

Das Projekt „TextGrid. Modulare Plattform für verteilte und kooperative wissenschaftliche Textdatenverarbeitung – ein Community-Grid für die Geisteswissenschaften“. Chancen und Perspektiven für eine neue Wissenschaftskultur in den Geisteswissenschaften¹

I. Einleitung

Digitale Publikationen für die Geschichtswissenschaft wurden in den vergangenen Jahren in großer Zahl mit unterschiedlichen Verfahren und Ergebnissen und in sehr heterogener Qualität erstellt. Dennoch ist der Bestand retrodigitalisierter oder digital erstellter Ressourcen im Vergleich zu gedruckten nach wie vor gering.² Sowohl Primärquellen wie auch Forschungsergebnisse haben in den Geisteswissenschaften in der Regel Gültigkeit für mehrere Forschergenerationen bzw. sind von langfristigem Interesse, während sie in einigen Bereichen der Naturwissenschaften sehr kurzlebig sein können und es daher auf die Schnelligkeit der Publikation ankommt. Obwohl es also unterschiedliche Voraussetzungen und Rahmenbedingungen gibt, können alle Wissenschaftszweige von digitalen Publikationen und damit verbundenen neuen Zugriffs- und Analysemöglichkeiten sowie neuen kollaborativen Arbeitsweisen profitieren. Doch werden dafür zunächst vor allem geeignete Kommunikationsformen und Infrastrukturmaßnahmen benötigt.

Dass es immensen Bedarf an digitalen Ressourcen gibt, zeigt eindrucksvoll die „Google Buchsuche“. Ein ausführlicher Beitrag wäre nötig, die Diskussionen darüber oder gar das Angebot selbst kritisch zu sichten und zu bewerten.³ Doch im Folgenden soll es nicht um diese Art der Massendigitalisierung, sondern um Strukturen und Arbeitsweisen gehen, die bewährte und neue Formen der Zusammenarbeit bei der Erarbeitung und Bereitstellung wissenschaftlicher digitaler Publikationen wirksamer als bislang unterstützen können. Zunächst werden das Projekt TextGrid, seine Rahmenbedingungen und Zielsetzungen sowie der derzeit erreichte Stand vorgestellt. Darüber hinaus werden prinzipielle Fragen und Forderungen an eine Forschungsinfrastruktur für die Geisteswissenschaften diskutiert.

II. TextGrid im Rahmen der D-Grid-Initiative

TextGrid ist ein vom Bundesministerium für Bildung und Forschung (BMBF) gefördertes Projekt des Forschungsverbunds D-Grid, in dem in der ersten Phase sechs⁴ Community-Projekte und ein Integrationsprojekt darauf zielen, eine nachhaltige Grid-Infrastruktur in Deutschland aufzubauen. Die Arbeitsbereiche bzw. Disziplinen sind breit gefächert, vertreten sind die Astronomie mit AstroGrid-D, die Klimaforschung mit C3-Grid, die Hochenergiephysik mit HEP-Grid, die Ingenieurwissenschaften mit InGrid, die medizinische Grundlagenforschung mit MediGrid und schließlich die Geisteswissenschaften bzw. genauer die Textwissenschaften mit TextGrid.⁵ Das Konzept des Grid spiegelt – ähnlich wie das World Wide Web – die Idee einer neuen Wissens- und Kommunikationskultur. Grid als Infrastruktur ermöglicht eine gemeinschaftliche Verwendung dezentral erstellter und dezentral verwalteter Ressourcen auf möglichst unkomplizierte Weise. Einbezogen sind dabei sowohl die Rechenleistung, der Speicherplatz, die Software, wie auch die ermittelten bzw. erarbeiteten Forschungsdaten selbst (Primär- und Sekundärdaten).⁶

Nach der Etablierung der D-Grid-Initiative Anfang 2003, die die Integration von Grid-Technologie in die wissenschaftliche Arbeit befördern sollte, wurde vom BMBF das übergreifende e-Science-Konzept entwickelt und für die Bereiche E-Learning, Wissensmanagement und D-Grid ausgeschrieben.

