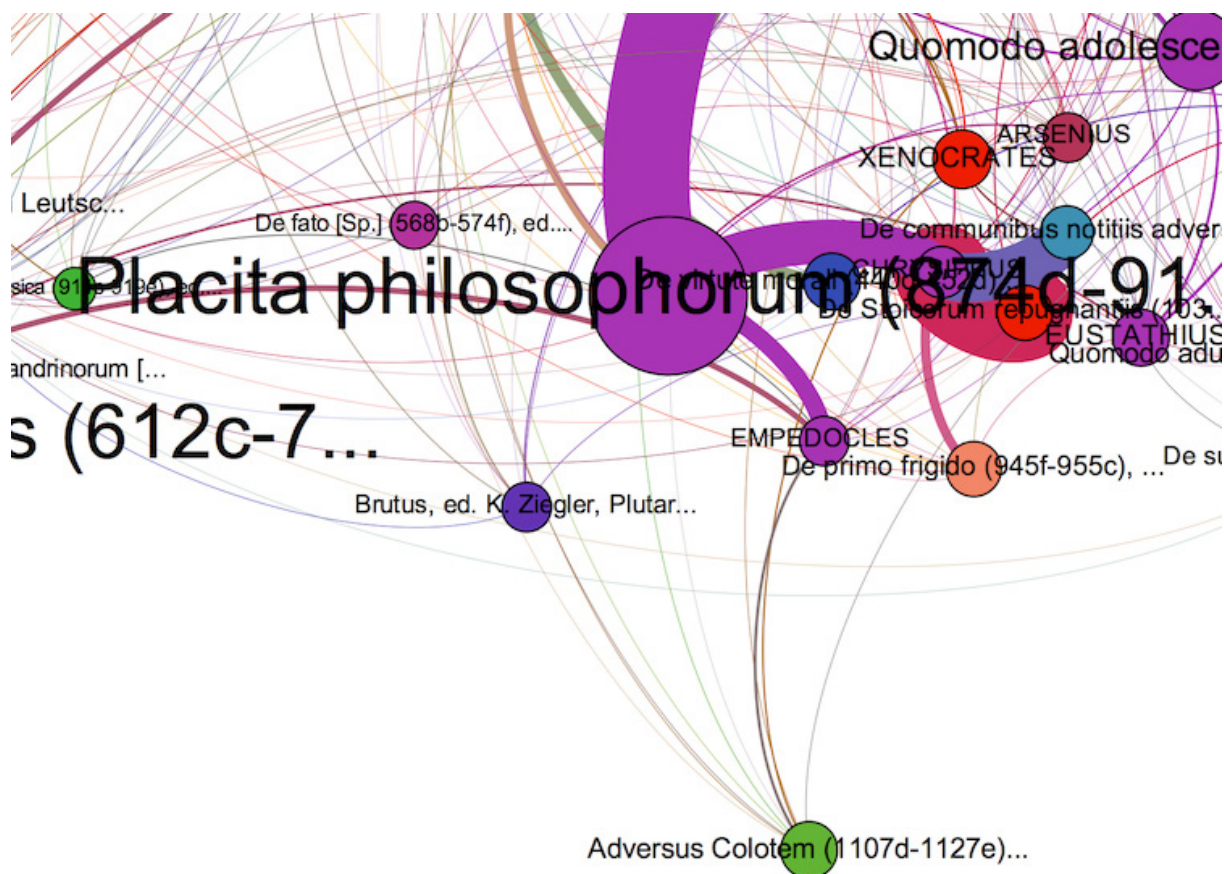


Abb. 11: Vergrößerte Detailansicht aus Abb. 10 mit der Beziehung zwischen den *Placita Philosophorum*, Empedokles und der Schrift Plutarchs *Adversus Colotem* (1107d–1127e)

Bei Plutarch findet sich in dem Werk *Adversus Colotem* eine zentrale Passage aus dem Fragment Empedokles B8. Dieses Fragment wird in der gesamten griechischen Literatur nur in diesem Werk Plutarchs und in den *Placita Philosophorum* (885 d6–9) mit allen 4 Versen in dieser Länge zitiert (hier nach Adv.Col.1111f.):

ἄλλο δέ τοι ἐρέω· φύσις οὐδενὸς ἔστιν ἐκάστου
 θνητῶν, οὐδέ τις οὐλομένη θανάτοιο γενέθλη·
 ἀλλὰ μόνον μῖξις τε διάλλαξις τε μιγέντων
 ἔστι, φύσις δ' ἐπὶ τοῖς ὀνομάζεται ἀνθρώποισι.

Stobaios zitiert diese Passage nicht, insofern verweist auch die Visualisierung für diesen Fall auf eine enge Verbindung zwischen den *Placita Philosophorum* und Plutarch unabhängig von Stobaios. Einerseits bestätigt dies die vielfach schon geäußerte Meinung, dass der Autor der *Placita Philosophorum* (sei es nun Plutarch oder ein anderer Autor) und Stobaios auf zum Teil unterschiedliche Vorlagen zurückgegriffen haben müssen.³² Andererseits gibt die visualisierte Form der Beziehung dieser Textpassagen, basierend auf der N-Gramm-Analyse und nicht auf einer editorischen Zuordnung, einen sehr sichtbaren und schnell zu erkennenden Hinweis auf einen textuellen Zusammenhang, der sich in der wissenschaftlichen Diskussion und der edito-

³² Die Differenzen zwischen dem Zitat in *Adversus Colotem* und in den *Placita* lassen sich demgegenüber leicht darauf zurückführen, dass Plutarch in *Adversus Colotem* entweder aus dem Gedächtnis zitiert hat oder entsprechend dem Kontext der Schrift *Adversus Colotem* durch eine modifizierte Wortwahl zugespitzt hat, während die *Placita* offenbar auf Exzerpierrearbeit beruhen. Dazu ausf. Verf. Die Arbeitsweise Plutarchs: Notizen, Zitate und *Placita* (erscheint in: Rheinisches Museum für Philologie).

rischen Kontextualisierung aufgrund der oben beschriebenen editorischen Praxis – in diesem Fall die lange Editions-geschichte von und Diskussion um ‚Aëtius‘ und die antike Doxographie – zu einer dermaßen komplexen Situation entwickelt hat, dass er nur schwer zu erkennen ist.³³

Fazit

Zweifellos reduzieren diese Formen der Visualisierung die Komplexität im Hinblick auf eine Analyse der Quellen nicht, sondern erhöhen sie eher. Andererseits ermöglichen die Visualisierungen eine Form der Übersicht und der Repräsentation, die einerseits zusammenfassend ist und andererseits Einsichten in die Struktur von Beziehungen gibt, die so in ihrer Gesamtheit nicht immer offen liegen, wenngleich sie in den Editionen, Kommentaren oder Monographien in der Quellenanalyse meist enthalten sind.

Diese Form der Verdichtung stellt eine neue und andere Möglichkeit zur Einsicht und Auswertung bereit, die durch ihre Visualität auf einer anderen Ebene angesiedelt ist als die textuelle Repräsentation von Quellen in Printpublikationen. Aus diesem Grund ist es erforderlich, die Voraussetzungen der Anordnung und Zuordnung sorgfältig zu überdenken und gegebenenfalls im Hinblick sowohl auf das Medium wie auch im Hinblick auf die Analyse-methode selbst bisherige Taxonomien neu zu gliedern. Der hier vorgelegte Versuch zeigt, dass Visualisierungen durchaus zur Erweiterung der Editions- und Interpretationsmethoden beitragen können und es sollte erwogen werden, diese im Rahmen von digitalen Editionen zur Erweiterung der Editionstechniken heranzuziehen.

