



.hist 2006

Geschichte im Netz: Praxis, Chancen, Visionen

ISSN: 1612-5940

Historisches Forum

10 • 2007 • Teilband I

Vom 22. bis zum 24. Februar 2006 veranstaltete der Kooperationsverbund Clio-online in Berlin die Tagung „.hist 2006: Geschichte im Netz – Praxis, Chancen, Visionen“. Dies war bereits der zweite große Kongress an der Humboldt-Universität zu Berlin und der Berlin-Brandenburgischen Akademie der Wissenschaften, der sich mit „Geschichte im Netz“ und mit den Möglichkeiten und Grenzen der Nutzung von Kommunikationsnetzwerken durch die Geschichtswissenschaft in theoretischer Reflexion wie praxisbezogener Diskussion beschäftigte. Der Austausch zwischen Geisteswissenschaftlern/innen, die mit netzbasierten Medien arbeiten, Archivaren/innen, Bibliothekaren/innen, Museumsmitarbeitern/innen und Informatikern/innen zielte auf die Überwindung disziplinärer und institutioneller Grenzen, was sich auch in den vorliegenden Bänden widerspiegelt.

Die Konferenz bot in zwölf Sektionen und Workshops, zwei Podiumsdiskussionen und an zahlreichen Projektständen mit insgesamt etwa hundert Vorträgen, Diskussionsrunden und Präsentationen mehr Informationen, als die zahlreichen Besucher/innen aufnehmen konnten. Umso mehr freuen wir uns, im vorliegenden Doppelband schriftliche Ausarbeitungen der meisten Referate, bereits über H-Soz-u-Kult veröffentlichte Querschnittsberichte zur Tagung sowie eine Zusammenfassung der Abschlussdiskussion vorlegen zu können. Die Teilbände 1 und 2 des Historischen Forums Nr. 10 bündeln mit insgesamt über 80 Beiträgen die Mehrzahl der auf der Tagung „.hist 2006“ präsentierten Vorträge und machen sie allen Interessierten zugänglich.

ISBN: 978-3-86004-205-2

Beiträge der Tagung .hist 2006

Herausgegeben für Clio-online von
Daniel Burckhardt, Rüdiger Hohls und Claudia Prinz



http://edoc.hu-berlin.de/e_histfor/10_I/
Veröffentlichungen von Clio-online, Nr. 2



Historisches Forum ist eine Reihe von Themenheften des von der Deutschen Forschungsgemeinschaft geförderten historischen Fachportals Clio-online (<http://www.clio-online.de>) und seiner Kooperationspartner. Die Reihe bündelt ausgesuchte Beiträge geschichtswissenschaftlicher Online-Foren und herausragende Artikel, Debattenbeiträge, Kontroversen und Berichte zu ausgewählten historischen Fragestellungen. Sie erscheint in Kooperation mit den Verbundpartnern von Clio-online und der Humboldt-Universität zu Berlin. Jedes Heft wird von einem oder mehreren Herausgebern redaktionell betreut und enthält außer einer Einführung in das Thema auch ergänzende Verweise auf die Forschungsliteratur und andere Informationsquellen. Die Veröffentlichung erfolgt über den Dokumenten- und Publikationsserver der HUB: http://edoc.hu-berlin.de/e_histfor/.

[Historisches Forum]

Historisches Forum. - Berlin: Clio-online
und Humboldt-Universität zu Berlin

Gesamttitel: Veröffentlichungen von Clio-online, Nr. 2

ISSN: 1612-5940

Erscheinungsweise: ca. 3 Hefte pro Jahr.

Bd. 10, Teilband I: **Geschichte im Netz: Praxis, Chancen, Visionen. Beiträge der Tagung .hist 2006** / hrsg. für Clio-online von Daniel Burckhardt, Rüdiger Hohls und Claudia Prinz / (Historisches Forum: Bd. 10, I) - Berlin: Clio-online und Humboldt-Universität zu Berlin, 2007
ISBN: 978-3-86004-205-2

Dieses Werk einschließlich aller Abbildungen ist urheberrechtlich geschützt. Es unterliegt den Nutzungsbedingungen des Dokumenten- und Publikationsservers der Humboldt-Universität Berlin (<http://edoc.hu-berlin.de>). Es darf und soll zu wissenschaftlichen Zwecken und zum Eigengebrauch kopiert und ausgedruckt werden. Die weiteren Rechte an den einzelnen Texten verbleiben bei den Autoren. Jede kommerzielle Nutzung der Dokumente, auch von Teilen und Auszügen, ist ohne vorherige Zustimmung und Absprache mit den Serverbetreibern und den redaktionell verantwortlichen Herausgebern ausdrücklich verboten.

Redaktionsschluss und letzte Überprüfung der Internet-Adressen:
15.02.2007

Geschäftsführende Herausgeber:

Rüdiger Hohls – Wilfried Nippel
in Verbindung mit Clio-online, H-Soz-u-Kult (Karsten Borgmann – Irmgard Zündorf) und Zeitgeschichte-online (Jürgen Danyel – Jan-Holger Kirsch).

Technische Leitung:

Daniel Burckhardt - Felix Herrmann

Verantwortliche Redakteurin für dieses Heft:

Claudia Prinz

Clio-online

Humboldt-Universität zu Berlin

Philosophische Fakultät I

Institut für Geschichtswissenschaften

Unter den Linden 6

D-10099 Berlin

Telefon: ++49-(0)30/2093-2541

E-Mail: prinzc@geschichte.hu-berlin.de

Umschlaggestaltung:

Kai Pätzke

© 2007 Clio-online

Historisches Forum

Veröffentlichungen von Clio-online, Nr. 2
ISSN: 1612-5940

Historisches Forum 10, I · 2007

Geschichte im Netz:
Praxis, Chancen, Visionen.
Beiträge der Tagung .hist 2006

Teilband I

Herausgegeben für Clio-online
von Daniel Burckhardt, Rüdiger Hohls und Claudia Prinz

unter Mitwirkung von
Sebastian Barteleit
Gudrun Gersmann
Peter Haber
Madeleine Herren
Patrick Sahle
Daniel Schlögl
Georg Vogeler
Claudia Wagner
Irmgard Zündorf



http://edoc.hu-berlin.de/e_histfor/



ISBN: 978-3-86004-205-2

INHALTSVERZEICHNIS

INHALTSVERZEICHNIS	1
Editorial: Geschichte im Netz – Praxis, Chancen, Visionen	7
<i>von Daniel Burckhardt, Rüdiger Hohls, Claudia Prinz</i>	
ERÖFFNUNG – GRÜßWORTE – PODIUMSDISKUSSION	24
Eröffnungsrede „,hist 2006: Geschichte im Netz“	25
<i>von Wilfried Nippel</i>	
Grußwort des Präsidenten des Bundesarchivs.....	29
<i>von Hartmut Weber</i>	
Geschichte im Netz – Praxis, Chancen, Visionen	33
<i>von Wolfgang Schmale</i>	
Grußwort des Präsidenten der Berlin-Brandenburgischen Akademie der Wissenschaften	53
<i>von Günter Stock</i>	
Podiumsdiskussion: „,Vom Nutzen und Nachteil virtueller Informationswelten für die Historie“	56
<i>Zusammenfassung von Daniel Schlögl</i>	
QUERSCHNITTSBERICHTE ÜBER DIE TAGUNG	68
Historische Quellenbestände – Digitalisierung und Suchstrategien	69
<i>von Stefan Gorßen</i>	
Mittelalter im WWW – Eine Bestandsaufnahme	76
<i>von Harald Müller</i>	

Themenportale zwischen Bibliothek und Fachwissenschaft	87
<i>von Stephanie Marra</i>	
Kollaboratives Schreiben, Lehren und Lernen	97
<i>von Daniel Burckhardt</i>	
ELEKTRONISCHES PUBLIZIEREN IN DEN GEISTESWISSENSCHAFTEN	104
Elektronisches Publizieren – Einleitung.....	105
<i>von Gudrun Gersmann</i>	
Sehepunkte und Zeitenblicke – Entwicklungslinien nach einem halben Jahrzehnt des elektronischen Publizierens	111
<i>von Michael Kaiser</i>	
« Zeithistorische Forschungen / Studies in Contemporary History » – Erfahrungen mit hybriden Publikationskonzepten	125
<i>von Jan-Holger Kirsch</i>	
Zur Akzeptanz des elektronischen Publizierens in den Geisteswissenschaften – Ergebnisse und Folgen der DFG- Studie „Publikationsstrategien im Wandel“	141
<i>von Johannes Fournier</i>	
Elektronisches Publizieren – Die Nutzerperspektive	158
<i>von Karsten Borgmann</i>	
Zur rechtlichen Problematik des elektronischen Publizierens.....	172
<i>von Karl-Nikolaus Peifer</i>	
Langzeitarchivierung elektronischer Publikationen im Dienste der Wissenschaft – Eine neue Aufgabe der Bibliotheken	191
<i>von Thomas Wolf-Klostermann</i>	

DIGITALISIERUNG UND STRATEGIEN DER LANGZEITARCHIVIERUNG.....	206
Die Deutsche Forschungsgemeinschaft und deutsche Initiativen zur Langzeitarchivierung für die Wissenschaft	207
<i>von Max Vögler</i>	
Die EUBAM-Gruppe und ihre Mitwirkung in den euro- päischen Kooperationsprojekten MINERVA und MICHAEL.....	216
<i>von Monika Hagedorn-Saupe</i>	
Die Digitalisierungsprojekte der Universitätsbibliotheken in Frankreich – PERSEE und andere Angebote	227
<i>von Frédéric Blin und Valérie Néouze</i>	
kopal – ein digitales Archiv zur dauerhaften Erhaltung unserer kulturellen Überlieferung	244
<i>von Thomas Wollschläger</i>	
Das Projekt nestor – Kooperative Strukturen für die Über- lieferungsbildung – Beispiele aus dem Bestand maschinen- lesbarer Daten des Bundesarchivs.....	258
<i>von Andrea Hänger und Karsten Huth</i>	
Langfristige Verfügbarkeit von Quelleneditionen im digitalen Zeitalter – Text Encoding Initiative, Geschichtswissenschaft und Bibliotheken	275
<i>von Stefan Cramme</i>	
Verteilte Digitale Inkunabelbibliothek – Ein Baustein zur Gesamtdigitalisierung aller Inkunabelausgaben.....	294
<i>von Timo Steyer</i>	
Zeitschriften der Aufklärung im Netz – Retrospektive Digitalisierung wissenschaftlicher Rezensionsorgane und Literaturzeitschriften	308
<i>von Sabine Rahmsdorf</i>	

Die digitalisierten „Jahresberichte für deutsche Geschichte“ aus der Zwischenkriegszeit – Ein Spiegel der Geschichtswissenschaft zwischen 1925 und 1938	322
<i>von Daniel Schlögl</i>	
VERFAHREN DER QUELLENSCHLIEßUNG UND SUCHSTRATEGIEN	330
Ein Tor zu vielen Quellen – Das BAM-Portal.....	331
<i>von Frank von Hagel</i>	
Quellenschließung für die Montangeschichte – Ein Werkstattbericht aus dem montan.dok	350
<i>von Stefan Przigoda</i>	
Vom Nutzen der Strukturen – Archivische Recherche- strategien im Internet.....	368
<i>von Anke Löbnitz und Jessica von Seggern</i>	
Struktur meets Schlagwort – der Clio-online Findmittelkatalog.....	393
<i>von Sebastian Barteleit und Robert Zepf</i>	
Manuscripta mediaevalia im digitalen Verbund.....	411
<i>von Robert Giel</i>	
Die Regesta Imperii Online	422
<i>von Andreas Kuczera</i>	
Die Pressearchive von HWWA und ZBW –Retrodigita- lisierung der Altbestände von 1900 bis 1930	430
<i>von Thomas Sergej Huck und Max-Michael Wannags</i>	
Der virtuelle Lesesaal des Archivs für Zeitgeschichte (ETH Zürich) – Online-Präsentation und Retrieval von Archivgut	446
<i>von Daniel Nerlich</i>	

Quellenarbeit im Projekt „Preußen als Kulturstaat“ – Strukturierte Informationserfassung mit dem „Archiv-Editor“	455
<i>von Alexander Czmiel und Bärbel Holtz</i>	
QUELLEN IM NETZ	466
Quellen im Netz, Vorstellung und Zusammenfassung der Sektion.....	467
<i>von Patrick Sahle und Georg Vogeler</i>	
New Methods for the Analysis of Digitized Medieval Latin Charters	482
<i>by Michael Gervers and Michael Margolin</i>	
Das Monasterium-Projekt – Ein Beispiel für kollaborative und inkrementelle Editionen	501
<i>von Karl Heinz</i>	
Darstellung mehrschichtiger, komplex-strukturierter Quellen – Die computergestützte dynamische Edition	514
<i>von Matthias P. Perstling</i>	
Kognitive Karten des Mittelalters – Digitale Erschließung mittelalterlicher Weltkarten.....	539
<i>von Günther Görz</i>	
Historische Filmeditionen im Internet-Zeitalter.....	573
<i>von Christopher N. Carlson</i>	
Zeithistorische Online-Editionen: „Die Kabinettsprotokolle der Bundesregierung“ und Retrodigitalisierung der Edition „Akten der Reichskanzlei – Weimarer Republik“.....	589
<i>von Jörg Filthaut</i>	

WÖRTERBÜCHER UND LEXIKA.....	606
Das Wörterbuch-Netz: Verfahren – Methoden – Perspektiven	607
<i>von Thomas Burch und Andrea Rapp</i>	
Vom Nachschlagewerk zum virtuellen Informationssystem –	
Perspektiven historischer Lexikografie im Internet.....	628
<i>von Robert Charlier</i>	
Wörterbücher als digitale Ressourcen für Mensch und	
Maschine – Die Wörterbuchprojekte der Berlin-	
Brandenburgischen Akademie der Wissenschaften.....	644
<i>von Gerald Neumann</i>	
Bayerisches Musiker-Lexikon Online	653
<i>von Josef Focht</i>	
PERSONENREGISTER FÜR BAND I UND II.....	662

EDITORIAL: GESCHICHTE IM NETZ – PRAXIS, CHANCEN, VISIONEN

von Daniel Burckhardt, Rüdiger Hohls, Claudia Prinz

Zwischen dem 22. und 24. Februar 2006 veranstaltete der Kooperationsverbund Clio-online in Berlin die Tagung „hist 2006: Geschichte im Netz – Praxis, Chancen, Visionen“. Nach einer Vorgängertagung im Frühjahr 2003, die unter dem Leitthema „Geschichte und Neue Medien“ stand¹, war „hist 2006“ der zweite große Kongress an der Humboldt-Universität zu Berlin und der Berlin-Brandenburgischen Akademie der Wissenschaften, der sich mit „Geschichte im Netz“ und mit den Möglichkeiten und Grenzen der Nutzung von Kommunikationsnetzwerken durch die Geschichtswissenschaft beschäftigte, gleichzeitig aber auch die Realisierbarkeit und Wünschbarkeit kritisch zu hinterfragen und am Praxisalltag zu messen suchte.

