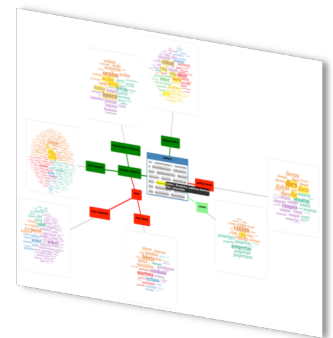
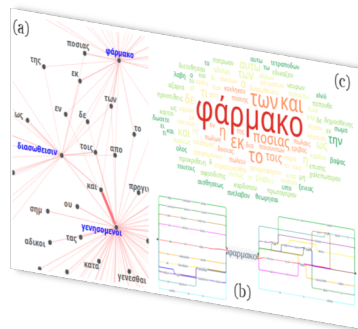
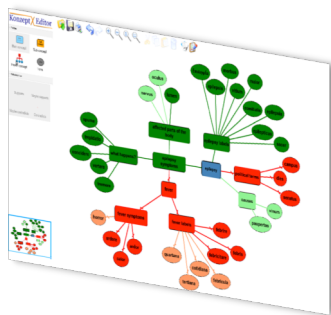


**Exploring Concept Change and Transfer in Antiquity**  
**Moderne Informationstechnologien in den Altertumswissenschaften**  
**eXChange**

**BMBF Förderschwerpunkt:**  
**„Förderung von Forschungs- und Entwicklungsvorhaben aus dem Bereich der eHumanities“**

exchange



Abschlussbericht 2016

Vorgelegt von:

Prof. Dr. Charlotte Schubert, Lehrstuhl für Alte Geschichte, Historisches Seminar, Universität Leipzig (Koordinatorin)

Prof. Dr. Marcus Deufert, Lehrstuhl für Latinistik, Institut für Klassische Philologie und Komparatistik, Universität Leipzig

Prof. Dr. Gerhard Heyer, Abteilung für Automatische Sprachverarbeitung, Institut für Informatik, Universität Leipzig

Prof. Dr. Foteini Kolovou, Lehrstuhl für Byzantinische und Neugriechische Philologie, Universität Leipzig

Prof. Dr. Geric Scheuermann, Abteilung für Bild- und Signalverarbeitung, Universität Leipzig

Prof. Dr. Kurt Sier, Lehrstuhl für Gräzistik, Institut für Klassische Philologie und Komparatistik, Universität Leipzig

## INHALTSVERZEICHNIS

I. KURZE DARSTELLUNG	4
1. AUFGABENSTELLUNG	4
2. VORAUSSETZUNGEN	4
3. PLANUNG UND ABLAUF DES VORHABENS	5
4. WISSENSCHAFTLICHER UND TECHNISCHER STAND	8
5. ZUSAMMENARBEIT MIT ANDEREN STELLEN	9
II. EINGEHENDE DARSTELLUNG	11
1. VERWENDUNG DER ZUWENDUNG UND ERZIELTE ERGEBNISSE	11
2. WICHTIGSTE POSITIONEN DES ZAHLENMÄßIGEN NACHWEISES	18
3. NOTWENDIGKEIT UND ANGEMESSENHEIT DER GELEISTETEN ARBEIT	18
4. ERGEBNISSE/ WECHSELWIRKUNGEN/ TRANSFER	19
5. BEKANNT GEWORDENER FORTSCHRITT BEI ANDEREN STELLEN	24
6. ERFOLGTE/ GEPLANTE VERÖFFENTLICHUNGEN	24

Gemeinsamer Sach- und Schlussbericht für das Verbundprojekt eXChange: Exploring Concept Change and Transfer in Antiquity Moderne Informationstechnologien in den Altertumswissenschaften

Projekt im Förderprogramm „Förderung von Forschungs- und Entwicklungsvorhaben aus dem Bereich der eHumanities“ des BMBF

Koordinatorin

Prof. Dr. Charlotte Schubert

Teilprojekte

Automatische Sprachverarbeitung	Prof. Dr. Gerhard Heyer Thomas Efer, M.Sc.	01UG1248
Bild- und Signalverarbeitung	Prof. Dr. Gerek Scheuermann Dipl. Inf. Stefan Jänicke / M. Sc. Muhammad Faisal Cheema	01UG1248
Wissenschaftssprache Medizin und sprachliches Handeln in der griechischen Literatur	Prof. Dr. Kurt Sier Eva Wöckener-Gade	01UG1248
Wissenschaftssprache Medizin und sprachliches Handeln in der byzantinischen Literatur	Prof. Dr. Fotini Kolovou Dariya Rafiyenko	01UG1248
Wissenschaftssprache Medizin und sprachliches Handeln in der römisch-lateinischen Literatur	Prof. Dr. Marcus Deufert Judith Blumenstein / Christoph Weilbach	01UG1248
Der historische Kontext: Wissenschaftssprache in Politik und Medizin	Prof. Dr. Charlotte Schubert Dr. Roxana Kath / Dr. Michaela Rücker / Friedrich Meins / Oliver Bräckel	01UG1248

**Projektlaufzeit:** Juli 2012 – September 2015

## I. Kurze Darstellung

### 1. Aufgabenstellung

Im Rahmen des Projektes sollten durch die Kombination von Text Mining, Visual Analytics und historischer Diskursanalyse Bedeutungsverschiebungen semantischer Räume nach Ort und Zeit visualisiert werden, um die Wechselwirkungen zwischen Wissenschaft, Alltag und Politik in der Antike zu untersuchen.

Ziel des Projektes war der Aufbau eines Recherchesystems, durch das Begriffe, die sowohl in medizinischen Fachtexten als auch in politischen und alltäglichen Texten verwendet werden, auf ihren Kontext in argumentativen Strategien hin analysiert werden können. Während Methoden und Techniken zur automatischen Verarbeitung und inhaltlichen Erfassung großer Mengen an Textkollektionen heute weit verbreitet sind, ist die Erfassung von semantischem Inhalt noch in einem weitgehend experimentellen Stadium.

Die gewonnenen Ergebnisse des Projektes können beispielgebend als neuer, methodischer Zugang für andere geistes- und sozialwissenschaftliche Disziplinen dienen, um diskursive Prozesse in ihrer Verzeitlichung zu repräsentieren.

### 2. Voraussetzungen

Grundlage des Projektes war die im BMBF-Projekt eAQUA etablierte Zusammenarbeit der Abteilung für Automatische Sprachverarbeitung (ASV) und der klassischen Altertumswissenschaften an der Universität Leipzig. Die vorhandene Expertise wurde durch die Einbindung der Abteilung für Bild- und Signalverarbeitung (BSV) und des Lehrstuhls für Byzantinistik erweitert.

Aus dem Projekt eAQUA konnten zahlreiche Werkzeuge zur Beantwortung spezifischer textbasierter Fragestellungen nachgenutzt werden: Auf technischer Seite war dies u. a. die bereits umgesetzte Anbindung diverser Korpora und der entsprechenden Trägerformate an eine gemeinsame Datenbank. Daneben konnten viele entwickelte Verfahren als qualitative Messlatte oder konzeptionelle Ausgangsbasis für die benötigten Realtimeverfahren verwendet werden. Von besonderem Interesse waren die Vorarbeiten zur Verknüpfung von Termen, die eine ähnliche lexikalische Umgebung aufweisen. Generell konnte aus dem Projekt auf breite Erfahrungen in der automatischen Analyse antiker Texte zurückgegriffen werden. Das Projekt eXChange war daher in einem Umfeld angesiedelt, in dem bereits ein produktiv für Forschung und Lehre genutztes Text Mining-gestütztes Recherchesystem existierte (Projektergebnis des eAQUA-Verbundvorhabens). Daraus ergab sich die technologische und methodische Ausrichtung des Vorhabens auf darüber hinausgehende innovative Zugangsmöglichkeiten zu Korpora, bei denen eigens entwickelte Text Mining-Ansätze neue Formen explorativer Suche und Abfrage ermöglichen. Im Gegensatz zur klassischen Funktion vieler Rechercheoberflächen, Dokumente in Form eines Auffinde- und Referenzwerks zu präsentieren, kann der Schwerpunkt der in eXChange umgesetzten Arbeiten in der Schaffung ergänzender Analysefunktionen gesehen werden.

Eine der wichtigsten Voraussetzungen war die Bereitstellung der digitalen Textkorpora: Die historisch arbeitenden Altertumswissenschaften (Alte Geschichte, Klassische Philologie, Epigraphik, Papyrologie) greifen heute alle auf digitalisierte Textkorpora zurück, die in unterschiedlichsten Formen (digitalisierte Bibliotheken im Internet, auf CD ROM) und

Formaten (z.B. verschiedene ASCII –basierte Beta-Code-Varianten und UNICODE-basierte UTF-Formate) vorliegen.

Datengrundlage der entwickelten Tools sind: das Korpus Thesaurus Linguae Graecae (TLG), umfangreiche Textsammlungen des Perseus-Projektes, die Bibliotheca Teubneriana Latina (BTL 1–4, erschienen 1999–2006, Onlineversion 2009) und die Patrologia Latina Database (PLD). Die Datenbasis für den Bereich der literarischen Texte umfasst somit fast den kompletten erhaltenen Textbestand vom 8. Jahrhundert v. Chr. bis zum Fall von Byzanz 1453.

Hinzu kamen von der Seite der Altertumswissenschaften eine ganze Reihe von Metadaten, die auf der Basis der Zielstellung zusammengestellt und in die Tools eingepflegt wurden (Angaben zu Lebenszeit und Aufenthaltsorten griech. u. lat. Autoren; ein griech.-lat. Wortkatalog medizinischen Fachvokabulars sowie fachlich qualifizierte Stoppwortlisten und Normalisierungsregeln). Diese wurden im Projektverlauf händisch auf der Basis altertumswissenschaftlicher Lexika, Fachliteratur und eigener Quellenrecherchen erstellt.

### 3. Planung und Ablauf des Vorhabens

Die ersten Monate des Projektes wurden für die Installation der notwendigen Infrastruktur, die interdisziplinäre Verständigung über die Anforderungen an das zu entwickelnde Recherchetool und die Auswahl und Aufbereitung geeigneter Metadatenbestände genutzt.

Zentraler Bestandteil der Arbeiten waren von Projektbeginn an kontinuierliche interdisziplinäre Arbeitstreffen mit allen oder einem Teil der Projektpartner, um Zielsetzungen zu definieren, konkrete Anforderungen an die zu entwickelnde Infrastruktur zu formulieren, entwickelte Anwendungen zu evaluieren und wichtige Informationen auszutauschen. Hierunter fielen auch zahlreiche Workshops, die der Vorstellung und der Arbeit an einzelnen Tools wie z. B. dem Konzepteditor gewidmet waren:

#### Liste der Arbeitstreffen und Workshops

2012

11. Juli	Kick-Off-Treffen der altertumswissenschaftlichen Projekte: Diskussion zum Umgang mit Metadaten
01. August	Kick-Off-Treffen zwischen Projektkoordination u. ASV: Infrastruktur u. Metadaten
20. September	Kick-Off-Treffen zwischen Projektkoordination, ASV u. BSV: Metadaten/-editor
18. Oktober	Kick-Off-Treffen aller Projektpartner: Metadaten-Tool, Wortlisten
06. November	Koordinierung der Arbeit an den Wortlisten
12. November	Workshop Digital Humanities III
28. November	Arbeitstreffen zur Anforderungsanalyse
04. Dezember	Arbeitstreffen zur Kategorienbildung bei Wortlisten

13. Dezember	Arbeitstreffen zur Visualisierung, Implementierung der TLG-Daten, Vorbereitung des Workshops vom 20.03.2013
--------------	---

2013

17. Januar	Arbeitstreffen: Workflows
5. März	Besprechung zur eXChange-Webseite
7. März	Vorbesprechung zum Workshop eXChange am 20. März 2013
14. März	Generalprobe Workshop
20. März	Workshop mit Vertretern des Beirates und des Projektträgers
10. Juli	Arbeitstreffen zum neuen Recherchesystem
22. Juli	Interdisziplinäres Sommerfest der BMBF Projekte eXChange & ePol
24. Oktober	Arbeitstreffen: Arbeitstand und weitere Planungen
11. November	Vorbesprechung zum Projektworkshop „Visualisierung“
13. Dezember	Projektworkshop: Visualisierung und Konzepteditor

2014

Januar	Workshop „New Visual Hermeneutics“ an der HSU Hamburg
31. Januar	projektinterner Workshop: „Workflows“
08. Mai	2. Release des Konzepteditors & projektinterner Workshop
Juli	Präsentation auf der The European Summer University in Digital Humanities 2014 (ESU, Leipzig)
Juli	Teilnahme an der DH 2014 (Lausanne, Schweiz)
17. Oktober	projektinterner Workshop: Userinterface für die Konzeptsuche
Oktober– November	Arbeitstreffen: Bewerbung für die DH 2015 in Sydney (Australien) sowie für die DHd-Tagung 2015 in Graz (Österreich)
24.–28. November	eXChange-Nutzerstudie „Visualisierung“

2015/16

Januar	Workshop „New Visual Hermeneutics“ an der HSU Hamburg
23.–27. Februar	DHd 2015 „Von Daten zu Erkenntnissen“, in Graz
17. März	eXChange Projektmeeting

5. Mai	eXChange Projektmeeting
17. Juni	eXChange Projektmeeting
30. Juni	eXChange Projektmeeting
Juli	Workshop „New Visual Hermeneutics“ an der HSU Hamburg
29. Juni–3. Juli 2015	Teilnahme an der DH 2015 (Sydney, Australien)
September	Gewährung einer kostenneutralen Verlängerung um 4 Monate
17. Oktober	eXChange Projektmeeting
15.–17. Februar 2016	Fachtagung „Postclassical Greek: Intersections of Philology and Linguistics“

#### Informatik:

Der Abteilung ASV oblag im Projekt die textzentrierte Datenaufbereitung im Umfeld antiker Korpora, die Konzeption und Umsetzung eines flexiblen realtime-fähigen Text-Mining-Systems, sowie die Definition und Realisierung von Schnittstellen für den Zugriff auf diese Daten und Verfahren innerhalb eines web-basierten explorativen Recherchesystems. Die Arbeiten hatten entsprechend der auf die Untersuchung eines Konzeptwandels und Sprachtransfers ausgerichteten fachlichen Fragestellungen einen besonderen Fokus auf Begriffe und deren Kontexte.

Als erster Arbeitsschritt war die Erstellung eines Metadateneditors geplant. Hierfür wurden die bereits in Vorgängervorhaben als Spreadsheet erhobenen Metadatenätze gesichtet, transformiert und dedupliziert. Dabei wurde die Anwendung von Topic Maps als semantischer Repräsentationsform untersucht, was aber nicht zuletzt wegen eingeschränkter Verfügbarkeit aktueller Implementierungen verworfen wurde. Ein erster Prototyp eines Metadateneditors (auf Grundlage der Graphdatenbank OrientDB) wurde erstellt, die Fertigstellung jedoch so weit verschoben, bis eine zweckmäßige Einordnung in ein Gesamtsystem und die damit verbundenen (und noch zu definierenden) Workflows möglich war. Gleichzeitig wurden grundlegende infrastrukturelle Arbeiten (Servereinrichtung, Backuplösung, Einrichtung eines Versionskontrollsystems, etc.) vorgenommen. Ebenfalls in einer frühen Phase des Projekts begannen die Arbeiten zur Integration der Korpusdaten und zum Finden einer für Echtzeit-Abfragen geeigneten Repräsentationsform der Daten. Während im Projektantrag die Korpuserschließung und Ermöglichung parallelverarbeitender korpusstatistischer Verfahren noch als sehr getrennte Aufgaben wahrgenommen wurden, zeigte sich schnell, dass beide Problemfelder nur durch gemeinsame Betrachtung und Bearbeitung zufriedenstellend gelöst werden können. Letztendlich zeigte sich, dass mit einer technologischen Festlegung auf Graphdatenbanken sowohl eine geeignete Repräsentation der Textdaten als auch ein direkter Anschluss der Metadaten, eine parallele Datenanfrage und ein flexibler web-basierter Zugriff umgesetzt werden kann. Zusätzlich zu den im Antrag vorgesehenen Arbeiten wurden in prototypischen Versuchen auch grammatikalisch-morphologische Ressourcen (aus dem Perseus-Projekt) eingebunden, und mit den aus dem Textgrid-Repository (Zeno-Sammlung) erhaltenen und automatisch extrahierten Wörterbuchlemmata sowie dem normalisierten Vokabular der untersuchten Korpora verknüpft. In der konsolidierten Text Mining-Umgebung für das Recherchebackend wurden in der letzten Projektphase zahlreiche Webservice-Schnittstellen

angereichert, mit denen verschiedene Abfrageszenarien (insbesondere die innerhalb der interaktiven Visualisierungen benötigten lokal fokussierten Aggregationen) effizient asynchron beantwortet werden können.