Damit soll die deutsche Wissenschaft Teilhabe an und Sichtbarkeit in der global vernetzten Wissenschaftslandschaft erhalten. Primärdaten, Software und Forschungsergebnisse sollen unmittelbar, nachhaltig und effizient verfügbar gemacht werden.

In den Naturwissenschaften, insbesondere den ressourcenintensiven Bereichen wie z.B. der Physik oder der Klimaforschung ist das Grid-Paradigma seit längerem etabliert, andere Wissenschaftszweige beginnen nun ebenfalls die Vorteile zu entdecken und für sich zu nutzen. Mit TextGrid haben auch die Geisteswissenschaften Anteil an aktueller Hightech-Entwicklung, so dass eine neue Infrastruktur für textbasierte Forschungen entstehen kann: Eine grid-fähige Workbench schafft die Voraussetzung für die verteilte und kollaborative Bearbeitung, Analyse, Annotation, Edition und Publikation von Texten. Verschiedene Strategien erlauben einen integrativen Zugriff auf heterogene und verteilte Daten; dazu gehören beispielsweise standardisierte Metadaten oder textsortenspezifische, auf TEI⁷ basierende TextGrid-Codierungsformate sowie spezifische Ontologien. Die für weitere Projekte offenen Schnittstellen ermöglichen Synergien mit anderen Initiativen in der wissenschaftlichen Textdatenverarbeitung. TextGrid ist derzeit auch international das einzige Projekt, das Grid-Technologie für die Geisteswissenschaften nutzbar macht.⁸

Das TextGrid-Konsortium vereinigt verschiedene Institutionen mit unterschiedlichen Kompetenzen, um den angestrebten Zielen eine breite Basis zu verschaffen. Die Federführung liegt bei der Niedersächsischen Staats- und Universitätsbibliothek Göttingen, die insbesondere ihre Kompetenzen in Projektmanagement und nachhaltiger Infrastruktur-Entwicklung einbringt. Zu den fachwissenschaftlichen Partnern gehören das Institut für Sprach- und Literaturwissenschaft der TU Darmstadt, das Kompetenzzentrum für elektronische Erschließungs- und Publikationsverfahren in den Geisteswissenschaften an der Universität Trier sowie das Kompetenzzentrum für EDV-Philologie an der Universität Würzburg; an der Schnittstelle von Fachwissenschaft, Korpustechnologie und Linguistischer Datenverarbeitung arbeitet das Institut für Deutsche Sprache in Mannheim. Die Fachhochschule Worms liefert den Input im Bereich Web-Service-Technologie, während zwei mittelständische Unternehmen, die Saphor GmbH und die DAASI International GmbH – beide in Tübingen ansässig – technologisches und Grid-Knowhow einbringen.⁹

III. Projektziele und derzeitiger Projektstand

Erklärtes Projektziel ist die Schaffung integrierter Instrumente und Werkzeuge, die zum einen die wissenschaftliche Textbearbeitung (im weitesten Sinne) wirksam unterstützen, die zum anderen auch den Transfer von e-Science-Methoden netzbasierten Arbeitens in die Geisteswissenschaften ermöglichen. TextGrid, im Februar 2006 offiziell gestartet, gliedert sich in sechs Teilprojekte bzw. Arbeitspakete, die im folgenden kurz vorgestellt werden: Gestützt auf Reports über die Nachnutzbarkeit vorhandener Software und Tools (AP 1) werden einzelne Module für die Grid-Anwendung adaptiert oder neu entwickelt (AP 2) und in die Integrationsplattform bzw. Middleware-Umgebung integriert (AP 3). Ihre konkrete Leistungsfähigkeit wird durch Einsatz in Musterapplikationen demonstriert (AP 4). Schwerpunkt des AP 5 ist der Beitrag von TextGrid für die Bildung eines Semantic Grid. Projektmanagement und Öffentlichkeitsarbeit sind Aufgabe von AP 6. Die einzelnen Arbeitspakete bauen aufeinander auf und sind infolge unterschiedlicher Laufzeiten und Starttermine unterschiedlich weit fortgeschritten.

