

Fachbeitrag

Hans-Günter Schmidt

Kallimachos: Digital Humanities als Auftrag der Universitätsbibliothek Würzburg

DOI 10.1515/abitech-2016-0035

Zusammenfassung: KALLIMACHOS baut an der Universitätsbibliothek Würzburg ein Digital-Humanities-Zentrum mit den Arbeitsschwerpunkten OCR, digitale Edition, Textmining und quantitative Analyse auf. Das Zentrum bietet eine technische und soziale Infrastruktur, die Geisteswissenschaftler bei der Beantwortung von Forschungsfragen unterstützt und innovative Werkzeuge, Methoden, prototypische Arbeitsabläufe und Dienste entwickelt. In sechs Beispielprojekten aus Papyrologie, Geschichte, Philologie, Philosophie und Pädagogik werden die Neuentwicklungen erprobt, die Fruchtbarkeit der neuen Verfahren vermittelt und Bausteine für ein künftiges Forschungsdatenmanagement geschaffen. Das BMBF unterstützt den Aufbau des Zentrums bis Ende 2017 im Rahmen der Förderlinie „eHumanities“.

Schlüsselwörter: Digitalisierung, Forschungsinfrastruktur, Digital Humanities

Kallimachos: Digital humanities as a university library's task in Würzburg

Abstract: Kallimachos is to set up the Würzburg Center for Digital Humanities. The new center with its special focus on OCR, digital edition, natural language processing, and text analytics will provide a technical and social infrastructure to support research in the humanities by offering a wide portfolio of digital services and by developing innovative tools, workflows, and digital methodologies for the entire process chain of text-oriented research. Six model projects from disciplines as diverse as papyrology, history, German, English, French, Latin and Ancient Greek philology, philosophy, and education illustrate the capability and usefulness of the project's assets. Kallimachos is recipient of a grant from Germany's Federal Ministry of Education and Research (line of funding "eHumanities", 2014–2017).

Keywords: Digitization, Research Infrastructure, Digital Humanities

1 Die UB: „Digitale Rüstkammer“ für die Geisteswissenschaften?

Dass die Bibliothek die älteste Alliierte und „Rüstkammer“ insbesondere der Geisteswissenschaften ist, ist eine Binsenweisheit. Dass sie dann, wenn die Geisteswissenschaften digital werden, auch zur digitalen Rüstkammer werden muss, mittlerweile ebenso. Dabei hat das Thema Digital Humanities in der bibliotheksinternen Diskussion neue, auch hochschulpolitische Relevanz erhalten im Rahmen der sich zuletzt deutlich konkretisierenden Diskussion um Forschungsdatenmanagement und Nachhaltigkeit, die durch neue Leitlinien der Hochschulrektorenkonferenz,¹ der DFG,² der eHumanities-Förderlinie des BMBF³ und der EU im Programm HORIZON 2020⁴ initiiert wurde. Tatsächlich können die Forderungen nach Nachvollziehbarkeit, Interoperabilität und quasi ewiger, unbegrenzter Nachnutzbarkeit von Forschungsdaten auch und gerade in den Geistes- und Kulturwissenschaften nur eingelöst werden, wenn schon relativ früh auf den Datenproduktionsprozess Einfluss genommen wird bzw. die Daten-

1 <https://www.hrk.de/positionen/gesamtliste-beschluesse/position/convention/management-von-forschungsdaten-eine-zentrale-strategische-herausforderung-fuer-hochschulleitungen/> (Empfehlung vom 13.5.2014) und <https://www.hrk.de/positionen/gesamtliste-beschluesse/position/convention/wie-hochschulleitungen-die-entwicklung-des-forschungsdatenmanagements-steuern-koennen-orientierung/> (Empfehlung vom 10.11.2015). Alle in diesem Beitrag zitierten Netzressourcen wurden am 13.6.2016 letztmals aufgerufen.

2 <http://www.allianzinitiative.de/de/handlungsfelder/forschungsdaten/grundsätze.html> bzw. http://www.dfg.de/service/presse/pressemitteilungen/2013/pressemitteilung_nr_06/ und http://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/richtlinien_forschungsdaten.pdf (Appell vom 30.9.2015).

3 <https://www.bmbf.de/de/informationsinfrastrukturen-745.html>; https://www.bmbf.de/pub/forschungsinfrastrukturen_geistes_und_sozialwissenschaften.pdf.

4 http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf (Guidelines on Data Management, 15.2.2016) bzw. <http://www.kowi.de/Portaldata/2/Resources/horizon2020/wp/h2020-wp1617-infrastrukturen.pdf> (Arbeitsplan 2016–2017).