Im Zentrum des Kongresses sollte die Doppeldeutigkeit des „Netzes“ stehen: das Netz als technische Infrastruktur, das neue Formen des Arbeitens, Forschens, Lehrens, Kommunizierens und Publizierens ermöglicht, und das Netz als sozialer und wissenschaftlicher Raum, der durch Adaptionen, Erfahrungen, Anwendungen und Optionen der Technik eine fortwährende Umgestaltung erfährt. Begriffe wie Web, Kommunikation, Information oder Wissensgesellschaft haben Eingang in unseren alltäglichen Sprachgebrauch gefunden. Anknüpfungspunkte bestehen daher nicht nur zur technischen Seite historischer Netzangebote, sondern auch zu Beziehungsaspekten, zu Fragen der Wahrnehmung, Kognition und Konstruktion

¹ Die Ergebnisse der ersten Tagung sind in zwei Bänden veröffentlicht: Burckhardt, Daniel; Hohls, Rüdiger; Ziegeldorf, Vera (Hgg.): Geschichte und Neue Medien in Forschung, Archiven, Bibliotheken und Museen. Tagungsband .hist 2003, Teilband I & II, Berlin 2005 (= Historisches Forum, Band 7 • 2005 • Teilband I: <http://edoc.hu-berlin.de/histfor/7_I/>; Teilband II: <http://edoc.hu-berlin.de/histfor/7_II/>).

von Wirklichkeiten und Geschichtsbildern neben solchen, wie sich das Verhältnis von Erinnern und Vergessen, von Geschichte und Gegenwart durch die neuen Speicher- und Verarbeitungsqualitäten des Internets verändert.

Die Tagung „hist 2006: Geschichte im Netz“ wandte sich bezüglich der Referenten/innen und Teilnehmer/innen vor allem an drei Zielgruppen:

- An Geisteswissenschaftler/innen, primär, aber nicht nur, aus den Geschichtswissenschaften, die im Rahmen ihrer Forschungs- und Lehrtätigkeit netzbasierte Medien und Techniken einsetzen und somit über praktische Erfahrungen verfügen.
- An Archivare/innen, Bibliothekare/innen sowie Museumsmitarbeiter/innen, die sich mit digitalen Methoden der Archivierung, Erschließung und Bereitstellung von historischen Quellen und Artefakten, Publikationen und Fachinformationen beschäftigen.
- An Informatiker/innen, die im Bereich der Informationsverarbeitung in den Geisteswissenschaften tätig sind.

Die Zielgruppen bilden mittelbar auch die Konzeption von Clio-online als einem Verbund von Partnereinrichtungen aus Fachwissenschaft und Fachinfrastruktureinrichtungen ab.² Die beabsichtigte Überwindung disziplinärer und institutioneller Grenzen spiegelte sich auf der Tagung und auch im vorliegenden Band wider. Der Kongress sollte ein Forum zum Bilanzieren, Reflektieren und Fortbilden bereitstellen, in dem Berichte aus der Praxis realistische Bezugspunkte zum Alltag und Erfahrungsraum der Zuhörer/innen liefern. Durch die Präsentation neuer Entwicklungen sollten Chancen aufgezeigt werden. Nicht zuletzt sollte den Teilnehmern/innen auch ein Rahmen geboten werden, um sich über Zukunftsperspektiven

2 Zu den Verbundpartnern von Clio-online zählen: Bayerische Staatsbibliothek München; Bibliothek für Bildungsgeschichtliche Forschung des Deutschen Instituts für Internationale Pädagogische Forschung; Bundesarchiv; Berlin-Brandenburgische Akademie der Wissenschaften, Akademievorhaben: Jahresberichte für deutsche Geschichte; Niedersächsische Staats- und Universitätsbibliothek Göttingen; Staatsbibliothek zu Berlin – Preußischer Kulturbesitz; Zentrum für Zeithistorische Forschung Potsdam. Weitere Informationen unter <<http://www.clio-online.de>>.

auszutauschen.³

Die Konferenz bot in zwölf Sektionen und Workshops, zwei Podiumsdiskussionen und über zahlreiche Projektstände mit etwa hundert Vorträgen, Diskussionsrunden und Präsentationen den angereisten Experten/innen, Anwendern/innen und Projektmitarbeitern/innen aus der Geschichtswissenschaft, aus Bibliotheken, Archiven und Museen einen Überblick zu aktuellen Projekten. Letztlich kann die Tagung als Leistungsschau des DFG-Referats „Wissenschaftliche Literaturversorgungs- und Informationssysteme“ (LIS)⁴ betrachtet werden, auf dessen Förderung mehr als 80 Prozent der vorgestellten Projekte beruhen, worauf Manfred Thaller in der abschließenden Podiumsdiskussion hinwies. Zusätzlich war die DFG noch mit Vorträgen von Fachreferenten der LIS-Gruppe [Fournier und Vögler, Bd. I; Goebel, Bd. II] und durch einzelne Mitglieder von Fachausschüssen [Gersmann und Thaller, Bd. I] in die Planung und Durchführung von Sektionen eingebunden. Dieser Umstand weist darauf hin, welche außerordentliche Bedeutung die Förderpraxis der DFG für die Entwicklung der elektronischen Fachinformationsangebote in den Geisteswissenschaften hierzulande hat. Deshalb verwundert es auch nicht, wenn insbesondere ausländische Wissenschaftler/innen gelegentlich auf die Unterschiede in der Förderpraxis in vergleichenden Studien hinweisen. Nationale Besonderheiten, wie sie die DFG als Selbstverwaltungskörperschaft der deutschen Wissenschaft zur Förderung der Forschung und des Forschungsnachwuchses darstellt, hinterlassen ihre Spuren eben auch, wenn es um die Entwicklung globaler Kommunikation oder um die Strukturen von Informationssystemen der Wissensgesellschaft geht. Allerdings fanden entsprechende Entwicklungen im nicht-deutschsprachigen Ausland wie etwa Persée in Frankreich [Blin/Néouze, Bd. I] oder das DEEDS-Projekt in Kanada [Gervers/Margolin, Bd. I] mit Ausnahme der international ausge-

3 Auf den Call for Papers im Frühjahr 2005 gingen weitaus mehr Vorschläge ein als im späteren Programm berücksichtigt werden konnten. Allerdings wurden nicht für alle Aspekte und Themenbereiche, die das Programmkomitee zu berücksichtigen gedachte, Vorschläge eingereicht. So gab es beispielsweise kaum Vorschläge zum Themenkomplex „Ethik - Recht - Ökonomie“; vgl. <<http://hsozkult.geschichte.hu-berlin.de/termine/id=4190>>.

4 Vgl. <<http://www.dfg.de/lis/>>.

richteten Sektion über Geographische Informationssysteme in den Geschichtswissenschaften [Bodenhamer; Kunz; Berman, Bd. I] im Rahmen der Tagung nur punktuell Beachtung.

Das Interesse am Programm der Tagung war groß: Nach über 400 Voranmeldungen konnten eine Woche vor Beginn der Tagung keine weiteren Teilnehmerregistrierungen entgegengenommen werden. Das dichte Programm mit mehreren Parallelsessionen forderte auch von den Anwesenden eine Beschränkung auf einen kleinen Ausschnitt des Gesamtprogramms. Erste Eindrücke von der Tagung lieferten die vier Querschnittsberichte von S. Gorißen, H. Müller, S. Marra und D. Burckhardt, die im Frühjahr 2006 über H-Soz-u-Kult veröffentlicht wurden und in Band I nachzulesen sind. Umso mehr freuen wir uns, im vorliegenden Doppelband immer noch recht zeitnah schriftliche Ausarbeitungen der meisten Referate und sowie eine Zusammenfassung der Abschlussdiskussion vorlegen zu können.

Der vorliegende Doppelband bündelt mit insgesamt über 80 Beiträgen die Mehrzahl der auf der Tagung „hist 2006“ präsentierten Vorträge. Ein Großteil der Referenten/innen ist unserem Wunsch nachgekommen, ihre Vorträge zu Artikeln auszuarbeiten; allerdings sahen sich dazu einige aus verschiedenen Gründen auch nicht in der Lage. So eigneten sich nicht alle Beitragsformate wie Round Tables oder Projektpräsentationen zur schriftlichen Ausarbeitung, teilweise waren Referenten/innen auch aus zeitlichen Gründen verhindert, ihren Tagungsbeitrag zu einem Artikel auszuarbeiten. Wie schon bei der Organisation des Kongresses selbst haben uns einige der Sektionsleiter/innen aktiv bei der Durchsicht und beim Redigieren der Beiträge ihrer Sektionen unterstützt; nur so war es überhaupt möglich, diesen umfangreichen Doppelband fertig zu stellen.

Die in den Bänden enthaltenen Beiträge dokumentieren den Stand der Entwicklung im Februar 2006 und wurden nur vereinzelt aktualisiert oder ergänzt. Die Gliederung des Doppelbandes spiegelt weitgehend das Tagungsprogramm; wobei in Einzelfällen die Beiträge von Sektionen oder Workshops für die vorliegende Publikation in inhaltlich verwandte Kapitel verschoben wurden. Die Beiträge sind unter den Kapitelüberschriften „Elektronisches Publizieren in den Geisteswissenschaften“, „Digitalisie-

rung und Strategien der Langzeitarchivierung“, „Verfahren der Quellenerschließung und Suchstrategien“, „Quellen im Netz“ und „Wörterbücher und Lexika“ im ersten Teilband, „Virtuelle Fachbibliotheken für die historische Forschung“, „Historische Fach- und Themenportale“, „Historical GIS online“, „Collaboratories. Über das gemeinschaftliche Schreiben von Geschichte“ und „E-Learning-Projekte als kollaborativer Entwicklungsprozess“ im zweiten Teilband für die vorliegende Publikation gruppiert worden.

Ausgangslage

Das Internet ist auch aus dem Arbeitsalltag in den Geschichtswissenschaften nicht mehr wegzudenken. Von Beginn an wurde die Nutzung des Internets im Fach von einer Diskussion um die Vor- und Nachteile begleitet. Diese Diskussion wird uns wohl auch noch länger begleiten, da sich in dieser Frage unterschiedliche Diskursstränge vermischen. In der Praxis wird die Debatte selten zwischen den Netzaktivisten/innen unter den Historikern/innen, die häufig ein ungebrochen euphorisches Verhältnis zu ihrer Tätigkeit haben, und Vertretern/innen der traditionellen Geschichtswissenschaft geführt, die dem Treiben der meist jüngeren Kollegen/innen mit Skepsis, Unverständnis und Zurückhaltung begegnen. Die meiste Literatur zum Thema „Geschichte und Internet“ enthält eher praktische Anleitungen für den Umgang mit einzelnen Aspekten des Webs als theoretische Reflexionen über die wissenschaftlichen Implikationen der Veränderung.⁵ Einen theoretischen Reflexionsversuch, intellektuelle Konsequenzen aus der Veränderung aufzuzeigen, hat jedoch Wolfgang Schmale in seinem Eröffnungsvortrag [Bd. I] unternommen.

Die Debattenlage ist auch deshalb verworren, weil sich die Digitalisierung des Arbeitsplatzes des Historikers und die Etablierung des Internets als Ort für historische Recherchen, der Vermittlung und des Publizierens mit einer, wie Winfried Schulze in der Podiumsdiskussion [Bd. I] heraus-

⁵ Exemplarisch: Trinkle, Dennis A.; Merriman, Scott A. (Hgg.): *The history highway: A 21st century guide to Internet resources*, Armonk, N.Y., 4. Aufl. 2006; Epple, Angelika; Haber, Peter; Jucker-Kupper, Patrick (Hgg.): *Vom Nutzen und Nachteil des Internet für die historische Erkenntnis*, Zürich 2005.

gestellte, parallel verlaufenden dramatischen Veränderung in der Geschichtsschreibung selbst vermischen. Der vormals eingeschränkte Zitierhorizont ist im Zuge der Globalisierung, die der Internetnutzung voranging, durch diese aber noch einmal verstärkt wurde, einem weltweiten bibliografischen Erfahrungshorizont gewichen. Die Pluralisierung der Themen und Methoden, die Entkanonisierung und der Dominanzverlust bestimmter Denkmodelle werden mit der wachsenden Bedeutung des Internets für den Lehr- und Forschungsbetrieb in Verbindung gebracht, wobei sich die medienspezifische Informationskompetenz nicht in gleicher Weise entwickelt hat.

Im Rückblick lassen sich laut Peter Haber drei Phasen mit unterschiedlichen Angeboten und Nutzungsverhalten unterscheiden. In den ersten Jahren zwischen 1995 und 1998 waren neben der E-Mail-Kommunikation die wichtigsten Ressourcen Bibliothekskataloge, die nicht selten nur über umständliche Telnet-Verbindungen zu erreichen waren. In einer zweiten Phase wurde das Web an Universitäten, von Forschungsinstituten, Museen, Archiven und engagierten Einzelpersonen als Distributionskanal für eigene Forschungsergebnisse, Quellenbestände, für Linkkataloge sowie als wichtiges Forum für die Fachkommunikation und für erste eigene E-Learning-Angebote entdeckt. Das vergleichsweise unstrukturierte Web der 1990er Jahre hat sich durch die Umstellung vieler Angebote auf datenbankbasierte Content Management Systeme inzwischen in Ansätzen in ein maschinenlesbares Web verwandelt. Zwar fehlt die systematische Ergänzung der für menschliche Leser/innen aufbereiteten Webinhalte um semantische Metainformationen. Die strukturierte Datenhaltung erleichtert aber die automatisierte Auswertung und Klassifikation von Seiteninhalten und ermöglicht Mashups, also Verknüpfungen zwischen verschiedenen Wissensbasen. Ontologisch komplexe Wissensbeziehungen werden unter anderem geknüpft durch Rückgriff auf Klassifikationen, Schlagwortregister, Regelwerke, Lexika und Wörterbücher, deren Informationen mit den gespeicherten Inhalten verbunden werden können [Burch/Rapp; Charlier; Neumann, Bd. I].

Seit etwa drei Jahren tauchen zudem vermehrt Dienste im Web auf, die das Netz nicht nur als kostengünstige Plattform zur Verbreitung vor-

mals gedruckter Angebote nutzen, sondern über Einbezug der Besucher/innen versuchen, die eigenen Angebote zu kommentieren, aktualisieren und zu erweitern. In diesem Web 2.0, im Gegensatz zum Read-only Web der ersten zehn Jahre manchmal auch Read-write Web genannt, werden also nicht mehr nur die Angebote der Gutenberg-Galaxie in das digitale Medium übersetzt, sondern konsequent die neuen Möglichkeiten und spezifischen Vorteile des Mediums WWW ausgelotet [Haber, Bd. II]. Gemeinsam sind den Blogs und Wikis des Web 2.0 niedrige Einstiegshürden, die die Grenze zwischen Nutzern/innen und Anbietern/innen verwischen.