AP 1 – Inhaltliche Studie mit Empfehlungen über die Nachnutzbarkeit internationaler Editonstools: Dieses Teilprojekt kann bereits entsprechende, über die TextGrid-Homepage öffentlich zugängliche Reports zum Thema Text Processing und Linking aufweisen. Einschlägige Literatur wur-

de auf Anwendungsfälle hin gesichtet, aus denen sich Anforderungen an die TextGrid-Workbench ableiten lassen. Ferner wurden relevante Werkzeuge aus diesen Anwendungsbereichen evaluiert und auf ihre Nutzbarkeit für TextGrid hin überprüft.

AP 2 – Entwicklung Community-spezifischer Werkzeuge (Annotations- und Analyse-Tools):¹⁰ Das Teilprojekt entwickelt eine Grid-fähige Workbench für die Erstellung, Bearbeitung, Annotation und Analyse von XML-kodierten Textdaten. Die Workbench wird aus einem erweiterbaren Satz von Modulen aufgebaut. Hiermit entsteht den Textwissenschaften ein umfangreiches Instrumentarium zur Digitalisierung, Verarbeitung, Publikation und nachhaltigen Bereitstellung wissenschaftlicher Quellentexte. Sie schließt alle Textwissenschaften zu einer Community zusammen, die ihre Ressourcen gemeinsam erstellt und einander zugänglich macht. Auch für dieses Arbeitspaket liegt ein öffentlich zugänglicher Report vor. Ferner konnte bei der Zwischenevaluierung im Februar 2007 eine erste Demoversion gezeigt werden, die folgendes Anwendungsszenario ausführt: Ein Beispieltext wird durch einen Tokenizer in Wortformen zerlegt, die durch XML-Tags gekennzeichnet sind, und anschließend mittels Lemmatizer morphologisch analysiert sowie auf die Grundform abgebildet (Infinitiv bzw. Nominativ Singular). Diese Grundform wird einer Wörterbuchsuche übergeben und die entsprechenden Fundstellen in den angeschlossenen digitalen Wörterbüchern im Netz aufgerufen. Alle Module sind eingebettet in eine auf Eclipse basierende Workbench, deren integrierter Workflowmanager das Zusammenspiel koordiniert und die Definition neuer Arbeitsabläufe ermöglicht. Analyseergebnisse werden als XML/TEI ausgegeben bzw. gespeichert. Diese Demo-Musteranwendung wird voraussichtlich noch im Laufe des Jahres 2007 online präsentiert werden können.¹¹

AP 3 – Anbindung der Community-Tools und Vorschläge für Entwicklungen an der Integrations-Plattform: Die in den anderen Arbeitspaketen erstellten Softwarebausteine werden über eine TextGrid-spezifische Middleware-Infrastruktur an die übergreifende Grid-Infrastruktur angebunden.¹² Hierzu muss eine Schnittstelle zu den Tools spezifiziert und ebenso implementiert werden wie eine Schnittstelle zu den Diensten des Grid. Dabei ist darauf zu achten, dass die Entwicklungen bei der Integrationsplattform, die allen D-Grid-Community-Projekten zur Verfügung stehen wird, spezifische Anforderungen und Funktionalitäten von TextGrid berücksichtigt.

AP 4 – Entwicklung der Community Muster-Applikation: Die Nutzbarkeit und Leistungsfähigkeit der Grid-fähigen Werkzeuge wird anhand von Musterapplikationen erprobt und in Tests mit Anwendern unterschiedlicher Profile evaluiert, so dass die Ergebnisse iterativ in den Entwicklungsprozess der Software zurückfließen. Ein Schwerpunkt liegt derzeit bei der Digitalisierung und Erschließung des Campe'schen Wörterbuchs.¹³

AP 5 – Semantic Web und TextGrid = Semantic TextGrid: Im Rahmen von TextGrid werden hochstrukturierte Textmaterialien aus verschiedenen digitalen Wörterbüchern des Deutschen und genuin philologische Methoden und Technologien eingesetzt, um einen übergreifenden gemeinsamen Einstieg in die verschiedenen Wörterbuchinformationen zu ermöglichen, z.B. über ein „Standard-Lemma“ einer „Meta-Lemmaliste“. Diese dient darüber hinaus als Basis-Ontologie für die Erschließung geisteswissenschaftlicher Primärquellen. Derzeit ist ein Webservice in der internen Testphase, der eine flexible, integrierte Suche über zehn digitalisierte Wörterbücher ermöglicht.¹⁴

AP 6 – Projektmanagement und Öffentlichkeitsarbeit: In dieses Teilprojekt fällt die formale und administrative Leitung des Gesamtprojektes. Ein weiterer Schwerpunkt besteht darin, national und international über TextGrid und seine Teilergebnisse zu informieren, wobei besonderer Wert auf den Informationsaustausch mit der eigenen Community gelegt wird.