produzenten mit entsprechender Kompetenz, intuitiv nutzbaren Werkzeugen und geeigneten Workflows in die Lage versetzt werden, Forschungsdaten homogen und normierbar zu erzeugen.⁵ Alles andere, etwa das bloße Einsammeln von Daten bei Abschluss eines Drittmittelprojektes, führt in der Regel allenfalls zu „Bitstream-Preservation“ mit flachen Metadaten, nicht aber zur echten „Data Curation“, mit der eine Institution zur Treuhänderin vielfältig vernetzter Forschungsdaten auch über Forschergenerationen und politische und biographische Wechselfälle hinaus wird. Die Entwicklung in den USA, die der in Deutschland um gut ein Jahrzehnt voraus ist, hat hier insbesondere die Rolle von Universitätsbibliotheken betont,⁶ die als zentrale Einrichtungen einer Universität in der Regel über deutlich mehr organisatorische Stabilität verfügen als isolierte Lehrstuhl- oder Fachbereichsinitiativen und die über ihre Fachreferenten bzw. „Subject Librarians“ und Teilbibliotheken deutlich näher an den Desideraten ihrer Kunden in Forschung und Lehre sind, als etwa nationale, vom Alltag der Datenproduzenten entkoppelte Datenzentren. Die digitalen Wege der einzelnen Bibliotheken in der Praxis sind notwendigerweise so unterschiedlich wie die Institutionen und Benutzer, denen sie dienen sollen, auch wenn die Zielstandards der Daten sich international immer weiter verfestigen. Kallimachos, das Zentrum für digitale Edition und quantitative Analyse an der Universitätsbibliothek Würzburg, ist eine Antwort der Julius-Maximilians-Universität auf Problemstellungen rund um digitale Forschungsdaten für einen Teilbereich der Geisteswissenschaften, die Textwissenschaften.

5 HORIZON 2020 (wie Fußnote 4), S. 5 fasst die Anforderungen unter den Schlagwörtern „(1) Discoverable, (2) accessible, (3) assessable and intelligible, (4) useable beyond the original purpose for which it was collected, (5) interoperable to specific quality standards“ zusammen. Vgl. auch die aus Interviews mit Wissenschaftlern gewonnenen Ergebnisse bei Andorfer, Peter: *Forschen und Forschungsdaten in den Geisteswissenschaften. Zwischenbericht einer Interviewreihe*. Göttingen 2015. 10–12 (DARIAH-DE Working Papers 10. urn:nbn:de:gbv:7-dariah-2015-3-8) und, breiter, disziplinenübergreifend und dokumentationsstark, den Abschlussbericht von bwFDM-Communities vom 30.9.2015 (<https://bwfdm.scc.kit.edu/bzw>. <http://bwfdm.scc.kit.edu/downloads/Abschlussbericht.pdf>).

6 Vgl. z. B. <https://www.cni.org/>.

2 Kallimachos: der „große Schritt“ für Würzburg durch ein BMBF-Projekt

Der Name des neuen Zentrums geht auf Kallimachos von Kyrene, Bibliothekar der antiken Bibliothek von Alexandria und einer der Begründer der wissenschaftlichen Philologie, zurück. Sein Name setzt den Akzent, der dieses neue Zentrum prägen soll: Kallimachos war Dichter und Literaturwissenschaftler, zugleich aber auch Verfasser der ersten wissenschaftlichen Bibliothekskataloge, der „Pinakes“, die Literatur nicht nur nach Signaturen, sondern auch geordnet nach Literaturgattungen, nach Autoren und zugehörigem Werk-Korpus, nach charakteristischen Textanfängen, in ihren Bezügen zu anderen Werken und nach Gesamtzeilenzahl, also auch unter quantitativen Gesichtspunkten erschlossen.⁷ Somit ist der Name *Kallimachos* für ein Zentrum, das Bibliothek und digitale Geisteswissenschaften, Edition und quantitative Analyseverfahren zusammenführen soll, treffend.

Die digitalen Geisteswissenschaften – oder, vorsichtiger gesagt, EDV-Verfahren in den Geisteswissenschaften – können dabei in Würzburg auf eine längere Tradition seit den späten 1980er Jahren zurückschauen: Auf mehr als zwanzig Jahre vertieftes Engagement, insbesondere in der Germanistik um Werner Wegstein,⁸ zehn Jahre der Kooperation zwischen Universitätsbibliothek und dem Lehrstuhl Informatik II (Jürgen Albert) mit über achtzig gemeinsam betreuten Praktika, Diplomarbeiten, Promotionen⁹ und nicht zuletzt dem außerordentlich erfolgreichen DFG-Projekt „Libri Sancti Kiliani digital“ zur Digitalisierung mittelalterlicher Handschriften in den Jahren 2010–2013/14.¹⁰ Besonders erwähnt werden müssen das Engagement des Lehrstuhls Informatik VI (Frank Puppe, Andreas Hotho) im Bereich Dokumentenanalyse und Information Extraction,¹¹ viele einzelne Initiativen an Lehrstühlen und Instituten sowie der 2009 im Zentrum institutionalisierte

7 Vgl. Blum, Rudolf: „Kallimachos und die Literaturverzeichnis bei den Griechen. Untersuchungen zur Geschichte der Biobibliographie.“ In *Archiv für Geschichte des Buchwesens* 18 (1978): 169 ff., zu den „Pinakes“ 224–244. Jochum, Uwe: „The Alexandria Library and its aftermath.“ In *Library History* 15 (1999) 5–12.

8 Vgl. z. B. Wegstein, Werner: „TextGrid – eine Geschichte.“ In *TextGrid: Von der Community – für die Community*. Hrsg. v. Heike Neuroth u. a. Glückstadt 2015. 23–28. DOI: <http://dx.doi.org/10.3249/webdoc-3947>.

9 http://www.bibliothek.uni-wuerzburg.de/fileadmin/ub/user_upload/pdf-Dateien/Sondersammlungen/prakt-UB-info2.pdf.

10 <http://libri-kiliani.eu/> bzw. <http://vb.uni-wuerzburg.de/ub/pdfs/lskd/LSKD-Digitalisierung-2013.pdf>.

11 <http://www.is.informatik.uni-wuerzburg.de/>.