Altertumswissenschaften:

Das erste Projektjahr war für die altertumswissenschaftlichen Teilprojekte geprägt von den Einstiegsarbeiten. Dazu gehörten das Aufarbeiten des aktuellen Forschungsstandes, die formale und inhaltliche Analyse der projektrelevanten Texte, die Entwicklung von Konzepten für die Textdatenbanken und die Überlegungen zu Aufbau und Struktur der für die Arbeit benötigten Werkzeuge. Die z. T. händische Auswahl des Materials für die Datenbanken wurde durch die Projektmitarbeiter kontinuierlich evaluiert und korrigiert, um eine möglichst homogene Datenbasis zu erhalten. Für die Entwicklung des dynamischen Recherchetools haben die MitarbeiterInnen ihre bisherige Arbeitsweise bei Recherche und Hypothesenbildung beschrieben und übertrugen diese auf digitale Workflows. Dies geschah auf der Grundlage einer detaillierten Beschreibung von fachwissenschaftlichen Fallstudien.

Seit dem zweiten Projektjahr standen den Teilprojekten erste Versionen der Tools zur Verfügung. Wichtigste Aufgabe war im Folgenden die Evaluierung dieser Werkzeuge an ausgewählten Forschungsfragen, die weitere Kontrolle des Datenmaterials und die Kommunikation von Fehlern, Problemen und Verbesserungsvorschlägen an die Informatik.

Außerdem wurden auf der Basis von PHI 5 die Metadaten für 372 lateinische Autoren und Werke (v. a. Lebenszeit und Orte) im Format Excel-xlsx bereitgestellt. In den Verhandlungen mit dem De Gruyter Verlag erreichte die Koordinatorin, dass das Projekt die vorhandenen Metadaten der BTL verwenden darf (die Implementierung erfolgte Anfang 2014). Durch die Kooperation mit Textgrid konnte das Projekt zudem die Metadaten der beiden Wörterbücher Pape (gr.) und Georges (lat.) erhalten und integrieren. Dadurch ist derzeit bereits eine inverse Suche im eXChange-Prototypen möglich. Auf der Basis der Autoren Galen und Celsus haben die Mitarbeiter der altertumswissenschaftlichen Teilprojekte einen umfangreichen griechischen und lateinischen Wortkatalog erstellt (Celsus 770 Einträge; Galen 4077 Einträge). Dieser umfasst medizinisches Fachvokabular in den Sprachen Altgriechisch – Latein – Englisch – Deutsch und ermöglicht dadurch eine genauere (auch) sprachübergreifende Suche.

Das dritte Projektjahr war durch die Evaluation der entwickelten Tools gekennzeichnet. Daraus resultierten weitere Modifizierungen der Werkzeuge und deren praktische Anwendung auf altertumswissenschaftliche Forschungsfragen.

Für eine detaillierte Darstellung der Arbeit in den Teilprojekten sei auf die Zwischenberichte des Projektes (2012–2015) verwiesen.

#### 4. Wissenschaftlicher und technischer Stand

Die letztlich im Projekt gefundene Lösung knüpft auf technischer Ebene an bestehende Softwaresysteme an. Hervorzuheben ist das Graphdatenbanksystem Titan, welches es erlaubt eine Reihe verschiedener „Speicherbackends“ zu verwenden, die sowohl lokal als auch auf verteilten Rechnerclustern angesiedelt sein können. Dadurch ist die Skalierungsfähigkeit des Systems nicht beschränkt. Für eine Vielzahl von Graphdatenbanken existiert eine gemeinsame Schnittstellenbeschreibung namens Tinkerpop, deren Nutzung sogar einen kompletten Austausch des Datenbankprogramms mit geringem Aufwand ermöglicht. Die Laufzeitumgebung für Titan und Tinkerpop ist Java, wodurch eine weitest gehende Plattformunabhängigkeit erreicht wird. Für die agile Webentwicklung wurde die



Programmiersprache Ruby verwendet, welche ein reichhaltiges Ökosystem für diese Aufgabe bietet (u. a. das verwendete Webframework „Sinatra“). Zur Verknüpfung beider Aspekte des Backends wurde die Variante JRuby verwendet, welche eine Reimplementierung der Ruby-Sprache in Java darstellt. Dadurch werden Technologiebrüche vermieden und es wird eine optimale Kommunikation der Komponenten gewährleistet. Für JRuby steht mit „Pacer“ eine effiziente Ruby-Komponente für die Abfrage von Graphstrukturen über Tinkerpop zur Verfügung.

Auf wissenschaftlicher Ebene wurde an zahlreiche existierende Bemühungen angeknüpft, um Netzwerkstrukturen aus Texten zu generieren. Das besondere Ziel, die übliche „Ordered Hierarchy of Content Object“ als Dokumentenmodell aufzulösen, ist in Verbindung mit den darauf aufsetzenden Text Mining-Verfahren jedoch eine originäre im Projekt entwickelte Idee. Entsprechend sind die Vorarbeiten weniger relevant, insbesondere weil das Projekt wie bereits beschrieben eine komplementäre, und nicht ersetzende technologische Lösung für den Umgang mit antiken Korpora liefern sollte.

## 5. Zusammenarbeit mit anderen Stellen

1) Zur Dissemination und Erweiterung der im Projekt konzipierten Einsatzmöglichkeiten für Graphdatenbanken im Text Mining, auch über die Grenzen der Altertumswissenschaften hinweg, hat die ASV zahlreiche anwendungsfallbezogene Machbarkeitsstudien durchgeführt, z. B. im Bereich der Global and Area Studies (s. Thomas Efer und Ninja Steinbach-Hüther: Quantitative Analyses in Global and Area Studies using Graph-based Filtering of Heterogeneous Catalogue Data. In: Proceedings of INFORMATIK 2014, Bonn, Gesellschaft für Informatik, 2014) oder der Aufbereitung von Archivbeständen (s. Thomas Efer, Jens Blecher und Gerhard Heyer: Leipziger Rektoratsreden 1871–1933. In: Historical Corpora. Challenges and Perspectives, Narr, 2015). Weiterhin gab es eine enge Kooperation mit dem Leipziger CLARIN-D-Zentrum um die Möglichkeiten der Eingliederungen der neu entstandenen Technologien in die CLARIN-Infrastruktur abzuklären.

2) Die Lizenzvereinbarung mit dem Verlag De Gruyter, die durch die Koordinatorin Ch. Schubert mit dem Verlag De Gruyter ausgehandelt wurde, ist im Hinblick auf die aktuelle Urheberrechtsdiskussion in Deutschland und insbesondere für die Zukunft der Digital Humanities eine richtige Sensation. Neben einer Einigung über ein vereinfachtes Zugriffsverfahren (Authentifizierung ohne Shibboleth o. ä.), die Erlaubnis, die Datenbank zu speichern, auszuwerten und weiterzuverarbeiten, wurde folgende Vereinbarung getroffen: „De Gruyter gewährt der UL das Recht, für eigene Forschungszwecke und zur Lehre die unter 1.1. genannten XML-Daten zu ver- oder bearbeiten und mit zusätzlichen Metadaten anzureichern. Der UL ist es außerdem gestattet, die Datenmengen auszuwerten (Datamining) und für die Anwendung in den Projekten eXChange und eAQUA technisch vorzubereiten. Zu diesen Zwecken dürfen in Verantwortung der ausführenden Stelle Sicherheitskopien der XML-Daten erstellt werden.“

Dies ist insofern von Bedeutung, als dass beispielsweise die Frage der Verwendung von Visualisierungen, die via Text- und Datamining aus geschützten Daten hervorgehen, bisher lizenzrechtlich kaum geregelt ist. Wir regen daher an, dass derartige Vereinbarungen mit Lizenzträgern künftig auch vom BMBF als wegweisend mit unterstützt werden.

3) Durch die Kooperation mit Textgrid konnte das Projekt die Metadaten der beiden Wörterbücher Pape (gr.) und Georges (lat.) in XML-Format erhalten und integrieren.

4) Besonders wertvoll für die fachwissenschaftliche Methodologie war der im Austausch mit dem Hamburger Verbundprojekt „Postdemokratie und Neoliberalismus“ (ePol) entwickelte neue methodologische Ansatz. Die Beteiligten erarbeiteten eine Explikation einer „New Visual Hermeneutics“ für die Digital Humanities. Der Ansatz rekurriert epistemisch auf die empirische Phänomenologie von Don Ihde und inkludiert darüber hinaus methodische Ansätze aus den Sozialwissenschaften. Die Grundzüge einer neuen visuellen Hermeneutik als Forschungsansatz sollen den besonderen epistemischen Herausforderungen der Digital Humanities Rechnung tragen und vor allem den Fachwissenschaften einen methodischen Zugang über ihre eigenen fachspezifischen Forschungsansätze hinaus ermöglichen.

5) Im Projektverlauf neu gewonnene Kooperationspartner:

Steffen Albrecht (Medienzentrum der TU Dresden),

Prof. Dr. Oksana Arnold (FH Erfurt),

Dr. Elton Barker (Berlin; The Open University),

Prof. Andreas Blätte (Universität Duisburg-Essen),

Dr. Stylianos Chronopoulos (Freiburg),

Dr. Maria Effinger (UB Heidelberg),

Prof. Claudia Fantapié-Altobelli (HSU Hamburg),

Eddy Gouder (CTLO),

Bart Janssens (Brepols),

Prof. Dr. Klaus Jantke (Fraunhofer IDMT Erfurt),

Prof. Dr. Bernhard Jussen (Frankfurt),

Prof. Dr. Alexander Mehler (Frankfurt),

Prof. Paul Molitor (Halle),

Prof. Dr. Joachim Scharloth (TU Dresden),

Dr. Felix K. Maier (Freiburg),

Dr. Heike Neuroth (Göttingen)

Dr. Anna Novokhatko (Freiburg)

Prof. Dr. Stefan Pickl (Universität der Bundeswehr München).

Dr. Lutz Popko (Projekt „Strukturen und Transformationen des Wortschatzes der ägyptischen Sprache“, SAW Leipzig),

Prof. Dr. Jörg Rüpke (Erfurt),

Prof. Dr. Christoph Schäfer (Trier),

Dr. Leif Scheuermann (Erfurt/ Graz),

Prof. Dr. Wolfgang Spickermann (Max-Weber-Kolleg Erfurt/ Graz).

## II. Eingehende Darstellung

### 1. Verwendung der Zuwendung und erzielte Ergebnisse

Die Projektziele und Erfolge eines interdisziplinären Projektes sind sehr vielschichtig. Zur Erreichung der Projektziele waren auch methodische Instrumente nötig (Translatoren), die in einem Projektergebnis nur indirekt messbar sind. Im Folgenden werden nur die zentralen Projektergebnisse beschrieben, die methodisch neue Wege innerhalb der textbasierten Altertumswissenschaften eröffnen und die auch schon in der Scientific Community Aufmerksamkeit gefunden haben. Für eine detailliertere Darstellung der Teilergebnisse sei auf die Zwischenberichte des Projektes (2012–2015) verwiesen.

#### Liste der im Projektverlauf erzielten Ergebnisse

Datum	Bezeichnung & Bemerkung	Zugehörige Arbeitspakete
März 2013	Projekt-Webseite: <a href="http://www.exchange-projekt.de">http://www.exchange-projekt.de</a>	AP 4.6
März 2013	1. Projektworkshop	AP 4.3; 4.4; 4.6; 5
	Bereitstellung korpusübergreifender Wörterbücher (gr./lat): Pape (gr.): 98.295 Einträge Georges (lat.): 54.831 Einträge	AP 4.1
Juli 2013	Bereitstellung von Meta-Daten zu den lateinischen Autoren: Informationen (Lebensdaten und Orte) zu 372 lateinischen Autoren	AP 4.1; 4.2
Juli 2013	Bereitstellung der BTL-Daten: Lizenzvertrag über die Nutzung der Daten mit dem De Gruyter Verlag; Integration von Probedaten	AP 2.1
Juli 2013	Prototypen zur Korpusverwaltung	AP 2.3; 3.3
Dezember 2013	Prototyp(en) Visualisierung	AP 3.2; 3.3
Dezember 2013	Prototyp Konzepteditor	
Dezember 2013	Bereitstellung eines griechisch-lateinischen Wortkatalogs (medizinisches Fachvokabular gr./ lat./ dt./ engl.): Celsus (lat.): 770 Einträge Galen (gr.): 4077 Einträge	AP 4.2
August 2014	3. Release des Konzepteditors & Integration in das eXChange Recherchefrontend	AP 2.2; 3.3
17. Oktober 2014	projektinterner Workshop: Userinterface für die Konzeptsuche	AP 3.2; 3.3
Okt.–Nov. 2014	Bewerbung für die DH 2015 in Sydney (Australien) sowie für die DHd-Tagung 2015 in Graz (Österreich); Short Paper nach Peer-Review-	AP 4.4; 4.6

	Verfahren angenommen	
Okt.–Dez. 2014	Einbindung in die Lehre	AP 4.4; 4.6
24.–28. November 2014	eXChange-Nutzerstudie „Visualisierung“	AP 3.3; 4.4; 4.5
	1. Versionen einer fachlich qualifizierten Stoppwortliste und von Normalisierungsregeln	AP 2.3
Januar 2015	Workshop „New Visual Hermeneutics“ an der HSU Hamburg; Fortführung der Kooperation	AP 4.3; 4.5; 4.6
23.–27. Februar 2015	DHd 2015 „Von Daten zu Erkenntnissen“, in Graz; Präsentation des angenommen Papers	AP 3.3; 4.6
März 2015	Bereitstellung eines Prototypen zur Ergebnisausgabe; Anzeige einer Ergebnisauswahl im Textmodus	AP 3.2; 4.4
April 2015	Bereitstellung eine Betacodeconverters (Betacode.js)	AP 2.1
29. Juni–3. Juli 2015	Teilnahme an der DH 2015 (Sydney, Australien); Präsentation der angenommen Paper	AP 3.3; 4.6
Okt.–Dez. 2015	Einbindung in die Lehre	AP 4.4; 4.6
25. November 2015	Bereitstellung des Metadateneditors	AP 2.2
August 2016	Sonderheft „eXChange“ in Digital Classics online	AP 4.4; 4.5; 4.6
2016	Veröffentlichung des Recherchebackend-Systems „Kadmos“; inklus. Anbindung (Anbindung des Konzepteditors und des Metadateneditors an die Datenbanken	AP 3.2.; 3.3; 3.4

Die technologischen und algorithmischen Beiträge der Informatik haben in Kooperation mit den Altertumswissenschaftler\_Innen als Ergebnis ein System hervorgebracht, das sich durch die folgenden Merkmale auszeichnet:

- Die Möglichkeit zur Verwaltung mehrerer Korpora unterschiedlicher (und selbst gemischter) Sprachen.
- Flexible Metadaten-Verknüpfung innerhalb der Graphstruktur.
- Onomasiologisch geleitete und fachsprachenabhängige Definition von Konzepten zur Verfeinerung der aktuellen Sichtweise, expliziten Erfassung von semantischen Zusammenhängen und als Basis für eine „Query Expansion“.
- Definition und Optimierung von Services für nichtlineare explorative Suchen (z. B. über im Projekt durch die BSV erstellte interaktive Visualisierungen).