Ein Blick auf das Programm der Tagung sowie das Inhaltsverzeichnis der beiden vorliegenden Bände zeigt, dass der Schwerpunkt unserer Konferenz weniger solche Zukunftsvisionen als die Projekt-Praxis betraf. Dem Vorschlag von Georg Vogeler und Patrick Sahle [Vogeler/Sahle Bd. I] folgend, scheint es beim jetzigen Umsetzungsstand angebracht, nach der Konvergenz der verschiedenen Ansätze und Abläufe zu fragen. Konvergieren die verschiedenen Projekte in inhaltlicher Hinsicht? Konvergieren sie in technischer Hinsicht? Führt die an verschiedenen Orten und innerhalb unterschiedlicher institutioneller Kontexte vorangetriebenen Projekte zu einem pragmatischen Konsens von technischen und inhaltlichen Methoden? Reichen diese gemeinsamen Standards und Paradigmen für den Übergang von einzelnen Projektinseln zu einem interoperablen Gesamtbestand? Oder ketzerisch gefragt: Sind viele Projekte in erster Linie temporäre Baustellen, von denen neben der Einsicht, was nicht oder nur schlecht funktioniert, nur wenig bleiben wird? Diese pessimistische Sicht muss jedoch einer virtuellen Gegentagung vorbehalten bleiben, der Wolfgang Schmale im Eröffnungsvortrag schon mal den Titel „Geschichte im Netz: Praxis, Verluste, Untergang“ gab.

Praxis

In der Praxis haben das Internet und die neuen digitalen Kommunikationsmedien die wissenschaftliche Fachinformationslandschaft mittlerweile nachhaltig verändert und dabei auch die Geistes- und Kulturwissenschaften massiv erfasst. Die Suche nach einschlägiger Literatur, die Kommunikati-

on mit den Kollegen/innen, die Lektüre von Tagungsankündigungen, selbst die Rezeption von Literatur finden zunehmend häufiger in der virtuellen Welt statt. Dennoch repräsentieren das Buch und die gedruckte Zeitschrift bis dato in den Geschichtswissenschaften das Hauptmedium der wissenschaftlichen Kommunikation, zumal traditionelle Rezeptionsformen und die verbreitete Skepsis gegenüber der Qualität und ‚flüchtigen‘ Form des elektronischen Publizierens einem beschleunigten Medienbruch entgegenstehen.

Den Schritt von der Rezeption elektronischer Inhalte zur Produktion elektronischer Publikationen gehen viele Geisteswissenschaftler/innen hingegen noch zögernd, obwohl die Distribution von Forschungsergebnissen über das Internet einen weltweiten Leserkreis und damit höchste Sichtbarkeit für die Autoren/innen verspricht. Dass elektronische Publikationen derzeit nicht in allen Fächern dieselbe Anerkennung finden ist nur zu verstehen, wenn man sich vergegenwärtigt, dass jeder Publikation eine Vielzahl verschiedener Funktionen – Verbreitung von Information, Sicherung von Priorität, Qualitätskontrolle, Kanonbildung und Zuweisung von Reputation – zukommen. Steht bei der Publikationstätigkeit etwa in den Naturwissenschaften und der Medizin die Verbreitung von Information und Sicherung von Priorität im Vordergrund, werden schnelle und weltweit verfügbare elektronische Publikationsformen bevorzugt. In den Geisteswissenschaften dürfte dagegen die Zuweisung von langfristigem Renommee die für Autoren/innen wichtigste Funktion sein. Bei elektronischen Publikationen war im Gegensatz zu Büchern und Artikeln bei gut eingeführten Verlagen und Zeitschriften bislang kaum abschätzbar, welche Anbieter oder Formate diese an jegliche Publikation geknüpfte Funktion der Zuteilung wissenschaftlichen Renommees dauerhaft sichern können. Zunehmend wird aber auch im Internet eine Unterscheidung zwischen weniger ernst zu nehmenden und seriösen Informationsanbietern – etwa Universitäten und akademische Selbstorganisationen – möglich, die der unterschiedlichen Reputation von Verlagen in Zeitschriften im Printbereich entspricht [Thaller, Bd. I].

Andererseits lässt sich bereits heute absehen, dass die vielfältigen Transformationen im Prozess des wissenschaftlichen Arbeitens – die

Kommunikation per E-Mail, der Rückgriff auf digitale Quellen, die Möglichkeit der unmittelbaren Einbindung auch multimedialer Komponenten in Publikationen – auf längere Sicht deutliche Auswirkungen auf die Publikationskultur haben werden: Dass kommunikative Prozesse, die im Vorfeld jeglicher wissenschaftlicher Publikation stattfinden, zunehmend stärker über elektronische Medien abgewickelt werden, führt zu einer zunehmend größeren Vertrautheit mit dem Internet, aus der wiederum eine zunehmende Akzeptanz auch elektronisch publizierter Forschungsergebnisse resultiert. Damit wird sich die elektronische Publikation langfristig auch in den Bereichen der Geschichtswissenschaften, die Netz-Veröffentlichungen heute noch überwiegend kritisch sehen, etablieren.

Optimistisch stimmt, dass sich dafür in den letzten Jahren gewisse Einsichten durchgesetzt sowie „Best-Practice“ Richtlinien herauskristallisiert haben, die stark auf unmittelbaren Erfahrungen im Fach selbst und weniger auf ideologischen Prämissen oder kommerziellen Interessen einzelner Anbieter beruhen:

- Vermeidung von proprietären Datenformaten. Tiff oder JPEG bei Bilddaten, XML als Metagrammatik für strukturierte Texte oder Metadaten, wobei sich erst in wenigen Bereichen breit anerkannte Doctypes wie TEI, METS oder EAD⁶ gegenüber ad-hoc Modellierungen durchgesetzt haben [Cramme, Bd. I].
- Unbestritten ist auch der Kanal, über den diese digitalen Daten präsentiert und ausgetauscht werden. Das Internet ist zum primären Kanal für die Verbreitung von Forschungs- und Metadaten geworden.
- Abgelöst wurden damit die Diskussionen über die Haltbarkeit von CD-ROMs oder über den Wechsel der Diskettenformate durch Konzepte zur Langfristarchivierung von in Standardformaten abgelegten Daten, die sowohl von Bibliotheken, Archiven als auch Rechenzentren vorangetrieben werden [Wollschläger; Hänger/Huth Bd. I]. Dabei ist sicherzustellen, dass die Projekte ihren Nutzen für die Geschichts-

6 Text Encoding Initiative (TEI): <<http://www.tei-c.org/>>; Metadata Encoding and Transmission Standard (METS): <<http://www.loc.gov/standards/mets/>>; Encoded Archival Description (EAD): <<http://www.loc.gov/ead/>>.

wissenschaft und für die kulturelle Selbstvergewisserung der Gesellschaft entfalten können und nicht zu Datenfriedhöfen verkommen [Vögler, Bd. I].

- Konsens scheint auch, dass mit öffentlichen Mitteln erstellte Inhalte für Forschung und Unterricht frei zugänglich sein sollen. Einschränkungen werden nur dann akzeptiert, wenn Persönlichkeits- oder Urheberrechte primär an Bild- und Ton-Daten dies zwingend erfordern [Peifer, Bd. I]. Allerdings fehlen Geschäftsmodelle zur Weiterführung nach Ablauf des Förderzeitraums, die diesem Anspruch auf Offenheit der Projektergebnisse genügen.
- Die Grenzen zwischen Archiven, Bibliotheken, Museen und Universitäten sind fließend geworden [von Hagel, Bd. I]. Dies zeigt sich nicht nur im kooperativen Charakter vieler Projektanträge [Heuberger, Bd. II]. Die digitale Repräsentation der Daten scheint in produktiver Weise nivellierend auf die unterschiedlichen institutionellen Traditionen zu wirken. So finden sich beispielsweise in vielen virtuellen Fachbibliotheken neben genuin bibliothekarischen Kernmodulen auch wissenschaftliche Elemente sowie Basisangebote für die universitäre Lehre [Enderle; Wirtz/Griese; Hein-Kircher, Bd. II]. Bibliotheken, Rechen- oder Medienzentren und Herausgeber aus historischen Instituten agieren gemeinschaftlich in der Rolle von Verlegern [Gersmann, Bd. I].
- Es herrscht weniger ein Mangel als ein Überangebot von geeigneter Software zur Umsetzung der Projektziele [Rettinger/Schrade, Bd. II]. Die konkrete Wahl zwischen den Projekten ist deshalb eher von individuellen Präferenzen als von harten technischen Randbedingungen bestimmt.

Chancen

Die Mehrzahl der auf die Geisteswissenschaften spezialisierten Fachverlage hierzulande verhielt sich gegenüber dem Internet sehr lange abwartend. Potentiale haben sich stärker in den Forschungs- und Infrastruktureinrichtungen entwickelt: Die Materialbezogenheit historischer Forschung, also der Rückgriff auf vorhandene gedruckte Publikationen, handschriftliche wie gedruckte Quellen sowie Materialien in anderen medialen Formen,

sorgt dafür, dass der mediale Wandel in den Geisteswissenschaften in erster Linie von den historische Ressourcen bewahrenden, erschließenden und bereitstellenden Bibliotheken, Archiven und Museen getragen wird. Die Fachwissenschaft entdeckt die mediale Beschaffenheit ihres Gegenstandes erst wieder neu, weshalb Chancen der netzbasierten Publikations- und Vermittlungsmöglichkeiten in Lehre und Forschung auf der Tagung praxisnah reflektiert wurden. Wolfgang Schmale sieht darin Chancen für eine Systemtransformation, für die Transformation der Geschichtswissenschaft als institutionalisiertes System.

Für Manfred Thaller und Winfried Schulze ergibt sich eine große Veränderung für das Fach aus der anstehenden vollständigen Digitalisierung der Zeitschriften- und Buchproduktion [Podiumsdiskussion, Bd. I]. Stehen momentan noch das gedruckte Buch oder die gedruckte Zeitschrift – gelegentlich durch Online-Angebote ergänzt – im Zentrum des Publikationsprozesses, wird künftig das digitale Angebot der Standard sein. Bei längeren Texten werden daraus vermutlich auch in Zukunft in vielen Fällen Bücher generiert. Anders bei den Zeitschriftenartikeln: Da bei diesen das gedruckte Exemplar primär dem Gewinnstreben der Verlage entgegenkommt, rückt hier die digitale Ausgabe ins Zentrum [Kaiser; Kirsch, Bd. I]. Offen ist, wie zukünftig die Nutzungsmodelle für diese elektronischen Fachzeitschriften aussehen werden. Denkbar sind sowohl kommerzielle Angebote, die über Abonnemente, Pay-Per-View-Zahlungen oder Nationallizenzen finanziert werden, als auch frei zugängliche Open Access-Angebote, die primär von den Förderinstitutionen bzw. aus dem Wissenschaftsbereich getragen werden [Fournier, Bd. I].

Durch die inzwischen in großer Zahl realisierten (Retro-)Digitalisierungsprojekte rücken die Vernetzung und Integration sowie der Aufbau übergeordneter Portale und Suchmaschinen zunehmend ins Zentrum. Digitale Wörterbücher und Editionen sollten für die Recherche und Weiterverarbeitung zur Verfügung stehen. Dabei geht es nicht nur darum, passende Dokumente zu bestimmten Suchworten zu ermitteln („Information retrieval“), sondern mit Methoden des „Data minings“ Gemeinsamkeiten und Unterschiede zwischen verschiedenen maschinenlesbaren Beständen zu ermitteln [Burch/Rapp, Bd. I]. Besonders produktiv scheint

die neudeutsch „Mashup“ genannte (Re-)Kombination von Inhalten aus verschiedenen Beständen und die Integration mit externen Zusatzangeboten. So könnte zum Beispiel ein historiografischer Text Hinweise liefern, die eine andere digitale Quelle, zum Beispiel eine Karte, verständlich machen [Vogeler/Sahle, Bd. I]. Dass dies keine reine Zukunftsmusik ist, zeigen die „Projekte des Monats“ der BBAW⁷, in denen unter anderem ein georeferenzierter Zugang zu einer kunsthistorischen Datenbank per Google Maps (Census), die Integration von mehr als einem Dutzend an verschiedenen Standorten aufgebauten Wörterbüchern in ein gemeinsames Wörterbuchportal sowie die wechselseitige Verknüpfung von Datenbank und Scan-Galerie (Ritter-Katalog der Leibnizforschung) realisiert wurden [Neumann; Czmiel/Holtz, Bd. I].

Einigkeit bestand unter den Tagungsteilnehmern/innen darüber, dass die in großem Umfang laufende Digitalisierung historisch relevanter Buchbestände für die Geschichtswissenschaften große Chancen eröffnen wird. In nicht zu ferner Zukunft wird die essentielle historische Überlieferung komplett digital zur Verfügung stehen. Während für die Digitalisierung der Buchbestände auf vorhandene Metadaten aus Bibliothekskatalogen zurückgegriffen werden kann, ist zu erwarten, dass die parallel massenhaft in Web in digitaler Form verfügbaren unedierte Quellen die Historiker/innen vor neue Herausforderung stellen werden. Historische Themenportale, die ausgewählte und teilweise kommentierte Quellenkorpora zu einzelnen Regionen, Themen oder Epochen bereitstellen, zeigen Lösungsansätze auf [Hein-Kircher, Beyer-Thoma, Kellner, Bd. II].

Visionen

Kulturgeschichtlich werden die Phänomene Internet und Web häufig als Medienrevolution bezeichnet, deren Geschichtsmächtigkeit vergleichbar sei mit dem althistorischen Übergang von der Oralität zur Schriftlichkeit sowie mit dem Übergang zum Buchdruck zu Beginn der Neuzeit. Medienrevolutionen stehen in einem interdependenten Verhältnis zur parallelen Veränderung der sie ermöglichenden Kulturen und Gesellschaften. Das

7 Vgl. <<http://www.bbaw.de/pom/rueckblick.html>>.

Charakteristische der gegenwärtigen Medienrevolutionen scheint nicht die zentrale Bedeutung von Wissen und Information zu sein, sondern die Anwendung dieses Wissens und dieser Information zur Erzeugung neuen Wissens und zur Entwicklung von Geräten bzw. Systemen zur Informationserschließung und -verarbeitung sowie zur Kommunikation. In der öffentlichen Wahrnehmung und Diskussion verschwimmen nicht selten die Visionen der Planer mit den wahrscheinlicheren Szenarien und realen Gegebenheiten. So scheinen die riesigen Digitalisierungsprojekte im Ausland, nicht zuletzt Google Books in den Vereinigten Staaten und Gallica in Frankreich, die leistungsfähigen Metasuchmaschinen der Bibliotheksverbände und die wachsenden historischen Quellen- und Fachdatenbanken eine „Informationswelt ohne Grenzen“ zu ermöglichen.