Neben diesen Arbeitspaketen, die das Projekt strukturieren, sind horizontale Arbeitsgruppen eingerichtet worden, die grundlegende Querschnittprobleme bearbeiten. In der AG Architektur werden die informatischen und technologischen Grundlagen der Architekturkonzeption unter Berücksichtigung der fachwissenschaftlichen Anforderungen erarbeitet. Die AG Textformate definiert das TextGrid-Kernformat auf der Basis von TEI. Die AG Archive diskutiert die Schnittstellen und Anbindungsstrategien externer Archive an TextGrid. Die AGs zu Wörterbüchern und zu linguistischen Korpora fokussieren die Themen aus der AG Textformate und der AG Archive dokumentenspezifisch.

Die AG Print schließlich wurde eingerichtet, da im Laufe der Projektarbeit deutlich wurde, dass dem Modul für anspruchsvollen wissenschaftlichen Satz (z.B. von komplexen Editionen oder von Wörterbüchern) besondere Aufmerksamkeit zukommt, da vorhandene Lösungen unbefriedigend oder nicht OpenSource sind. Sie formuliert daher einen eigenen Drittmittelantrag, um eine gridfähiges OpenSource-Modul für wissenschaftlichen Satz einbringen zu können. Dieses Modul ist für die geisteswissenschaftliche Community besonders wichtig, da gerade Editionen und Nachschlagewerke meist für mehrere Forschergenerationen Gültigkeit behalten und daher nach wie vor (auch) in Buchform publiziert werden.

TextGrid bietet also den Kern einer geisteswissenschaftlich ausgerichteten Kommunikations- und Produktionsplattform, die homogenen und standardisierten, genau definierten und kontrollierten Zugriff auf Werkzeuge und Ressourcen bei zugleich größtmöglicher Flexibilität gestattet. Institutionen oder Wissenschaftler sollen ihre Texte unter der TextGrid-Oberfläche und mit den entsprechenden Möglichkeiten anbieten können, ohne ihre Identität zu verlieren; gewährleistet wird dies durch das dezentrale Modell.

IV. Kommunikations- und Wissenschaftskulturen

Nach einhelligem Urteil weisen die unterschiedlichen Fachdisziplinen je spezifische „Fachkulturen“ auf. Geisteswissenschaftliche Forschung im Allgemeinen wie auch die textwissenschaftliche im Besonderen ist traditionell in besonderer Weise an Personen und die individuelle Einzelleistung des Forschers gebunden, während die Natur- und Lebenswissenschaften wesentlich stärker auf Forschungsverbände und Projektteams setzen. Doch auch in den Geisteswissenschaften geraten diese Arbeitsweisen verstärkt in den Blick,¹⁵ insbesondere angesichts einer erneuten Hinwendung zu Empirie, die befördert wurde durch neue Möglichkeiten der Verfügbarmachung und Erschließung großer, fast unbegrenzt scheinender Datenmengen im digitalen Medium. Doch nicht allein die Verfügbarkeit der Ressourcen, auch die Forschungsmöglichkeiten und -fragen und damit die Methoden ändern sich durch neue Technologien.