³³ Vgl. dazu die in Anm. 25 genannten Werke von Mansfeld / Runia.

Literatur:

Andurand et al (2015): A. Andurand / L. Jégou / M. Maisonobe / R. Sigrist, Les mondes savants et leur visualisation, de l'Antiquité à aujourd'hui, in: La visualisation des données en histoire, Revue Histoire et Informatique / Geschichte und Informatik, Band 18/19 (2015), S. 54–94.

Bastian et al. (2009): M. Bastian / S. Heymann / M. Jacomy, Gephi: an open source software for exploring and manipulating networks. International AAAI Conference on Weblogs and Social Media 2009.

(URL: <http://www.aaai.org/ocs/index.php/ICWSM/09/paper/view/154>, zuletzt aufgerufen am 24.09.15)

Bédier (1928): J. Bédier, La tradition manuscrite du «Lai de l'ombre». Réflexions sur l'art d'éditer les anciens textes, Romania 54 (1928), S. 161–196.

Bottler (2014): H. Bottler, Pseudo-Plutarch und Stobaios: Eine synoptische Untersuchung, Göttingen 2014.

Bremmer (1988): J. Bremmer, Aëtius, Arius Didymus and the Transmission of Doxography, Mnemosyne 51 (1988), S. 154–160.

Bubenhofer et al. (2015): N. Bubenhofer / J. Scharloth, / D. Eugster, Rhizome digital: Datengeleitete Methoden für alte und neue Fragestellungen in der Diskursanalyse, in: Zeitschrift für Diskursforschung, Sonderheft Diskurs, Interpretation, Hermeneutik (1. Beiheft, 2015), S. 144–172.

Bubenhofer (2014): N. Bubenhofer, GeoCollocations – Diskurse zu Orten: Visuelle Korpusanalyse, in: Sondernummer Mitteilungen des Deutschen Germanistenverbandes 1/2014: Korpora in der Linguistik – Perspektiven und Positionen zu Daten und Datenerhebung, S. 45–59.

(URL: <http://www.bubenhofer.com/publikationen/PreprintBubenhofer2014.pdf>, zuletzt aufgerufen am 24.09.2015)

Büchler et al. (2010a): M. Büchler / A. Geßner / T. Eckart: Unsupervised Detection and Visualisation of Textual Reuse on Ancient Greek Texts, in: JDHC 1, 2 (2010).

(URL: <https://letterpress.uchicago.edu/>, zuletzt aufgerufen am 11.09.2015)

Büchler et al. (2010b): M. Büchler / A. Geßner / G. Heyer / T. Eckart, Detection of Citations and Textual Reuse on Ancient Greek Texts and its Applications in the Classical Studies: eAQUA Project. Presented at: DH2010: Digital Humanities 2010 Conference; 2011 Jul 05–10; London 2010.

(URL: <https://www.linguistik.hu-berlin.de/institut/professuren/korpuslinguistik/events-en/nehdfg/pdf/buechler-reuse>, zuletzt aufgerufen am 11.09.2015)

Diels (1879): H. Diels, Doxographi Graeci, Berlin 1879.

Eder (2014): M. Eder, Stylometry, network analysis, and Latin literature. In Proceedings of the Digital Humanities, Lausanne 2014. S. 457–458.

(URL: <http://dharchive.org/paper/DH2014/Poster-324.xml>, zuletzt aufgerufen am 24.09.2015)

Endert et al. (2014): A. Endert / M. Shahriar Hossain / N. Ramakrishnan / Chr. North / P. Fiaux / Chr. Andrews, The human is the loop: new directions for visual analytics, *Journal of Intelligent Information Systems* 42 (2014), S. 1–25.

Jacomy et al. (2014): M. Jacomy / T. Venturini / S. Heymann / M. Bastian, ForceAtlas2, a Continuous Graph Layout Algorithm for Handy Network Visualization Designed for the Gephi Software, in: *PLoS One* 2014.
(DOI: 10.1371/journal.pone.0098679)

Jänicke et al. (2015): S. Jänicke / G. Franzini / M. F. Cheema / G. Scheuermann, On Close and Distant Reading in Digital Humanities: A Survey and Future Challenges. A State-of-the-Art (STAR) Report, in: (Proceedings) EuroVis 2015: The EG/VGTC Conference on Visualization, Cagliari 2015, S. 25–29.
(DOI:10.2312/eurovisstar.20151113)

Kaden (2009): B. Kaden, Alles auf einer Karte. Franco Morettis Modellvorschlag für die Literaturwissenschaft und was ihm folgen könnte, *Libreas* 14 (2009) (Rezension zu: F. Moretti, Kurven, Karten, Stammbäume. Abstrakte Modelle für die Literaturgeschichte. Frankfurt / Main 2009).
(URL: <http://libreas.eu/ausgabe14/013kad.htm>, zuletzt aufgerufen am 09.09.2015; URN: urn:nbn:de:kobv:11-10097095)

Kath/Schaal/Dumm (2015): R. Kath / G. Schaal / S. Dumm, New Visual Hermeneutics, in: *Zeitschrift für germanistische Linguistik* 43/1 (2015), S. 27-45.

Koplenig (2014): A. Koplenig, Prinzipielle Probleme der Anwendung statistischer Signifikanztests in der Korpuslinguistik. Vortrag, gehalten auf der 1. Jahrestagung der Digital Humanities im deutschsprachigen Raum (DHD 2014).
(URL: https://www.conftool.pro/dhd2014/index.php?page=browseSessions&path=adminSessions&form_session=1, zuletzt aufgerufen am 09.05.2014)

Kramer (2011): J. Kramer, *Von der Papyrologie zur Romanistik*, Berlin / New York 2011

Keim et al. (2010): D. Keim / J. Kohlhammer / G. Ellis / F. Mansmann, *Mastering the Information Age – Solving Problems with Visual Analytics*, Eurographics Association, Goslar 2010.

Lebedev (1988 / 2013): A. Lebedev, Aëtius, Arius Didymus and the Transmission of Doxography, *Philosophie et culture. XVIIe Congrès mondial de philosophie. Actes/Proceedings, vol. III, Montréal 1983 (Éditions Montmorency 1988; rev. Fassung mit Postscript 2013)*, S. 813–817.
(URL: http://iph.ras.ru/uplfile/histph/publ/lebedev_did-the-doxo.pdf, zuletzt aufgerufen am 24.09.2014)

Maas (1960): P. Maas, *Textkritik*. 4. Auflage. B. G. Teubner, Leipzig 1960.

Mansfeld / Runia (1997): J. Mansfeld / D. Runia, *Aëtiana. The Method and Intellectual Context of a Doxographer Vol. I: The sources*, Leiden 1997.

Mansfeld / Runia, (2009): J. Mansfeld / D. Runia, *Aëtiana. The Method and Intellectual Context of a Doxographer. Vol. II (2 Bde.): The Compendium*, Leiden 2009.

Mansfeld / Runia (2010): J. Mansfeld / D. Runia, *Aëtiana. The Method and Intellectual Context of a Doxographer Vol. III: Studies in the Doxographical Traditions of Ancient Philosophy*, Leiden 2010.

McGann (1983): J. McGann, *A Critique of Modern Textual Criticism*, Chicago 1983.

McGann (1991): J. McGann, *The Textual Condition*, Princeton 1991.