Häufig werden die weit und hoch abgesteckten Erwartungshorizonte nicht eingelöst, weshalb Vision und Skepsis in den Diskussionen häufig eng beieinander liegen. Trotz langjähriger Digitalisierungsvorhaben sind immer noch vergleichsweise wenige Inhalte frei im Netz zugänglich, häufig werden nur Meta-Informationen über in Druckprodukten verfügbare Inhalte angeboten. Hoffnung verspricht die Stärkung der Open Access-Bewegung, zugleich jedoch krankt der Ausbau der wenigen vorhandenen geschichtswissenschaftlichen Portale an unsicheren Finanzierungsmodellen. Erst die Sachzwänge, die im Rahmen der Umstellung auf Bachelor- und Master-Studiengänge eine stärkere Normierung der Curricula erzwangen, um die als Ziel vorgegebenen Erfolgsquoten zu erreichen, sorgen für den auch außerhalb der Konferenz beobachteten Boom der Online-Lehre. Angebote, die Lehrstoff vermitteln und praktische Fertigkeiten trainieren sollen, müssen anders aufgebaut sein als solche, die Überblickswissen vermitteln oder Diskussions- oder Kommunikationsfähigkeit stärken wollen. Möglicherweise auch deshalb ist an den Beiträgen der Sektion „E-Learning-Projekte als kollaborativer Entwicklungsprozess“ das hohe Maß an Selbstreflexion über Sinn und Form der vorgestellten Angebote auffällig [Herren; Pfanzer Sausgruber, Bd. II].

Wenn zum Schluss doch noch Visionen zum Thema „Geschichte im Netz“ formuliert werden sollen, dann scheint es vorteilhaft, mit Nutzerorientierung sowie Teamarbeit zwei Begriffe in den Vordergrund zu stellen,

die stärker mit einem Mentalitätswandel der Akteure als mit neu zu entwickelnden Technologien verknüpft sind.

Von wenigen Ausnahmen abgesehen [Borgmann, Bd. I; Gugerli, Beyer-Thoma, Bd. II] sind die Berichte im vorliegenden Band primär aus der Anbieter-Perspektive geschrieben. Wie Ralf Goebel [Bd. II] in seinem Beitrag ausführt, bedeutet Kundenorientierung in der „Google-Welt“ aber, dass die Benutzer/innen mit dem Service zufrieden sind, nicht die Anbieter. Da die Nutzer/innen im Netz nicht mehr räumlich an eine einzelne Universität gebunden sind, wird immer weniger die Größe des Bestandes, die Exzellenz der Sammlung oder die Gesamtzahl der Benutzer/innen über die Bedeutung eines Anbieters entscheiden als vielmehr die Qualität der angebotenen Dienstleistungen. Bei der Konzeption muss auch bedacht werden, wem man die Informationen bereitstellen will. Sind Forscher, Studierende oder interessierte Laien die Zielgruppe [Müller, Bd. I]? Da viele Besucher/innen aber über Links von anderen Seiten sowie Trefferlisten von Suchmaschinen auf ein Angebot stoßen, kann die tatsächliche Besuchergruppe deutlich von der ursprünglich avisierten Zielgruppe abweichen. Ein nutzerorientiertes Angebot sollte auf solche Veränderungen reagieren können.

Kaum ein anderer Begriff wurde im letzten Jahr von deutschen Internet-Nutzerinnen und -Nutzern bei Google so oft nachgefragt wie „Wikipedia“.⁸ Eine breite Öffentlichkeit hat inzwischen die freie Online-Enzyklopädie als nützliche Anlaufstelle bei Recherchen im Netz entdeckt. Entsprechend präsent war die Wikipedia auch auf der Tagung .hist 2006. In verschiedenen Sektionen und Einzelgesprächen wurde wiederholt die Frage nach Verlässlichkeit und Zitierbarkeit der im ständigen Wandel befindlichen Wikipedia-Artikel gestellt. Das Spektrum der Inhalte sowie die fachliche Qualität der einzelnen Beiträge mögen umstritten sein. Die Art und Weise, wie die Texte durch über 50.000 Autoren/innen erstellt, diskutiert und überarbeitet werden, ist aber unbestritten ein faszinierender Prozess ohne direkten Vorläufer [Voß, Bd. II]. Offen bleibt, ob solche neue Formen kollektiver Autorschaft eine Entwicklung, die in den Naturwissen-

8 Vgl. <<http://www.google.com/press/zeitgeist/archive2006.html>>.

schaften seit Jahren üblich ist, die Mitautorschaft eines ganzen Teams unter der Federführung des Gruppenleiters/der Gruppenleiterin, auch in den Geisteswissenschaften vorantreibt. Die minutiöse Buchführung von Wiki-Systemen über sämtliche Text-Eingaben und -Änderungen könnte ein wichtiger Faktor für die Akzeptanz von Kollektiv-Autorschaften werden [Hodel, Bd. II]. Vermutlich wird bei Monografien und Aufsätzen aber ein namentlich ausgewiesener Autor der Regelfall bleiben. Quellsammlungen, Editionen, Bibliografien sowie neue Formen der historischen Fachinformation und des E-Learnings werden in den nächsten Jahren aber vermutlich mehr und mehr kollaborativ und netzgestützt erstellt [Haber, Bd. II]. Das in derselben Sektion vorgestellte WIKINGER-Projekt der Kommission für Zeitgeschichte zeigte, dass solche Veränderungen keine reine Visionen bleiben, sondern bereits vor der Tagung Teil der Wissenschaftspraxis geworden sind [Hummel/Burtscheidt, Bd. II].

Danksagungen

Zum Schluss möchten wir uns bei allen Institutionen und Personen bedanken, die durch ihre großzügige Unterstützung die Tagung und den vorliegenden Band erst ermöglicht haben. Nur einige können hier genannt werden: Hauptmittelgeber war die Deutsche Forschungsgemeinschaft im Rahmen der Projektförderung von Clio-online. Für zusätzliche finanzielle Mittel danken wir der Forschungsabteilung und dem Institut für Geschichtswissenschaften der Humboldt-Universität zu Berlin (HU). Die HU und die Berlin-Brandenburgische Akademie der Wissenschaften stellten kostenlos ihre Räumlichkeiten und technische Infrastruktur für die Tagung zur Verfügung; zusätzlich durften wir auf die Infrastruktur der Staatsbibliothek zu Berlin – Preußischer Kulturbesitz zurückgreifen. Bei der Öffentlichkeitsarbeit wurden wir vom Presse- und Öffentlichkeitsbüro der HU tatkräftig unterstützt, der Computer- und Medienservice der HU stellte das Tagungssystem zur Verfügung und betreut die Umsetzung der vorliegenden E-Publikationen auf dem Dokumentenserver der Universität. Den Mitarbeiterinnen und Mitarbeitern in den genannten Einrichtungen danken wir sehr für ihre engagierte Unterstützung. Hauptverantwortlich für Planung, Durchführung und Abrechnung der Tagung „hist 2006“ war Claudia

Wagner. Ihr gilt unser ganz besonderer Dank.

Organisation, Infrastruktur und Finanzen sind notwendige Voraussetzungen für das Zustandekommen jeder größeren Tagung. Entscheidend für das Gelingen sind letztendlich aber stets die präsentierten Ideen und Inhalte. Allen Sektionsleitern, Referentinnen und Autoren danken wir deshalb an dieser Stelle noch einmal ganz herzlich für ihre Beiträge und Berichte auf der Tagung und ihre Unterstützung und Geduld beim Zustandekommen dieses Bandes.

Berlin, im März 2007

von Wilfried Nippel

Meine Damen und Herren,
ich darf Sie im Namen von Clio-online herzlich zu unserer Tagung begrüßen. Zugleich überbringe ich Ihnen die Grüße des Präsidiums der Humboldt-Universität. Der Vizepräsident für Forschung wäre gerne gekommen, muss heute Abend aber einer unabweislichen anderen Verpflichtung nachkommen. Die von ihm übermittelten guten Wünsche sind mehr als diplomatische Höflichkeit, sondern entsprechen der Praxis des Hauses; unsere Universitätsleitung hat in den vergangenen Jahren die Arbeit von Clio-online stets unterstützt und, was natürlich besonders wichtig ist, Mittel zur Fortführung des Projektes nach Ablauf der Finanzierung durch die Deutsche Forschungsgemeinschaft (DFG) in Aussicht gestellt. Auch deshalb sind wir zuversichtlich, dass die von einem liebevollen Spötter schon einmal prophezeite notwendige Umbenennung in „Clio-offline“ nicht ansteht.

Das Verbundprojekt Clio-online startete im Mai 2002 mit seiner Arbeit und hat in den zurückliegenden Jahren erfolgreich verschiedene Teilprojekte und Webportale für die historischen Wissenschaften in Angriff genommen. Es wird, wie erwähnt, dankenswerterweise für die Anschubphase von der DFG finanziert. Es handelt sich um einen Verbund, an dem neben der Humboldt-Universität die Berlin-Brandenburgische Akademie der Wissenschaften (mit dem Vorhaben Jahresberichte für Deutsche Geschichte), die Bibliothek für Bildungsgeschichtliche Forschung, das Bundesarchiv, die Staatsbibliotheken in Berlin und Göttingen und das Zentrum für Zeithistorische Forschung in Potsdam beteiligt sind. Die in den letzten Jahren praktizierte gute Zusammenarbeit hat sich auch bei den aufwändigen Vorbereitungen für diese Tagung wieder bewährt, bei denen alle

Partner ihre finanziellen und organisatorischen Beiträge geleistet haben, wofür ich herzlich allen Beteiligten danken möchte. Wir freuen uns auch sehr, dass die abschließende Podiumsdiskussion im Leibnizsaal der Akademie der Wissenschaften stattfinden kann, der nicht nur wunderschön, sondern – normalerweise – auch teuer ist. Unser Dank gilt auch dem Akademie-Verlag und dem Verlag Vandenhoeck & Ruprecht, die unsere Tagung finanziell unterstützen. Das ist auch ein symbolisches Zeichen dafür, dass es um Kooperation und nicht um Konfrontation zwischen aus öffentlichen Mitteln finanzierten Internetangeboten und privatwirtschaftlich arbeitenden Verlagen geht.

Unsere jetzige Tagung ist die zweite große Tagung, die Clio-online veranstaltet. Die erste Tagung 2003 war ein großer Erfolg – wegen des Ertrags der Vorträge und Diskussionen, aber auch des großen Medienechos wegen. Die Vielzahl der für diese Tagung angemeldeten Teilnehmer/innen, die Expertise der zahlreichen in- und ausländischen Referenten/innen aus unterschiedlichen institutionellen Kontexten und die thematische Breite des Programms lassen uns auf eine Wiederholung dieses Erfolges hoffen.

„Geschichte im Netz“ soll ein Forum für den Erfahrungsaustausch von Fachleuten aus der Geschichtswissenschaft, der Informatik, den Bibliotheken, Museen, Archiven und Verlagen darstellen, die auf je unterschiedliche Weise an der Erschließung und Präsentation von Materialien und an der Diskussion von Forschungsfragen in diesem Medium beteiligt sind. Die grundsätzliche kulturkritische Frage, ob sich ein solch „flüchtiges“ Medium überhaupt für ernsthafte Wissenschaft eignet und/oder die Buchkultur zerstört, hat sich erledigt. Umso mehr gilt es die Fragen der Sicherheit und Stetigkeit der bereit gestellten Informationen, ihrer optimalen Zugänglichkeit für unterschiedliche Nutzergruppen und die Möglichkeiten ihres Einsatzes auch im akademischen Unterricht zu diskutieren.

Nach mir wird jetzt der Präsident des Bundesarchivs, Hartmut Weber, zu Ihnen sprechen. Anschließend folgt der Vortrag von Herrn Schmale. Beide Redner haben aktiv auch an der vorangehenden Tagung 2003 mitgewirkt. Herr Weber referierte damals unter anderem über das Thema „BundOnline 2005 und was Archivbenutzer davon haben“ und stellte die ambitionierten Projekte des Bundesarchivs vor, wie zum einen die Über-

tragung der Findmittel und ausgewählter Quellen in elektronische Präsentationen erfolgen soll und welche Perspektiven sich hinsichtlich der langfristigen Verfügbarkeit elektronischer Dokumente im Bundesarchiv ergeben. Herr Weber hat damals berichtet, dass „unspektakuläre“ Archivalien des Bundesarchivs von Archivnutzer/innen nur circa alle 40 bis 45 Jahre eingesehen würden. Andererseits veranschlagte er für die Digitalisierung der häufig oder regelmäßig nachgefragten Archivalien, die circa ein bis zwei Prozent des Bestandes des Bundesarchivs ausmachten, ein etwas 30 Jahre umfassendes Projekt, wenn es bei den vorhandenen Techniken und Förderinstrumenten bliebe. Wir schätzen uns glücklich, dass noch im Jahr 2003 mit dem Bundesarchiv und der Berliner Staatsbibliothek im Rahmen von Clio-online ein wesentlich kleineres und überschaubareres Projekt in Angriff genommen wurde, das den Aufbau eines elektronischen Findmittelkatalogs zum Ziel hat.¹

Wir freuen uns, dass wir Wolfgang Schmale, Professor für Neuere und Neueste Geschichte an der Universität Wien, für den nachfolgenden Eröffnungsvortrag gewinnen konnten. Herr Schmale ist mit zahlreichen Arbeiten zur Verfassungs-, Sozial-, und Kulturgeschichte der Neuzeit, mit Schwerpunkten auf Frankreich wie zum interkulturellen Austausch (so unter anderem beim Thema Menschenrechte) hervorgetreten; er hat zudem Untersuchungen zu einer historischen Anthropologie, darunter zur sozialen Konstruktion von Körperlichkeit und Männlichkeit, vorgelegt.

2003 hat Herr Schmale hier in Berlin eine Sektion mit dem Titel „Historisches Lehren und Lernen über das Netz (E-Learning)“ geleitet.² Ein Jahr später wurde das von Herrn Schmale geleitete Web-Projekt *pastperfect* mit dem Mediaprix 2004 ausgezeichnet, was für die Qualität dieses lehrorientierten Angebots spricht. Beteiligt ist Herr Schmale zusammen mit einer Reihe weiterer Wiener Kollegen/innen an *Geschichte Online*,

1 Vgl. dazu Barteleit, Sebastian; Zepf, Robert, Struktur meets Schlagwort – der Clio-online Findmittelkatalog, in diesem Band.

2 Vgl. dazu die Beiträge des Kapitels „Lehre und Wissensvermittlung“, in: Burckhardt, Daniel; Hohls, Rüdiger; Ziegeldorf, Vera (Hgg.), Geschichte und Neue Medien in Forschung, Archiven, Bibliotheken und Museen. Tagungsband .hist 2003, Teilband II (Historisches Forum 7/2005), S. 407ff., vgl. <http://edoc.hu-berlin.de/histfor/7_II/PDF/HistFor_7-2005-II.pdf>.

einem Forschungsprojekt zu E-Learning im Fach Geschichte. Clio-online hat sich bisher auf die forschungsbezogene Informationsvermittlung konzentriert, aber vielleicht lassen sich ja Schnittflächen und Synergien zwischen den Wiener und Berliner Projekten ausmachen. Herr Schmale steht somit für jene Verbindung von fachwissenschaftlicher Exzellenz und kreativer Nutzung neuer technischer Möglichkeiten, die unser aller Ziel darstellt.