Abb. 1: Projektlogo mit Teilprojekten und Modulen.

Studiengang Digital Humanities mit dem Lehrstuhl für Computerphilologie und Neuere deutsche Literaturwissenschaft (Fotis Jannidis), der als Projektpartner in Text-Grid und DARIAH zu den profiliertesten DH-Standorten in Deutschland gehört.¹² Damit baut Kallimachos auf einem breiten Fundament an Einzelinitiativen und über einen längeren Zeitraum erworbenen Erfahrungen auf. Zugleich aber – und dies ist für die Praxis nicht minder wertvoll – gab es durchaus auch zwiespältigere Erfahrungen: Isolierte, an einzelne Personen gebundene Initiativen, steckengebliebene Institutionalisierungsversuche und Ausgründungsprojekte mit schwieriger Geschäftsbasis, proprietäre Sonderwege, Entwicklungsszenarien mit kleineren Firmen, die zunächst „trivial“ erscheinende Problemstellungen im vorgegebenen Zeit- und Kostenrahmen dann doch nicht stemmen konnten und das Erlebnis enger Spielräume in nationalen und regionalen Verbänden. Für die Einsicht, dass sich ein solches Zentrum nicht allein mit Interesse und Idealismus schaffen, geschweige denn verstetigen lässt, war Würzburg insofern ebenso ein guter Ort.

Ausschlaggebend dafür, dass das bisherige breite Engagement nun auf eine neue Qualität gehoben werden kann, ist die Förderung durch das Bundesministerium für Bildung und Forschung im Rahmen seiner Förderlinie „eHumanities“.¹³ Als eines von bundesweit drei geförderten Projekten unterstützt das Ministerium 2014–2017 den

Aufbau von Kallimachos zum regional basierten Zentrum für elektronische geisteswissenschaftliche Basisdienstleistungen und zum Forschungszentrum in den Bereichen OCR, digitale Edition und quantitative Analyse mit überregionalem Anspruch. Kernidee von Kallimachos ist, die vielfältigen nationalen und internationalen Retrodigitalisierungsinitiativen der letzten zwei Jahrzehnte, die eine wahre digitale Bilderflut beschert haben, zusammenzufassen: DFG-geförderte Digitalisierungsprojekte, das Google-Books-Projekt und v. a. m. haben die rein visuelle Verfügbarkeit von Quellentexten für Forschung und Lehre bedeutend erweitert, einfacher und komfortabler gemacht, und dies ist zweifellos von großem praktischen Wert. Doch ein funktionaler Mehrwert durch Verfahren der Informatik beginnt erst dort, wo aus dem reinen Pixelbild auch eine tiefer digital erschließbare, interpretierbare, weiterzuverarbeitende Ressource wird. Hier setzt Kallimachos an, wobei zentraler Bezugspunkt quantitative Verfahren sind: Quantitativ basierte Trainingsverfahren, Sprachmodelle und regelbasierte Auswertung im Postprocessing bei der OCR-Weiterverarbeitung von Scans, Unterstützung von Editionen durch quantitative Verfahren bei der Markup-Erstellung und beim Crowdsourcing sowie quantitative, etwa stilometrische Auswertungsstrategien (statistische Verfahren, die es erlauben, stilistische Unterschiede sichtbar und messbar zu machen) bei der Analyse von Texten. Die Erstellung einer Edition, die früher magistral zwischen zwei Buchdeckeln im Regal stand, ist nicht die Endstation, sondern eine Etappe zu weiterführenden Analysen, die mit spezifischen Textmining-Strategien unterstützt werden. Traditionell qualitative Ansätze der Textwissenschaften sollen quantitativ unterfüttert und damit bereichert werden, und dies in jedem Stadium des philologischen Arbeitsprozesses – kurz, wenn man einen schillernden, überbeanspruchten, aber in diesem Fall

¹² <http://www.germanistik.uni-wuerzburg.de/lehrstuehle/computerphilologie/startseite/> bzw. <http://www.germanistik.uni-wuerzburg.de/lehrstuehle/computerphilologie/studium/>. Der Studiengang hatte in den Aufbaustudiengängen „Linguistische Informations- und Textverarbeitung“ und „EDV-Philologie“ seit 2003 bereits institutionalisierte Vorläufer.

¹³ Vgl. <http://www.pt-dlr-gsk.de/de/992.php> bzw. <http://www.bmbf.de/foerderungen/bekanntmachung.php?B=804>.

nicht unzutreffenden Begriff bemühen will: Big Data für die Geisteswissenschaften.