Moretti (2007): F. Moretti, *Graphs, Maps, Trees*, London / New York 2007.

Manovich (2001): L. Manovich, *The Language of New Media*, The MIT Press 2001.

Manovich (2008): L. Manovich, *Data Visualisation as New Abstraction and Anti-Sublime*, Berlin August 2002, <http://manovich.net/index.php/projects/tag:Article/display:list> (zuletzt aufgerufen am 10.09.2015); gedruckt in: L. Manovich, „Data Visualization as New Abstraction and as Anti-Sublime.“ *Small Tech: The Culture of Digital Tools*, in: B. Hawk, D. Reider, & O. Oviedo (Eds.), *Electronic Mediations Ser. 22*, Minneapolis 2008.

Rohrdantz (2010): C. Rohrdantz / S. Koch / C. Jochim / G. Heyer / G. Scheuermann / T. Ertl / H. Schütze / D. A. Keim, *Visuelle Textanalyse. Informatik-Spektrum 33, 6 (2010)*, S. 601–611. (DOI: 10.1007/s00287-010-0483-x)

Sahle (2013): P. Sahle, *Digitale Editionsformen. Teil I: Das typografische Erbe, Teil II: Befunde, Theorie und Methodik, Teil III: Textbegriffe und Recodierung*, Norderstedt 2013. (URN: 7 = urn:nbn:de:hbz:38-53510, 8 = urn:nbn:de:hbz:38-53523, 9 = urn:nbn:de:hbz:38-53534)

Scharloth et al. (2013): J. Scharloth / D. Eugster / N. Bubenhofer: *Das Wuchern der Rhizome. Linguistische Diskursanalyse und Data-driven Turn*, in: D. Busse / W. Teubert (Eds.), *Linguistische Diskursanalyse. Neue Perspektiven*. Wiesbaden 2013, S. 345–380.

Schubert (2010): Ch. Schubert, *Zitationsprofile, Suchstrategien und Forschungsrichtungen*, in: *Das Portal eAQUA – Neue Methoden in der geisteswissenschaftlichen Forschung I*, Working Papers Contested Order No.1, Oktober 2010, S. 42–55. (URL: <http://journals.ub.uni-heidelberg.de/index.php/eaqua-wp>, zuletzt aufgerufen am 11.09.2015; DOI: <http://dx.doi.org/10.11588/ea.2010.0>)

Schubert (2012): Ch. Schubert, *Zitate und Fragmente: Die kulturelle Praxis des Zitierens im Zeitalter der Digitalisierung*, in: *Das Portal eAQUA – Neue Methoden in der geisteswissenschaftlichen Forschung III*, Working Papers Contested Order No.7, Nov 2012, S. 3–30. (URL: <http://journals.ub.uni-heidelberg.de/index.php/eaqua-wp>, zuletzt aufgerufen am 11.09.2015; DOI: <http://dx.doi.org/10.11588/ea.2012.2>)

Schubert / Weiß (2015): Ch. Schubert / A. Weiß, *Die Hypomnemata bei Plutarch und Clemens: Ein Textmining-gestützter Vergleich der Arbeitsweise zweier ‚Sophisten‘*, *Hermes 143 (2015)*, S. 447–471.

Timpanaro (1963): S. Timpanaro, *La Genesi del metodo del Lachmann*, Firenze 1963.

West (1973): M. L. West, *Textual Criticism and Editorial Technique, applicable to Greek and Latin texts*, Stuttgart 1973.

Weitere Ressourcen (zuletzt aufgerufen am 18.02.2016):

eAQUA, Extraktion von strukturiertem Wissen aus Antiken Quellen für die Altertumswissenschaft, URL: <http://www.eAQUA.net>

Gephi - The Open Graph Viz Platform,
<http://gephi.github.io/>

Autorenkontakt³⁴

Prof. Dr. Charlotte Schubert

Universität Leipzig
Historisches Seminar
Lehrstuhl für Alte Geschichte
Email: schubert@uni-leipzig.de

URL: <https://www.gko.uni-leipzig.de/historisches-seminar/seminar/alte-geschichte/professur.html>

³⁴ Die Rechte für Inhalt, Texte, Graphiken und Abbildungen liegen, wenn nicht anders vermerkt, bei den Autoren.

Projektankündigung

Mehrsprachiges Online Wörterbuch zum Fachwortschatz der Verwaltungssprache des griechisch-römisch-byzantinischen Ägypten

Reinhold Scholl

Abstract: The multilingual online dictionary for the language of the administration in the Greco-Roman-Byzantine Egypt replaces, updates and expands the 100-year-old work of Friedrich Preisigke, links to the digital full text, allows the composition of special subject groups and provides a basis for the history of a technical language.

1 Ausgangslage

Seit Bestehen des Faches Papyrologie ist diese jüngste altertumswissenschaftliche Grundlagendisziplin durch ihre innovativen und kreativen Fachvertreter stets darum bemüht, nützliche Hilfsmittel für ihre Arbeiten zu entwickeln.

Bei der Schaffung der analogen Hilfsmittel ist besonders der Telegraphendirektor in Straßburg, Friedrich Preisigke, zu nennen, der gleich mehrere dieser heute noch unersetzlichen Hilfsmittel entwickelt hat. Diese waren im vordigitalen Zeitalter der Geisteswissenschaften einzigartig und sind es bis heute.

Zu diesen Hilfsmitteln zählen:

- Fachwörter des öffentlichen Verwaltungsdienstes Ägyptens: In den griechischen Papyrusurkunden der ptolemäisch-römischen Zeit, Göttingen 1915.

- Namenbuch, enthaltend alle griechischen, lateinischen, ägyptischen, hebräischen, arabischen und sonstigen semitischen und nichtsemitischen Menschnennamen, sowie sie in griechischen Urkunden (Papyri, Ostraka, Inschriften, Mumienchildern usw.) Ägyptens sich vorfinden: Mit einem Anhang von Enno Littmann, enthaltend die in diesem Namenbuche vorkommenden abessinischen, arabischen, aramäischen, kanaanäischen und persischen Namen, Heidelberg (Selbstverl. d. Hrsg.), 1922.

- Wörterbuch der griechischen Papyrusurkunden: mit Einschluß der griechischen Inschriften, Aufschriften, Ostraka, Mumienchilder usw. aus Ägypten, hrsg. von Friedrich Preisigke, u. Emil Kiessling u.a., I, A-K, Berlin 1925; II, Λ-Ω, Berlin 1927; III, besondere Wörterliste, Berlin 1931; IV, A-Ziv., Berlin 1944 —. Supplement 1 (1940 – 1966), hrsg. von E. Kiessling, Amsterdam 1969 – 1971. Supplement 2 (1967 – 1976), hrsg. von . H.-A. Rupprecht und A. Jördens, Wiesbaden 1991. Supplement 3 (1977 – 1988), ed. H.-A. Rupprecht und A. Jördens, Wiesbaden 2000.

Diesen Hilfsmitteln lag unter anderem der Gedanke zugrunde, dass die Alltagssprache, die wir in den griechischen Papyri antreffen, sich durchaus von der Literatursprache der Antike unterscheidet. Neben dieser Alltagssprache haben sich auch eigene Fachsprachen entwickelt, wie z.B. die Verwaltungssprache. Bestimmte Wörter haben in einem bestimmten Kontext und bisweilen in einer bestimmten Zeit jeweils eine völlig andere Bedeutung.