Prof. Dr. Wilfried Nippel ist Professor für Alte Geschichte an der Humboldt-Universität zu Berlin. Er ist Mitglied der Berlin-Brandenburgischen Akademie der Wissenschaften.

GRÜßWORT DES PRÄSIDENTEN DES BUNDESARCHIVS

von Hartmut Weber

Wie alle Wissenschaften lebt die Geschichtswissenschaft von der Kommunikation – der Kommunikation innerhalb der eigenen Disziplin und über die Grenzen des Fachs hinweg. Die Geschichtswissenschaft ist in besonderem Maße darauf angewiesen, mit den Geschichtsquellen zu kommunizieren, die sich nicht zuletzt in den Archiven finden. Unabhängig von Zeit und Ort und der Erinnerungsfähigkeit von Zeitzeuginnen und Zeitzeugen ermöglicht das Archivgut die Auseinandersetzung mit dem Geschehenen. Nicht nur das „Vetorecht der Quellen“, um den kürzlich verstorbenen Historiker Reinhard Kosellek zu zitieren, lässt die Historikerinnen und Historiker und im Falle des Bundesarchivs insbesondere die Zeithistorikerinnen und Zeithistoriker die Lesesäle der Archive aufsuchen, um einen Zugang zu „Zeugen“ und Relikten der näheren oder fernerer Vergangenheit zu erhalten. Erst diese Formen der Kommunikation ermöglichen Wissenschaft im Sinne eines Austausches von Gedanken und Thesen und die belastbare Auseinandersetzung mit den Thesen anderer.

Kommunikation entsteht in unserer aufgabenteiligen Gesellschaft zwangsläufig aber auch zwischen denen, die Archivgut sichern, erschließen und zugänglich machen und denen, die es nutzen, zwischen Archivarinnen und Archivaren und Historikerinnen und Historikern also. Dies ist auch gut so, weil sich im Idealfall diese Kommunikation für beide Seiten befruchtend auswirkt. Und seit sich die Archive auch bewusst als öffentliche Dienstleistungseinrichtungen verstehen, ist diese Kommunikation ein Gebot der Kundenorientierung. Als Kooperationspartner der Humboldt-Universität im Kooperationsverbund Clio-online freue ich mich sehr, dass diese Tagung mit einer Fülle hochinteressanter Beiträge zustande gekommen ist, die aus der Sicht des Bundesarchivs nicht nur die Funktion einer

Leistungsschau hat, sondern zugleich Markt- und Bedarfsforschung im Sinne der angesprochenen Kundenorientierung ermöglicht. So möchte ich Sie, meine Damen und Herren, auch im Namen des Bundesarchivs ganz herzlich begrüßen.

Mit dem Einzug der Informationstechnik und vor allem mit der Vernetzung durch das Internet haben sich diese Kommunikationswege entscheidend verändert. Archive, Bibliotheken und Museen haben sich daran gemacht, im Internet die Erschließungsinformationen zu ihren Beständen mit dem Mehrwert an Navigations-, Recherche- und Verknüpfungsmöglichkeiten bereit zu stellen, die das Medium bietet. Digitales Kulturgut im Netz ist keine Vision mehr. Von nahezu jedem Ort der Erde kann man so gleichzeitig auf Informationen zugreifen, die bislang nur ortsgebunden in den jeweiligen Institutionen und zu bestimmten Öffnungszeiten einzusehen waren. Kommunikation untereinander und Konsultation der im Internet bereitgestellten Informationen kann nun weitgehend grenzenlos und zu jeder Zeit stattfinden. Diese grundsätzlich grenzenlose Verfügbarkeit von kulturellen Inhalten bedeutet aber auch für Archive, Bibliotheken und Museen eine neue Ausrichtung ihrer Arbeit.

So implizieren die großen Digitalisierungsstrategien der letzten Jahre einen zunehmend größeren geografischen Bezugsrahmen – sei es das amerikanische „Google Print“, das französische „Gallica“ oder die jüngste Initiative i2010 der Europäischen Kommission zur Schaffung einer Europäischen Digitalen Bibliothek. Nicht primär der Ort oder die Region, sondern zumindest die Nation oder gar transnationale oder internationale Zusammenhänge werden hier in den Blick genommen. Damit begeben sich die Archive wie die Bibliotheken und Museen in einen neuen internationalen Arbeitskontext, der Chancen bietet, aber natürlich auch Ressourcen bindet. Die heterogene Archivlandschaft macht sich auf den Weg zum Archivsystem. Damit eröffnen sich auch der Geschichtswissenschaft neue Möglichkeiten, die in Ansätzen wie der transnationalen Geschichtsschreibung verfolgt werden. Das viel strapazierte Schlagwort von der Globalisierung trifft hier sicherlich zu. Allerdings verweist der freie Blick über Grenzen hier auf die Chancen, die sich auch für die Geschichtswissenschaft bieten: Die Gegenwart sieht vielleicht nicht das „Ende der Geschich-

te“, vielleicht auch nicht den „Kampf der Kulturen“, aber sie sieht eine Geschichtswissenschaft, die sich verstärkt einer neuen Weltgeschichte zuwenden wird. Eine Geschichtsschreibung, die den Vergleich zwischen Phänomenen, zwischen Regionen und Staaten, zwischen Kulturen und Religionen stärker in den Blick nimmt.

Meine Damen und Herren, auch wenn die Vision, von seinem heimischen Arbeitsplatz Zugriff auf *alle* Informationen und Quellen zu erhalten, in den nächsten Jahren nicht Realität wird und es auch kritisch hinterfragt werden sollte, wie sinnvoll ein solch totaler Zugriff tatsächlich wäre, so können Historikerinnen und Historiker doch bereits in weitaus größerem Maße als noch vor wenigen Jahren wichtige Arbeitsschritte an ihrem heimischen Schreibtisch erledigen. Die Vorbereitung von Forschungsaufenthalten wird durch die Bereitstellung von Erschließungsinformationen und Digitalisaten entscheidend verbessert, die Kommunikation mit Fachkollegen wird durch E-Mail, Mailinglisten und Internetpublikationen ebenfalls wesentlich erleichtert.

Nun ist aber das Erreichte zwar gut, aber nicht so als könnte es nicht noch verbessert werden: Deshalb sind wir hier! In den nächsten zwei Tagen werden wir über den erreichten Stand, die Chancen und möglicherweise auch Grenzen, in jedem Fall aber die weiteren Perspektiven der Entwicklung diskutieren. Das Bundesarchiv bringt sich in diesen Kongress mit Beiträgen aus seinen unterschiedlichen Arbeitsfeldern ein: In Regierung und Verwaltung wird in unserem hybriden Zeitalter neben den herkömmlichen Akten mit elektronischen Aufzeichnungen gearbeitet. Nicht allein das Bundesarchiv ist vor die Aufgabe gestellt, auch diese digitalen Aufzeichnungen von bleibendem Wert als Archivgut auf Dauer zu sichern und zugänglich zu machen. Beim letzten Kongress habe ich eher nur von Vorhaben berichten können, was die Bereitstellung von Erschließungsinformationen im Internet angeht. Inzwischen bieten wir unter anderem die komplette Beständeübersicht und über 500 Online-Findbücher im Internet an und man kann nun vor diesem Hintergrund über Recherchestrategien diskutieren. In einem mehrjährigen Programm zur systematischen Retrokonversion vorhandener herkömmlicher Findmittel ist vorgesehen, in den nächsten Jahren eine Million weiterer Titelaufnahmen im Internet bereit zu

stellen. Und schließlich stellen wir hier die kontinuierlich wachsende Online-Ausgabe der Kabinettsprotokolle der Bundesregierung vor. Demnächst werden die 23 Bände der Akten der Reichskanzlei der Weimarer Zeit mit entsprechendem Mehrwert für die Nutzung der Forschung im Internet zur Verfügung stehen. Auf dem Erreichten werden wir uns aber nicht ausruhen. Denn eins scheint sicher: Die Entwicklung geht weiter und zumindest für das Bundesarchiv möchte ich sagen, wir werden die Entwicklung nicht nur weiter begleiten, das Bundesarchiv wird seinen Beitrag dazu leisten, sie zu gestalten.

Prof. Dr. Hartmut Weber, geboren 1945, ist seit 1999 Präsident des Bundesarchivs. Studium der Geschichte und Germanistik, Ausbildung für den höheren Archivdienst, unterschiedliche Aufgaben beim Hauptstaatsarchiv Stuttgart, beim Generallandesarchiv Karlsruhe und bei der Landesarchivdirektion Baden-Württemberg. Mitglied nationaler und internationaler Fachgremien. Lehrauftrag an der Staatlichen Akademie der Bildenden Künste Stuttgart.

GESCHICHTE IM NETZ – PRAXIS, CHANCEN, VISIONEN

von Wolfgang Schmale

Einleitung

Über den Titel unseres Kongresses kann man lange nachdenken: Die drei Stichwörter *Praxis*, *Chancen*, *Visionen* sind ausgesprochen positiv: pragmatisch positiv (*Praxis*), hoffnungsvoll positiv (*Chancen*), euphorisch positiv (*Visionen*). Das pragmatische Wort *Praxis* ließe sich genauso gut auch als Plattform für eine andere Perspektivenfolge verwenden: Dächte man statt an *Chancen* an *Verluste* und statt an *Visionen* an *Untergang* (*der Wissenschaft*), so lautete der Titel der virtuellen Gegentagung: *Geschichte im Netz: Praxis, Verluste, Untergang*. Beides – den nonchalanten Optimismus des Tagungstitels wie den raunzenden Pessimismus des virtuellen Titels (einer Gegentagung) – müssen wir im Auge behalten.

Was aber soll *Geschichte im Netz* heißen, was kann es heißen? Das Tagungsprogramm legt die Interpretation nahe, dass es sich um bestimmte Praxisformen von *Geschichtswissenschaft* und deren Potenziale handeln soll. *Geschichte* lässt sich freilich genauso gut als *Darstellung oder Erzählung von Geschichte im Netz*, als *Auseinandersetzung mit Geschichte im Netz* verstehen. Es könnte aber zum Beispiel auch bedeuten, dass sich *Geschichte im Netz* – sprich WWW oder kurz Web – vollzieht: Ein interessanter Gedanke, doch worauf könnte er sich beziehen: auf die *Praxis*? Fällt er in das Ressort von *Chancen* und *Visionen*? Und was wäre das für eine *Geschichte*, die sich im Netz vollzieht?

Ich selber stehe auf der Seite der Befürworter der Nutzung des Web für die vielfältigsten geschichtswissenschaftlichen Zwecke; außerdem bin ich der Überzeugung, dass wir erst am Anfang stehen, das heißt, dass vor allem die *Chancen*, und noch mehr die *Visionen* unsere Aufmerksamkeit beanspruchen sollten. Zunächst, der *Praxis* mich zuwendend, stelle ich

mich auf den Standpunkt der Interpretation von *Geschichte im Netz* als „Nutzung des Web für die vielfältigsten geschichtswissenschaftlichen Zwecke“.¹

Geschichte im Netz: Praxis

Nüchtern betrachtet ist die Nutzung des Netzes für geschichtswissenschaftliche Zwecke eine Minderheitengeschichte. Die meisten Seiten, die irgendwie mit Geschichte zu tun haben, entstammen nicht den Geschichtswissenschaften. Sie wurden von interessierten Laien verfasst, zum Beispiel von solchen, die als Freizeitbeschäftigung mittelalterliche Kostüme, oder das, was sie für mittelalterlich halten, herstellen und irgendwelche Alltagssituationen nachstellen, oder was sie für mittelalterlichen Alltag halten, und das Ganze als Homepage ins Netz stellen. Den Professionisten/innen fällt nun mal – man mag das bedauern – als erstes die in der Regel unfreiwillige Komik auf, die solchen Seiten oftmals eigentümlich ist. Zweifellos sind solche Seiten für die Wissenschaft von Interesse, zum Beispiel für die Erforschung von Geschichtskulturen. Keineswegs grundsätzlich sind Seiten nichtwissenschaftlichen Ursprungs immer zu belächeln; es gibt Seiten schulischen Ursprungs, die ganz hervorragend sind; zuverlässige heimatkundliche Seiten, mal sehr gute, mal grausig schlechte kommunale Seiten zur Ortsgeschichte. Während es in den meisten Fällen in Ordnung geht, die Seiten so schnell wieder zu vergessen, wie man sie mit der Suchmaschine gefunden hat, gibt es Fälle, wo sich eine genauere Beschäftigung lohnt oder sogar zwingend erforderlich ist. In der Wissenschaft berüchtigt sind Seiten mit verstecktem oder offenem rechtsradikalen Inhalt, insbesondere zum Nationalsozialismus, aber es gibt hinreichend viele scheinbar völlig harmlose Seiten, da im gebügelten Cut der Faktografie herkommend, in denen die Kunst des suggestiven Weglassens Regie geführt hat. Die Ausübung dieser Künste steht naheliegenderweise im Zusammenhang mit offiziellen Feierlichkeiten. So beging Österreich 2005 ein offizielles Gedankenjahr (60 Jahre Kriegsende, 50

¹ Für kritische Lektüre des Textes danke ich: Jakob Krameritsch, Marion Romberg, Martin Gasteiner (alle Wien).

Jahre Staatsvertrag, 10 Jahre EU-Mitgliedschaft), eine gute Gelegenheit für die Erstellung von Homepages über die österreichische Zeitgeschichte zur Verwendung in den Schulen. Bei manchen Seiten musste man nicht nur zu dem Schluss kommen, dass die Erfindung der Didaktik vermutlich frühestens im nächsten Jahrtausend stattfinden wird², sondern dass vor allem weggelassen wurde, was die Bezeichnung *Gedankenjahr* für den kritischen Kopf beinhaltet: Innehalten, kritische Distanz nehmen, reflektieren – zum Beispiel über den Umgang mit dem Nationalsozialismus und den Nazis in den Nachkriegsjahren oder über die Auseinandersetzung mit der nationalsozialistischen Vergangenheit etwa seit den 1970er Jahren in Österreich. Was ich hier kritisiere, stellt kein österreichisches Spezifikum dar, sondern ist generalisierbar.

Besonderer Aufmerksamkeit bedürften Computerspiele und Webgames, insbesondere solche, die den Zweiten Weltkrieg als Kulisse verwenden. Spiele wie *Hearts of Iron 2*³ beinhalten hunderte Seiten an historischen Erläuterungen und, im konkreten Fall, 20.000 Fotos aus der Kriegszeit; andere Spiele bieten Faksimiles historischer Urkunden, Zeitzeugeninterviews und mehr an – ganz zu schweigen vom filmischen Charakter der Spiele, der den Eindruck eines historisch-authentischen Kriegsgeschehens vermittelt.⁴ Selbst wenn nicht zu erwarten ist, dass die Spieler/innen – in der Regel Männer zwischen 15 und 45 – das alles lesen, ist auch das *Geschichte im Netz*, eine ganz spezielle Form, in der die suggestive Kunst selektiver Faktografie besonders tragend zu sein scheint und mit der sich die Geschichtswissenschaft im übrigen künftig etwas genauer befassen sollte.