Dies lässt sich gut anhand von Beispielen nachvollziehen. In den Philologien stellen insbesondere „Massenwörter“, hochfrequente Verben,¹⁶ und u. a. auch Pronomina, eine Herausforderung für den Bearbeiter – sei es eines Wörterbuchs, sei es einer Grammatik – dar. Die übersichtliche Sortierung, Darstellung und intelligente Kookkurrenzanalyse des Bestands bringt nicht allein eine Rationalisierung und Beschleunigung des Arbeitsprozesses, sondern in erster Linie einen entscheidenden Qualitätsfortschritt für die linguistische Analyse, da nunmehr erstmals verlässliche Aussagen über Massenphänomene mit vertretbarem zeitlichen Aufwand möglich sind.¹⁷ Eine andere Herausforderung sind beispielsweise Editionen reich überlieferter Texte – man denke z.B. nur an Bibelhandschriften –, wo zum einen EDV-Technologie bei der Bereitstellung der Quellen durch qualitativ hochwertige Imagescans, beim automatischen Kollationieren der Texte oder bei der Registererstellung zum Ein-

satz kommt,¹⁸ wo zum anderen jedoch auch Fortschritte durch dezentrale, kollaborative Verfahren in großen, internationalen Projektteams erreicht werden können.¹⁹

Zusammenfassend lässt sich sagen, dass dezentrale und kollaborative Arbeitsweisen immer dann sinnvoll anwendbar sind, wo Aspekte der Bereitstellung, Erschließung und Kommentierung von Primärquellen zum Tragen kommen und wo Massendaten bearbeitet werden müssen. Die Fülle und Vielfalt unseres kulturellen Erbes sowie die Anforderungen bei dessen Bewahrung, Erschließung und Erforschung machen demnach eine neue wissenschaftliche Kultur der Zusammenarbeit sowie den Einsatz von Spitzentechnologie nicht allein wünschenswert, sondern dringend erforderlich. Die Geisteswissenschaften brauchen also eine koordinierte e-Humanities-Infrastruktur, an der die Fachwissenschaft zur Formulierung der spezifischen Bedürfnisse maßgeblich beteiligt sein sollte. Die forschende, interpretatorische Einzelleistung bleibt davon unberührt und behält ihren Wert. Doch auch die kollaborative Leistung in großen (internationalen) Projektteams muss stärker als bisher ihren Platz im Bewertungssystem der Geisteswissenschaften erhalten – dies erfordert durchaus eine Änderung der bisherigen Wissenschaftskultur.

V. Chancen und Perspektiven

Festzuhalten ist, dass kollaborative Arbeitsweisen auch in den Geisteswissenschaften in bestimmten Bereichen keineswegs neu, sondern seit langem etablierte Praxis sind, gerade in den angesprochenen Arbeitsfeldern Edition, Kommentierung, Erstellung von Grundlagenwerken (Wörterbücher, Lexika, Grammatiken).²⁰ Von hier aus besteht also die Chance, vorhandene Strukturen durch Technologie zu optimieren und auszubauen und somit auch in andere Bereiche zu übertragen und zu etablieren. Bibliotheken und Archiven kann in diesem Zusammenhang eine besondere Rolle zukommen: einerseits bieten sie mit ihren Beständen den Ausgangspunkt wissenschaftlicher Beschäftigung, andererseits können sie durch Bereitstellung weiterer Forschungsinfrastruktur zu einem zentralen Knotenpunkt der vernetzten Wissenschaftslandschaft werden.

Andrea Rapp

Zur Autorin

Dr. ANDREA RAPP, Akademische Rätin im Fach Germanistik, Ältere deutsche Philologie, sowie Geschäftsführerin des Kompetenzzentrums für elektronische Erschließungs- und Publikationsverfahren in den Geisteswissenschaften an der Universität Trier

Kontakt: andrea.rapp@uni-trier.de

Website des Projekts: <http://www.textgrid.de>

Anmerkungen

- 1 Im vorliegenden Beitrag wird der traditionelle Begriff „Geisteswissenschaften“ im Sinne Diltheys – nicht zuletzt, da wir 2007 im entsprechenden Wissenschaftsjahr sind und durchaus in Abgrenzung zu den traditionellen „Naturwissenschaften“ – beibehalten; die Diskussion um die Definition der „Kulturwissenschaften“ oder die „Humanwissenschaften“ kann hier zunächst ausgeklammert bleiben.
- 2 Dies liegt nicht unbedingt allein an Ressentiments gegenüber Digitalisaten – obwohl solche auch vorhanden sind –, sondern ist auch eine Frage der Kapazitäten.
- 3 Die massenhafte Bereitstellung von Literatur soll nicht herabgewürdigt werden; dennoch muss betont werden, dass im wissenschaftlichen Zusammenhang der Fachwissenschaftler entscheiden sollte, was wie publiziert wird.