3 Anwendungsfälle, Produkte und Leistungen in strukturierten Workflows

Dieser strategische Grundgedanke wird im Projektverlauf an sechs Anwendungsfällen erprobt,¹⁴ von der antiken Papyrologie im Teilprojekt „Anagnosis“ (Michael Erler, Holger Essler) über die synoptische Edition von Sebastian Brants „Narrenschiff“ in seinen deutschen, lateinischen, französischen und englischen Textausgaben bis 1500 bzw. 1509 (Teilprojekt „Narragonien“, Leitung Brigitte Burrichter und Joachim Hamm) und bis zur Erschließung von Europas größter Sammlung von Schul-Wandkarten und bildern an der Forschungsstelle Historische Bildmedien Würzburg (Ina Uphoff, Andreas Dörpinghaus). Weitere Teilprojekte erkunden, anhand deutscher und französischer Textkorpora aus dem TextGrid Repository und dem Projekt Gutenberg, Möglichkeiten der automatischen Erkennung von Romanfiguren, deren Attributen und Beziehungen untereinander, sowie Methoden der automatischen Klassifikation von narrativen Techniken, Erzählformen und literarischen Gattungen (Fotis Jannidis, Stefan Evert, Frank Puppe, Andreas Hotho). Ein weiteres Thema ist das bessere Verständnis und die Weiterentwicklung von stilometrischen Verfahren, die auf Delta-Maßen basieren, so etwa zur Identifizierung anonymer Übersetzer in einem so spezialisierten und komplexen Untersuchungskorpus wie lateinischen Übersetzungen von arabischen philosophischen Texten des 12. Jahrhunderts (Dag-Nikolaus Hasse, Stefan Evert). Die „Use Cases“ demonstrieren damit an sehr unterschiedlichen Medienformen aus unterschiedlichen Kulturepochen und -räumen die Fruchtbarkeit der innovativen Editions- und Textanalyseverfahren.

Im Kontext dieser sechs forschungsinduzierten Teilprojekte, die zu Kompetenzaufbau und zu einer lebendigen Kooperationskultur unter den Projektpartnern in Würzburg führen sollen, werden für die breitere Community verschiedene Werkzeuge entwickelt, außerdem Softwarekomponenten, die sich an „Novizen“ und Experten gleichermaßen wenden, und letztlich das Know-how, wie

man damit die spezifischen Anliegen und aktuellen Themen der Wissenschaft unterstützen kann. Alle im Projekt eingesetzten oder zu erarbeitenden Werkzeuge sind dabei auch außerhalb der spezifischen Würzburger Umgebung verwendbar bzw. bedienen Schnittstellen zu anderen standardisierten Infrastrukturen und deren Datenmodellen, z. B. TextGrid mit dem Format TEI, CLARIN mit TCF, UIMA mit DKPro, METS/MODS (nach Definition DFG-Viewer) aus der Bibliothekswelt, PageXML im OCR-Bereich und CoNLL für annotierte Textmining-Daten. Die Communities, die sich bereits gebildet haben, sind damit fest im Blick. Alle Softwareentwicklungen sind ausnahmslos als Open-Source-Komponenten, alle im Projekt erarbeiteten Text- und Bilddaten als Open-Access-Ressourcen unter Creative Commons-Lizenzen konzipiert.

Zu den zu entwickelnden Softwarebausteinen für den Arbeitsbereich Edition gehören ein OCR-Modul, ein Editor und ein Wiki-Modul, die die Erfassung von Texten mit ihren Varianten besonders einfach machen sollen und durch Einbeziehbarkeit von Hilfskräften und Integration in reguläre Seminare und Übungen zugleich den Transfer der Projektthemen und -techniken in die Lehre der kooperierenden Lehrstühle erleichtern sollen. Das OCR-Modul „AnyOCR“, das am Deutschen Forschungszentrum für Künstliche Intelligenz (DFKI) in Kaiserslautern entwickelt wird, soll deutliche Fortschritte bei Segmentierung und Text-Bild-Zuordnung von Handschriften und Frühdrucken erbringen und Erkennungsraten von 95 Prozent und mehr auch bei komplexen historischen Druckwerken wie Inkunabeln möglich machen. Die Entwicklungsarbeiten erfolgen hier im Kontext der Teilprojekte „Anagnosis“ und „Narragonien“. Eine Arbeitsgruppe an der Universitätsbibliothek Würzburg begleitet den Prozess mit der Evaluation von Trainingsverfahren für die bereits etablierten OCR-Engines OCRopus und Tesseract, insbesondere durch Aufbau von verlagsspezifischen Typeninventaren,¹⁵ elektronischen Wörterbüchern und Sprachmodellen für das Postprocessing. Die Module „Synoptischer Editor“ und „Wiki“ sollen hingegen der Korrektur und Nachverarbeitung des OCR-Outputs dienen. Sie sind betont intuitiv gehalten und sollen als einfachere, voraussetzungslosere Editionswerkzeuge neben XML-Editoren wie den TextGrid Editor oder kommerzielle Produkte wie den oXygen-Editor treten. Die Module sind komplett webbasiert, erfordern also von den Anwendern keinerlei lokale Installation, und werden in die rollen- und gruppenbasierte Nutzer-, Redaktions- und Archivierungsverwaltung des lokalen Würzburger Workflowsystems WÜsyphus-2 eingebaut.

¹⁴ Vgl. <http://kallimachos.de>; hier breitere Darstellungen der Teilprojekte, ihre bereits publizierten Ergebnisse und Hinweise auf die jeweiligen Projektmitarbeiterinnen und -mitarbeiter.

¹⁵ Vgl. den Beitrag von Felix Kirchner und Marco Dittrich in diesem Heft.