Noch einmal: Geschichtswissenschaftliche Seiten stellen im Netz eine Minderheit dar, auch wenn die Geschichte als Fach, wie es Michael Nentwich⁵ herausgearbeitet hat, zu den eifrigsten Produzenten im Web gehört.

² Vgl. z.B. die Website zur Ausstellung „Österreich ist frei“, die einen „Unterrichtsleitfaden“ anbietet: <<http://www.oesterreichistfrei.at/lehr1.htm>> (letzter Aufruf 05.04.2006).

³ Zur Information: <<http://www.heartsofiron2.com/>> (05.04.2006).

⁴ Ich beziehe mich auf eine im Entstehen begriffene Diplomarbeit von Daniel Hufler, Wien, über „Der Zweite Weltkrieg in Computerspielen“.

⁵ Nentwich, Michael, Cyberscience. Research in the Age of Internet, Wien 2003.

Unser Kongress .hist 2006 spiegelt die gegenwärtigen Kategorien von web-basierten geschichtswissenschaftlichen Angeboten umfangreich wider. Es geht um digitalisierte Primär-Quellen, um Inventare aller Art, um Bibliotheks- und andere Kataloge aller Art, um Linkkataloge oder Link-Bibliotheken, um Sekundärquellen aller Art, um webadäquate geschichtswissenschaftliche Einführungen, um historische Darstellungen, um Datenbanken, um E-Learning und E-Didaktik. Natürlich geht es auch um verschiedene technische Lösungen und um Design. Auf *Praxis* im Sinne der Frage: „Was finden wir eigentlich im Netz?“, das heißt auf *Praxis* im vordergründigen Sinn, brauche ich unter Verweis auf die kommenden Tage nicht detailliert einzugehen.

Die gegenwärtige *Praxis* geschichtswissenschaftlicher Seiten im Netz, um deren kritische Reflexion es mir geht, ist beschreibbar als komplexer Prozess, dessen wesentliches Merkmal die Systemmodifikation ist. Es handelt sich um eine *Praxis*, die als Subtext oder auch subversiver Sprechakt aus der vordergründigen *Praxis* herausgelesen werden muss. Der Systembegriff bezieht sich auf die Geschichtswissenschaft als institutionalisiertes Kommunikationssystem. Modifiziert werden die Kommunikationsmedien selber wie auch ihr systemischer Zusammenhang, modifiziert wird das traditionelle System geschichtswissenschaftlicher Kommunikation. Umgekehrt moduliert das traditionelle System, das ein System ohne Internet und Web gewesen war, auch das Netz. Den Zustand, den wir gegenwärtig haben, die gegenseitige Beeinflussung alter und neuer Medien in der Geschichtswissenschaft, sehe ich vorerst noch als ein Verbund-System, wenn auch mit zwei deutlich von einander unterscheidbaren Teilsystemen.

Die *Systemmodifikation* besteht darin, dass einerseits traditionelle Medien beibehalten, aber unter dem Einfluss der so genannten Neuen Medien verändert werden, während andererseits neue Medien eingeführt, aber unter dem Einfluss der alten Medien moduliert werden. In der einschlägigen Forschung wird dies als *Remediation* bezeichnet.⁶ Eine Analyse des

6 Vgl. beispielsweise Bolter, Jay David; Grusin, Richard, *Remediation. Understanding New Media*, Cambridge/Mass 1999; Kümmel, Albert; Scholz,

Seitenlayouts und einiger didaktischer Vorgehensweisen bei geschichtswissenschaftlichen Lehrbüchern – als Beispiel – ergibt, dass Seitengestaltungsprinzipien und proto-hypertextuelle Techniken, die für eine Website kennzeichnend sind, für Bücher übernommen werden. Ein prominenter Prototyp hierfür ist der Band *Frühe Neuzeit des Oldenbourg Geschichte Lehrbuch*, dessen Konzept von Anette Völker-Rasor entwickelt wurde.⁷ Umgekehrt lässt sich behaupten, dass viele Seiten im Netz nichts anderes darstellen, als die medienadäquat – manchmal auch nicht medienadäquat – aufbereitete Publikation eines Aufsatzes oder sonstigen Textes bis hin zur kompletten Monografie.

Ein Wort zu dem Adjektiv proto-hypertextuell: Der Begriff *Hypertext* ist von Anfang an für das Web reklamiert worden; eine Web-Adresse beginnt deshalb mit der berühmten Abkürzung *http* (hypertext transport protocol). Die aktuelle Hypertextforschung empfindet dies aber nicht unbedingt als glücklichen Umstand, weil zumindest der wissenschaftliche Hypertext sich wissenschaftstheoretisch stark von dem unterscheidet, was konventionell in Gestalt des Web auf Grund seiner theoretisch endlosen Verlinkung als Hypertext bezeichnet wird.⁸ Eine durchschnittliche Website, die die durchschnittliche Verlinkungstechnik anwendet, hat nur eine proto-hypertextuelle Qualität.

Modifizierende Beeinflussungen in beide Richtungen zwischen den auf *Alten Medien* und dem auf *Neuen Medien* basierenden Säulen des Verbund-Systems existieren auch hinsichtlich der Schreibformate. Aus vielerlei Gründen erfordert eine Website die Granulierung oder Portionierung der Inhalte. Idealerweise wird dabei nicht einfach ein langer Text – auf die Multimedialität, die ja mehr ist als nur Text im Designer Gewand des Web, komme ich später zu sprechen – granuliert oder portioniert, sondern Texte werden von vorneherein medienadäquat geschrieben. In

Leander; Schumacher, Eckhard (Hgg.), *Einführung in die Geschichte der Medien*, Paderborn 2004.

7 Völker-Rasor, Anette (Hg.), *Oldenbourg Geschichte Lehrbuch. Frühe Neuzeit*, München 2000.

8 Vgl. Krameritsch, Jakob, *Geschichte(n) im Netzwerk. Hypertext und dessen Potenziale für die Produktion, Repräsentation und Rezeption der historischen Erzählung*. Dissertation Universität Wien 2005.

Bezug auf das Netz bedeutet medienadäquat: kurz und bündig, präzise, auf das Wesentliche konzentriert, zugleich „an den Enden offen“ für weitere Vernetzung zu schreiben. Das sind Tugenden, die man auch gerne im gedruckten Buch oder Aufsatz verwirklicht sähe, im Web heißt das aber, dass die Rezeptionseinheit bildschirmadäquat zu sein hat, sprich bei Texten, aus nicht mehr als rund 1.500 Zeichen bestehen kann. Die erforderliche Granulierung der Inhalte kommt einer Modularisierung gleich. Die Systemmodifikation berührt somit zentral die wissenschaftlichen Schreibtechniken. Die angesprochene Modularisierung reicht noch weiter: Module werden als informationelle Einheiten erst sinnvoll, wenn sie mit anderen Modulen zu Netzwerken informationeller Einheiten verwoben oder verknüpft werden. Die Tätigkeit des Webens bzw. Knüpfens wird zu einem Gutteil auf die User/innen verlagert; Module sind nicht nur Bausteine unkomplexer Rezeptionsprozesse, wie oft unterstellt wird, sondern Bausteine individueller Sinnbildung.

Eine Systemmodifikation im Sinne der Rückwirkung von Eigenarten der Web-Publikation auf gedruckte wissenschaftliche Publikationen ist hinsichtlich deren Lebensdauer zu diagnostizieren. Vielen wissenschaftlichen Websites oder Web-Publikationen wird die Volatilität bzw. Fluidität ihrer Inhalte vorgeworfen. Oft gleichen sie mehr einem Diskussionsstand als einem gesicherten Ergebnis, ganz zu schweigen vom Verschwinden aus dem WWW. Dieselbe Beobachtung kann natürlich positiv als Stärke des Netzes gewertet werden; ich komme darauf zurück. Parallel zur Verbreitung des Web in den Geschichtswissenschaften ist eine Multiplizierung klassischer Publikationsformate zu beobachten: Noch nie gab es so viele Fachzeitschriften – mit begrenzter oder sehr begrenzter Reichweite! –, noch nie gab es so viele Sammelbände, noch nie gab es so viele Monografien mit immer kleineren Auflagen. Die Erfindung des Publishing-On-Demand ist ebenso aufschlussreich wie symptomatisch. Vieles von dem, was gedruckt wird, stellt auch nichts anderes als einen Diskussionsstand dar. In gewissem Sinne stellt wissenschaftliche Forschung nie etwas anderes als einen Diskussionsstand dar, aber es gibt doch sehr unterschiedliche Reifegrade von Argumenten. Kurz: Der *Vorwurf*, der oft gegen das Netz erhoben wird, muss sich bei selbstkritischer Betrachtung innerhalb

der Geschichtswissenschaften auch gegen gedruckte Publikationen wenden – oder wird, parallel zum Netz, positiv als produktive Stärke gesehen, als Entwicklung weg von der autoritativen, manchmal auch autoritären Lehrhaltung, weg von der autoritativen Meistererzählung hin zu offeneren Formen der Auffindung wissenschaftlicher *Wahrheit*. Ob die Entwicklung des WWW die Geschichtswissenschaft auf diesen neuen, gelegentlich als Weg demokratisierter wissenschaftlicher Erkenntnisproduktion bezeichneten Weg geführt hat, also die Ursache dafür ist, ist derzeit schwer nachzuweisen, zumindest aber verstärkt sie diesen Trend, und zwar nicht nur im Web, sondern auch im traditionellen System.

Traditionelle und Neue Medien beeinflussen sich gegenseitig und erleichtern damit den gleichzeitigen Umgang mit dem traditionellen wie dem neuen System. Bezeichnend ist die Tendenz, denselben Wissensbestand sowohl mit Hilfe der alten wie der neuen Medien zu publizieren. Dieses gegenseitige *backing* ist durchaus erfolgreich und beweist, dass beide Systemteile oder -säulen inzwischen aufeinander angewiesen sind. Der gleichzeitige Umgang mit beiden Systemteilen entspringt keiner Wahlfreiheit, sondern erfolgt zwangsweise. Vorläufig weisen beide einen je spezifischen Mehrwert auf. Trotz zahlreicher inhaltlicher Doppelungen sind die beiden Systeme ja nicht austauschbar. Wer den gleichzeitigen Umgang nicht beherrscht oder nicht will, wird sehr schnell als Staub hinweggepusht.

Chancen

Unter „Geschichte im Netz: Chancen“ ist eine Systemtransformation zu verstehen, die Transformation der Geschichtswissenschaft als institutionalisiertes System. Im Folgenden ist somit von zwei Systemen, von einem alten und einem neuen zu sprechen. Die Transformation lässt sich an sechs Beobachtungsfeldern festmachen:

1. An erster Stelle steht die Verflüssigung und Beschleunigung der Kommunikation, die hier im Kontext der zunehmenden Verlagerung der geschichtswissenschaftlichen Wissensbestände in das Netz zu sehen ist. Nicht nur laufen Suchprozesse nach *Informationen* aller Art schneller und schneller ergebnisreich ab, sondern auch die Zurverfü-

gungstellung, die Rezeption, die Verarbeitung von wissenschaftlichem Wissen beschleunigt sich. Anders ausgedrückt: Zumindest potenziell beschleunigt sich Forschung, beschleunigt sich der Weg zu einem Forschungsergebnis, beschleunigt sich die Einspeisung neuer Forschung in den Kommunikationsprozess. Ich verweise diese Feststellung bewusst in das Kapitel *Chancen*, denn noch wird dieses Potenzial des Netzes nur ansatzweise genutzt, und zwar hauptsächlich aus zwei Gründen: Zum einen entsprechen viele ins Netz eingespeiste Wissensbestände, insbesondere Primär- und Sekundärquellen, in puncto wissenschaftlicher Aufbereitung nicht den Anforderungen der Quellenkritik, der Überprüfbarkeit und der Verlässlichkeit. Viele Seiten sind nicht anders als rudimentär, schlimmstenfalls als Müllhalde, auf die aber doch manch schönes unversehrtes Stück gelangt ist, zu bezeichnen. Oft fehlen die erforderlichen Metadaten. Zum anderen, und daraus resultieren zum Teil die im vorigen Satz skizzierten Zustände, genießt das wissenschaftliche Publizieren im Netz bei einem sehr großen Teil der Historiker/innen nur ein geringes Ansehen, das mit der klassischen Publikation von gedruckten Monografien, Aufsätzen oder Quelleneditionen nicht mithält.

Andererseits verbessert sich die Qualität der Seiten im Netz zunehmend, die aktive und passive E-Medienkompetenz steigt und ist selber Gegenstand intensiver Debatten, die Masse der für die Geschichtswissenschaft lebensnotwendigen Wissensbestände, die in zuverlässiger Qualität im Netz verfügbar gehalten wird, steigt dynamisch an. Es ist schwer vorherzusagen, wann der kritische Punkt erreicht wird, an dem sich die Verhältnisse umgedreht haben, das heißt ab dem die in alten Medien zugänglichen Informationen und Wissensbestände vorwiegend den Charakter historischer Primärquellen erhalten und beispielsweise Bibliotheken nur mehr die Funktion eines historischen Archivs erfüllen.