Im Zeitalter der digitalen Geisteswissenschaften sind ursprünglich analoge Wörterbücher digitalisiert und einige davon auch online gestellt worden. Die neuesten und aktuellsten auf dem Gebiet der Altertumswissenschaften sind folgende:

Logeion¹ bietet die gleichzeitige Suche und Anzeige in mehreren griechischen Wörterbüchern (DGE, LSJ sowie Spezialwörterbüchern) und Lateinwörterbüchern (Lewis-Short, DuCange). Es gibt außerdem Angaben zu Häufigkeiten und gängigen Wortverbindungen. Es erfolgt auch eine Verlinkung zu dem Lemma in dem ursprünglichen Wörterbuch.

DGE en línea² online ist die digitale Ausgabe der sieben bisher als Druck veröffentlichten Bände des Wörterbuchs Griechisch-Spanisch (Diccionario Griego-Español), das bisher von α – ἔξαρως geht. Es ist in XML kodiert.

The Online Liddell-Scott-Jones Greek-English Lexicon³ in Verbindung mit dem kostenpflichtigen Thesaurus Linguae Graecae⁴ verlinkt die Belegstellen für die Bedeutung eines Wortes mit dem Volltext in dieser Textdatenbank. D.h. die Idee, ein Wörterbuch mit einer Volltextdatenbank zu verbinden, ist nicht neu. Wer allerdings keine Lizenz für TLG besitzt, hat von diesem Service nichts. Bei diesem Wörterbuch wurde von Perseus die Sprache XML und das Dokumentenformat TEI/DTD übernommen und weitergeführt. Die Supplemente der neuesten Druckausgaben wurden jedoch nicht berücksichtigt.

LSJ:⁵ Henry George Liddell. Robert Scott. A Greek-English Lexicon, revised and augmented throughout by Sir Henry Stuart Jones, with the assistance of Roderick McKenzie. Oxford. Clarendon Press 1940. Eine Verlinkung der Belegstellen erfolgt auch hier, aber auf das frei verfügbare Textcorpus der literarischen Texte bei Perseus. Da bei Perseus aber noch nicht alle griechischen Texte digital zur Verfügung stehen, hat man in vielen Fällen nur die Angabe der Belegstelle, aber keine Verlinkung zum Volltext. Zudem ist dieses Wörterbuch in dieser Form nicht mehr aktuell, sondern es muss auf das oben genannte LSJ Online zurückgreifen.

Das LBG⁶ (Lexikon zur byzantinischen Gräzität) ist ebenfalls online zugänglich, unabhängig von einer TLG Lizenz. Es werden lediglich die Belegstellen angegeben, ohne Verlinkung zum Volltext und die Datenbank ist in XML ausgeführt.

Ein für Papyrologen nützliches und online zugängliches Spezialwörterbuch ist LMPG online:⁷ Es handelt sich um die digitale Version des Buches *Léxico de magia y religión en los papiros mágicos griegos* von Luis Muñoz Delgado aus dem Jahr 2001 als Anhang V des DGE. Es

1 <http://logeion.uchicago.edu/about.html>.

2 <http://dge.cchs.csic.es/xdge>.

3 <http://www.tlg.uci.edu/lj/#eid=1&context=lsj>.

4 <http://www.tlg.uci.edu/>.

5 <http://www.perseus.tufts.edu/hopper/text?doc=Perseus%3atext%3a1999.04.0057>.

6 <http://stephanus.tlg.uci.edu/lbg>.

7 <http://dge.cchs.csic.es/lmpg>.

basiert auf XML im Einklang mit den Empfehlungen für die TEI-P5 Struktur. Auch hier sind die Bedeutungen bereits mit den Belegstellen verlinkt, allerdings nicht zu dem Volltext. Es wird auch nicht aktualisiert.

Ebenfalls hilfreich ist PAWAG⁸ (Poorly Attested Words in Ancient Greek).

Das Manko bei allen diesen Online-Wörterbüchern liegt darin, dass zwar die Belegstellen aufgeführt werden, diese aber nicht alle direkt im Volltext einsehbar sind. Ein weiterer Nachteil ist der, dass bisweilen zusätzlich ein modernes Wörterbuch benötigt wird, um Nuancen, die in einer modernen Sprache bzw. Übersetzung ausgedrückt werden, zu verstehen. Hinzu kommt, dass die altgriechischen Wörterbücher im vorletzten bzw. Anfang des letzten Jahrhunderts erstellt wurden, wobei die modernen Sprachen sich seitdem weiterentwickelt haben, so dass manche Übersetzungsvorschläge veraltet oder gar mittlerweile missverständlich sind. Preisigkes Nachschlagewerke sind außerdem in Frakturschrift gedruckt (vgl. Abb. 1), was für ausländische und jüngere Wissenschaftlerinnen und Wissenschaftler und Studentinnen und Studenten die Benutzung erschwert. Schließlich sind einige dieser ambitionierten Wörterbücher nicht vollständig, und eine Fertigstellung ist bislang nicht vorgesehen.

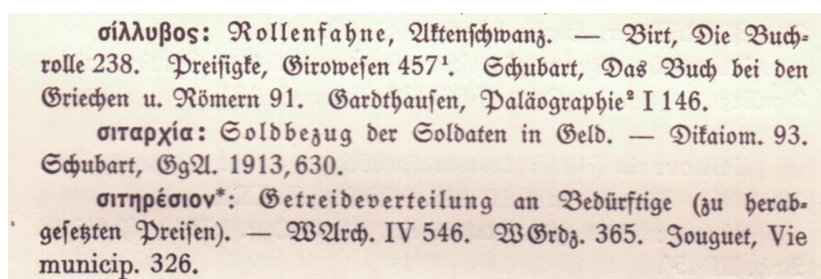


Abb. 1: Ausschnitt aus Friedrich Preisigke, Fachwörter des öffentlichen Verwaltungsdienstes Ägyptens: In den griechischen Papyrusurkunden der ptolemäisch-römischen Zeit, Göttingen 1915, S. 156.

Schon mit der Eroberung Alexanders des Großen im Jahr 332 v. Chr. hat in Ägypten die griechische Sprache als Verwaltungssprache Einzug gehalten und sie ist es auch über die römische Zeit hinweg bis zur arabischen Eroberung des Landes 641 n. Chr. geblieben. Das sind fast tausend Jahre. In dieser Zeit hat sich natürlich die griechische Sprache weiterentwickelt, auch unter dem Einfluss des Ägyptischen in Form des Demotischen und Koptischen sowie unter dem des Lateinischen.⁹

Die griechische Verwaltungssprache hat sich als Fachsprache im Rahmen der neu entstehenden Verwaltung in Ägypten langsam entwickelt, etabliert und weiterentwickelt. So sind mit dem Beginn der römischen Herrschaft in Ägypten über das lateinische Fachvokabular ebenfalls Einflüsse auf die griechische Verwaltungssprache zu konstatieren. Es war notwendig, da die Verwaltungssprache innerhalb Ägyptens weiterhin Griechisch blieb, nun griechische Termini für römische Institutionen zu bilden. Eine Untersuchung dazu liegt von Hugh J. Mason¹⁰ vor, der die entsprechenden Begriffe auflistet und erklärt. Bisweilen wurden in diesem Prozess lateinische Begriffe als Lehnwörter einfach transliteriert, manchmal übersetzt und gelegentlich ein in etwa entsprechender Ausdruck im Griechischen übernommen, der damit

⁸ <http://www.aristarchus.unige.it/pawag/searchForm.php>.