- Nicht-destruktive Normalisierung (d. h. mit kompletter Rekonstruierbarkeit der Originalformen) in Kombination mit stellenweiser Grundformverknüpfung zu Wörterbuchressourcen.
- Abarbeitung von Text Mining-Verfahren als vorberechnungsfrei; Echtzeitanalyse (Nach der Importphase zum Aufbau der Graphstruktur kombinieren die Verfahren darauf statuslose Traversierungsszenarien.)
- Clusterfähigkeit durch Einsatz eines neuartigen Graphdatenbanksystems

Die im Projekt entwickelte innovative Form der Korpusrepräsentation und die dadurch ermöglichten graphbasierten Text Mining-Verfahren sind Gegenstand der Dissertationsschrift von Thomas Efer. Sie wird im Herbst 2016 eingereicht.

Geisteswissenschaftler\_Innen, die mit historischen Textkorpora arbeiten, stellen in der Regel Keyword-Suchanfragen (= Stichwortsuche), um relevante Textpassagen abzurufen. Das größte Problem der Keyword-basierten Suche, ist die geringe Genauigkeit der Ergebnisse. Die Nutzer\_In muss die Abfrage immer wieder reformulieren bzw. verfeinern, um die Qualität der Ergebnisse zu verbessern oder die Anzahl der Suchergebnisse zu reduzieren. Darüber hinaus ergibt die einfache Stichwortsuche in der Regel unvollständige Ergebnismengen, da sie unterschiedliche Wortformen oder Synonyme nicht berücksichtigt. Es gibt zahlreiche Alternativen zur Stichwortsuche (wie z. B. „Topic Models“), aber die meisten dieser Methoden geben den Wissenschaftler\_Innen keine hinreichende Kontrolle über den Suchvorgang. Mit automatisierten Suchverfahren, verlieren die Wissenschaftler\_Innen wiederum einen großen Vorteil einer Stichwortsuche, die Fähigkeit, iterativ Suchergebnisse zu verbessern.

Vorrangiges Ziel des Projektes „eXChange“ war es daher, die Zahl und die Genauigkeit der Ergebnisse zu erhöhen, die durch traditionelle Stichwortsuchen erreicht werden, indem eine sog. „Concept Search“ entwickelt und eingesetzt wurde. Eine weitere wichtige Anforderung der Geisteswissenschaftler\_Innen war ein Suchwerkzeug, das sie den Suchvorgang nach Wunsch steuern lässt.

Im Projekt wurde eine nutzergesteuerte Suchumgebung für „Konzepte“ entwickelt, die den Geisteswissenschaftler\_Innen die volle Kontrolle über die Entwicklung der Konzepte sowie die Analyse der Suchergebnisse gibt und die Möglichkeit, graduell die Suchergebnisse durch iteratives Modifizieren der zugrunde liegenden „Konzepte“ zu verbessern.

Die Projektpipeline wurde in drei Phasen unterteilt:

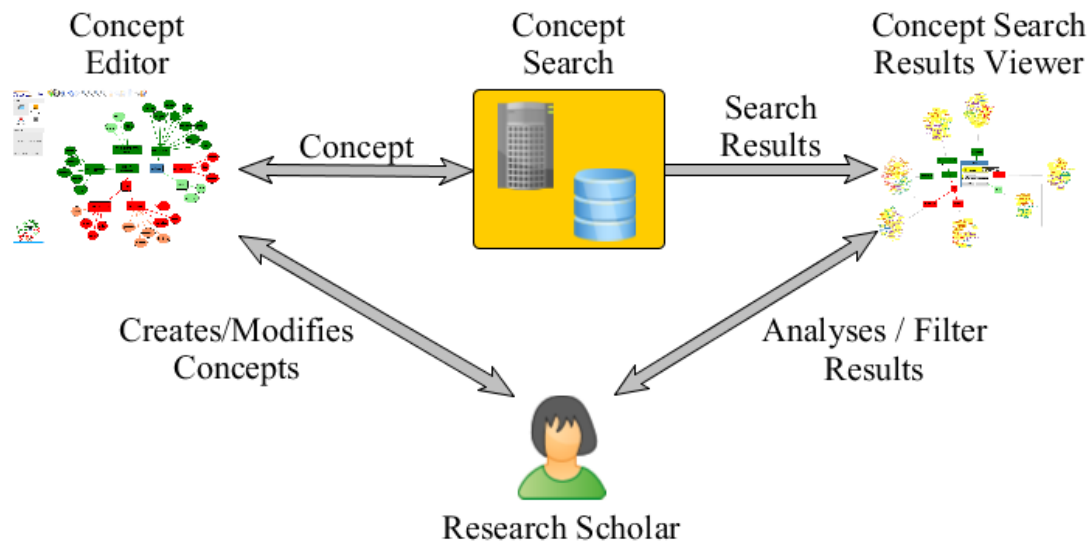


Abbildung 1: Concept Search Pipeline

- Concept Editor, ermöglicht es Geisteswissenschaftler\_Innen nach dem Vorbild von Mind Mapping Werkzeugen, Konzepte entlang ihrer Forschungsfragen zu modellieren.
- Concept Search, Suchalgorithmus, der alle Texte zu dem modellierten Konzept entsprechend ihrer Relevanz abrufen.
- Concept Search Results Viewer, Bereitstellung und Visualisierung der Ergebnisse der Konzeptsuche durch interaktive Statistiken der auftretenden Begriffe mit Hilfe von TagPies.

### 1. Concept Editor

Der Konzepteditor erlaubt es Geisteswissenschaftler\_Innen, ihre Ideen von Konzepten auf dem Bildschirm zu modellieren. Nach der Evaluierung von Prototypen und Feedback von den Geisteswissenschaftler\_Innen wurde die Entwicklung einer ersten Version des Konzepteditors abgeschlossen. Diese Version wurde in einem Workshop final evaluiert. Die Geisteswissenschaftler\_Innen konnten direkt mit dem Werkzeug arbeiten (hands on) und brachten weitere Empfehlungen ein, die in das Tool integriert wurden. Schließlich wurde der Konzepteditor in das Frontend des eXChange-Systems integriert.

Abbildung 2 zeigt einen Screenshot des Konzepteditors mit dem Hauptkonzept „Epilepsie“. Rechteckige Knoten repräsentieren Teilkonzepte, während ovale Knoten konkrete Terme bezeichnen. Die Knoten können per Drag-and-Drop angelegt werden und die Geisteswissenschaftler\_Innen können das Konzept entsprechend ihrer Ideen strukturieren, indem sie zusammengehörige Knoten mit Linien verbinden. Die Knotenfarben zeigen an, ob ein Teilkonzept und seine verbundenen Terme zwingend zu dem Hauptkonzept gehören sollen (grün) oder nicht (rot). Jeder Term kann mit allen möglichen Schreibweisen und Wortformen, die in der Datenbank enthalten sind, verbunden werden. Die Geisteswissenschaftler\_Innen können aber auch selbst eine Auswahl der gewünschten Wortformen und Schreibweisen treffen.

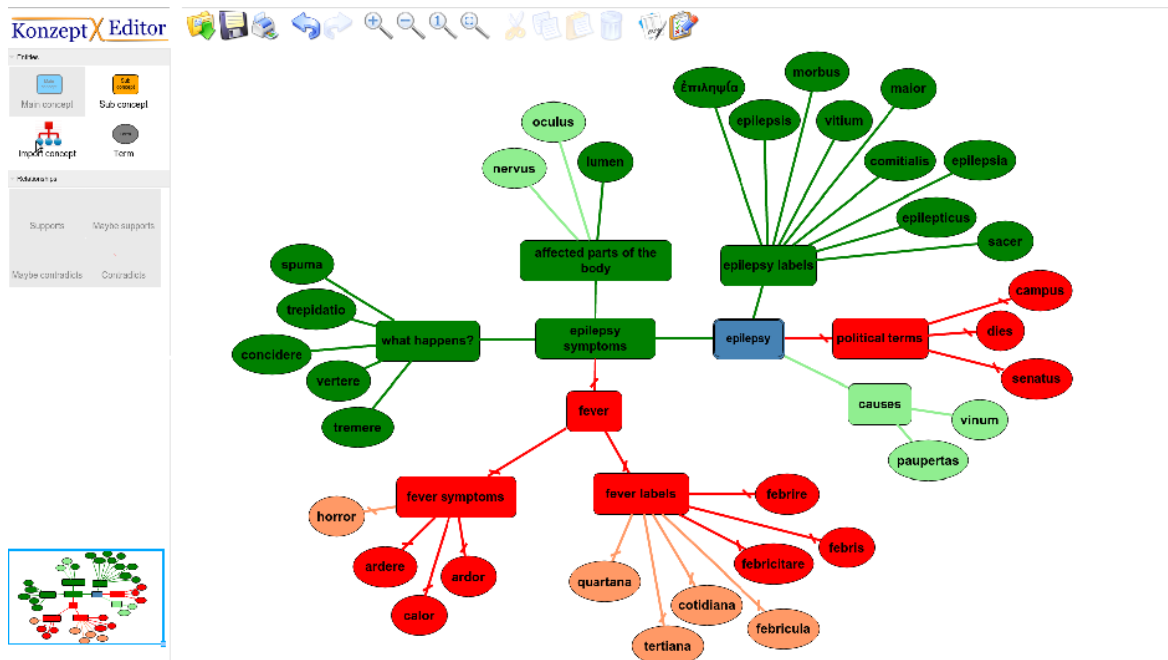


Abbildung 2: Concept Editor Interface mit dem Konzept „Epilepsie“

## 2. Concept Search

Nach der Modellierung wird die Konzeptsuche auf den konstruierten Modellen durchgeführt. Die Konzeptsuche ist ein zweistufiger Prozess. Zunächst wird eine hierarchische Stichwortsuche auf der Grundlage des modellierten Konzepts ausgeführt. Alle Terme im Konzeptmodell werden in der Datenbank nachgeschlagen und die Ergebnisse in einer Matrix zusammengefasst. In einem zweiten Schritt wird jedes Suchergebnis dem am besten passenden Knoten in dem Konzeptmodell (Baum/ Zweig) zugeordnet. Statt nur eine Liste der Suchergebnisse anzuzeigen, werden alle Ergebnisse in der Konzepthierarchie abgebildet, die die Wissenschaftler\_Innen selbst kreiert haben und kennen. Dies erleichtert die Beurteilung der Relevanz einer gefundenen Textstelle für die jeweilige Forschungsfrage.

## 3. Concept Search Results Viewer

Der Concept Search Results Viewer unterstützt die Analyse der gefundenen Textstellen in zwei Distant Reading-Ansichten und einer Close Reading-Ansicht. Die erste Ansicht zeigt die Suchergebnisse in einer hierarchischen Ansicht. Abbildung 3 zeigt die Gesamtansicht des Konzepts „Epilepsie“. In dieser Ansicht werden alle gefundenen Texte (inkl. Autor und Titel des Textes) den entsprechenden Knoten des Konzeptes zugeordnet. Die Schriftgröße spiegelt den während der Konzeptsuche ermittelten Rang der Ergebnisse. Die Suchergebnisse werden den am besten passenden Knoten zugeordnet.

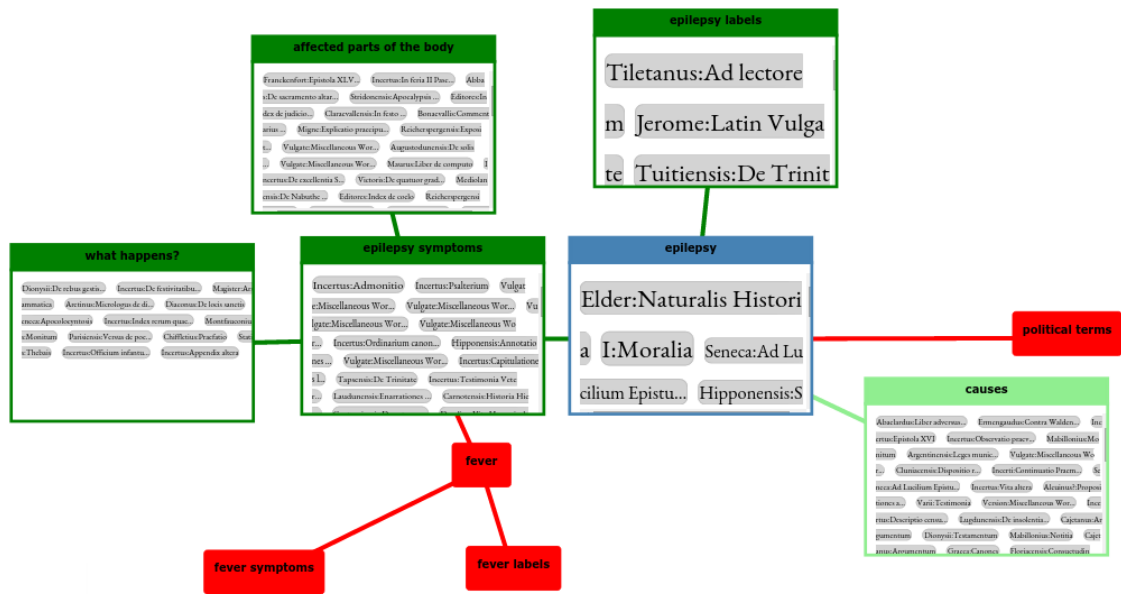


Abbildung 3: Concept Search Results Viewer mit der Anzeige der Ergebnisse für das Konzept „Epilepsie“

In einer zweiten Ansicht (Abbildung 4) können die Beziehungen zwischen den einzelnen Texten und Wortformen interaktiv analysiert werden. Um den Explorationsprozess zu unterstützen, wird eine detaillierte Ansicht angeboten, die Beziehungen zwischen einer Teilergebnismenge (Texte eines einzelnen Knotens) und den Wortformen des Textes zeigt. Um Platz auf dem Bildschirm zu sparen, werden derweil alle anderen Knoten mit Ausnahme der ausgewählten reduziert. Die Terme des ausgewählten Textes werden visuell hervorgehoben. Dies hilft der Anwender\_In, sich auf die jeweilige Aufgabe zu konzentrieren und interaktiv die Ergebnisse betrachten.

Schließlich kann die Nutzer\_In auf einzelne Ergebnisse klicken, um die enthaltenen Textpassage(n) anzuzeigen. Die Wissenschaftler\_In kann durch den Text navigieren und Notizen machen (Abbildung 5).



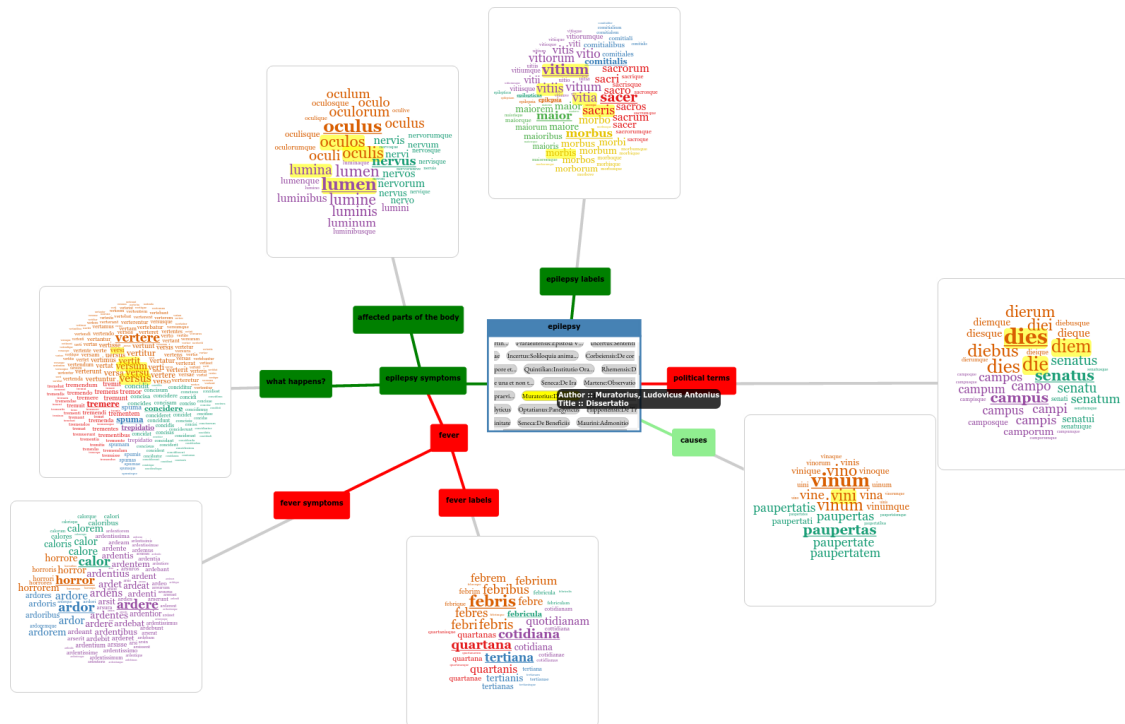


Abbildung 4: Concept Search Results Viewer

Start verse :: vol. 2, lib. 10, p. 372, lin. 28, lin. 1  
 No. of verses :: 2

sic igitur breviter coepat Iuppiter loqui: si mora praesentis leti tempus que **caduco** oratur iuveni me  
 que hoc ita ponere sentis, tolle fuga Turnum atque instantibus eripe fati: ne videretur Iuppiter totum  
 denegasse poscenti, concedit quod fatorum cursum vexare non posset et uxoris postulata non  
 laederet. patior, inquit, ut illum interim ingruenti iam subtrahas morti quem **caducum** esse manifestum  
 est.