Der Zeitpunkt, von dem an die wesentlichen *Forschungsimpulse* im Netz passieren, weil das Netz der von allen Mitgliedern des Systems der Geschichtswissenschaften bevorzugte Ort der Publikation und Kommunikation geworden ist, lässt sich nicht genau vorhersagen, ich

- reihe ihn aber unter die *Chancen* und nicht unter die *Visionen* ein.
2. Forschung in einem Netz-basierten System der Geschichtswissenschaft bedeutet nicht das Gleiche wie im traditionellen System. Während das Publikum, das Fachbibliotheken aufsucht, einigermaßen präzise bestimmt werden kann, während die sozioprofessionellen Gruppen, die wissenschaftliches Wissen verwenden, im traditionellen System ebenfalls relativ genau bekannt sind, herrschen im Netz potenziell andere Verhältnisse. Die Charakterisierung als *potenziell* erscheint angebracht, da derzeit wissenschaftliche Seiten vorwiegend von denselben sozioprofessionellen Gruppen genutzt werden wie im alten System. Zugriffsstatistiken beweisen außerdem, dass entgegen der Bezeichnung *world wide* nationale Nutzungsgewohnheiten vorherrschen; Deutsche konsultieren überwiegend deutsche und deutschsprachige Seiten, dann englischsprachige, dann andere. Das gilt sinngemäß für Franzosen wie Spanier und Engländer. Die Barrierelosigkeit des Netzes erweist sich in der Praxis eher als Chance, denn als tatsächlich genutzter Freiraum. Prinzipiell fallen aber viele Mauerringe hinweg, die im alten System den Elfenbeinturm der Wissenschaft sicherten und die Systemhoheit bei den Angehörigen des Systems beließen. Um es deutlich zu sagen: die Barrierelosigkeit und Offenheit des Netzes ist aus der Sicht der Geschichtswissenschaft nur dann positiv zu werten, wenn *sie* die Qualitätsstandards setzt und durchsetzt, wenn *sie* die Inhalte schafft. Da sie das in einem offenen System macht oder täte, besteht die Chance, den Markt für geschichtswissenschaftliches Wissen deutlich zu erweitern und mehr gesellschaftlichen Einfluss zu nehmen.
- Forschung in einem netzbasierten System bedeutet deshalb etwas anderes als im traditionellen System, weil das Publikum, das freien Zugang zu den Ergebnissen hat, diffus ist. Der Informationsfluss und seine Auswirkungen, die Verwendung, auch missbräuchliche Verwendung wissenschaftlicher Ergebnisse ist wesentlich weniger kontrollierbar als im System der Printmedien. Das wirkt nicht nur auf die mediale Aufbereitung von Forschungsergebnissen, sondern mittelfristig auf die Forschungsthemen zurück. Der Anteil so genannter relevanter Themen

und anwendbarer Forschung wird auch in der Geschichtswissenschaft steigen – jedenfalls solange nicht ein Schritt vollzogen wird, der technisch möglich ist, aber das Entwicklungspotenzial des Netzes abtöten würde, nämlich einen geschichtswissenschaftlichen Webspaces einzurichten, in den nur ein kontrollierter, eventuell sogar kostenpflichtiger Zugang vorgesehen ist. Es wäre technisch überhaupt kein Problem, den Elfenbeinturm der Wissenschaft im Web neu aufzubauen und sich ganz auf die unstrittigen Vorteile der Geschwindigkeit und leichten Verfügbarkeit von Wissensbeständen im Netz zu beschränken. In der Perspektive der raunzenden Pessimisten unserer virtuellen Gegentagung wäre das vermutlich eine Chance.

3. Eine wesentliche *Chance* besteht im multimedialen Charakter des Mediums *Netz*. Multimedialität hat drei Gesichter: Multimediale Darstellungsweise, Nutzung multimedialer Quellen, Inter- oder Multidisziplinarität. Alle drei Aspekte sind prinzipiell auch mit dem Modus *world wide* oder international, sagen wir: mit dem Modus der Globalität verbunden.

Die Geschichtswissenschaft hat sich in den letzten 25 Jahren zunehmend mit anderen als Text- oder schriftlichen Quellen befasst. Eine der Entwicklungstendenzen lässt sich mit dem Schlagwort „historische Bildwissenschaft“ beschreiben. In der Tat stellen ikonografische Quellen heute neben den Text- und schriftlichen Quellen die wichtigste benutzte Quellengruppe dar und finden sich infolge dessen neben Texten am häufigsten in geschichtlichen Seiten im Netz. An dritter Stelle steht inzwischen die Analyse von Fotografien und von Filmen, an weiteren Positionen folgen Videos und zum Beispiel Web-Games als audiovisuelles Quellenmaterial. Tonquellen wie Musik werden von der Geschichtswissenschaft – ich meine nicht die historische Musikwissenschaft – noch sehr vernachlässigt. Sieht man von Copyright-Fragen und kommerziellen Hürden ab, die insbesondere von den Inhabern der Rechte an Filmen und Tondokumenten aufgebaut werden, dann erleichtert das Netz auf Grund der ihm inhärenten Multimedialität die gleichzeitige Verwendung unterschiedlichster Quellentypen für die Forschung, da sie sich im Original, das heißt Ton als Ton, Film als

Film bzw. Videofile usw. wiedergeben lassen. Vielleicht führt über das so genannte elektronische Papier, in dem sich auch Ton und Video *drucken* lassen, ein Weg zum multimedialen *Buch*, aber das bleibt abzuwarten. Das Netz macht das Forschen mit unterschiedlichsten Quellentypen auch deshalb attraktiv, weil diese in immer größerer Zahl im Netz zur Verfügung stehen, und fördert wenigstens indirekt die Bereitschaft zu Inter- oder Multidisziplinarität.

Multimedialität bedeutet also Inter- oder Multidisziplinarität. Da die Zielgruppen im Netz heterogen sind und sich theoretisch nicht um Disziplingrenzen scheren, da das Medium *Netz* gewissermaßen laut am Bildschirm aufstöhnt, wenn es nur mit Text und nicht wenigstens auch mit Bildquellen gefüttert wird, da es erst dann zufrieden grunzt, wenn auch Tonquellen und Movies enthalten sind, verfügt das Netz über eine Art struktureller Gewalt, die der Inter- und Multidisziplinarität zu Gute kommt.

Um in diesem Kontext von *Chancen* zu sprechen, reicht es aber bereits aus, darauf hinzuweisen, dass das Netz den geschichtswissenschaftlichen Blick wieder erweitert. Die Verfügbarkeit unterschiedlichster Quellentypen und die vergleichsweise unkomplizierte technische Möglichkeit, sie miteinander verbunden in eine wissenschaftliche Analyse einzubauen, fördert die Entwicklung der Geschichtswissenschaft in Richtung einer inter- oder multidisziplinären historischen Kulturwissenschaft, die im Idealfall im Modus der Globalität gespielt wird.

4. Die Systemtransformation, der ich die *Chancen* zuordne, äußert sich somit einmal in einem Beziehungsnetz mit den Knoten Multimedialität, Inter- oder Multidisziplinarität, historische Kulturwissenschaften, jeweils im Modus der Globalität oder dem einer begrenzteren transnationalen oder einsprachigen Ebene. Sie greift jedoch noch weiter und tiefer. Aus der strukturellen Gewalt des Netzes resultiert ein im Lauf der Zeit immer größer werdender Druck bezüglich inter- und multidisziplinären Forschens und Darstellens. Obwohl Interdisziplinarität seit Jahrzehnten gefordert wird, setzt sie sich im alten System nur sehr langsam durch; alle, die interdisziplinär arbeiten, wissen ein Lied davon zu singen, wie gering die Wertschätzung ist, wenn es darauf an-

kommt: Interdisziplinär arbeitende Wissenschaftler/innen sind dann plötzlich keine richtigen Historiker/innen mehr, sie gelten als nicht eindeutig zuordenbar – und nichts schadet der beruflichen Karriere mehr als eine solche Einstufung. Im Netz ist Inter- und Multidisziplinarität Pflicht. Je größer das Gewicht des Netzes im System Geschichtswissenschaft wird, umso stärker wird sich die Beherrschung seiner Eigenheiten karrierefördernd auswirken.

Das genannte Beziehungsnetz ist letztendlich um zahlreiche Komponenten zu erweitern: Ich erwähnte schon, dass im Netz medienadäquate Schreibtechniken gefordert sind, die mit einer inhaltlichen Granulierung und Modularisierung Hand in Hand gehen. Diese Vorgänge wiederum korrespondieren mit der Zielgruppenoffenheit des Netzes und mit der hohen Bedeutung, die der aktiven Sinnbildung durch die Nutzer/innen des Netzes selbst zukommt, bzw. mit dem Bedeutungsverlust, den autoritatives oder autoritäres Wissen erleidet.

5. Nicht ausgeführt habe ich bisher den Aspekt des elektronischen Lernens, das E-Learning. Es kommt in der Geschichtswissenschaft erst langsam in Gang. Es zählt zu den Chancen, da es der Effektivitätssteigerung des Lehrens und Lernens dienlich ist. An den Universitäten und besonders in den geistes- und kulturwissenschaftlichen Fächern wird E-Learning vorrangig als *Blended Learning* betrieben, das heißt es wird eine Mischung aus Präsenz- und Online-Lehre praktiziert. E-Learning, egal in welcher genauen Gestalt – ob nun mittels einer Lernplattform (Lernmanagement System) oder in Gestalt der Internetgestützten Lehre (IGL) –, erhöht grundsätzlich den Anteil kommunikativer Operationen und fördert das selbstgesteuerte Lernen. Darin spiegeln sich die bereits erwähnten Netz-typischen Elemente wie Verflüssigung der Kommunikationsabläufe und Erhöhung der Verantwortung des Individuums für die Wissens- und Sinnbildung. E-Learning und IGL greifen zentrale Elemente aus der heute bestimmenden Alltagskultur auf, die Schüler, Studierende, an Geschichte interessierte andere Erwachsene mitbringen. Viele Elemente sind mit Schnelligkeit verbunden: schnell an Informationen gelangen, schnell Informationen verknüpfen können, schnell zu Ergebnissen gelangen, Probleme und

Fragestellungen schnell lösen, schnell zu Entscheidungen kommen. Das ist unser Alltag! Die Nutzung von Web und Internet in der Lehre wie im Fall der IGL greift diesen alltagskulturellen Habitus auf und moduliert ihn in einen wissenschaftlichen Habitus um. Das Netz ermöglicht es Studierenden zum Beispiel, Forschungstechniken und die Ergebnisse, zu denen diese führen, live im Zeitraffertempo vorzuführen. Das Netz als elektronisches Medium ermöglicht es, die Lösungen von Forschungsproblemen auch mit größeren Studierendengruppen effektiv durchzuspielen. Diese Aussage gilt schon jetzt, gehört dennoch in den Bereich der Chancen, da längst nicht alle Hörsäle mit W-Lan ausgestattet sind oder PC-Übungsräume in ausreichender Größe und Menge vorhanden sind, wo alle Studierenden am Notebook oder PC mit Internetanschluss die Forschungs- und Lösungstechniken mitvollziehen bzw. selber durchspielen können. Da gerade Nobelpreisträger nicht müde werden, auf den Zusammenhang von Wissenschaft und *Fun* hinzuweisen, fühle ich mich legitimiert zu sagen, dass der Einsatz des Netzes den Fun-Faktor in der Geschichtswissenschaft auf produktive Weise wieder erhöht und im Übrigen die traditionelle zivilisatorische Grenze zwischen Wissenschaft und Alltag öffnet.

Die Wahl der Mittel für E-Learning ist nicht gleichgültig. Lernplattformen stellen geschützte Lern- und Lehrbereiche dar; das hat durchaus gute pädagogische und didaktische Gründe, mit Bezug auf das Netz und seine Potenziale sind aber die, im Fach Geschichte derzeit auch noch überwiegenden, unterschiedlichen Formen der IGL vorzuziehen, da diese im Netz öffentlich zugänglich sind – von der Möglichkeit, Seiten durch Passwort abzuschirmen, wird nur ausnahmsweise Gebrauch gemacht. Lernplattformen, so genannte Lernmanagement Systeme, stützen ungewollt das Elfenbeinturmprinzip. IGL durchbricht dieses Prinzip. Geschichtswissenschaft im Netz – auch für Lehr- und Lernzwecke – kann ein riesiges gesellschaftliches Potenzial entwickeln, wenn die gegebene Chance, Geschichtswissenschaft sehr breit zu vermitteln, genutzt wird: E-Learning ist eine zentrale Möglichkeit, da der Anwendungsbereich nicht auf Universität und Schule begrenzt ist, sondern genauso gut in allen möglichen Spielarten der

Erwachsenen- und beruflichen Bildung zu Hause ist, und weil es darüber hinaus sich die Tatsache offener bzw. heterogener Zielgruppen im Netz zu Nutzen machen kann.

Diejenige Wissenschaft, die hier beherzt zugreift, das heißt entsprechende Kräfte, Energien und Gelder in die professionelle und umfassende Nutzung aller Potenziale des Netzes leitet, hat aufgrund der rasant wachsenden Nutzungsraten des Web, die sich arithmetisch auch als Bedeutungszuwachs rechnen lassen, die Chance, ihre Platzierung im System der Wissenschaften und der Wissenschaftspolitik zu verbessern. Praktisch bedeutet das zum Beispiel, multifunktionale virtuelle historische Mehrebenen-Lernobjekte zu erarbeiten. Als Modell ist *pastperfect*⁹ zu nennen, das hypertextbasiert die Geschichte des Zeitalters Karls V. zum Inhalt hat und exakt ein solches multifunktionales virtuelles historisches Mehrebenen-Lernobjekt darstellt.¹⁰

6. Das Netz verändert im System der Geschichtswissenschaft die Beziehungen zwischen Individuum, Kollektiv und wissenschaftlichem Wissen. Als Ort autoritativen Wissens wird das Individuum geschwächt, es wird gestärkt als Teil eines riesigen Wissensnetzwerkes, weil ihm eine größere Verantwortung im Prozess der Sinnbildung zugewiesen wird. In der Geschichtswissenschaft als Kollektiv wird die dominante Stellung von Einzelpersonen geschwächt; langfristig wird das weitreichende Auswirkungen auf sämtliche Institutionen der Geschichtswissenschaft haben. An diesem Punkt ist nun der Technik, genauer gesagt, der Software, kurz Aufmerksamkeit zu widmen. Auch das traditionelle System basiert auf unterschiedlichen Techniken. Im Netz kommen andere Techniken zur Anwendung, aber der Unterschied liegt nicht bei *Technik-basiert* versus *nicht Technik-basiert*. Unter Weglassung aller Details zeige ich mit dem Finger nur auf einen bestimmten Wandel, der sich im Netz vollzieht: Hat bisher bei der Erstellung von Websites HTML-Software dominiert, setzen sich allmählich Content

9 Vgl. <www.pastperfect.at>.

10 Geleitet von Jakob Krameritsch und Wolfgang Schmale; das Projekt wurde mit dem Medidaprix-Förderpreis 2004 sowie dem Medidaprix-Publikumspreis 2004 ausgezeichnet.

Management Systeme durch. CMS-Lösungen sind *die* Teamwork-Software schlechthin – jedenfalls derzeit. Ein CMS wie der Hypertextcreator, den wir in Wien für die Anwendung in den historischen Kulturwissenschaften entwickelt haben, bündelt die hier mehrfach aufgezählten Potenziale des Netzes. Die Systemtransformation im Sinne der Multimedialität und Multidisziplinarität erstreckt sich nicht nur auf die Geschichtswissenschaft im Kontext der historischen Kulturwissenschaften, sondern – ich paraphrasiere Michel Houellebecq¹¹ – die ‚Kampfzone weitet sich aus‘ in das Feld der Programmierertechnik, des Mediendesigns, ja, der Medienkunst. Damit komme ich von den Chancen zu den Visionen.