⁹ Vgl. Lex.Lat.Lehnw. = Lexicon der lateinischen Lehnwörter in den griechischsprachigen dokumentarischen Texten Ägyptens mit Berücksichtigung koptischer Quellen, hrsg. von I.-M. Cervenka-Ehrenstrasser zusammen mit J. Diethart, Wien 1996 ff. (MPER N.S. XXVII); bisher publiziert Fasc. 1, alpha (1996) und 2, beta-delta (2000); siehe auch H. Förster, Wörterbuch der griechischen Wörter in den koptischen dokumentarischen Texten, Berlin 2002 und Database and Dictionary of Greek Loanwords in Coptic (DDGLC): <http://research.uni-leipzig.de/ddglc/>.

¹⁰ Greek terms for Roman institutions: A lexicon and analysis, Toronto 1974 (=American Studies in Papyrology 13).

eine Bedeutungsveränderung durchmachte (transliteriert: *πραίτωρ* – *praetor*; übersetzt: *imperator* – *αὐτοκράτωρ* oder : *fasces* – *ῥάβδοι*). Da ein Großteil dieser lateinischen Termini mit Verwaltung zu tun hat, werden diese Wörter ebenfalls berücksichtigt, falls sie nicht schon sowieso bei Preisigke im Fachwörterbuch aufgenommen sind.

Für die römische Zeit wird das umfangreiche Werk von Nicolas Hohlwein¹¹ herangezogen, unter anderem auch für die französische Fachterminologie. Bereits Preisigkes Fachwörterbuch hatte diejenigen Lemmata, die sich auch in Hohlweins umfangreichen Werk finden, mit einem hochgestellten Stern (*) versehen (Vgl. Abb. 1).

2 Ziel des Forschungsprojektes

Das vom Sächsischen Ministerium für Wissenschaft und Kunst auf Vorschlag der Sächsischen Akademie der Wissenschaften zu Leipzig geförderte Projekt ist als ein Pilotprojekt für ein allgemeines Wörterbuch für papyrologische Texte gedacht. Das neue Fachwörterbuch wird zunächst das oben bereits genannte „Fachwörterbuch“ von Friedrich Preisigke ersetzen, das im Jahr 2015 100 Jahre alt wurde. Seitdem sind viele neue Papyrusfunde hinzugekommen und die Zahl der Texte hat sich rasant vermehrt. Das neue, digitale Fachwörterbuch soll in Bezug auf die Belegstellen und die weiterführende Literatur für die jeweiligen Bedeutungen – manchmal hat ein und dasselbe Wort mehrere Bedeutungen – aktualisiert, sprachlich modernisiert und um den Verwaltungswortschatz des byzantinischen Ägypten erweitert werden, der bisher nicht systematisch in diesem Fachwörterbuch erfasst worden ist. Außerdem wird die Übersetzung der Lemmata mehrsprachig, das heißt zunächst auf Deutsch, Französisch, Englisch, Italienisch und Spanisch erfolgen. Andere Sprachen können später noch hinzugefügt werden. Damit werden auch die einzelnen Wissenschafts- und Konferenzsprachen gestärkt, Präzision und Verständlichkeit sowie schließlich auch die Multilingualität und Multikulturalität der Altertumswissenschaften und der Wissenschaften gefördert.

Zu jedem Lemma werden Schreibvarianten, Synonyme, Antonyme und Hyperonyme sowie Hyponyme ermittelt und auch verwandte Begriffe aufgenommen. Die Herkunft des Wortes wird ebenfalls, besonders, wenn es aus einer anderen Sprache stammt, eruiert. Der Benutzer des Wörterbuches soll auf diese Weise in die Lage versetzt werden, sich eigene sach- und themenbezogene Unterwörterbücher zu schaffen. Zu jedem Lemma werden auch die feststehenden Formeln, Phrasen und Redewendungen aufgeführt.

Neben der Übersetzung der Begriffe wird es in vielen Fällen eine Sacherklärung geben, wenn die Übersetzung des Begriffes nicht selbsterklärend ist, sondern selbst einer Fachsprache entstammt. Dies ist besonders bei juristischen und originär verwaltungssprachlichen Begriffsbildungen der Fall (Beispiele: *Agio*, *Eviktionsgarant*, *Liturgie*, *Anachoresis*, *Aktenschwanz* (vgl. Abb. 1)).

Preisigke hatte in seinem Fachwörterbuch an einigen Stellen auch auf weiterführende Literatur, die allerdings heute in vielen Fällen veraltet ist, verwiesen. Solche aktuelleren Verweise wird es auch in diesem Online-Wörterbuch geben.

¹¹ L'Égypte romaine, recueil des termes techniques relatifs aux institutions politiques et administratives de l'Égypte romaine, suivi d'un choix de textes papyrologiques, (Académie royale de Belgique, Classe des Lettres et des Sciences morales et politiques et Classe des Beaux-Arts, Mémoires, deuxième série, tome VIII) Bruxelles 1912.

Jedes Lemma ist auf sein Vorkommen und richtige Übersetzung oder Erklärung bei Preisgabe zu überprüfen. Es gibt beispielsweise Lemmata, für die es keine Belege gibt, die z.B. nur erschlossen sind. Es sind für die unterschiedlichen Bedeutungen, die ein Wort zu einer bestimmten Zeit, in einer bestimmten Region und in einem bestimmten Kontext haben kann, die Belegstellen zu ermitteln und zu erfassen, und zwar mit Datum und Ortsangabe, um auch regionale und zeitliche Schwerpunkte und Varianten ausfindig zu machen und dem Nutzer diese Informationen anbieten zu können. Dazu werden die entsprechenden Links – das gilt in erster Linie für den frühesten und spätesten Beleg – zum Volltext bei dem Papyrological Navigator¹² gelegt. Dort stehen die Volltexte der griechischen dokumentarischen Papyri der Duke Data Bank of Documentary Papyri online frei zur Verfügung. Bei neueren noch nicht digitalisierten Papyri werden die Verweise auf die jeweilige Edition vermerkt und der Link wird sobald wie möglich nachgetragen.

Neben den digitalen Texten in papyri.info stehen die von Dieter Hagedorn regelmäßig veröffentlichten „Wörterlisten aus den Registern von Publikationen griechischer und lateinischer dokumentarischer Papyri und Ostraka“ zur Verfügung.¹³ Diese Liste enthält die Lemmata und die Belegstellen der Wörter aus den jüngst publizierten Papyri und Ostraka, deren Originaltexte noch nicht in papyri.info zu finden sind. Diese werden ebenfalls aufgenommen, wobei allerdings eine Verlinkung zum Volltext erst später vorgenommen werden kann.

3 Informationstechnologisches

Als informationstechnologische Grundlage des Fachwörterbuches werden die papyrologischen Anwendungen von MyCoRe genutzt, die bereits im Rahmen von DFG Projekten wie dem als Verbundprojekt konzipierten Papyrus-Projekt Halle-Jena-Leipzig¹⁴ und das von diesen drei Papyrussammlungen unter der Federführung Leipzigs erstellten Papyrusportal Deutschland,¹⁵ das alle digitalisierten Papyrussammlungen in Deutschland unter einem Dach vereint, eingesetzt werden. Diese Datenbankstruktur wird weitergenutzt und auf das zu entwickelnde Wörterbuch angepasst. Das Fachwörterbuch wird mit dem Papyrusportal vernetzt werden. Besonderes Augenmerk gilt dabei der Suche in verschiedenen Sprachen.

Grundlage ist der Aufbau einer XML-basierten Datenbank, die sowohl die griechischen Lemmata erfasst als auch die Übersetzungen dieser Lemmata in mehrere moderne Sprachen anbietet.