Notes

Is Relevant?  
 Yes  No

Previous Sentence Next Sentence

Abbildung 5: Close Reading-Ansicht von Textpassagen für ein ausgewähltes Suchergebnis

#### 4. Fazit

Wir haben eine neuartige Forschungsumgebung für GeisteswissenschaftlerInnen entwickelt, die diese dabei unterstützt, historische Texte auf Konzepte zu untersuchen, die sie nach eigenen Ideen modelliert haben. Das Systemdesign ermöglicht eine iterative Anpassung der Konzepte, so dass eine graduelle Verbesserung der Suchergebnisse erreicht wird.

## 2. Wichtigste Positionen des zahlenmäßigen Nachweises

Position Gesamtfinanzierungsplan	Posten	Entstandene Ausgaben bis einschl. 2016 in EUR
0812	Beschäftigte (E12–E15)	977.024,35
0822	sonst. Beschäftigungsentgelte	77.833,14
0843	sonst. allg. Verwaltungsausgaben	10.716,43
0846	Dienstreisen	13.352,25
0850	Gegenstände über 410 €	14.843,20
Summe:		1.093.769,37

## 3. Notwendigkeit und Angemessenheit der geleisteten Arbeit

Die im Projekt geleisteten Arbeiten waren notwendig und angemessen. Dabei ging es um die Bereitstellung der Datenbasis für die jeweiligen Rechenoperationen, wie auch um die Anwendung, Evaluation und Weiterentwicklung der Tools. Die wissenschaftlichen Mitarbeiter waren notwendig zur fachlich gesicherten Betreuung und Koordination der Arbeiten, der Anleitung der Hilfskräfte, der Sichtung, Aufbereitung und Kontrolle des Datenmaterials, der Kommunikation aufgetretener Probleme an die ASV und BSV, der wissenschaftlichen Evaluation der entwickelten Werkzeuge. Die eingestellten Hilfskräfte waren notwendig zur Recherche aktueller Forschungsliteratur, zur Erstellung zusätzlicher Metadateninformationen und zum Einpflegen des Datenmaterials.

Die durchgeführten Arbeiten ergaben sich zum Großteil direkt aus dem wissenschaftlichen Arbeitsstand und waren jeweils geeignet, den angestrebten Erkenntnisgewinn über die Recherchegegenstände und die untersuchten Technologien zu erreichen. Die nicht direkt wissenschaftlich notwendigen Arbeiten waren solche, die eine Nachnutzbarkeit der Verfahren und Software sicherstellen sollten. Gemessen an dem eingeforderten hohen Innovationsgrad des Projekts waren die Arbeiten, auch wenn sie sich auf unbekanntem technologisch/methodischem Terrain bewegt haben, stets angemessen, da eine konservativere Orientierung nur den aktuellen Stand von Forschung und Technik reproduziert hätte.

#### 4. Ergebnisse/ Wechselwirkungen/ Transfer

Die im Projekt entwickelten Softwarekomponenten, die einen Entwicklungsstand aufweisen, der über einen reinen Laborprototypen hinausgeht, werden als Open Source Software zur Verfügung gestellt.

Kleinere Nebenprodukte sind bereits veröffentlicht: Betacode.js ist eine im Browser lauffähige Javascript-Komponente, mit der sich automatisch griechische Texte zwischen Betacode- und Unicode-Notationen konvertieren lassen (<http://exchange-projekt.github.io/betacode.js/>). Der „BTL Viewer“ ist ein klassisches Recherchesystem, welches einen volltextindizierten Zugriff auf die von DeGruyter zu beziehenden XML-Daten der „Bibliotheca Teubneriana Latina“ (<https://github.com/exchange-projekt/btl-viewer>) ermöglicht.

Die Veröffentlichung des Recherchebackend-Systems „Kadmos“ steht derzeit noch aus (<http://kadmos.textgraph.science/>). Sie wird einhergehen mit wissenschaftlichen Veröffentlichungen und Dokumentationsmaterial für Nutzer (Installationsanleitungen für Systemadministratoren, Instruktionen für den Datenimport, Verwendung bestehender Schnittstellen, Erweiterungsmöglichkeiten für das Programmieren etc.). Grundlegende Überlegungen zum zugrundeliegenden Datenmodell wurden bereits veröffentlicht (Efer, Thomas: Text Mining with Graph Databases – Traversal of Persisted Token-Level Representations for Flexible On-Demand Processing. In: Autonomous Systems – Proceedings of the 8th GI Conference, VDI Verlag, 2015).

Die Forschungsergebnisse in Form von Publikationen sind sowohl über die Webseite des Projektes (<http://exchange-projekt.de/index.html>) als auch über die Webseite des Open Access Online Journals „Digital Classics Online“ (<https://journals.ub.uni-heidelberg.de/index.php/dco/index>) der scientific community zugänglich gemacht worden. Für den August ist ein Sonderheft mit altertumswissenschaftlichen Use-Cases zur Darstellung der Nutzung der im Projekt entwickelten Werkzeuge geplant (s. u. die Abstracts der Beiträge).

Damit ist der Grundstein gelegt für künftige weitere Forschungsvorhaben und neue Kooperationen. Im Vordergrund steht allerdings zunächst die Verstärkung der entwickelten Verfahren in den beteiligten Disziplinen durch die Kommunikation mit interessierten Fachkollegen, die Förderung und Schulung der Nachwuchswissenschaftler und die Einbindung in die Lehre: Im Bereich der Altertumswissenschaften haben bereits Seminare zur Thematik stattgefunden (bspw. das Seminar „Digital Classics“ im WS 2013/2014) um so die Motivation zum verstärkten Einsatz von computer literacy zu erhöhen. Insgesamt halten wir den Einsatz der neuen Methoden in der Lehre – neben der Verbreitung in der Fachcommunity durch Publikationen und Vorträge – für einen wesentlichen Faktor, um die Nachhaltigkeit der Ergebnisse zu gewährleisten.

Als eine abschließende Maßnahme soll der im Rahmen des eXChange-Projektes gekaufte Server (Dual Opteron QuadCore, Rechnungsdatum 02.10.2012, Anschaffungswert 10.087,15€) an das Historische Seminar der Universität Leipzig (Prof. Schubert) mit einem vollfunktionsfähigen Portal sowie allen Datenbanken übergeben werden. Für die technisch-administrative und physische Übergabe der Hardware werden zwei der am Lehrstuhl für Alte Geschichte beschäftigten Informatiker abgestellt, die auch für den Weiterbetrieb des Servers sowie die Datenhaltung und System- bzw. Portaladministration zuständig sein werden. Dadurch sollen die im Projekt entwickelten Tools weiterhin der Forschung zur Verfügung stehen.

## Workshops

Das Verbundprojekt hat während seiner Laufzeit neben zahlreichen Arbeitstreffen und kleineren Workshops auch einen großen Workshop ausgerichtet, an dem neben Mitgliedern des Beirates auch Gäste aus anderen altertumswissenschaftlichen Instituten (u. a. Durham/UK, Erfurt, Graz, Münster, München) sowie aus den Sozialwissenschaften (Hamburg, Osnabrück) und von DARIAH-DE/TextGrid teilgenommen haben. Ziel des Workshops war zum einen die Kommunikation und Diskussion über den aktuellen Arbeitsstand der Teilprojekte, die Kombinationsmöglichkeiten der bereits erstellten Tools untereinander sowie über die anstehenden Herausforderungen. Zum anderen waren die Präsentationen bereits dazu gedacht, die Fortschritte in die Fachgemeinschaft zu tragen und dadurch auch weitere Kooperationspartner bzw. Möglichkeiten für Anschlussprojekte zu gewinnen.

Vorträge & Abstracts des 1. Projektworkshops: 20. März 2013 in Leipzig; 10 – 18 Uhr (Bibliotheca Albertina)

- Thomas Efer: eXChange: Dynamische Metadaten und konzeptbasierte Anfragen an große Textbestände

Eine methodische Verzahnung von Text Mining, Visual Analytics und historischer Diskursanalyse ermöglicht neue Blickwinkel auf existierende Korpora digitalisierter griechischer und lateinischer Texte des Altertums, deren Entstehung bis ins 8. vorchristliche Jahrhundert zurückreicht. Neben der multilingualen und stark diachronen Ausdehnung der umfangreichen Textdaten ergeben sich durch die Einbettung in den Forschungskontext weitere Besonderheiten: So sind die Metadaten, die zu einzelnen Dokumenten und den ihnen zugeordneten Autoren erfassbar sind, teilweise lückenhaft, spekulativ und innerhalb der Wissenschaftscommunity kontrovers diskutiert. Nicht zuletzt ist die Zuordnung von Autoren oft selbst Teil historischer Hypothesen. Im Projekt werden daher dynamische Metadatenfacetten zugelassen: eine Änderung aller Einträge und Zuordnungen kann jederzeit erfolgen und verändert „in Echtzeit“ die Sicht auf das Gesamtkorpus. Entsprechend muss die Erfassung und Verwaltung dieser oft pro Nutzer unterschiedlich gesetzten Metainformationen als separate Aufgabe des Wissensmanagements behandelt werden. In unmittelbarer Konsequenz der flexiblen Facettierung ergeben sich enorme technische Herausforderungen für das Text Mining, besonders in der Indizierung und Strukturierung der Datenbasis und der Gestaltung effizienter Zugriffsmethoden.

Aufsetzend auf den Text Mining-Methoden wird in enger Kooperation mit der Bild- und Signalverarbeitung (BSV) eine Open-Source-Rechercheanwendung entwickelt, mit deren Hilfe Hypothesen zum Konzeptwandel und -transfer aufgestellt, formalisiert, überprüft und angepasst werden können. Dabei sollen Hypothesenbausteine und Analyseergebnisse expliziert werden und in geeigneter Form auch kollaborativ bearbeitet werden können. Basis dafür ist die Adaption konzeptbasierter Korpusabfragen, welche eine domänenspezifische Konzeptualisierung einzelner Begrifflichkeiten über Sprachgrenzen hinweg ermöglicht.

Die Beschäftigung mit dieser Abstraktion der einfachen Suche nach Zeichenketten geht einher mit einer explorativen und gleichzeitig kritischen Auseinandersetzung mit den erfassten Texten. Die direkte Nutzerinteraktion mit Konzepten und den durch sie beschriebenen Belegstellen, ergänzt durch unterstützende statistische Kennzahlen, erlaubt ein Arbeiten im Stil der Visual Analytics, welches die wissenschaftliche Hypothesenbildung auf qualitative wie quantitative Weise fördert.

Als technologische Grundlage der Text-Mining-Methoden werden aktuelle Werkzeuge des Big-Data-Ökosystems sowie alternative Datenbanksysteme auf ihre Eignung für Textdaten hin untersucht und ggf. angepasst.

## Workflows

- Übertragung von geisteswissenschaftlichen Arbeitsabläufen bei Recherche und Hypothesenbildung auf digitale Workflows auf der Grundlage einer detaillierten Beschreibung von Fallstudien.
- => Entwicklung eines Recherchesystems, das es Forschern erlaubt ihre (Hypo)thesen in einer strukturierten Forschungsumgebung zu testen und anzupassen.

- Michaela Rücker: Vergleich der Wissenschaftssprache in Politik und Medizin im Athen des 5./4. Jh.s v. Chr. unter Anwendung der eAQUA-Differenzanalyse

Die Medizingeschichte ist Teil unserer Kulturgeschichte. Man kann an ihr erkennen, wie eine Gesellschaft funktioniert. Die Art und Weise, wie eine Gesellschaft mit fundamentalen Problemen von Körper, Geist, Tod und Krankheit umgeht, spiegelt die Gesellschaft insgesamt wider. Die sich im 5. Jh. v. Chr. entwickelnde medizinische Fachliteratur verwendet Begriffe, Metaphern und Argumentationsstrategien, die in zentralen Bereichen aus der antiken Alltagsanschauung und -ethik, der Politik und der Philosophie stammen. Umgekehrt finden sie auch Eingang aus der Medizin in den Sprachgebrauch von Politik, Poesie, Philosophie etc. Kann man aus der Verwendung dieser Metaphern auf eine sich etablierende Wissenschaftssprache schließen, deren Vokabular so eingängig war, dass ein Transfer keine Begriffserklärungen benötigte, d. h. der Zuhörer, die Metaphern auch in dem Sinne verstand, in dem sie gemeint waren?

In dem gehaltenen Vortrag wurden von M. Rücker die Verwendung medizinischer Metaphern und Termini am Beispiel der attischen Reden und hier spezifisch an denen des großen Rhetors Demosthenes beleuchtet und ein Modell eines möglichen Workflows vorgestellt. Die Auswahl der Texte erfolgte zum einen mit Blick auf die Entwicklung des Metadatentools: Handelt es sich hierbei doch um eine überschaubare Menge gut datierbarer Texte, so dass zuverlässige Metadaten erstellt werden können. Zum anderen fallen diese Texte in eine politisch sehr interessante Zeit und spiegeln vor allem eine gewisse Alltagsebene wider.

Mögliche Forschungsfragen sind:

- 1) Waren sich die Redner möglicherweise der hinter den Metaphern stehender Konzepte bewusst?
- 2) Erfuhr das „medizinische Konzept“ in den Reden eine Umdeutung?
- 3) Haben medizinische Metaphern auch losgelöst vom medizinischen Kontext auctoritas, d. h. war das Konzept auch ohne die dahinter stehende technē der Ärzte wirksam?
- 4) Inwieweit wirkte sich die persönliche Erfahrung des Rhetors auf die Verwendung medizinischer Metaphern aus?

Für die Analyse wurde mit der Differenzanalyse ein Werkzeug gewählt, das im Vorgängerprojekt eAQUA entwickelt wurde. Dieses Werkzeug dient der Ermittlung von diskriminierenden Termen in einem Referenz- und einem Analysekorpus, indem die unterschiedlichen Verteilungen von Wortformen bzw. Wortkombinationen ausgewertet

werden. Bei der vorliegenden Fragestellung sind vor allem die Worte interessant, die in beiden Korpora auftauchen, analysiert wurde das Werk des Demosthenes (Analysekorpus) im Vergleich zum Korpus Hippocraticum (Referenzkorpus).

- Eva Wöckener-Gade: Heilen, Hexen und Vergiften: Zur Entwicklung des Konzeptes φάρμακον in der frühgriechischen Literatur

Wie kaum ein anderes Wort ist φάρμακον verbunden mit der Entwicklung von medizinischer Terminologie in der griechischen Antike. Schon in der Ilias finden sich φάρμακα in der Hand von Ärzten an fast allen Textstellen, an denen es um Heilung geht. In der Odyssee hingegen werden sie (vor allem von zauberkundigen Frauen) zur Verhexung verwendet, aber auch bereits als Gifte, um unliebsame Gegner aus dem Weg zu schaffen.