Narragonien: Sebastian Brants Narrenschiff

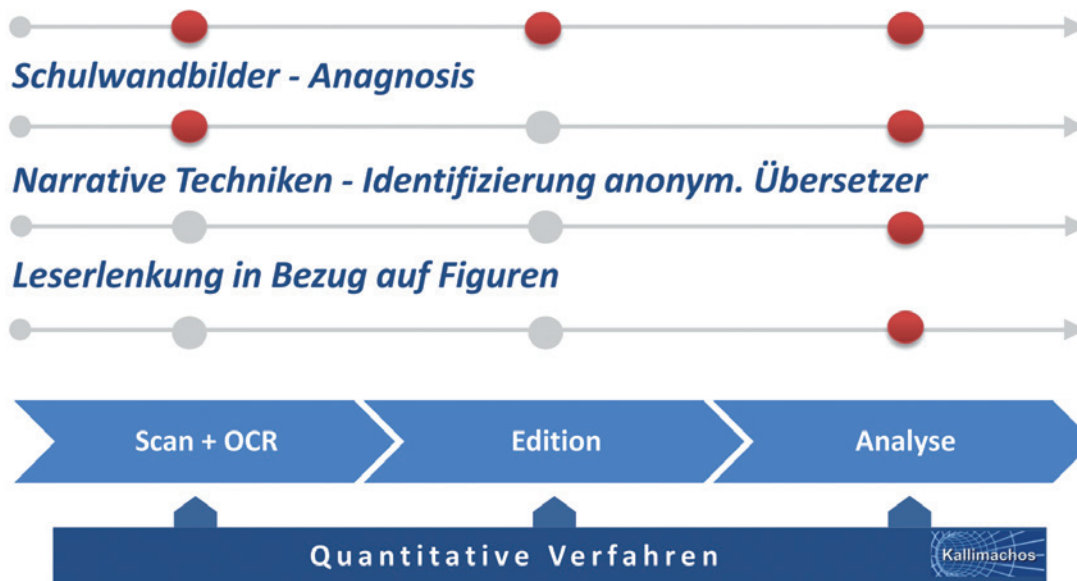


Abb. 2: Teilprojekte, Arbeitsschwerpunkte.

Sie zeichnen sich zudem durch volle Integration von Zeichensätzen der Medieval Unicode Font Initiative¹⁶ aus. Im Zusammenspiel mit einem bereits seit 2012 an der Universitätsbibliothek eingesetzten Metadateneditor, der die Pflege von bibliographischen Daten zu Digitalisaten ermöglicht, sollen in der sich entwickelnden Workflowkette Projektergebnisse größtenteils dynamisch im Portal Virtuelle Bibliothek Würzburg¹⁷ publiziert werden können. Das auf der Basis von Semantic MediaWiki entwickelte Wiki-System wird zudem künftig nicht nur als Nachpflege- und Crowdsourcing-Komponente, sondern in anderer Ausprägung auch als leicht adaptierbare Startplattform für Projektideen und kleinere Datenbankprojekte in der Konzeptions- und Antragsphase im „Werkzeugkasten“ des Kallimachos-Zentrums vorgehalten werden. Inwieweit durch eine Speicherung von Daten aus Kallimachos in das bayerische Langzeitarchivierungssystem ExLibris Rosetta¹⁸ bei Projektende zusätzliche Datensicherheit erreicht werden kann, ist noch offen.

Über die Unterstützung der Texterfassung hinaus wird ein Modul „Schnittstelle Repositories-Datenanalyse“ Analyse- und Annotations-Verfahren in Natural-Language-Processing Pipelines erleichtern. Dieses Modul un-

terstützt, aufbauend auf den Textanalyse-Use-Cases, die Zusammenstellung eines Korpus zu analysierender Texte aus dem TextGrid-Repository bzw. WÜsyphus-2 sowie die Aufbereitung der gewählten Texte mit ihren Metadaten für eine Analyse mit UIMA, die Durchführung der Analyse und schließlich eine Einbindung der Analyseergebnisse in TextGrid, etwa durch die Rückübertragung von Annotationen durch UIMA nach TEI. Diese Teilschritte können ggf. in angepasster Form von anderen Projekten nachgenutzt werden. Zugleich werden Textwissenschaftlern prototypische Arbeitsabläufe mit schon etablierten Open-Source-Tools an die Hand gegeben. Sie sollen ihnen ermöglichen, ihre Forschungsziele in ihrer eigenen Umgebung zu verfolgen – nach ihren eigenen Vorstellungen, für sie selbst durchschaubar und erweiterbar und nach den eigenen Datenschutz- oder Transparenz-Vorstellungen –, können aber auch für Studierende zum Ausprobieren und „Spielen“ bereit stehen. So sollen langfristig sowohl Experten als auch Neueinsteiger im Bereich Datenanalyse von den Vorteilen der automatischen Analyseverfahren profitieren, etwa wenn es um die Erkennung grammatischer Fälle oder von Eigennamen in einem Text geht. Auf dem Weg zu Standardisierung und geläufigerer Verwendung von Textmining-Verfahren soll damit eine große Etappe zurückgelegt werden.

Generell ist der didaktische Aspekt zentral für Kallimachos. Potentielle Nutzer sollen mit den entwickelten Modulen und Workflows auch etwas anfangen können. Zunächst decken die inhaltlichen Teilprojekte von Kal-

¹⁶ <http://folk.uib.no/hnooh/mufi/>.

¹⁷ <http://vb.uni-wuerzburg.de>.