Das Wörterbuch wird wie die anderen genannten Papyrusprojekte in Kooperation mit dem Universitätsrechenzentrum durchgeführt, das auch die Langzeitarchivierung der Daten sicherstellt.

¹² <http://papyri.info>.

¹³ <http://www.papyri.uni-hd.de/WL/WL.html>.

¹⁴ <http://papyri.uni-leipzig.de>.

¹⁵ <http://www.papyrusportal.net>.

4 Bedeutung

Es gibt bisher kein mehrsprachiges Wörterbuch für die papyrologischen Texte oder überhaupt für Altgriechisch, geschweige denn für die Verwaltungsfachsprache. An dieser Stelle möchte das neue Fachwörterbuch einsetzen, indem es die Übersetzungen der griechischen Wörter in gängige Wissenschaftssprachen anbietet. Gleichzeitig ist es umgekehrt auch als modernsprachlich-altgriechisches Wörterbuch nutzbar.

Durch die Nähe und Verbindung sowie Verlinkung mit dem Lehnwortprojekt Database and Dictionary of Greek Loanwords in Coptic (DDGLC) und der angestrebten Einbindung des Wörterbuchs auch in die digitalen Projekte und Nachfolgeprojekte von eAQUA (Extraktion von strukturiertem Wissen aus Antiken Quellen für die Altertumswissenschaft)¹⁶ am Lehrstuhl für Alte Geschichte der Universität Leipzig würden auch fächerübergreifende Synergien freigesetzt. Das Papyrusprojekt Halle-Jena-Leipzig und das Papyrusportal und das neue Fachwörterbuch können als untereinander vernetzbare Module angesehen werden.

Sobald erste Ergebnisse vorliegen werden diese auch online zur Verfügung gestellt. Das Projektteam ist aufgeschlossen gegenüber Übersetzungsvorschlägen anderer Papyrologen, die regelmäßig über eine bereits bestehende Mailinglist über die Fortschritte unterrichtet werden, zumal für manche selten bezeugte Wörter jeder neue Beleg eine Bedeutungsverschiebung mit sich bringen kann. Es hat sich bereits bei den früheren Projekten eine Feedback-Funktion bewährt, die auch hier zum Einsatz kommen soll. Eine Online-Version ist damit gegenüber der gedruckten Version im Vorteil, weil auf Änderungsvorschläge und neue Belege schnell reagiert werden kann. Das Wörterbuch lässt sich auch im Hinblick auf Epigraphik und Numismatik sowie die gesamte Antike mit ihrer Textproduktion erweitern.

Autorenkontakt¹⁷

Prof. Dr. Reinhold Scholl

Historisches Seminar
Alte Geschichte
Universität Leipzig
Beethovenstr. 15
04107 Leipzig

Tel.: 0341-97-37073
e-mail: scholl@uni-leipzig.de

URL: <https://www.gko.uni-leipzig.de/historisches-seminar/seminar/alte-geschichte/apl-professor.html>

¹⁶ <http://www.eaqua.net>.

¹⁷ Die Rechte für Inhalt, Texte, Graphiken und Abbildungen liegen, wenn nicht anders vermerkt, bei den Autoren.

Project Announcement

Annotating and Editing with Canonical Text Services (CTS) Project funded by the Andrew W. Mellon Foundation: 2016–2017

Christopher Blackwell, Charlotte Schubert

1 Goals

1. We plan to coordinate ongoing work on two specific projects: CITE/CTS and the DFG-funded project eComparatio (<http://ecomparatio.net/>).¹ CITE (Collections, Indices, Texts, and Extensions) is a protocol for identification and retrieval of digital data using concise citations in URN form.² The „Text” component of CITE is the well-developed CTS protocol (Canonical Text Services). CTS allows identification of passages of text using canonical citation:

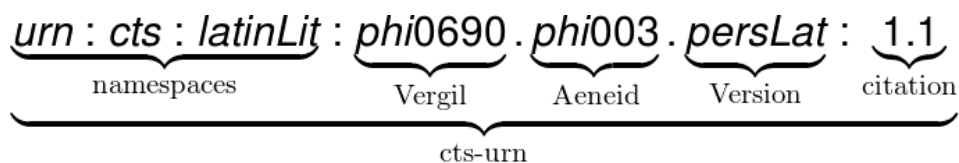


Fig. 1: The CTS-URN captures the semantics of „text” as „an ordered hierarchy of citation objects.”

CTS URN citations effectively capture the semantics of a text, and because CTS specifies a *service* for retrieving passages of text based on URNs, it facilitates comparison of different versions. For example, to compare Book 1, line 1 of two editions of the *Aeneid*, a scholar could simply request the text identified by these two URNs, which differ only in the *version-component*:

urn:cts:latinLit:phi690.phi003.deLaCerde1612:1.1
urn:cts:latinLit:phi690.phi003.ribbeck1860:1.1

eComparatio, developed at the University of Leipzig as part of the eAQUA suite of resources for philological commentary, is an online tool that implements sophisticated algorithmic approaches to analyzing *textual differences*. These may include „vertical differences”, in which one edition of a text has additional passages of text, or „horizontal differences” in which two texts have the same number of citation-units, but differ in the specific content of those citation units, or combinations of both.

¹ Login necessary. For a short description: <http://www.eaqua.net/>.

² W3C/IETF URI Planning Interest Group (21 September 2001). “URIs, URLs, and URNs: Clarifications and Recommendations 1.0”. W3C.

To realize fully the potential of this work, scholars need to be able to do the following:

- To capture canonical citations of passages of text more (or less) precisely than those defined by a text's citation scheme. For example, „The first letter *upsilon* in Book 1, line 2 of the *Iliad* as it appears on Manuscript A,” or „The first word (μῆνιν) of *Iliad* 1.1, and the first word (οὐλομένην) of *Iliad* 1.2, but *nothing in between*, because those two words are the first noun-phrase of the text.
- To capture and cite canonically an *analysis* of passages of text. In other words, if Manuscript M of Thucydes' *History* has at 1.10 οὐν („and so”), but all the other manuscripts have γοῦν („and so, at least”), we not only want to see this difference (which we will find and show with the tool eComparatio) and comment on it (which eAQUA allows), but to capture and *cite* the scholarly object that consists of three things: the differing editions of the text, the specific point of difference, and the scholarly interpretation.

CITE/CTS provides the flexibility for creating these citations; we are proposing to develop workflows that allow eComparatio to process texts based on URN citations, and to express its analytical finding by means of URNs. (Realization: C. Blackwell)

2. The German partners in the project will implement as a use case „Plutarch's *Life of Pericles* as a Network of Sources”, by which the specific benefit with regard to the technical, philological, and historical research will be demonstrated: In combination with the eAQUA tools, the *Life of Pericles* will be annotated with parallel passages, quotations and *testimonia* (as provided by the eAQUA citation tool), and textual variants (as provided by eComparatio), by storing variants and parallels via CITE-URNs, connected to the corresponding CTS-URNs. (Realization: Ch. Schubert)

3. Workshops with colleagues from the different research fields from the USA and from Germany in Leipzig and Greenville, an international exchange of American and German students, joint sections with expert conferences, publications in international journals are to ensure the visibility and dissemination.