Dem Begriff, einer so genannten vox media, liegen also von Beginn der literarischen Tradition an verschiedene Konzepte zugrunde, die die Leser\_In voneinander scheiden muss. Im Vortrag wurde – ausgehend von den Quellen – der Versuch unternommen, diese Konzepte näher zu definieren sowie sie und ihre Entwicklung (bis hin zu ihrer Aufnahme in die medizinische Fachterminologie) miteinander zu vergleichen. Schließlich wurde die Übertragbarkeit der aufgefundenen Konzepte auf den lateinischen Begriff venenum diskutiert.

- Dariya Rafiyenko: Wissenschaftssprache und sprachliches Handeln in der byzantinischen Literatur

Das in Byzanz durch die Galenischen und Hippokratischen Schriften tradierte medizinische Wissen (Hunger II 1978, 287–290) diente nicht nur als eine die menschliche Physis pflegende Wissenschaft, sondern auch als ein Weg zum Philosophieren, es fungierte als eine Brücke zwischen Hyle und Nous, zwischen Materiellem und Spirituellem. Die Untersuchung der menschlichen Physis, eine in der ärztlichen Praxis ausgeübte Beschäftigung mit der Schöpfung, sollte mithilfe der Physiologie ermöglichen, die Natur Gottes zu erfassen. Insofern war das medizinische Wissen für die Eliteschicht der gebildeten Byzantiner, die sowohl mit der antiken als auch mit der christlichen Literatur vertraut waren, wohl selbstverständlich. Zahlreiche Belege dafür bieten z. B. Prokopius, Michael Psellos, Anna Komnene, Johannes Tzetzes, Michael Italikos, Michael Choniates, um nur einige zu nennen.

Der Vortrag erläuterte anhand der Frage, wie sich das Medizinwissen im sprachlichen Handeln abbildet, den möglichen Workflow für den Projektteil. Das geschah am Beispiel der Vernetzung des Wortes φάρμακον mit anderen Begriffen wie φίλτρον und ἰυγξ sowie der Bedeutungsverschiebungen der hinter ihm stehenden Begriffe (Gift, Zauberei, Besprechung, Mittel, Medikament) gezeigt werden. Die von der Antike geerbte semantische Polyvalenz der Bedeutung (negativ „Gift“ – neutral „Zaubermittel“, „Mittel (wofür)“ – positiv „Heilmittel“) kann bis in die byzantinische Zeit hinein nachgewiesen werden. Dabei sind im Hinblick auf die besondere Stellung der Medizin in der byzantinischen Geisteswelt bestimmte Tendenzen in der Verschiebung der Bedeutung zu beobachten. So scheint im Laufe der Zeit die negative Bedeutung des Wortes φάρμακον zu verschwinden. In theologischen Schriften und hagiographischen Texten ab dem 3./4. Jh. scheint nur noch die neutrale oder die positive Bedeutung „(Heil)mittel“ erhalten geblieben zu sein (vgl. Lampe s.v., vgl. auch die Komposita φαρμακοθήκη, φαρμακολύτρια), eine positiv-neutrale Bedeutung scheint ebenfalls in der mittel- und spätbyzantinischen Epistolographie etabliert zu sein. Als φάρμακον werden mittels der amphoteroglossia (nach Roilos 2005) z. B. der Brief eines

guten Freundes, die Lösung einer problematischen Situation mit Gottes oder Kaisers Hilfe, eine Steuerentlastung u.a. bezeichnet (Timplallexi 2002, 140). Dabei bleibt die Bedeutung „Zaubermittel“, „Behexung“ zu jeder Zeit erhalten, was auch beispielsweise in volksmedizinischen oder epistolographischen Texten zu beobachten ist. Auf der Grundlage von byzantinischen, nicht-medizinischen Texten wie etwa von juristischen Texten, wird die Vernetzung des Wortes mit weiteren Begriffen untersucht, wobei besondere Aufmerksamkeit den Bedeutungsverschiebungen zwischen unterschiedlichen literarischen Gattungen sowie dem Transfer des Wortes in neue Bereiche gewidmet ist. Mittels der Analyse der sprachlichen Strategien wird nach historischen und sozialen Kontexten gesucht, die zu Bedeutungsverschiebungen des Wortes φάρμακον führten.

- Stefan Jänicke: Dynamische Visualisierung zur Navigation auf großen lateinischen und griechischen Textkorpora

Aktuelle Workflows der am Projekt eXChange beteiligten Geisteswissenschaftler haben gezeigt, dass man nach einer Stichwortsuche mitunter tausende Ergebnisse in Form von Texten bzw. Textstellen erhält. Zudem hat sich ein Bedürfnis nach interaktiven Visualisierungen herauskristallisiert, um die gefundenen Ergebnisse zu analysieren. Daher wollen wir der Anwender\_In eine dynamische Suche auf den zur Verfügung stehenden lateinischen und griechischen Textkorpora mit Hilfe von interaktiven, visuellen Schnittstellen ermöglichen. Dafür transformieren wir gegebene Informationen zu den Daten, sogenannte Metadaten (z.B. Autor, Sprache, Zeit, Ort), in verschiedene Visualisierungen um verschiedene Sichten auf die Daten zu erlauben. Die Nutzer\_In kann den Datenbestand anschließend dynamisch explorieren und mit Hilfe verschiedener Filter eingrenzen. Wir wollen damit ermöglichen, nutzerspezifische Hypothesen zu prüfen und eventuell neue Hypothesen zu generieren.

- Judith Blumenstein / Christoph Weilbach: Entwurf eines Konzeptmodells am Beispiel morbus comitialis (Epilepsie)

Um die Vorstellung und Begrifflichkeit einer Erscheinung wie der Epilepsie im antiken Rom zu erfassen, gibt es in der Klassischen Philologie herkömmliche Rechercheverfahren. Mittels rückläufiger Wörterbücher wird zunächst das lateinische Begriffsspektrum erfasst. Die zu den Begriffen ausgewiesenen Belegstellen werden in den lateinischen Texten eingesehen und mit Hilfe von Kommentaren interpretiert.

Dieses herkömmliche Suchverfahren ist nicht vollständig. Georges liefert zu Epilepsie beispielsweise nur das klassische „morbus comitialis“, während Schenkl nur „epilepsia“ kennt. Zeitliche und gattungsspezifische Aspekte werden aus den Wörterbüchern nicht ersichtlich. Solche Defizite soll die informatische Konzeptsuche überwinden helfen. Mit ihrer Hilfe sollen Konzeptcontainer entworfen und bespeist werden. Als Container wird dabei ein Ordnungssystem vorgestellt, das die textlichen Erscheinungsformen eines Begriffes systematisch nach bestimmten variablen Kategorien (wie syntaktische Verwendung, Attribute etc.) ordnet. Hierzu ist die Klassifizierung der einzuspeisenden Daten nach Autoren, Zeit, Gattung sowie weiteren Parametern (christlich/heidnisch, Dichtung/Prosa) erforderlich. Auf diese Weise können historische, gattungsspezifische und gruppenspezifische Unterschiede und Entwicklungen in der Begriffsverwendung analysiert werden. Daraus kann abschließend ein vollständiges Konzept der Krankheit „Epilepsie“ rekonstruiert werden.

- Roxana Kath: Überlegungen zur Gestaltung eines integrierten Workflows für ein dynamisches Recherchesystem

Der Vortrag demonstrierte die mögliche Unterstützung des geisteswissenschaftlichen Forschungsprozesses durch das zu entwickelnde eXChange-Recherchetool. Exemplarisch wurde nach einer Suchstrategie und Systematisierung für die Niederlagen der Römer in der Zeit der Republik gefragt, die es zulässt, auch aufgrund historischer Überarbeitung uneindeutige Niederlagen in den Texten zu identifizieren.

Hierzu wurden Wege der Konzeptbildung, Visualisierungsmöglichkeiten vordefinierter Konzepte sowie die Vor- und Nachteile von close reading und distant reading vor dem Hintergrund der historischen Kontextualisierung diskutiert. Als Ergebnis konnte festgehalten werden, dass ein zu entwickelndes Recherchetool sowohl close reading als auch distant reading integrieren sollte. Dessen Fähigkeit, live auf Veränderungen der Metadaten und/oder des Konzeptdesigns zu reagieren, würde schließlich die Analyse komplexerer Fragestellungen ermöglichen, die aufgrund der Materialfülle nicht mehr händig zu bearbeiten sind.

## 5. Bekannt gewordener Fortschritt bei anderen Stellen

Es liegen keine Ergebnisse von anderen Standorten vor, die insbesondere gegensätzliche Ergebnisse darstellen. Weitere Arbeiten zur Untersuchung des Konzeptwandels in großen und heterogenen mehrsprachigen Korpora mittels vorberechnungsfreier, parallelisiert berechenbarer Text Mining Verfahren sind nicht bekannt geworden. Die einzige im Ansatz vergleichbare technologische Neuerung stellt das am Deutschen Textarchiv entwickelte System DiaCollo dar. Dieses erlaubt es, ähnlich dem eXChange-System, über die Zeitachse hinweg (allerdings in fest definierten Zeitscheiben) Kollokationen zu verfolgen. Die Backendfunktionalität stützt sich jedoch auf eigens für diesen Anwendungsfall vorberechnete Indexstrukturen, welche es, anders als unsere Methode jedoch z.B. nicht erlauben zur Abfragezeit dynamisch die "Fensterbreite" für Kollokationen zu verändern oder nach beliebigen lokalen Mustern hin zu filtern. (Details s. Bryan Jurish: DiaCollo: On the trail of diachronic collocations. In K. De Smedt (Hrsg.), Proceedings of the CLARIN Annual Conference 2015, Wrocław, Polen, 14.-17. Oktober, S. 28-31, 2015.)

## 6. Erfolgte/ Geplante Veröffentlichungen

### Publikationen

#### Judith Blumenstein

- Visualizing the Results of Search Queries on Ancient Text Corpora with Tag Pies, erscheint in Digital Humanities Quarterly (zus. m. S. Jänicke, M. Rücker, D. Zeckzer, G. Scheuermann)
- Modeling Concepts to Improve the Search Capabilities on Ancient Corpora. In: Proceedings of the Digital Humanities 2015 (zus. m. M. F. Cheema, S. Jänicke, C. Weilbach, G. Scheuermann)
- Exploring Ancient Texts with a User driven Concept Search (zus. m. M. F. Cheema, C. Weilbach, S. Jänicke, G. Scheuermann)



Oliver Bräckel

- Visualisierung und Wissensrepräsentation: Die kleisthenische Phylenreform. In: Das Portal eAQUA – Neue Methoden in der geisteswissenschaftlichen Forschung V, Working Papers Contested Order No. 10 (2013), 1-8 (zus. mit Charlotte Schubert & Corina Willkommen).
- Perser, Meder und Parther im Spiegel der Quellen. In Digital Classics Online 2,3 (2016) i.E.

Muhammad Faisal Cheema

- Modeling Concepts to Improve the Search Capabilities on Ancient Corpora. In: Proceedings of the Digital Humanities 2015 (zus. m. J. Blumenstein, S. Jänicke, C. Weilbach, G. Scheuermann)
- Exploring Ancient Texts with a User driven Concept Search (zus. m. J. Blumenstein, C. Weilbach, S. Jänicke, G. Scheuermann)
- On Close and Distant Reading in Digital Humanities: A Survey and Future Challenges in EuroVis STAR, 2015. (zus. m. S. Jänicke, G. Franzini, M. F. Cheema, G. Scheuermann)
- A Directed Concept Search Environment to Visually Explore Texts Related to User-defined Concept Models (IVAPP, 2016)

Thomas Efer

- Text Mining with Graph Databases – Traversal of Persisted Token-Level Representations for Flexible On-Demand Processing. In: Autonomous Systems – Proceedings of the 8th GI Conference, VDI Verlag, 2015
- Leipziger Rektoratsreden 1871–1933. In: Historical Corpora. Challenges and Perspectives, Narr, 2015 (zus. J. Blecher und G. Heyer)
- Text Mining am Beispiel der Dramen Shakespeares. In: Shakespeare unter den Deutschen, Stuttgart, Franz Steiner Verlag, 2015 (zus. m. G. Heyer, J. Jost)
- Designing Close and Distant Reading Visualizations for Text Re-use, to appear in Computer Vision, Imaging and Computer Graphics. Theory and Application Communications in Computer and Information Science. (zus. m. S. Jänicke, M. Büchler, G. Scheuermann)

Stefan Jänicke

- GeoTemCo: Comparative Visualization of Geospatial-Temporal Data with Clutter Removal Based on Dynamic Delaunay Triangulations in Computer Vision, Imaging and Computer Graphics. In: Theory and Application Communications in Computer and Information Science 359 (2013) 160-175 (zus. mit Christian Heine & Gerik Scheuermann).
- Visualizing Uncertainty: How to Use the Fuzzy Data of 550 Medieval Texts? In: Proceedings of the Digital Humanities (2013) (zus. mit David Joseph Wrisley).

- Visualizing the Results of Search Queries on Ancient Text Corpora with Tag Pies, erscheint in *Digital Humanities Quarterly* (zus. m. J. Blumenstein, M. Rücker, D. Zeckzer, G. Scheuermann).
- Modeling Concepts to Improve the Search Capabilities on Ancient Corpora. In: *Proceedings of the Digital Humanities 2015* (zus. m. J. Blumenstein, M. F. Cheema, C. Weilbach, G. Scheuermann)
- Designing Close and Distant Reading Visualizations for Text Re-use, to appear in *Computer Vision, Imaging and Computer Graphics. Theory and Application Communications in Computer and Information Science*. (zus. m. T. Efer, M. Büchler, G. Scheuermann)
- On Close and Distant Reading in Digital Humanities: A Survey and Future Challenges in *EuroVis STAR, 2015*. (zus. m. G. Franzini, M. F. Cheema, G. Scheuermann)
- A Distant Reading Visualization for Variant Graphs in *Proceedings of the Digital Humanities 2015*. (zus. m. A. Geßner, G. Scheuermann)

#### Roxana Kath

- Aesthetics are (ir)relevant: Für eine Neue Visuelle Hermeneutik in den Geisteswissenschaften. In: *Zeitschrift für Politische Theorie* 5 (1) 2014, 97–120.
- Zeit für einen Paradigmenwechsel in der Politischen Theorie? Der Ansatz der neuen visuellen Hermeneutik. In: André Brodocz / Daniel Schulz / Julia Schulze Wessel (Hgg.): *Die Verfassung des Politischen. Festschrift für Hans Vorländer 2014*, 331–350 (zus. mit G. Schaal).
- New Visual Hermeneutics. In: *Zeitschrift für germanistische Linguistik* 43 (1) 2015, 27–51 (Themenheft „Maschinelle Textanalyse“; zus. mit G. Schaal & S. Dumm).
- New Visual Hermeneutics. In: *Cybernetics and Human Knowing* 23 (2) 2016 (engl.; zus. mit G. Schaal, S. Dumm), 51-75
- Editorial. In: *Digital Classics Online* 2,3 (2016; Sonderheft „eXChange: Exploring Concept Change and Transfer in Antiquity – Moderne Informationstechnologien in den Altertumswissenschaften“ i.E.)

#### Friedrich Meins

- Digitale Editionen in den Altertumswissenschaften? In: *Digital Classics Online* 2,1 (2016), 49-57. (DOI: <http://dx.doi.org/10.11588/dco.2016.1.24491>)
- Anacharsis in der Diskussion über das Kriterium der Wahrheit in Sextus Empiricus' *Adversus mathematicos*. In: *Digital Classics Online* 2,3 (2016) i.E.