¹⁸ <https://knowledge.exlibrisgroup.com/Rosetta>. Vgl. <http://www.bib-bvb.de/documents/10180/25f54196-2667-4b70-a35f-5f0823c6c420>.

limachos, die „Use Cases“, ein breites Spektrum ab. Ihr Erfolg wäre die beste Werbung für das Zentrum, auch gegenüber weniger technikaffinen Wissenschaftlern. Unterstützt wird Kallimachos dabei durch die ebenfalls BMBF-geförderte Nachwuchsforschergruppe „Computergestützte literarische Gattungsstilistik – CLIGS“¹⁹ (Leitung: Christof Schöch), mit der vermittelt über den Lehrstuhl Computerphilologie vielfältige Kooperationsbeziehungen bestehen. Die Bachelor- und Masterstudiengänge des Studienfaches Digital Humanities in Würzburg sind ohnehin das Fundament vor Ort und stellen einen weiteren Kanal dar, um Methodenkompetenz für digitale Edition und Textmining breiter zu vermitteln. Innerhalb der Universität Würzburg werden proaktiv Lehrstühle und Arbeitsgruppen angesprochen, und tatsächlich gibt es nach einem guten Jahr Projektdauer schon sechs weitere konkrete Würzburger Projektvorhaben, die mit Kallimachos kooperieren wollen. Für die Allgemeinheit werden bis zum Projektabschluss Ende 2017 E-Tutorials und ein Nutzer-Forum im Web zum Einstieg in die digitale Textwissenschaft zur Verfügung gestellt. Über Graduierten- und Sommerschulen und offene Workshops werden Hilfestellung, Anleitung und Anregung zur Übernahme und Weiterentwicklung der Projektinhalte gegeben. Zentrales Schaufenster von Kallimachos ist dabei die etablierte Würzburger Workshop-Reihe „PhilTag“, die alljährlich den Teilnehmern die Möglichkeit gibt, sich über neue Trends, Projekte und Technologien im Bereich der Digital Humanities zu informieren. Seit 2016 liegt die Ausrichtung des „PhilTags“ bei der Universitätsbibliothek Würzburg.²⁰

4 Kooperationsstruktur und Interdisziplinarität: Die UB als Garant der Nachhaltigkeit

Wie werden diese ehrgeizigen Ziele organisatorisch erreicht? Die Kooperationsstruktur von Kallimachos ist im Kern „würzburgisch“ geprägt: Sie baut auf bereits etablierter Zusammenarbeit an der Julius-Maximilians-Universität auf. An zwei entscheidenden Stellen hat sich das Würzburger Konsortium aber zusätzlich verstärkt: Mit dem Erlanger Sprachwissenschaftler und renommierten Fachmann für Korpus- und Computerlinguistik Stefan Evert²¹ und mit

Andreas Dengel²² vom Deutschen Forschungszentrum für Künstliche Intelligenz in Kaiserslautern für den Aufgabenbereich OCR-Verfahren. Neben der breiten interdisziplinären Aufstellung mit Partnern aus verschiedenen Alt- und Neuphilologien, Pädagogik und Philosophie erhält Kallimachos seine besondere Stabilität durch die Universitätsbibliothek, an der das Zentrum institutionell angesiedelt ist. Damit greift Kallimachos Organisationsvorbilder aus den USA, etwa das Center for Digital Scholarship an der Brown University Library (Providence, Rhode Island)²³ auf. Als zentrale Einrichtung der Universität garantiert die Universitätsbibliothek Nachhaltigkeit, und dies nicht nur formal: Die bereits in Personalunion verbundenen UB-Abteilungen „Digitalisierungszentrum“ und „Sondersammlungen – Handschriften und Alte Drucke“ vollziehen mit Kallimachos eine bewusste Hinwendung zu den Digital Humanities. Was traditionell analog betrieben wurde, im Sinne von Bereitstellung besonderer Bibliotheksbestände für Editionsprojekte und entsprechender inhaltlicher Beratung, wird nun auch digital. Dieses neue Engagement hat die Bibliothek mit der Schaffung bzw. Umwidmung von zwei etatisierten Funktionsstellen für Informatiker (E13 bzw. A13/14) untermauert. Sie unterstützen Digital Humanities-Initiativen kontinuierlich und sichern auch über das Projekt im engeren Sinn hinaus Softwarepflege und -fortentwicklung. Das digitale Engagement der Bibliothek ist insofern keine Eintagsfliege, sondern bewusste Positionierung in die Zukunft und bezieht für Organisation und laufendes Management auch weitere Stellenanteile aus Direktion, Personal- und Mittelverwaltung mit ein. Begleitet wurde Kallimachos dabei von der Neueinrichtung des Digitalisierungszentrums auf ca. 400 Quadratmetern in der Zentralbibliothek der UB Würzburg, in dem Scantech und Mitarbeiterbüros konzentriert werden konnten. Seit 2015 stehen damit deutlich günstigere räumliche Gegebenheiten für das Zentrum zur Verfügung. In die Zukunft gedacht robust wird die Organisationsstruktur dadurch, dass sich zusätzlich zwei Lehrstühle dauerhaft in Kallimachos positioniert haben: Der Lehrstuhl Informatik VI mit Frank Puppe und die Digital Humanities mit Fotis Jannidis. Beide werden zusammen mit dem Direktor der UB und dem Leiter des Digitalisierungszentrums auch über das Projekt hinaus im Steuerungsgremium von Kallimachos mitwirken und damit kontinuierlich aktuelle Entwicklungen aus ihren Forschungsbereichen in das Zentrum hineinragen.

Dass ein komplexes Konsortium von zehn Projektpartnern, zu denen weitere unterstützende Partner in Rechen-

¹⁹ Vgl. <http://cligs.hypotheses.org>.

²⁰ <http://kallimachos.de/kallimachos/index.php/Philtag>.

²¹ <http://www.linguistik.fau.de/wir-ueber-uns/personal.shtml/stefan-evert.shtml>.