2 Analysis, Tokenization, Graphs, and Canonical Citation

Any text is subject to a variety of analyses, such as paleography, morphology, prosody, syntax. Each of these demands a different view of the text, or more precisely, different divisions of the text into tokens. Palaeography views the text as characters and ligatures; syntax views the text as word-tokens; morphology views the text as prefixes, augments, stems, and ending; prosody views the text as syllables, which ignore word-boundaries. These analyses exist in parallel, but influence each other. A paleographic analysis may alter a single character, but in so doing completely change the syntactic graph of a sentence, or the meter of line of poetry.

Using CTS and the concept of *analytical exemplars* we hope to demonstrate precise, machine-actionable canonical citation that admits many different kinds of analysis, and allows us to address specifically conflicting, controversial analyses.

Use case: A text critic argues in favor of one manuscript variant over another; the distinction is a single character. Her argument is based not only on evidence from imagery of two manuscripts, but on a metrical and syntactic analysis of the text. She should be able to cite the regions-of-interest on the manuscript images, but also to *cite* the graphs of syntax, and the metrical analyses, that would result from accepting one variant over another. The citation of the graph of syntax must be different from a citation to the text, but must fully align, at every level, with the text.

Blackwell’s work at Furman will pursue workflows and technologies for canonically citable tokenizations, graphs, and alignment, in support of the work on Plutarch’s *Life of Pericles*.

3 Use-Case: Citation, Quotation, and Parallel Transmission in Plutarch’s *Life of Pericles* with CTS and eAQUA

Main goals

The main goal of this use-case is to show the possibilities of a combination of the citation tool developed during the project „eAQUA“, funded by the BMBF from 2008–2013, eComparatio (DFG funded 4/2014–3/2016) and the CTS/CITE–infrastructure provided by Blackwell and Neil Smith.

Blackwell will work on CTS-workflows for analytical records aligning text fragments at various levels of granularity, which is the heart of a critical apparatus. The eAQUA citation tool provides a full survey of the parallels between any given text and a given reference corpus, e.g. *The Perseus Digital Library*, based on a match of five words or more.³

CTS/CITE provides an efficient option to link the results of an analysis such as provided by the eAQUA citation tool to a text with CTS infrastructure.

eComparatio currently compares a text with specified exemplars – i.e. different editions of it – and generates an *apparatus criticus*. By implementing in eComparatio CTS URNs, which inherently align versions with their notional „work“ – any CTS-URN to a version of the *Iliad* is implicitly aligned with any other version, edition, or translation of the *Iliad* – we can extend the potential for new insights across a much wider corpus of texts.

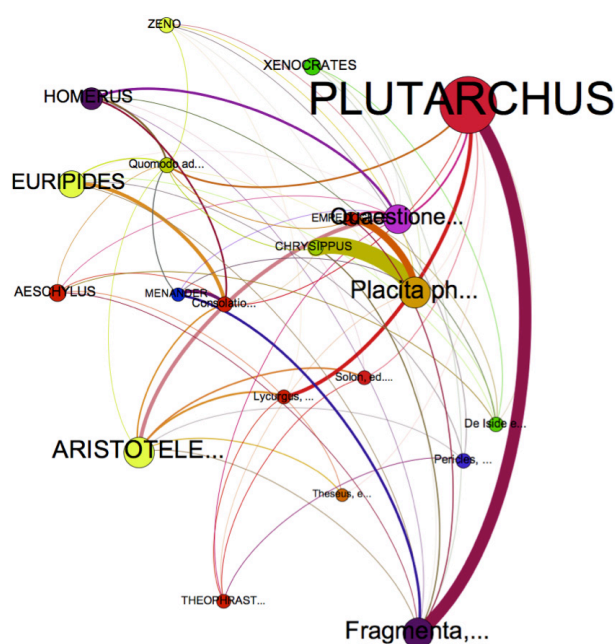


Fig. 2: The sources for Plutarch’s presentation of the *Life of Pericles*, drawn from data in eAQUA.

³ Ch. Schubert: Detailed description of eAQUA search portal, in: Ch. Schubert (Hg.): Neue Methoden der geisteswissenschaftlichen Forschung. Eine Einführung in das Portal eAQUA II (Working Papers Contested Order No.3). Leipzig 2011, ISSN 2191–3641, S. 33–53. eAQUA can draw on a variety of sources, e.g. the Thesaurus Linguae Graecae. In the United States text- and data-mining in non-commercial scientific context is possible without authorization based on the fair use doctrine text- and data-mining in non-commercial scientific context. This has been stated in the issue brief of the Association of Research Libraries as of June 2015. However our efforts will focus on digital libraries with content free from restrictive licensing.

Particular use for the scholarly community in classics and ancient history: diachronic and synchronic evaluation and automated creation of source networks.

Plutarch's *Life of Pericles* qualifies for being a use-case due to its relative brevity, making it a tractable problem. Its subject matter is the central figure of 5th century Athenian democracy in a very well-documented period of the history of Athens, with an explicit comparison to the Roman Fabius Maximus, victor over Hannibal in the Second Punic War. Furthermore, Plutarch is an author well known for his wealth of citations and parallel passages. He has already been the object of several studies based on analyses provided by the eAQUA citation tool.⁴ He also takes centre stage in ongoing scholarly debates about the habits of ancient authors with regard to their use and re-use of sources, citations, and quotations. Until now, metadata is required to create a network of sources. The example (above, created with Gephi) shows the authors which Plutarch quotes most frequently: In the works *Placita*, *Fragmenta*, *Questiones*, *Consolatio* and, most importantly, in the biographies of *Solon*, *Theseus* and *Pericles* especially *Homer*, *Euripides*, *Aeschylus*, *Theophrastus*, *Empedocles*. There are no possibilities yet, to point to such interdependencies on the level of single text passages and to visualize them, neither is there a standard for referencing.

A good example for the importance of studies in the use of sources in Plutarch in particular and the creation of new tools for equivalent surveys in general might be Plutarch's quotation in the *Life of Pericles* 26,4, the chapter, in which he describes the custom of the Athenians to mark their Samian prisoners' foreheads with the symbol of the *Samaina*, a characteristic Samian vessel with big prow, during the war of 441–439 B.C. Plutarch quotes the verse Σαμίῳν ὁ δῆμὸς ἔστιν ὡς πολυγράμματος („For oh! how lettered is the folk of the Samians!“).⁵ He also mentions that the tyrant Polycrates was the first to build this kind of ships, and that Aristophanes himself referred to the tattooing of the Samian prisoners with his verse.

The eAQUA citation-tool delivers five parallel passages, which show, on the one hand, a network of ancient text-transmission with its different steps of interdependence as well as various contexts, and, on the other hand, completely different appreciations by the modern fragmentary editions. (Those passages are neither treated in Ph. A Stadter's great commentary on the *Life of Pericles*, nor evaluated adequately yet.)

4 Text-Transmission in Antiquity

As one can see in the listing of the parallel passages, the saying has occurred first in the play *Babylonians* by Aristophanes (Hesych s.v. = fr. 64 Kock = 71 Kassel-Austin), of which only fragments have survived. Later, it has been picked up by Aristotle in his *Samian constitution* (fr.575 R). Plutarch, by putting the verse in the context of the Samian war and mentioning Polycrates, evokes a situation twice as shameful for the Samians: marked like slaves, with the sign of a tyrant.