#### Dariya Rafiyenko

- „Gicht“ im Altgriechischen. Korpus-basierte Studie („Gout“ in Ancient Greek. Corpus-based study). In: *Digital Classics Online* 2,1 (2016) 6–31. (URI: <http://journals.ub.uni-heidelberg.de/index.php/dco/article/view/24187/21868>)

Michaela Rücker

- Visualizing the Results of Search Queries on Ancient Text Corpora with Tag Pies. Erscheint in *Digital Humanities Quarterly*. (zus. m. St. Jänicke, J. Blumenstein, D. Zeckzer, G. Scheuermann)
- Editorial. In: *Digital Classics Online* 2,1 (2016) 1–5. (URI: <http://journals.ub.uni-heidelberg.de/index.php/dco/article/view/28178/21870>; zus. m. S. Brandt)
- Die Hüter des Hades. Zur Bedeutung und Interpretation der homerischen Kimmerier. In: *Digital Classics Online* 2,3 (2016) i.E.

Charlotte Schubert

- Zauberlehrling und Meister: Digital Humanities zwischen Informatik und Geisteswissenschaften? In: *20 Jahre Arbeitsgemeinschaft Geschichte und EDV*, hrsg. von Jörn Kobes, Kai Ruffing, Wolfgang Spickermann, Gutenberg 2013, 167-186.
- Visualisierung und Wissensrepräsentation: Die kleisthenische Phylenreform. In: *Das Portal eAQUA – Neue Methoden in der geisteswissenschaftlichen Forschung V, Working Papers Contested Order No. 10* (2013), 1-8 (zus. mit Oliver Bräckel & Corina Willkommen).
- Massive Open Online Courses (MOOC): Bildungserlebnis, Weiterbildung oder Geschäftsmodell? In: *VHD Journal* 2014, 16-19
- Aristodemos (Codex Parisinus Supplementum Graecum 607, fol. 83v–85r; 86v–87v): Ein neuer griechischer Atthidograph? In: *Klio* 2014; 96(1): 1–25; online: DOI 10.1515/klio-2014-0001, Juni 2014
- Brudermord und Fremdherrschaft, in: *Spektrum der Wissenschaft* 2/2014.
- Solon, in: *Oxford Bibliographies* 2014, <http://www.oxfordbibliographies.com/view/document/obo-9780195389661/obo-9780195389661-0162.xml>
- Cleisthenes, in: *Oxford Bibliographies* 2014, <http://www.oxfordbibliographies.com/view/document/obo-9780195389661/obo-9780195389661-0168.xml>
- Art. Anacharsis, in *The Encyclopedia of Ancient History* (ed. by R. Bagnall, K. Brodersen, C. Champion, A. Erskine, S. Huebner) Wiley-Blackwell 2015
- Close Reading and Distant Reading. Methoden der Altertumswissenschaften in der Gegenwart, in: <http://hsozkult.geschichte.hu-berlin.de/tagungsberichte/id=5625> (Tagungsbericht: 11.10.2014).
- Close Reading vs. Distant Reading – Quantität und Qualität: Bemerkungen zu epistemischen Anforderungen in den Digital Humanities, Editorial. In: *Digital Classics Online* 1, 2015 (URI: <http://journals.ub.uni-heidelberg.de/index.php/dco/issue/view/1954>)
- Die Hypomnemata bei Plutarch und Clemens: Ein Textmining-gestützter Vergleich der Arbeitsweise zweier ‚Sophisten‘. In: *Hermes* 143 (2015) (zus. m. A. Weiß)
- Digital Humanities: Laboratorium der Geisteswissenschaften oder Weg nach Atlantis? In: *Musikgeschichte zwischen Ost und West: von der ›musica sacra‹ bis zur Kunstreligion. Festschrift für Helmut Loos zum 65. Geburtstag*, hrsg. von St. Keym und St. Wünsche, Leipzig 2015, 747-758.

- Was ist ein Punkt? In: Lexikon der offenen Fragen, hrsg. von Jürgen Kaube/ Jörn Laakmann, Metzler Verlag, Stuttgart 2015
- Die Visualisierung von Quellennetzwerken am Beispiel Plutarchs. In: Digital Classics Online 2,1 (2016), 68–87. (URI: <http://journals.ub.uni-heidelberg.de/index.php/dco/article/view/23825/21867>)
- Annotating and Editing with Canonical Text Services (CTS) Project funded by the Andrew W. Mellon Foundation: 2016–2017. In: Digital Classics Online 2,1 (2016) 94–99. (URI: <http://journals.ub.uni-heidelberg.de/index.php/dco/article/view/28180/21872>; zus. m. Chr. Blackwell)
- Clio-Online Guide Alte Geschichte (zus. mit C. Willkommen, <http://guides.clio-online.de/guides/epochen/alte-geschichte/2016>)
- Das Datum des Phidias-Prozesses, die Aufstellung der Athena Parthenos und der Ausbruch des Peloponnesischen Krieges bei Philochoros, erscheint in: Mnemosyne 2016
- Galens Arbeitsweise: Textmining und Textanalyse (zus. mit Friedrich Meins), erscheint in: K.-H. Leven, Tagungsband Galen aus Pergamon, (Philosophie der Antike, hrsg. von W. Kullmann), 2016
- Heraklit und die ionischen Isonomien, in E. Fantino/ U. Muss/ Ch.Schubert/ K. Sier, Heraklit im Kontext (Beiträge zur Altertumswissenschaft, Berlin: De Gruyter 2016)
- Die Arbeitsweise Plutarchs: Notizen, Zitate und Placita (erscheint in: Rheinisches Museum für Philologie 2016)
- Der Protreptikos des Iamblich, in: S. Brandt/ Ch. Schubert: Textmining und Textanalyse zu Iamblichs Arbeitsweise, Digital Classics Books, Heidelberg 2016 (i.E.)
- Metropolenanspruch zwischen Tyrannis und Demokratie: Die attischen Hermen (zus. M. Martin Flashar), erscheint in: P. Niklas/ a. Hoffmann-Maxis, Metropolen, Berlin 2017 (i.E.)

#### Christoph Weilbach

- Modeling Concepts to Improve the Search Capabilities on Ancient Corpora. In: Proceedings of the Digital Humanities 2015 (zus. m. J. Blumenstein, M. F. Cheema, S. Jänicke, G. Scheuermann)
- Exploring Ancient Texts with a User driven Concept Search (zus. m. J. Blumenstein, M. F. Cheema, S. Jänicke, G. Scheuermann i.E.)
- Wie Laien und Fachleute über Medizinisches sprechen. Medizinische Äußerungen in den Briefen Ciceros, Senecas d. J. und Plinius‘ d. J. im Vergleich mit entsprechenden Ausdrücken und Formulierungen in der zeitgenössischen römischen medizinischen Fachliteratur, Diss. Leipzig (Abgabe voraussichtlich Oktober 2016).

#### Eva Wöckener-Gade

- Wann ist ein terminus technicus ein terminus technicus? – Das Beispiel ἀτρεμής im Corpus Hippocraticum. In: Digital Classics Online 2,3 (2016) i.E.
- Zum rätselhaften Ausdruck γυνίας im Mittelstück der Orestie des Aischylos. In: Hermes 2016 (i.E.)

- Roxana Kath: Editorial

- Eva Wöckener-Gade: Wann ist ein terminus technicus ein terminus technicus? – Das Beispiel ἀτρεμής im Corpus Hippocraticum

Gegenstand des Aufsatzes ist die im gräzistischen Teilprojekt von eXChange untersuchte Frage, inwiefern die antike griechische Fachsprache im Bereich der Medizin durch die vorhergehende Literatur und hier vor allem durch die Dichtung geprägt ist. Feststellbar und nachzuvollziehen ist eine solche Prägung am ehesten anhand der Übernahme bestimmter Terminologien aus der frühgriechischen Dichtung in die medizinischen Schriften, besonders in das Corpus Hippocraticum (CH). In verschiedenen Untersuchungen wie derjenigen zum Begriff φάρμακον konnte innerhalb des Projektes nachvollzogen werden, was sowohl in Arbeiten zu antiken Fachtexten als auch innerhalb der modernen Linguistik schon früher postuliert wurde, nämlich dass bei Übernahme eines allgemein gebräuchlichen Begriffes in eine Fachsprache dieser für gewöhnlich eine Verengung der Bedeutung erfährt. Mit Blick auf diesen als typisch zu charakterisierenden Übernahmemechanismus untersucht der Aufsatz den Gebrauch von Begriffen innerhalb des CH, bei denen eine medizinisch konnotierte Bedeutungsprägung schon in der früheren Dichtung vorliegt und die keine weitere Bedeutungsverengung erfahren haben. Das gewählte Beispiel ist das Adjektiv ἀτρεμής samt den ihm verwandten Wortformen.

- Michaela Rücker: Die Hüter des Hades. Zur Bedeutung und Interpretation der homerischen Kimmerier

Eine intensive Beschäftigung mit dem Volk der Kimmerier führt zu der in der Fachwelt immer wieder als „Kimmerierproblem“ formulierten Fragestellung, die Tim Bridgman als Titel seines Aufsatzes gewählt hat: „Who were the Cimmerians?“. Daran knüpft sich in vielen wissenschaftlichen Untersuchungen auch die Frage an, aus welchem Gebiet die Kimmerier kamen. Diese Überlegungen spielen für die Untersuchung eine eher sekundäre Rolle. Im Vordergrund steht mit Homer der Ausgangspunkt der Kimmerierbeschreibungen innerhalb der griechischen Quellen, seine Aussagen wurden von späteren Autoren aufgegriffen und interpretiert. Die unzähligen Versuche, entsprechende Informationen aus der Odyssee mit denen des im 7. Jh. v. Chr. in Kleinasien einfallenden Reitervolkes in Verbindung zu bringen, konnten bereits in der Antike kein klares Bild über dieses Volk liefern. Und auch die Interpretationen von Historikern und Philologen weisen immer wieder in ganz unterschiedliche Richtungen. Der Beitrag versucht das Bild der mythischen Kimmerier von dem der historischen zu lösen, um der Frage nach Herkunft und Lebensweise auszuweichen. Dem steht auch die Vermutung voran, dass Homers Beschreibung nicht beeinflusst war von einfallenden Reiterhorden aus dem Norden und dass Interpretationen über die realen Wohnsitze auch anders gedeutet werden können.

Hierbei liefert vor allem der in eXChange entwickelte Konzepteditor neue Wege zur Interpretation der bekannten Quellen. Der Aufsatz zeigt die Erstellung und Nutzung dieses Werkzeuges auf, verbunden auch mit den Schwierigkeiten im Umgang.

- Oliver Bräckel: Perser, Meder und Parther im Spiegel der Quellen

Im Mittelpunkt der Untersuchung steht die Anwendung von Visualisierungen in Form von Wordclouds für altertumswissenschaftliche Fragestellungen. Im Rahmen des Projektes eXChange wurde das Prinzip der Wordclouds in Form von TagPies auf die Textsammlung des TLG (Thesaurus Linguae Graecae) angewandt, um mit Hilfe dieser fast vollständigen Sammlung griechischer Autoren der Antike und des Mittelalters eine konkrete historische Fragestellung bearbeiten zu können. Der ausgewählte Untersuchungsgegenstand umfasst die Frage nach der Darstellung der drei Völker Perser, Meder und Parther in den griechischen Quellen und beginnt mit einer Reihe von Vorüberlegungen, die in die Untersuchung der Visualisierungen einfließen: In welchen verschiedenen Kontexten werden diese dargestellt und inwiefern überschneiden sich diese? Ebenso interessant ist die Frage, ob und in welcher Form diese Völker von den antiken Autoren zueinander in Beziehung gestellt wurden. Relevant ist dabei vor allem die Frage nach dem Fremdenbild der Griechen und den Einflüssen, die diese auf die jeweiligen Darstellungen hat.

- Friedrich Meins: Anacharsis in der Diskussion über das Kriterion der Wahrheit in Sextus Empiricus' Adversus mathematicos

Der Beitrag befasst sich mit der Frage nach der Rolle des Anacharsis bei Sextus Empiricus und dem Platz, den der Skythe in der Diskussion über das Kriterion einnimmt. Auf die Rolle, die die Figur des Anacharsis hier für die Argumentation des Sextus spielt, wurde in der modernen Forschung bisher nur recht vereinzelt eingegangen. Deshalb wird dieser Stelle mit Hilfe der digitalen Werkzeuge neue Aufmerksamkeit geschenkt.

## Qualifizierungsarbeiten

### 1. Promotion

Bräckel, Oliver: Politische Flüchtlinge im Römischen Reich.

Efer, Thomas: Graphbasierte Text-Mining-Verfahren (Einreichung im Herbst 2016)

Jänicke, Stefan: Close and Distant Reading Visualizations for the Comparative Analysis of Digital Humanities Data (abgeschlossen)

Meins, Friedrich: Untersuchungen zu den Antiquitates Romanae des Dionysios von Halikarnassos (abgeschlossen)

Weilbach, Christoph: Wie Laien und Fachleute über Medizinisches sprechen. Medizinische Äußerungen in den Briefen Ciceros, Senecas d. J. und Plinius' d. J. im Vergleich mit entsprechenden Ausdrücken und Formulierungen in der zeitgenössischen römischen medizinischen Fachliteratur

Wöckener-Gade, Eva: ἡ δ' ὑπόθεσις: Eine Untersuchung zu den Euripides-Hypothesen in ihrem Verhältnis zu den Dramen (abgeschlossen)

### 2. Habilitationsvorhaben:

Kath, Roxana: Anfänge nachhaltigen Denkens

Rücker, Michaela: Fremdvölker im Spiegel der Utopienliteratur

## Vorträge

Muhammad Faisal Cheema

- Modeling Concepts to Improve the Search Capabilities on Ancient Corpora (Digital Humanities 2015)
- A Directed Concept Search Environment to Visually Explore Texts Related to User-defined Concept Models (IVAPP, 2016)

Thomas Efer

- „Graph databases for the exploration of large corpora in the eHumanities“; Vortrag auf der International Conference of the German Society for Computational Linguistics and Language Technology, 27. September 2013 in Darmstadt
- 15.–16. Mai 2014, Mainz: Teilnahme am Symposium „Shakespeare unter den Deutschen“ der ADW Mainz (Akademie der Wissenschaften und der Literatur); Vortrag zur Anwendung statistischer und graphbasierter Text-Mining-Methoden für die Analyse von Kollektionen dramatischer Texte.
- 22. September 2014, Stuttgart: Teilnahme am Workshop DERM (Datenmanagement und Management elektronischer Ressourcen in Bibliotheken) im Rahmen der „INFORMATIK 2014“ (Jahrestagung der Gesellschaft für Informatik); Vortrag zu graphbasierter Integration und Analyse von bibliographischen Metadaten für quantitative fachwissenschaftliche Analysen.
- 21.–24. Oktober 2014, Rom: Teilnahme an der KDIR (International Conference on Knowledge Discovery and Information Retrieval); Vorstellung eines Papiers zum Text Mining historischer Korpusdaten, gemeinsam mit Christoph Kuras (CLARIN-D).

Stefan Jänicke

- “Visualizing Uncertainty: How to Use the Fuzzy Data of 550 Medieval Texts?”; Vortrag auf der Konferenz Digital Humanities 2013, 15.07.–19. Juli 2013 in Lincoln (USA)

Roxana Kath

- „eXChange: Exploring Concept Change and Transfer in Antiquity: Moderne Informationstechnologien in den Altertumswissenschaften“ (Vortrag im althistorischen Oberseminar, Leipzig, 11. Dezember 2012)
- „Neue methodische Zugänge zu Begriffsgeschichte und Quellenkritik in den Altertumswissenschaften am Beispiel des Portals eAQUA“; Vortrag im Rahmen der Großen Mommsen-Tagung „Landschaften - Umwelten in der klassischen Antike“, 30. Mai 2013 in Göttingen (zus. mit Charlotte Schubert & Michaela Rücker)
- „TextMining – Analyse großer Textdatenmengen“; Vorträge und Präsentationen im Rahmen der Langen Nacht des Wissens in Hamburg, 2. November 2013 (zus. mit Gary S. Schaal)
- Quotation and fragment in the age of digitization Experiences and research perspectives; Vortrag im Rahmen der „Culture & Technology“ – The European Summer University in Digital Humanities 2014 (zus. mit Ch. Schubert).