²² http://www.dfki.de/~dengel/content/index_ger.html.

²³ <http://library.brown.edu/cds/>.

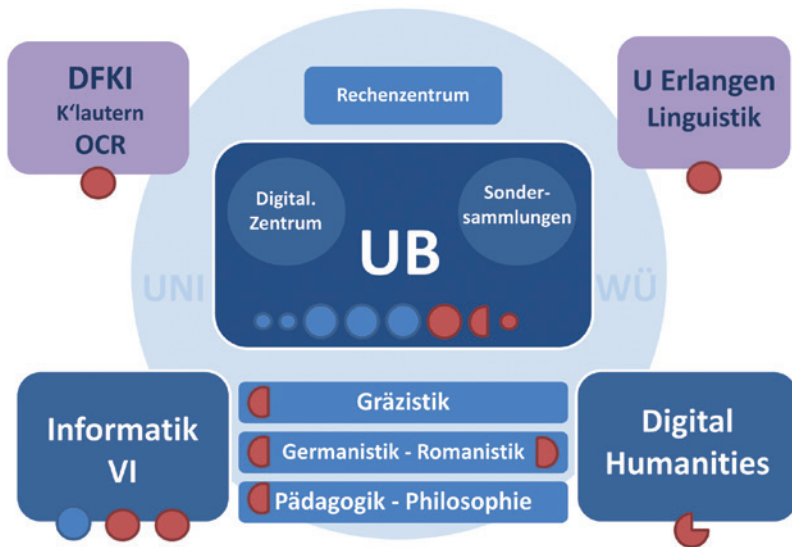


Abb. 3: Kooperationsstruktur (Stellen E13 und E9 – rot drittmittelfinanziert, blau dauerhaft in Kallimachos eingebracht; ohne Hilfskräfte und temporäre weitere Eigenleistungen).

zentrum und Zentralverwaltung der Universität hinzutreten, erheblichen Koordinations- und Fortbildungsbedarf in der Bibliothek, aber ebenso an den Lehrstühlen erfordert, versteht sich ebenso, wie dass der Aufbauprozess auf allen Seiten einen steten Lernprozess bzgl. der Spielräume, Bedürfnisse und Abhängigkeiten der Partner untereinander bedeutet. Die Arbeitsszenarien unterscheiden sich dabei von Arbeitspaket zu Arbeitspaket. Sie reichen, etwa bei der Entwicklung der technischen Module, von eher pflichtenheftbasierten, an „Scrum“²⁴ inspirierten, agileren Projektmanagementformen mit einem Hauch von Industrieprojekt bis zu eher experimentellen, laborartigen Kooperationskulturen in den forschungsorientierten Arbeitsformen. Nicht zuletzt gilt es mancherorts, an Lehrstühlen, in Zentralverwaltung und Rechenzentrum, aber auch – überraschend oder nicht – in der Universitätsbibliothek selbst, die „gute, alte UB“ in einer neuen, forschungsunterstützenden Rolle zu entdecken. Kallimachos soll so auch in einer neuen sozialen Kooperationsinfrastruktur münden und neue Solidaritäten innerhalb der Universität ausbilden helfen.

5 Ziele und Chancen: Mehrwerte für die Textwissenschaften

Mit Kallimachos soll Würzburg als verlässlicher Partner auf der Landkarte der konsolidierten Digital-Humanities-

Zentren erscheinen. Das Zentrum ist, wie bereits dargestellt, eingebettet in ein stabiles lokales Umfeld, mit Studiengängen und Nachwuchsgruppen in der Lehre, mit inhaltlich orientierten Projekten aus vielen geisteswissenschaftlichen Disziplinen in der Forschung. Die Projektergebnisse werden auf andere Forschungsinitiativen der traditionell stark editorisch und quellenorientiert arbeitenden Würzburger Geisteswissenschaften wirken, mit werbenden Effekten auch auf technologiefernere Fachvertreter. Das Projektkonsortium verbindet Anwender und Informatiker über Digital Humanities-Experten der Universität, die als Anwälte für die Forschungsziele der Anwender und Dolmetscher für deren technische Bedürfnisse gegenüber der Informatik fungieren. Den besonderen Akzent erhält Kallimachos jedoch durch die zentrale Rolle der Universitätsbibliothek, die Nachhaltigkeit, technische und personelle Kontinuität und organisatorische Routine sichert, und dies bei vergleichsweise geringem Overhead. Software, Workflows und im Projekt erarbeitete Editionen bzw. Korpora sind breit nachnutzbar, entsprechen internationalen Standards und Formaten. Sie sind interoperabel mit anderen in der Community benutzten Systemen. Durch die Verbindung mit TextGrid, DARIAH und durch den Erlanger Partner auch zu CLARIN ist sichergestellt, dass die Projektergebnisse sichtbar und breit nachnutzbar sein werden.

In die nationale und internationale Community will sich Kallimachos mit einer besonderen Kompetenz einbringen: Textmining und quantitative Auswertungsstrategien im spezifischen Fokus der deutschen Geisteswissenschaften, in allen Aspekten des digitalen Texts und

²⁴ Vgl. z. B. <https://de.wikipedia.org/wiki/Scrum>.

allen Stadien der Editionserstellung und Editionsanalyse. Damit wird spezielle Expertise für einen bedeutenden Aspekt textwissenschaftlicher Forschung aufgebaut; und dabei geht es in den Augen der Projektverantwortlichen um nichts Geringeres als um eine Schlüsselqualifikation für den Geisteswissenschaftler der Zukunft, auch außerhalb der universitären Welt.