⁴ Ch. Schubert: *Plutarchs Schrift De malignitate Herodoti (with Corina Willkommen)*, in: Schubert, Ch. / Klank, M. (Hrsg.): *Das Portal eAQUA – Neue Methoden in der geisteswissenschaftlichen Forschung III. Working Papers Contested Order No. 7*. Leipzig 2012, ISSN 2191–3641. Ch. Schubert / A Weiß, *Die Hypomnemata bei Plutarch und Clemens: Ein Textmining-gestützter Vergleich der Arbeitsweise zweier ‚Sophisten‘*, *Hermes* 143 (2015), S. 441–471; Ch. Schubert / F. Meins, *Galen und Plutarch*, erscheint in: K.-H. Leven, *Galen aus Pergamon*, (Philosophie der Antike, hrsg. von W. Kullmann) [in print].

⁵ Translation by Bernadotte Perrin.

The treatment of the passage in the ancient lexica shows wholly different contexts. Photius knows the context of the Samian war and names the *Nostoi* of Lysimachus, the mythographer and grammarian, and the *Histories* of Duris of Samos as his sources. In another passage,

Author (TLG)
0007 - PLUTARCHUS
012 - Pericles

Found : 5

Original

Σαμίον ὁ δῆμος ἐστὶν ὡς πολυγράμματος.

Source : PLUTARCHUS Biogr. Phil. (0007) (Chaeronensis - A.D. 1-2) : Pericles
Publication : Pericles, ed. K. Ziegler, *Plutarchi vitae parallelae*, vol. 1.2, 3rd edn. Leipzig: Teubner, 1964: 1-47.
Document citation : //26t/4/8 to //26t/4/10

Quotations

1. Photius <Σαμίον ὁ δῆμος> (ἐστὶν) <ὡς πολυγράμματος>:
Source : ARISTOTELES et CORPUS ARISTOTELICUM Phil. (0086) (Stagirites Pellaeus Atheniensis - 4 B.C.): *Fragmenta varia*
Publication : *Fragmenta varia*, ed. V. Rose, *Aristotelis qui ferebantur librorum fragmenta*. Leipzig: Teubner, 1886 (repr. Stuttgart: 1967)
Document citation : //8t/44t/575-578n/1n to //8t/44t/575-578n/2n
2. Photius Lex. v. [Σαμίον ὁ δῆμος ἐστὶν ὡς πολυγράμματος]:
Source : DURIS Hist. (1339) (Samius - 4-3 B.C.): *Fragmenta*
Publication : *Fragmenta*, ed. K. Müller, FHG 2. Paris: Didot, 1841-1870
Document citation : //59t50-68/111 to //59t50-68/211
3. Σαμίον ὁ δῆμος ἐστὶν ὡς πολυγράμματος:
Source : HESYCHIUS Lexicogr. (4085) (Alexandrinus - A.D. 5/6): *Lexicon (Π_Ω)*
Publication : *Lexicon (Π_Ω)*, ed. M. Schmidt, *Hesychii Alexandrini lexicon*, vols. 3-4. Halle: *n.p., 3:1861; 4:1862 (repr. Amsterdam: Hakkert, 1965)
Document citation : //sigma/150b/3t1 to //sigma/150b/3t1
4. πρὸς ταῦτα τὰ στίγματα λέγουσι καὶ τὸ Ἀριστοφάνειον ἠνίχθαι Σαμίον ὁ δῆμος ἐστὶν ὡς πολυγράμματος [Aristot.]
Source : MELISSUS Phil. (1494) (Samius - 5 B.C.): *Testimonia*
Publication : *Testimonia*, ed. H. Diels and W. Kranz, *Die Fragmente der Vorsokratiker*, vol. 1, 6th edn. Berlin: Weidmann, 1951 (repr. Dublin/Zurich: 1966): 258-267.
Document citation : //3/10 to //3/11 (Schema: Fragment/line) r
5. <Σαμίον ὁ δῆμος <ἐστὶν> ὡς πολυγράμματος>:
Source : PAUSANIAS Attic. (1569) (ca. A.D. 2): *Ἀττικῶν ὀνομάτων συναγωγή* Publication : *Ἀττικῶν ὀνομάτων συναγωγή*, ed. H. Erbse, *Untersuchungen zu den attizistischen Lexika* [Abhandlungen der deutschen Akademie der Wissenschaften zu Berlin, Philosoph.-hist. Kl. Berlin: Akademie-Verlag, 1950]
Document citation : //sigma/3*/1 to //sigma/3*/1

Fig. 3: Text-Transmission in Antiquity

the Athenians (403/2, cf. Ephorus FGrH 60 F106). Hesychius, too, (Σ 150) mentions the Samian war as well as the story about the alphabet, (cf. Fig. 3), whose tradition can be traced back from the Suda to Ephorus.

As we can see, at least three historical contexts of the quotation were known in antiquity: The one about the tyrant Polycrates, the Samian War, and the time after the Peloponnesian war.

Plutarch chooses the context of the Samian war (perhaps according to Duris), just mentions the tyrant Polycrates, and totally neglects the story about the adoption of the alphabet, which symbolizes the long-term good relations between Samos and Athens.⁶ The particularity of Plutarch's version and its own framing, which presents the Athenians from the perspective of the subjugated allies as tyrants, becomes apparent only through the citation-analysis of the parallel passages and their contexts. With the eComparatio tool, those can be visualized and examined for differences and cooccurrences.

Photius refers to the markings of former slaves relating to the granting of citizenship, as mentioned in the *Babylonians*. The passage from Photius is also found in the Suda (Σ 77 s.v. Σαμίον ὁ δῆμος), without the key-term πολυγράμματος, and even later in Michael Apostolius 15.32 with a short commentary: According to this, Aristotle (frg. 575 R, from Photius see below Fig. 3) told the story about the granting of citizenship to the Samian slaves in his *Samian constitution*. In Aristophanes' play, the verse was intended to make fun of the Samians, which were forced to grant the citizenship even to their former slaves after the tyranny of Polycrates, due to a lack of inhabitants.

An alternate version is handed over to us by the Suda, and attributed to the historian Andron of Ephesus: The Samians were the first to use the alphabet with 24 letters, introduced by Kallistratos, and later adopted by

⁶ Since he obviously knew the work of Ephorus, he must also have known the story about the alphabet.

5 The fragment in modern editions

Ascription of the fragment: Müller classifies the quote as a fragment of Duris (which seems justified by Photius' statement). Diels sees it as a *testimonium* about the fragments of Melissus, Erbse ascribes it to the atticist Pausanias.

Here, as well, the ascription to different authors leads to different contexts: Müller choses the frame of a historical work, Diels a specifically chronological context (the one of the Samian war, in which the philosopher Melissos was general of the Samians), Erbse sees it as a linguistic testimony.

As one can see from this example, an integration of the parallel passages as provided by the eAQUA citation tool would be highly useful. Performed for the whole *Life of Pericles*, a new quality of a digital *apparatus fontium* could be reached with this kind of diachronic and synchronic evaluation, hence the foundation of a more extensive annotation with CTS in terms both of a historical commentary and a commentary on the history of scholarship.

Autorenkontakt⁷

Prof. Christopher W. Blackwell

Furman University
Louis G. Forgione University Professor of Classics
Email: christopher.blackwell@furman.edu

URL: <http://www.furman.edu/academics/classics/about/Pages/FacultyandStaff.aspx>

Prof. Charlotte Schubert

Universität Leipzig
Historisches Seminar
Lehrstuhl für Alte Geschichte
Email: schubert@uni-leipzig.de

URL: <https://www.gko.uni-leipzig.de/historisches-seminar/seminar/alte-geschichte/professur.html>

⁷ Die Rechte für Inhalt, Texte, Graphiken und Abbildungen liegen, wenn nicht anders vermerkt, bei den Autoren.