### Dariya Rafiyenko

- 19. Februar 2015: „Krankheitsbezeichnungen in Byzanz: Neue Aspekte in der Untersuchung unter Verwendung von digitalen Methoden“, Arbeitstagung der Deutschen Arbeitsgemeinschaft zur Förderung Byzantinischer Studien (Heidelberg)
- 12. April 2015: „Gout and Its Formalization“, Workshop „To Mine and To Tie: Text Mining and Network Analysis for Historians“ (Bochum, <http://senereko.ceres.rub.de/en/hnrws2015-en/>; zus. m. H. Kahl)
- 15.–17. Februar 2016: „Corpus-based study of lexicon: gout and παιδεύω“, Postclassical Greek: Intersections of Philology and Linguistics (Mainz)

### Michaela Rücker

- „Neue methodische Zugänge zu Begriffsgeschichte und Quellenkritik in den Altertumswissenschaften am Beispiel des Portals eAQUA“; Vortrag im Rahmen der Großen Mommsen-Tagung „Landschaften - Umwelten in der klassischen Antike“, 30. Mai 2013 in Göttingen (zus. mit Charlotte Schubert & Roxana Kath)
- 23. Juli 2015: „Introduction“ (Digital Classics: Methods, Scholarly Communication and Genres of Scholarly Production. A workshop of the Project “Texte messen – Messungen interpretieren”, zus. mit Stelios Chronopoulos)
- 24. Juli 2015: „New Methods in the Humanities: TextMining in Ancient Greek Corpora“ (Digital Classics: Methods, Scholarly Communication and Genres of Scholarly Production. A workshop of the Project “Texte messen – Messungen interpretieren”, zus. mit Stelios Chronopoulos)

### Charlotte Schubert

- „Neue methodische Zugänge zu Begriffsgeschichte und Quellenkritik in den Altertumswissenschaften am Beispiel des Portals eAQUA“; Vortrag im Rahmen der Großen Mommsen-Tagung „Landschaften - Umwelten in der klassischen Antike“, 30. Mai 2013 in Göttingen (zus. mit Michaela Rücker & Roxana Kath)
- Quotation and fragment in the age of digitization. Experiences and research perspectives; Vortrag im Rahmen der International Summer School, „Digitization and its Impact on Society“ an der TU Dresden, 29. September bis 5. Oktober 2013; 30. September 2013 in Dresden
- Galen und seine Quellen. Neue Zugänge durch Methoden des Textmining; Vortrag auf dem Internationalen Symposium „Galen aus Pergamon. Medizin und Philosophie in der römischen Kaiserzeit“; Erlangen, 17.–18. Oktober 2013; 17. Oktober 2013 in Erlangen
- Digital Humanities zwischen Informatik und Geisteswissenschaften; Vortrag auf der Tagung „Open Access, E-Humanities & E-Science“, 23. und 24. Oktober 2013 in München
- Quotation and fragment in the age of digitization Experiences and research perspectives Vortrag im Rahmen der „Culture & Technology“ – The European Summer University in Digital Humanities 2014 (zus. mit R. Kath).
- Close Reading and Distant Reading. Methoden der Altertumswissenschaften in der Gegenwart; Sektion im Rahmen des 50. Deutschen Historikertages in Göttingen (Organisation & Moderation; zus. mit Tanja Scheer).



- 26. Februar 2015: eXChange: Exploring Concept Change and Transfer in Antiquity; Graz, DHd 2015 (zus. m. R. Kath, M. Rücker)
- 7. April 2015: Podiumsdiskussion „Digital Humanities und transnationale Geschichte“, (Einladung durch die AG Digitale Geschichtswissenschaften), Marburg
- 16. April 2015: Vortrag „Digital Humanities: Laboratorium der Geisteswissenschaften oder Weg nach Atlantis?“ (Einladung von der Kommission für Fachreferatsarbeit des Vereins Deutscher Bibliothekare e. V. mit Unterstützung der Universitätsbibliothek Heidelberg ); Heidelberg
- 11. Mai 2015: Akademientag, Vortrag „Neue Zugänge zur Alten Welt – Digitale Textanalyse“ (Einladung durch den Präsidenten der Union der Deutschen Akademien); Berlin
- 4. Juni 2015: Die Parmenides-Überlieferung: Eine Textmining-gestützte Untersuchung (Einladung zur Konferenz „Parmenidean Ways - The Poem and its Reception“); Leipzig

#### Konferenzteilnahmen & Weiterbildungen

Muhammad Faisal Cheema

- 29. Juni–3. Juli 2015: Digital Humanities 2015, annual international conference of the Alliance of Digital Humanities Organizations (ADHO), in Sydney, Australia

Thomas Efer

- 30.09.–5.10.2012 Herbstschule Information Retrieval in Wadern
- 06.12.–08.12 2012 Internationale Konferenz „Historical Corpora 2012“ in Frankfurt
- 08.–09. April 2013: Kickoff-Workshop der eHumanities-Verbundprojekte 2013 in Leipzig; Poster Vorstellung & fachlicher Austausch mit den anderen BMBF-geförderten eHumanities-Projekten
- 17.–20. Juni: EuroVis, i n Leipzig: Überblick über neue Entwicklungen der Informationsvisualisierung, Interdisziplinärer Austausch, Anregungen zu aktuellen Verfahren der Parallelverarbeitung
- 24.–27. September 2013: Jahrestagung der Gesellschaft für Sprachtechnologie & Computerlinguistik (GSCL2013) in Darmstadt: Teilnahme an den Tutorials „Natural Language Processing for Historical Texts“ und „How to Design Effective Visualizations for Natural Language Processing“
- 30. September: 2. CLARIN-D Disseminationsworkshop in Leipzig
- 18. Dezember 2013: WRoTe – Digital Humanities Dialog in Leipzig; Moderation einer Diskussionsgruppe zum Thema „Open Data, Wiederverwendbarkeit und projektübergreifende Synergien“
- 3. November 2014, Leipzig: Teilnahme am Workshop „Informatik und die Digital Humanities“ des Verbandes DHd (Digital Humanities im deutschsprachigen Raum); Dissemination der technologischen Ansätze und der praktizierten (interdisziplinären) Arbeitsweise des eXChange-Projekts insbesondere im Rahmen des sog. „Knowledge Café“ (Gruppendiskussionen in wechselnder Besetzung)

Stefan Jänicke

- 07. Juli –12. Juli 2014, Lausanne (Schweiz): Digital Humanities 2014
- 29. Juni–3. Juli 2015: Digital Humanities 2015, annual international conference of the Alliance of Digital Humanities Organizations (ADHO), in Sydney, Australia

Roxana Kath

- 08.–09. April 2013: Kickoff-Workshop der eHumanities-Verbundprojekte 2013 in Leipzig; Poster Vorstellung & fachlicher Austausch mit den anderen BMBF-geförderten eHumanities-Projekten
- 21. Juni 2013: TextGrid Nutzertreffen in Mainz
- 5.–7. Dezember 2013: Herrenhausen Conference: „(Digital) Humanities Revisited – Challenges and Opportunities in the Digital Age“, Herrenhausen Palace, Hannover
- 22. Januar – 23. Januar 2014, Hamburg: Workshop „New Visual Hermeneutics“
- 19. Juni 2014, Berlin: Auftaktveranstaltung des Forschungsverbundes Marbach – Weimar – Wolfenbüttel
- 01. September – 02. September 2014, Hamburg: Projektworkshop des BMBF Projektes ePOL „Text Mining in der Politikwissenschaft“ – Anwendung von Text Mining-Tools auf Demokratie bezogene Fragestellungen im Kontext der BR Deutschland (1949-2012).
- 13. Oktober 2014, Erfurt: „Webble“ Workshop des „Interdisciplinary Center of E-Humanities and Social Sciences Erfurt – Leipzig – Trier“ (ICE; Forschungsstelle am Max-Weber-Kolleg, Erfurt).

Dariya Rafiyenko

- 26.–27. April 2013: Workshop „Texte digital: XML und TEI in der geisteswissenschaftlichen Praxis“. Stuttgart
- 28.–29. Juni 2013: Workshop „Geisteswissenschaftliche Forschung in virtuellen Umgebungen: TextGrid“. Stuttgart
- 19.–20. Juli 2013: Workshop zu den Methoden der wissenschaftlichen Visualisierung, unter Leitung von Patrick Sahle (CCeH, Köln) und Ulrike Henny (CCeH, Köln). Leipzig
- 5.–9. August 2013: Workshop „Einführung in R“ im Rahmen der Sommerschule zur allgemeinen Sprachwissenschaft „Academia Salensis“. Salos (Litauen)
- 1.–2. Oktober 2013: „VisArgue Kickoff Workshop“. Konstanz
- 30. November 2013: Workshop „Byzantine Lexicography and the Digital Age“. Wien
- 07. Juli –12. Juli 2014, Lausanne (Schweiz): Digital Humanities 2014.
- 01. Dezember – 04. Dezember 2014, Leipzig: „New Approaches to Digital Critical Editing. Response talk“, Workshop „Greek and Latin in an age of Open Data“ (Lehrstuhl für Digital Humanities).

- 03. Dezember – 05. Dezember 2014, Rom: „The Scholarly Digital Edition and the Humanities. Theoretical approaches and alternative tools“ (organisiert von „DigiLab“, La Sapienza und „DiXIT Marie Curie training network“, <http://infolet.it/2014/12/04/1741/>).
- 15.-17. Februar 2016: In Kooperation zwischen dem Projekt „eXChange“ und der Universität Mainz wurde die Tagung „Postclassical Greek: Intersections of Philology and Linguistics“ (<http://www.rafiyenko.info/tagung/>) organisiert. Auf der Tagung wurden sowohl die im Rahmen des Projekts entwickelten Tools (vgl. den Vortrag von Prof. Schubert) als auch die erzielten Ergebnisse (vgl. den Vortrag von D. Rafiyenko) präsentiert.

#### Michaela Rücker

- 08.–09. April 2013: Kickoff-Workshop der eHumanities-Verbundprojekte 2013 in Leipzig; Poster Vorstellung & fachlicher Austausch mit den anderen BMBF-geförderten eHumanities-Projekten
- 21. Juni 2013: TextGrid Nutzertreffen in Mainz
- 30. September: 2. CLARIN-D Disseminationsworkshop in Leipzig
- 5.–7. Dezember 2013: Herrenhausen Conference: „(Digital) Humanities Revisited – Challenges and Opportunities in the Digital Age“, Herrenhausen Palace, Hannover
- 19. Juni, Berlin: Auftaktveranstaltung des Forschungsverbundes Marbach – Weimar – Wolfenbüttel.
- 13. Oktober, Erfurt: „Webble“ Workshop des „Interdisciplinary Center of E-Humanities and Social Sciences Erfurt – Leipzig – Trier“ (ICE; Forschungsstelle am Max-Weber-Kolleg, Erfurt).
- 23.–25. Juli 2015 Digital Classics: Methods, Scholarly Communication and Genres of Scholarly Production, A workshop of the Project „Texte messen – Messungen interpretieren“ in Freiburg

#### Charlotte Schubert

- 08.–09. April 2013: Kickoff-Workshop der eHumanities-Verbundprojekte 2013 in Leipzig;
- Poster Vorstellung & fachlicher Austausch mit den anderen BMBF-geförderten eHumanities-Projekten
- 30. September: 2. CLARIN-D Disseminationsworkshop in Leipzig
- 5.–7. Dezember 2013: Herrenhausen Conference: „(Digital) Humanities Revisited – Challenges and Opportunities in the Digital Age“, Herrenhausen Palace, Hannover; inklus. Leitung eines Workshops
- 23. September – 26. September 2014, Göttingen: „Gewinner und Verlierer“, 50. Deutscher Historikertag
- 13. Oktober 2014, Erfurt: „Webble“ Workshop des „Interdisciplinary Center of E-Humanities and Social Sciences Erfurt – Leipzig – Trier“ (ICE; Forschungsstelle am Max-Weber-Kolleg, Erfurt).
- 23.–27. Februar 2015: DHd 2015 „Von Daten zu Erkenntnissen“, in Graz

Eva Wöckener-Gade

- 05.–06. April 2013: Forschungsaufenthalt in Köln mit Besuch der Papyrussammlung und Austausch mit dem dortigen Kolleg
- 29. Mai – 01. Juni 2013: Besuch der 5. Internationalen Fayum-Konferenz in Leipzig
- 28.–29. Juli 2013: Besuch der 18. Aquilonia in Dresden
- 27. Juni – 28. Juni 2014, Dresden: „Aquilonia“, Tagung der norddeutschen Institute für Klassische Philologie.

#### Einladungen

- 19. Oktober 2012: Kooperationstreffen mit Prof. Alexander Mehler und Prof. Bernhard Jussen in Frankfurt/M.: Austausch über Projekte (eAQUA u. eXChange); Vereinbarung über den gegenseitigen Austausch von Metadateninformation (insb. latein. Autoren) und über die Nutzung des in Frankfurt aufgebauten lateinischen Wörterbuches im Rahmen von eXChange; Teilnehmer: Schubert, Kath, Rücker
- 29./30. November 2012: Jahrestagung der Arbeitsgemeinschaft Geschichte und EDV e.V. 2012 in Stuttgart: Vorstellung der Projekte eAQUA u. eXChange; Austausch über aktuelle Entwicklungen in den Digital Humanities; Begründung von Kooperationen insb. mit Prof. Christoph Schäfer (Trier) und Prof. Wolfgang Spickermann (Erfurt); Teilnehmer: Schubert, Kath, Rücker
- 04. Februar 2013: Kooperationsgespräch mit Vertretern der Bayerischen Staatsbibliothek München (BSB) in München; Teilnehmer: Schubert, Kath, Rücker
- 04. März 2013: 1. Kooperationsgespräch mit Vertretern des Max-Weber-Kollegs Erfurt in Erfurt; Teilnehmer: Schubert, Kath, Rücker
- 12. April 2013: 2. Kooperationsgespräch mit Vertretern des Max-Weber-Kollegs Erfurt in Erfurt; Teilnehmer: Schubert, Kath, Rücker
- 19. April 2013: Kooperationsgespräch mit Vertretern der FH Erfurt und des Fraunhofer IDMT Erfurt in Leipzig; Teilnehmer: Schubert, Kath, Rücker
- 10. Mai 2013: Kooperationsgespräch mit Vertretern des Medienzentrums der TU Dresden in Leipzig; Teilnehmer: Kath
- 24. Juni 2013: Treffen mit dem Verlag De Gruyter in Berlin; Verhandlung über die Bereitstellung der XML-Daten und Metadaten, Lizenzvereinbarungen und mögliche Kooperationen; Teilnehmer: Schubert, Kath
- 11. September 2013: Kooperationsgespräch mit Vertretern von Brepols und des CTLO in Leipzig; Teilnehmer: Schubert, Kath, Rücker
- 2014: Michaela Rücker wurde externe Beraterin des Projektes „Texte messen – Messungen interpretieren. Altertumswissenschaften und Digital Humanities als zukunftssträchtige Symbiose“ (<http://www.texte-messen.uni-freiburg.de>; Förderung: Heidelberger Akademie Wissenschaften).
- 02. Dezember 2014, Erfurt: Treffen des „Interdisciplinary Center of E-Humanities and Social Sciences Erfurt – Leipzig – Trier“ (ICE; Forschungsstelle am Max-Weber-Kolleg,

Erfurt) in Erfurt zur Vorbereitung eines gemeinsamen Panels im Rahmen der Dhd 2015 in Graz; Teilnehmer: Schubert, Kath, Rucker

## **Kurzfassung der Ergebnisse**

Projekt: „eXChange: Exploring Concept Change and Transfer in Antiquity – Moderne Informationstechnologien in den Altertumswissenschaften“

Projekt im Förderprogramm „Förderung von Forschungs- und Entwicklungsvorhaben aus dem Bereich der eHumanities“ des BMBF

Koordinatorin  
Prof. Dr. Charlotte Schubert

Disziplinäre Verortung  
Alte Geschichte/Byzantinistik/Gräzistik/Informatik/Latinistik

Im Rahmen des Projektes sollte das Verhältnis zwischen Wissenschaftssprache und alltäglichem Handeln insbesondere in der Politik von der Antike bis in die frühe Neuzeit untersucht werden.