Für den Würzburger Standort baut Kallimachos wesentliche Infrastruktur für Forschungsdatenmanagement in den Geisteswissenschaften auf; zunächst ganz vordergründig, indem die technischen, organisatorischen und personellen Voraussetzungen, die Kenntnisse und Fähigkeiten geschaffen werden, um auch künftig, in einem absehbar rein digitalen Forschungsumfeld, drittmittelfähig zu bleiben und flexibel und kurzfristig auf neue Forschungstrends und Anforderungen reagieren zu können. Kallimachos bringt für die Lehrstühle, deren Forschungsunternehmungen zunehmend den Zwei- bis Dreijahresrhythmen der Drittmittelförderung gehorchen müssen und dabei auch in der Lehre höheren Semestern Forschungserfahrungen vermitteln sollen (und oftmals von den Beiträgen der Studierenden und Hilfskräfte bei der Erarbeitung elektronisch nutzbarer Quellenkorpora sogar abhängig sind) an wesentlichen Fronten Entlastung. Tatsächlich überfordern die meisten Vergabebedingungen einschlägiger Drittmittellinien, etwa der DFG, mit ihren anspruchsvollen technischen Vorgaben und der damit einhergehenden mittel- und langfristigen Ressourcenbindung jeden isolierten Lehrstuhl, zumal in den meist nicht üppig dotierten Geisteswissenschaften. Von der Unterstützung bei der Antragsformulierung über die verwaltungstechnische Durchführung und Dokumentation von Drittmittelprojekten, die Nutzung laufend aktuell gehaltener technischer Basissysteme, -Workflows und Publikationsmöglichkeiten bis hin zur vertrauenswürdigen Archivierung weit über den Förderzeitraum eines Forschungsprojektes hinaus, besteht künftig die Option, sich an ein erfahrenes Kompetenzzentrum anzulehnen. Das Know-how wird dabei nicht als Arkanwissen spezialisierter Digital-Humanities-Lehrstühle begriffen, sondern als neue „digitale Hilfswissenschaft“, die nach einer Übergangszeit in den normalen Ausbildungs- und Kompetenzkanon des wissenschaftlichen „Mainstreams“ übergehen wird, auch wenn Wissenschaftler in der Startphase, in der sich Kallimachos gegenwärtig noch befindet, zunächst erst einmal „abzuholen“, „mitzunehmen“ und mit den neuen, digitalen Möglichkeiten (und Zwängen) vertraut zu machen sein mögen.

Nicht zuletzt ist Kallimachos aber die Antwort der Universitätsbibliothek Würzburg auf die digital motivierten Veränderungsprozesse in der geisteswissenschaftlichen

Forschungslandschaft. Die Bibliothek positioniert sich im Kontext des Projektes neu als forschungsunterstützende zentrale Einrichtung ihrer Universität, und zwar „bottom-up“, in der Unterstützung konkreter Forschungsanliegen ihrer Partner. Inwieweit sich dieser Weg als günstiger und schneller zum Ziel führend erweist als der andernorts eingeschlagene Weg über „top-down“ gedachte Forschungsdaten-Policies und politisch motivierte institutionelle Kooperationsvereinbarungen, bleibt abzuwarten. Für die Bibliothek hat der funktionale Veränderungsprozess in einem Teilbereich der Digital Humanities aber bereits begonnen. War sie bislang in erster Linie Hüterin historischer Quellenbestände und Spezialistin für die Speicherung und Verteilung von Information in ihrer publizierten Synthesform, wird sie nun zur Akteurin im Prozess der multiauktoriellen Datenerzeugung selbst und zur Treuhänderin von genetisch sich entwickelnden Themenplattformen, die die klassische „Trennung zwischen Edition als Quelle der forschenden Arbeit und der wissenschaftlichen Erkenntnis als Produkt“²⁵ verschwimmen lassen. Die Form der Darstellung wird dabei für die elektronische Interpretierbarkeit und Neukontextualisierbarkeit der Daten in der Zukunft fast ebenso wichtig wie deren Inhalt selbst. Die dafür notwendige Infrastruktur ist gegenwärtig ebenso Forschungsgegenstand wie die thematisch motivierte Forschung. Und damit mag mit Kallimachos die Universitätsbibliothek von morgen ein wenig wieder der Bibliothek von Alexandria von einst gleichen, die Daten- und Publikationssammlung, Kopier- bzw. Textdatenmigrations- und Erschließungszentrum und sozialer Raum der Forschung in einem war.

Autoreninformationen



Dr. Hans-Günter Schmidt
 Universitätsbibliothek Würzburg
 Leiter der Abteilung Handschriften
 und Alte Drucke
 Leiter des Digitalisierungszentrums
 Projektleiter Kallimachos
 Am Hubland
 D-97074 Würzburg
hans-guenter.schmidt@bibliothek.uni-wuerzburg.de
orcid.org/0000-0002-0511-9516

²⁵ Mittler, Elmar; Malte Rehbein: „Edition und Forschungsbibliothek – Chancen und Herausforderungen einer traditionsreichen Partnerschaft im digitalen Zeitalter.“ In *Digitale Bibliothek und Forschungsbibliothek, Bibliothek und Wissenschaft* 44 (2011): 9–22, hier: 11.