- Ein weiteres Problem liegt in der Qualität des Bereitgestellten; bei der Arbeit mit den von Google digitalisierten Büchern stößt man sehr rasch auf unvollständige, beschnittene, unscharfe, durch Hände verdeckte, verzerrte Beispiele. Auch Metadaten und Referenzierbarkeit sind durchaus ein diskussionswürdiges Problem.
- 4 Am 1.7.2007 wird das GDI-Grid – Geodateninfrastruktur-Grid mit einem neuen Community-Projekt starten.
 - 5 Weiterführende Informationen zu D-Grid, zum D-Grid-Integrationsprojekt DGI sowie zu den Community-Projekten sind stets aktuell über die D-Grid-Homepage verfügbar: <http://www.d-grid.de>.
 - 6 Eine sehr anschauliche und allgemeinverständliche Einführung in „Grid-Computing“ gibt das GridCafé: <http://gridcafe.web.cern.ch/gridcafe/>
 - 7 Text Encoding Initiative: <http://www.tei-c.org>
 - 8 Dies stellten auch die Gutachter bei der im Februar 2007 erfolgten Zwischenevaluierung fest. Mit Nachdruck wurden weitere Initiativen für die Geisteswissenschaften gefordert.
 - 9 Ausführliche Informationen zu TextGrid und zu den beteiligten Partnern finden sich unter <http://www.text-grid.de>, auch öffentliche Reports aus den TextGrid-Arbeitsgruppen sowie weitere Hinweise auf Berichte und Veröffentlichungen über TextGrid.
 - 10 Vgl. zu den Anwendungsszenarien, die die Basis für die Toolentwicklung bilden, die ausführliche Darstellung unter http://www.textgrid.de/fileadmin/TextGrid/TextGrid-Szenarien_061212.pdf
 - 11 Weitere spezifische Anwendungsszenarien und einzelne Werkzeuge werden aufgezeigt im Beitrag von Andreas Aschenbrenner, Tobias Blanke, Stuart Dunn, Martina Kerzel, Andrea Rapp und Andrea Zielinski: Von e-Science zu e-Humanities – Digital vernetzte Wissenschaft als neuer Arbeits- und Kreativbereich für Kunst und Kultur. In: Bibliothek. Forschung und Praxis, 31 (2007), S. 11–21; Preprint unter: http://www.bibliothek-saur.de/preprint/2007/ar2422_aschenbrenner_blanke.pdf
 - 12 Die Middleware sorgt bspw. für die Virtualisierung der vorhandenen Speicherressourcen in geographisch verteilten Knoten und erzeugt ein homogenes virtuelles Archiv. Sie vermittelt zwischen den generischen Grid-Diensten der Integrationsplattform und den Anforderungen für die TextGrid-spezifischen Dienste und sorgt für die Vernetzung mit anderen Grid-Projekten. Sie entlastet den Benutzer durch möglichst komfortable Schnittstellen von nicht-spezifischen Basisaufgaben.
 - 13 Joachim Heinrich Campe, Wörterbuch der deutschen Sprache. 1807–1811.
 - 14 Es handelt sich um die Wörterbücher, die das Trierer Kompetenzzentrum über das Wörterbuchnetz zur Verfügung stellt: <http://www.woerterbuchnetz.de>; mit der Implementierung des TextGrid-WebService ist in Kürze zu rechnen.
 - 15 Nicht zuletzt werden diese Arbeitsweisen von den Förderinstitutionen für die Geisteswissenschaften dezidiert eingefordert. Wie Lange feststellt, gibt es jedoch „derzeit kaum empirische Untersuchungen zu der Frage, ob (und wie) sich Verhaltensweisen von Geisteswissenschaftler/innen im Umgang mit dem WWW verändern“; Tanja Lange, Netzgestützte Kommunikation und Kooperation für Forschung (und Lehre) in den Geisteswissenschaften. In: Jahrbuch für Internationale Germanistik XXXVIII,1 (2006), S. 83–94, hier S. 87. Sie fordert dazu auf, „Arbeitsbereiche“ und „Arbeitsformen“ zu identifizieren, in denen kooperative Arbeitsweisen für Geisteswissenschaftler sinnvoll erscheinen (S. 88); auch Probleme und Kritik werden von ihr angesprochen (S. 94).
 - 16 Den umfangreichsten Artikel im Deutschen Wörterbuch der Brüder Grimm stellt das Verb ‚stehen‘, der über 320 Spalten umfasst – nicht wie lange Zeit vermutet wurde, der ‚Geist‘; der Grimm steht online unter <http://www.dwb.uni-trier.de>
 - 17 Zum Verfahren der Kookkurrenzanalyse des Instituts für deutsche Sprache Mannheim siehe <http://www.ids-mannheim.de/kl/projekte/methoden/ka.html>; zum iterativen Lemmatisierungsverfahren bei hochvarianten Texten des Mittelhochdeutschen siehe Ralf Plate, Ute Recker-Hamm, Elektronische Materialgrundlage und computergestützte Ausarbeitung eines historischen Belegwörterbuchs. Erfahrungen und Perspektiven am Beispiel des neuen Mittelhochdeutschen Wörterbuchs. In: Chancen und Perspektiven computergestützter Lexikographie, hg. v. Ingrid Lemberg, Bernhard Schröder und Angelika Storrer. Tübingen 2001 (Lexicographica; Series Maior). S. 155–177; sowie weitere Hinweise und Literaturangaben unter <http://www.mhdwb.uni-trier.de>.
 - 18 Zwei Beispiele seien hier herausgegriffen: Das von Peter Robinson initiierte Canterbury Tales Project hat Maßstäbe für die elektronische Analyse und Edition gesetzt, siehe dazu <http://www.canterburytalesproject.org/>; im Bereich der Altgermanistik erprobt das Parzival-Projekt von Michael Stolz neue Wege: Michael Stolz, Wolframs ‚Parzival‘ als unfester Text. Möglichkeiten einer überlieferungsgeschichtlichen Edition im Spannungsfeld traditioneller Textkritik und elektronischer Darstellung. In: Wolfram von Eschenbach – Bilanzen und Perspektiven.