Die medizinische Literatur der Antike als Wissenschaftssprache verwendet Begriffe, Metaphern und Argumentationsstrategien, die in zentralen Bereichen aus der antiken Alltagsanschauung und -ethik, der Politik und der Philosophie stammen. Umgekehrt sind Begriffe und Metaphern aus der Medizin in den Sprachgebrauch von Politik, Poesie, Philosophie etc. eingegangen.

Z.B.: metabole (μεταβολή)

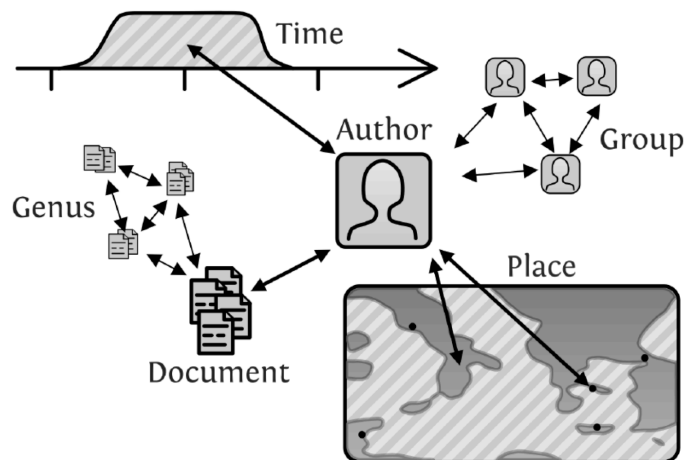
- politisch: Verfassungswechsel im Sinne der Revolution bzw. eines Umsturzes (Aristot. Ath. Pol. 13),
- medizinisch: Umschlag von Gesundheit in Krankheit (Hipp. Über die Umwelt 1), aber auch den therapeutisch induzierten Wechsel der Lebensweise (Hipp., Über die Umwelt 2)

In interdisziplinärer Zusammenarbeit wurde ein Recherchesystem aufgebaut, das Bedeutungsverschiebungen semantischer Räume nach Ort und Zeit visualisiert und durch das Begriffe, die sowohl in medizinischen Fachtexten als auch in politischen und alltäglichen Texten verwendet werden auf ihren Kontext in argumentativen Strategien hin analysiert werden können. Dies kann beispielgebend als neuer, methodischer Zugang für andere geistes- und sozialwissenschaftliche Disziplinen dienen, um diskursive Prozesse in ihrer Verzeitlichung zu repräsentieren. Eine methodische Verzahnung von Text Mining, Visual Analytics und historischer Diskursanalyse ermöglichte neue Blickwinkel auf existierende Korpora digitalisierter griechischer und lateinischer Texte des Altertums, deren Entstehung bis ins 8. vorchristliche Jahrhundert zurückreicht.

### **Dynamische Meta-Daten**

Neben der multilingualen und stark diachronen Ausdehnung der umfangreichen Textdaten ergaben sich durch den Forschungskontext weitere Besonderheiten: So sind die Metadaten, die zu einzelnen Dokumenten und den ihnen zugeordneten Autoren erfassbar sind, teilweise lückenhaft, spekulativ und innerhalb der Wissenschaftscommunity kontrovers diskutiert. Nicht zuletzt ist die Zuordnung von Autoren oft selbst Teil historischer Hypothesen. Im Projekt wurden daher dynamische Metadatenfacetten zugelassen: eine Änderung aller Einträge und Zuordnungen kann jederzeit erfolgen und verändert “in Echtzeit” die Sicht auf das Gesamtkorpus. Entsprechend muss die Erfassung und Verwaltung dieser oft pro Nutzer unterschiedlich gesetzten Metainformationen als separate Aufgabe des Wissensmanagements behandelt werden. In unmittelbarer Konsequenz der flexiblen Facettierung ergaben sich

enorme technische Herausforderungen für das Text Mining, besonders in der Indizierung und Strukturierung der Datenbasis und der Gestaltung effizienter Zugriffsmethoden.



© Th. Efer

### Konzeptsuche

Geisteswissenschaftler\_Innen, die mit historischen Textkorpora arbeiten, stellen in der Regel Keyword-Suchanfragen (= Stichwortsuche), um relevante Textpassagen abzurufen. Das größte Problem der Keyword-basierten Suche ist die geringe Genauigkeit der Ergebnisse. Die Nutzer\_In muss die Abfrage immer wieder reformulieren bzw. verfeinern, um die Qualität der Ergebnisse zu verbessern oder die Anzahl der Suchergebnisse zu reduzieren. Darüber hinaus ergibt die einfache Stichwortsuche in der Regel unvollständige Ergebnismengen, da sie unterschiedliche Wortformen oder Synonyme nicht berücksichtigt. Es gibt zahlreiche Alternativen zur Stichwortsuche (wie z. B. „Topic Models“), aber die meisten dieser Methoden geben den Wissenschaftler\_Innen keine hinreichende Kontrolle über den Suchvorgang. Mit automatisierten Suchverfahren verlieren die Wissenschaftler\_Innen wiederum einen großen Vorteil einer Stichwortsuche, nämlich die Fähigkeit iterativ Suchergebnisse zu verbessern.

Vorrangiges Ziel des Projektes „eXChange“ war es daher, die Zahl und die Genauigkeit der Ergebnisse zu erhöhen, die durch traditionelle Stichwortsuchen erreicht werden, indem eine sog. „Concept Search“ entwickelt und eingesetzt wurde. Eine weitere wichtige Anforderung der Geisteswissenschaftler\_Innen war ein Suchwerkzeug, das sie den Suchvorgang nach Wunsch steuern lässt.

Im Projekt wurde eine nutzergesteuerte Suchumgebung für „Konzepte“ entwickelt, die den Geisteswissenschaftler\_Innen die volle Kontrolle über die Entwicklung der Konzepte sowie die Analyse der Suchergebnisse gibt, und die Möglichkeit, graduell die Suchergebnisse durch iteratives Modifizieren der zugrunde liegenden „Konzepte“ zu verbessern.

### Konzepteditor

Der Konzepteditor erlaubt es Geisteswissenschaftler\_Innen, ihre Ideen von Konzepten auf dem Bildschirm zu modellieren. Nach der Evaluierung von Prototypen und Feedback von den Geisteswissenschaftler\_Innen wurde die Entwicklung einer ersten Version des Konzepteditors abgeschlossen. Diese Version wurde in einem Workshop final evaluiert. Die Geisteswissenschaftler\_Innen konnten direkt mit dem Werkzeug arbeiten (hands on) und brachten weitere Empfehlungen ein, die in das Tool integriert wurden.

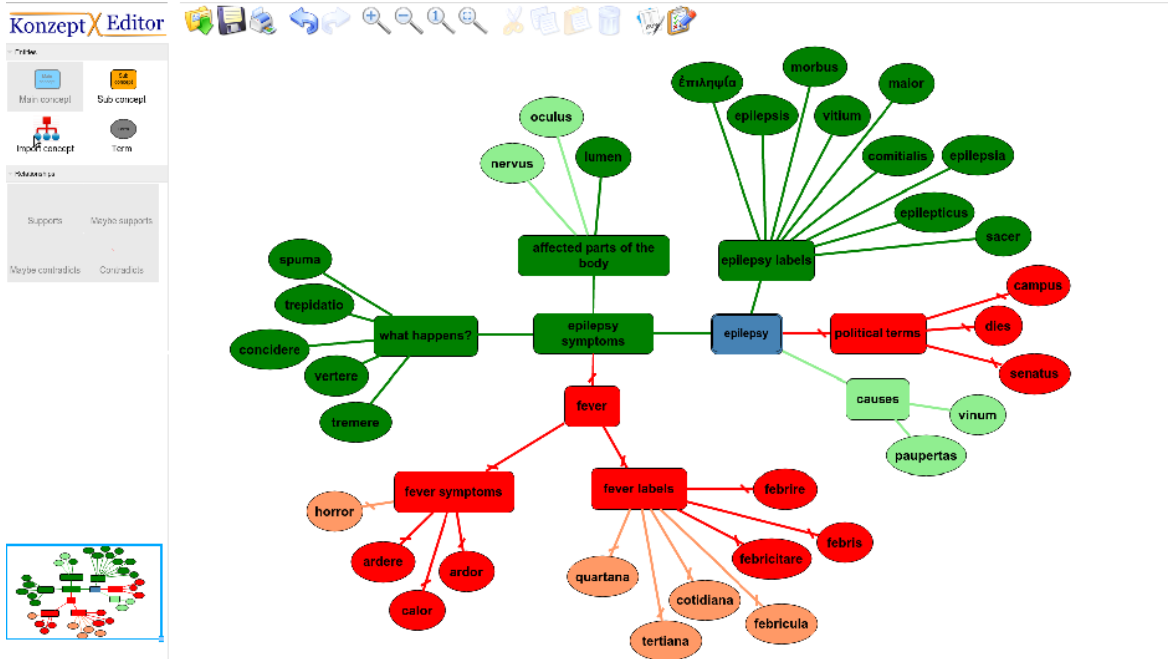


Abbildung 1: Concept Editor Interface mit dem Konzept „Epilepsie“

**Ergebnisanzeige**

Der Concept Search Results Viewer unterstützt die Analyse der gefundenen Textstellen in zwei Distant Reading-Ansichten und einer Close Reading-Ansicht. Die erste Ansicht zeigt die Suchergebnisse in einer hierarchischen Ansicht. Abbildung 2 zeigt die Gesamtansicht des Konzeptes „Epilepsie“. In dieser Ansicht werden alle gefundenen Texte (inkl. Autor und Titel des Textes) den entsprechenden Knoten des Konzeptes zugeordnet. Die Schriftgröße spiegelt den während der Konzeptsuche ermittelten Rang der Ergebnisse. Die Suchergebnisse werden den am besten passenden Knoten zugeordnet.

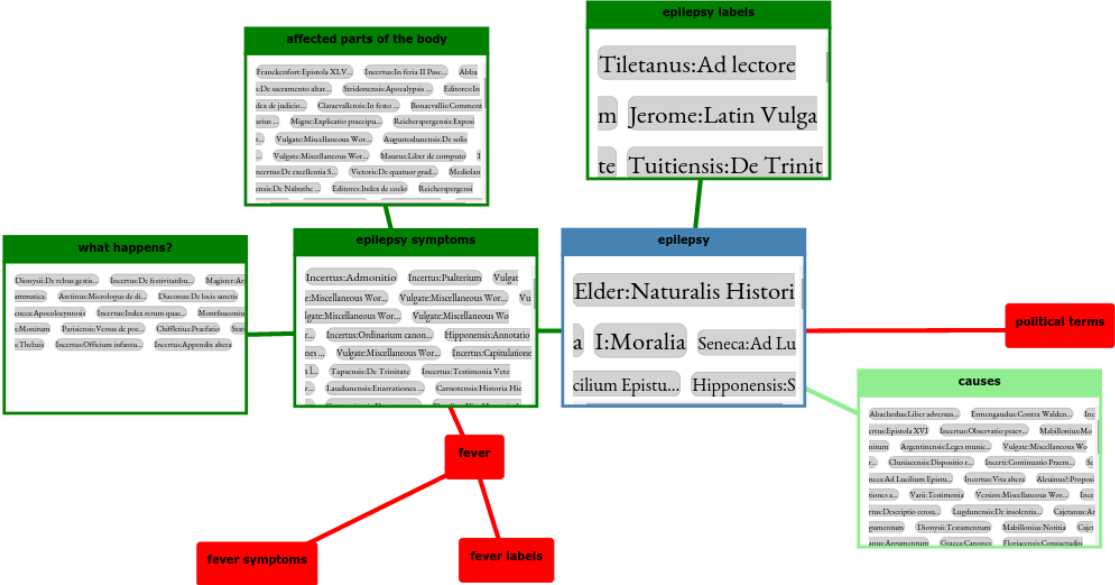


Abbildung 2: Concept Search Results Viewer mit der Anzeige der Ergebnisse für das Konzept „Epilepsie“



In einer zweiten Ansicht (Abbildung 3) können die Beziehungen zwischen den einzelnen Texten und Wortformen interaktiv analysiert werden. Um den Explorationsprozess zu unterstützen, wird eine detaillierte Ansicht angeboten, die Beziehungen zwischen einer Teilergebnismenge (Texte eines einzelnen Knotens) und den Wortformen des Textes zeigt. Um Platz auf dem Bildschirm zu sparen, werden derweil alle anderen Knoten mit Ausnahme der ausgewählten reduziert. Die Terme des ausgewählten Textes werden visuell hervorgehoben. Dies hilft der Anwender\_In, sich auf die jeweilige Aufgabe zu konzentrieren und interaktiv die Ergebnisse betrachten.

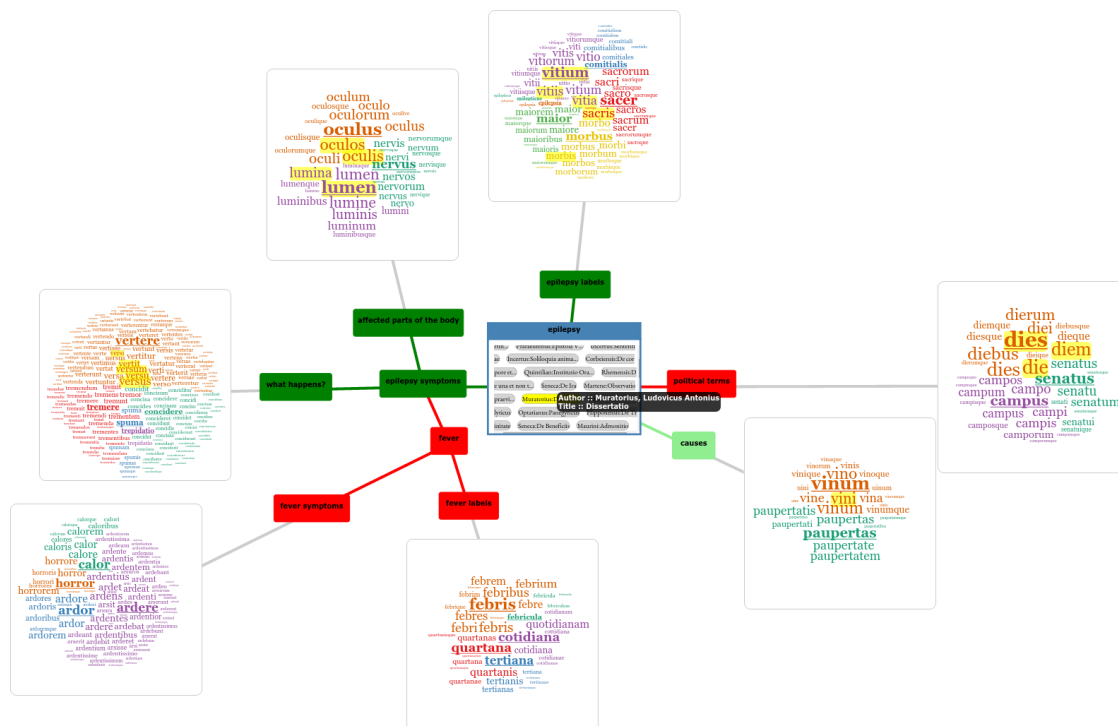


Abbildung 3: Concept Search Results Viewer

Schließlich kann die Nutzer\_In auf einzelne Ergebnisse klicken, um die enthaltenen Textpassage(n) anzuzeigen. Die Wissenschaftler\_In kann durch den Text navigieren und Notizen machen.

## Fazit

Wir haben eine neuartige Forschungsumgebung für Geisteswissenschaftler\_Innen entwickelt, die diese dabei unterstützt, historische Texte auf Konzepte hin zu untersuchen, die sie nach eigenen Ideen modelliert haben. Das Systemdesign ermöglicht eine iterative Anpassung der Konzepte, so dass eine graduelle Verbesserung der Suchergebnisse erreicht wird. Die entwickelten Softwarekomponenten werden dem Antrag entsprechend als Open Source Software zur Verfügung gestellt.