Eichstätter Colloquium 2000, hg. v. Wolfgang Haubrichs, Eckart C. Lutz und Klaus Ridder, Berlin 2002 (Wolfram-Studien 17), S. 294–321; zahlreiche weitere Veröffentlichungshinweise unter <http://www.parzival.unibas.ch>

- 19 Auch Lange bescheinigt den Editionswissenschaften eine bereits vollzogene „Neuorientierung“ in den Arbeitsweisen (wie Anm. 15), S. 88. Hervorgehoben sei hier noch die Max Planck-Gesellschaft, deren Institute eine strikte OpenAccess-Politik vertreten und die dezentral Quellen und Tools verfügbar machen; vgl. z.B. das Archimedes-Projekt zur Entwicklung der Mechanik unter http://archimedes2.mpiwg-berlin.mpg.de/archimedes_templates oder das europäische ECHO-Projekt <http://echo.mpiwg-berlin.mpg.de/home>
- 20 Auch in den so genannten „kleinen Fächern“, die in der Regel global verstreut sind, ist die internationale Kommunikation und Vernetzung stark ausgeprägt. Dies zeigt sich z.B. gut in der Papyrologie, wo auch die weltweit verstreut liegenden Bestände bereits in großem Umfang digitalisiert sind.

Auszug aus:

**Jahrbuch der historischen Forschung
in der Bundesrepublik Deutschland 2006**

Herausgegeben von der
Arbeitsgemeinschaft historischer Forschungseinrichtungen
in der Bundesrepublik Deutschland e.V.

© 2007 Oldenbourg Wissenschaftsverlag GmbH, München
oldenbourg.de

Das Werk einschließlich aller Abbildungen ist urheberrechtlich geschützt.
Jede Verwertung außerhalb der Grenzen des Urheberrechtsgesetzes ist ohne
Zustimmung des Verlages unzulässig und strafbar. Dies gilt insbesondere für
Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung
und Bearbeitung in elektronischen Systemen.