Visionen

Kulturgeschichtlich wurden die Phänomene *Internet* und *Web* als Medienrevolution charakterisiert. Diese Medienrevolution hat sehr schnell höchste historische Weihen erhalten, insoweit sie bezüglich ihrer Geschichtsmächtigkeit auf eine Stufe mit dem althistorischen Übergang zur Schriftlichkeit in der von dem Philosophen Karl Jaspers zuerst so genannten *Achsenzeit*¹² sowie mit dem Übergang zum Buchdruck im 15./16. Jahrhundert gestellt wurde. Wir wissen alle, dass *Revolution* als kulturwissenschaftlicher Begriff nicht in erster Linie auf die Schnelligkeit fundamentaler Transformationen als Merkmal rekurriert, sondern auf die *Fundamentalität* der Transformationen selbst. Zwar zeichnet sich das Netz wesentlich durch die technische Beschleunigung jedweder Art von Kommunikation und der Wahrnehmung von – allgemein gesprochen – Zeichen aus, aber als kulturelles Phänomen entwickelt es sich eher gemächlich, nicht anders als der Buchdruck. Zweifellos gibt es in Bezug auf den Buchdruck und das Internet jeweils einen historischen *Moment*, der im weiteren kulturellen Rezeptionsprozess symbolische Kraft erlangt hat. Im Fall des Buchdrucks ist klar, dass dieser *Moment* in der Person Gutenbergs und des Drucks der 42-

11 Houellebecq, Michel, *Extension du domaine de la lutte*, Paris 1994. Houellebecq beschäftigt sich in seinen Romanen nicht zuletzt mit der Bedeutung elektronischer Medien aus gesellschaftskritischer und zukunfts-kritischer Sicht.

12 Jaspers, Karl, *Vom Ursprung und Ziel der Geschichte*, München 1949.

zeitigen Bibel 1454 besteht; schon berühmte Zeitgenossen wie der spätere Papst Enea Silvio Piccolomini sahen das so. Im Fall des Internet ist das schwieriger, weil sich die Sache zunächst eher im Stillen entwickelte. Die meisten Bücher über das Internet favorisieren das Jahr 1991 mit der so genannten Erfindung des WWW durch Tim Berners-Lee vom CERN in Genf, ich halte allerdings dafür, dass wir uns für den entscheidenden *Moment*, wie schon im Falle Gutenbergs, am Papst, diesmal einem zu dem Zeitpunkt, um den es geht, amtierenden, orientieren: 1995 erlaubte Papst Johannes Paul II. die Freischaltung der ersten Website des Vatikan.¹³ 1995 ist das Jahr, post quem das Web wahrhaftig die Welt eroberte. Diskussionswürdig bleibt die Frage, wann die fundamentalen Transformationen, um die es geht, erreicht waren bzw. erreicht sein werden. Die zitierten *Momente* sind qualitativer, aber nicht quantitativer Natur; die Rede von fundamentalen Transformationen visiert jedoch quantifizierbare Phänomene an, das Erreichen der so genannten kritischen Masse, für die ein bestimmtes Medium lebenswichtig wird.

Medienrevolutionen sind Teil einer umfassenden Veränderung einer Kultur oder von Kulturen im Plural. Kultur – ich gehe von einem semiotischen Kulturbegriff aus – ist als ein sehr komplexer Code zu verstehen, der den Sinn oder Widersinn und die Bedeutungen, die menschlichem Tun und Handeln und Denken inhärent sind, codiert – manchmal eindeutig, zumeist aber zwei- und mehrdeutig! Die Zeichen dieses Codes sind vielfältig, auf einer mittleren Ebene werden sie komposit und bilden Subcodes aus: Gemeint mit dieser kompositen Ebene sind Texte, Bilder, musikalische Kompositionen, architektonische Gebilde, Gestik, Mode, Inszenierungen (die mehrere Subcodes miteinander verbinden), nicht zuletzt das WWW.

Wie in der Semiotik sind Bezeichnendes und Bezeichnetes im Konnex zu betrachten. Was bedeutet somit das *Netz* – hier in der doppelten Bedeutung von Internet und Web – kulturell und kulturgeschichtlich? Das Netz bedeutet zunächst einmal selbst einen Subcode, der bestimmte Transformationen codiert.

13 Zehnder, Matthias W., *Geschichte und Geschichten des Internets*, Kilchberg 1998, S. 80.

Um was für ein *Was?* handelt es sich, das in der kulturhistorischen Perspektive das Netz als Sprache oder Subcode übersetzt? Das Was ist die nicht-essentialistische, nämlich hybride, fluide, volatile, hypertextuelle Zivilisation, die unsere Gesellschaft zunehmend charakterisiert und die teilweise als Praxis, teilweise als Chance bzw. Schicksal, teilweise aber als Vision – manche würden sagen: Horrorvision – besteht. Hybridität, Fluidität, Volatilität, Hypertextualität, sind in erster Linie global verbreitete großstädtische, metropole Realitäten, die jedoch über ihren direkten Wirklichungsraum hinaus zivilisationsbildend wirken. Der Zusammenhang zwischen dem *Netz* als Codierung und der Zivilisation, dem Was, das es codiert, ist keineswegs nur metaphorisch. Die Welt, in der das *Netz* eine signifikante Bedeutung erlangt hat und von der aus es andere Räume und Köpfe erobert, ist die Zivilisation der metropolen Stadt, in der die sozialen, ökonomischen, religiösen, politischen, kulturellen und anderen Gefüge einen tiefgreifenden Wandel durchmachen, indem sie vom Gefüge zum Netz bzw. Hypertext mutieren.

Der Wandel bezieht sich infolgedessen auf die Auflösung von stabilen Gebundenheiten, Gebundenheiten an soziale, ökonomische, religiöse, politische, kulturelle, berufliche Gruppen. Gebundenheit oder anders ausgedrückt: Loyalität, wird durch Fluidität und Volatilität, oder anders ausgedrückt: durch zeitlich begrenzte und nur noch relative Loyalitäten ersetzt. Das menschliche Individuum nimmt eine neue Position ein: Es handelt sich nicht einfach um die Fortsetzung eines historischen Individualisierungsprozesses, sondern um den Beginn eines neuen Typus. Historisch gehören stabile Gebundenheiten oder Loyalitäten und essentialistische Konzeption des Individuums ebenso zusammen wie in der neuen Zivilisation zeitlich begrenzte und nur noch relative Loyalitäten mit dem Typus der Hybridität korreliert sind.

In der pessimistischen Sichtweise wird dieser Wandel als Verlust interpretiert; dabei wird aber übersehen, dass an die Stelle der essentialistischen Konstellation eine neue tritt, die sich mit dem Wort *Kohärenz* charakterisieren lässt. Hybridität, Fluidität, Volatilität, Hypertextualität sind erforderlich, um in einer wie noch nie in Fluss geratenen Welt Kohärenzen zu erzielen. Essentialismus kann sich sehr kontraproduktiv auswir-

ken, wenn er Kohärenzbildung verhindert. Es ist unrichtig, Hybridität, Fluidität und anderes mehr mit Wertelosigkeit, Amoralität, Verlorenheit, Identitätsverlust und anderen negativen Begriffen in Verbindung zu bringen.

Wie dem auch sei: In Bezug auf das Thema *Geschichte im Netz* bleibt festzuhalten, dass das Netz (Internet und Web) den anskizzierten zivilisatorischen Wandel codiert. Das ist – und damit komme ich auf die begrifflichen Abwägungen zu Beginn des Vortrages zurück – die ‚Geschichte, die sich im Netz vollzieht‘.

Die Veränderung unserer Zivilisation bringt eine durchschlagende Veränderung geschichtswissenschaftlicher Grundpositionen mit sich. Unsere Alltagswelt, oder sagen wir: die Gegenwartsgeschichte, funktioniert mehr und mehr nach dem Hypertext-Muster, verständlich wird uns aber auch vergangene Geschichte, wenn wir sie wie einen Hypertext sehen lernen, so dass deren Repräsentation im Medium des Hypertextes adäquat erscheint. Die grundsätzliche Position, von der aus wir *Geschichtswissenschaft* betreiben, ändert sich radikal. Die Fragen, die wir *an die Geschichte* stellen, das ist ein banaler Lehrsatz, sind an unsere Gegenwart rückgebunden, genauer: an unsere gegenwärtige Positionierung, die sich darin ausdrückt, dass wir, die einen mehr, die anderen weniger, immer mehr Elemente des vorhin beschriebenen hybriden, fluiden, volatilen, hypertextuellen Individuums annehmen. Wir brauchen nur die Ansätze, Konzepte und Theorien der letzten zehn bis fünfzehn Jahre Revue passieren lassen, die die Geschichtswissenschaft heute prägen: Kulturgeschichte – oft im weiteren Gewand der *cultural studies* und der historischen Kulturwissenschaften; Kommunikationstheorien, integrationsgeschichtliche Ansätze, also Ansätze, die Kohärenzbildungen untersuchen, und die vom zeitgeschichtlichen Feld der Erforschung der europäischen Integration sowohl auf die antike, wie die mittelalterliche, wie die (früh-)neuzeitliche Geschichte ausgedehnt wurden; Beziehungsgeschichte(n), die je nach Fokus Kulturtransferforschung, *histoire croisée*, *entangled history*, Internationale Beziehungen genannt werden; historisch-semiotische Ansätze; nicht zuletzt der Netzwerkbegriff als Begriff mit theoretischem Anspruch, der zunehmend zur kategorialen Einordnung historischer Untersuchungsgegenstände

verwendet wird.¹⁴

Als Vision ergibt sich ein neues Beziehungssystem, das aus den Komponenten Netzwerk-Zivilisation, Konzeptualisierung der Geschichtswissenschaft als Wissenschaft geschichtlicher Kohärenzen im Netzwerk und ihrer Zeitschichten, Web als Netzwerkmedium und dem hypertextuellem, fluidem, volatilem Individuum besteht.

Das *Netz* als Code wird sich kaum mehr statisch wie derzeit noch präsentieren: Die in Zukunft *typische* historische Website wird eher dem Typus des Webgame entsprechen: Die erwähnte multidisziplinäre und multimediale Herangehensweise, gegebenenfalls im Modus der Globalität, hat ja den tieferen Sinn, Geschichte nicht mehr selektiv, sondern komposit zu sehen und zu verstehen. Multidisziplinarität und Multimedialität als Forderung sind in dem Maße gewachsen, wie sich die metropole Netzwerkzivilisation entwickelt hat, aus der unsere an die Gegenwart rückgebundenen historischen Fragestellungen resultieren. Das Potenzial von Multidisziplinarität und Multimedialität wird erst dann ausgeschöpft, wenn diese Ansätze im Verbund mit der möglichen Technik genutzt werden, um Geschichte, das heißt geschichtliche Kohärenzen gewissermaßen im Fluss, fließend darzustellen und ebenso zu verstehen. Das Netz codiert die gegenwärtige transformierte Zivilisation besser als die *alten Medien*, und es codiert unsere neuen Sichtweisen auf Vergangenes adäquater.

Wir wissen nicht, wohin uns dieser Fluss führt, aber so wie vor circa 2.700 Jahren die Europa, nachdem sie ihre anfängliche Angst überwunden hatte, beherzt auf dem Rücken des Stiers ausharrte und das Meer durchquerte, kann ich zum Abschluss nur rufen: „Alle Historiker/innen auf ins Netz!“

14 Vgl. zur Begriffsbildung: Schulz, Matthias, Netzwerke und Normen in der internationalen Geschichte. Überlegungen zur Einführung, in: Historische Mitteilungen der Ranke-Gesellschaft 17 (2004), S. 1-14. In Bezug auf Neue Medien vgl. Castells, Manuel, Das Informationszeitalter. Band 1: Der Aufstieg der Netzwerkgesellschaft. Band 2: Die Macht der Identität. Band 3: Jahrtausendwende, Opladen 2001, 2002, 2003.

Prof. Dr. Wolfgang Schmale ist Professor für Geschichte der Neuzeit an der Universität Wien, Institut für Geschichte. Seine geschichtswissenschaftliche Forschungsschwerpunkte sind Geschichte und Neue Medien, Europa, Frankreich, Menschenrechte, Männlichkeit, Kulturtransfer. Details unter: <http://www.univie.ac.at/igl.geschichte/mitarbeiter/mitarbeiterangaben_schmale.htm>.

GRUßWORT DES PRÄSIDENTEN DER
BERLIN-BRANDENBURGISCHEN AKADEMIE DER WISSENSCHAFTEN

von Günter Stock

Meine sehr verehrten Damen und Herren,
ich darf Sie sehr herzlich zu unserer Podiumsdiskussion mit dem Thema *Vom Nutzen und Nachteil virtueller Informationswelten für die Historie* begrüßen. Diese gemeinsame Veranstaltung unseres Akademievorhabens *Jahresberichte für deutsche Geschichte* mit dem Fachportal *Clio-online* bildet den Abschluss der mehrtägigen Tagung *.hist 2006. Geschichte im Netz – Praxis, Chancen, Visionen*.

Das hier an der Berlin-Brandenburgischen Akademie der Wissenschaften betriebene Akademievorhaben *Jahresberichte für deutsche Geschichte* ist einzigartig, nicht nur – aber auch –, weil es von externen Gutachtern so gesehen wird, die ihm bescheinigt haben, dass es zur „internationalen Vorhut“ und damit zur Spitze unter den geschichtswissenschaftlichen Bibliografien gehört. Die *Jahresberichte* können auf eine 125-jährige Tradition in der historischen Fachinformation zurückblicken. Das für ihre Gründung wesentliche Ziel, nämlich der durch einzelne Wissenschaftlerinnen und Wissenschaftler nicht mehr zu bewältigenden, enormen Informationsflut in vernünftiger Weise beizukommen und diese einem qualifizierten Zugang zuzuführen, gilt heute mehr denn je. Insoweit ist es auch nicht verwunderlich, sondern nur konsequent, dass die *Jahresberichte* bereits sehr früh den Weg der EDV-gestützten Arbeitsweise beschritten und damit die Basis für das heutige digitale Informationsangebot, also für die frei zugängliche Online-Datenbank, geschaffen haben.

Der intensive Zugriff auf diese Datenbank zeigt, dass die *Jahresberichte für deutsche Geschichte* zu den wissenschaftlichen Vorhaben und Projekten zählen, die bei der Integration der digitalen Technologie in ihre

Arbeitsabläufe und bei der Nutzung des Internets für die Bereitstellung der Ergebnisse am weitesten fortgeschritten sind. Daher ist es für unsere Akademie besonders erfreulich zu sehen, dass dieses Vorhaben zu einem elementaren Bestandteil der digitalen historischen Fachinformationslandschaft geworden ist. Die Mitbegründung des Kooperationsverbundes Clio-online ist nur eine logische Konsequenz dieser Entwicklung.

Nicht nur aufgrund dieses Projekts, aber auch mit durch dieses und andere Projekte veranlasst, hatte die Berlin-Brandenburgische Akademie der Wissenschaften bereits frühzeitig eine Initiative namens „*Telota*“ (= *the electronic life of the academy*) ins Leben gerufen. Sie fungiert als Klammer für alle Aktivitäten der Akademie, die der Förderung von Forschung, Kommunikation und Präsentation durch elektronische Medien dienen – eine Aktivität, die unsere Internetpräsenz kontinuierlich erhöht und uns zunehmend besser in die Lage versetzt, über wissenschaftliche Inhalte weithin und fast mühelos, auf jeden Fall ohne Zeitverzug zu kooperieren und zu kommunizieren.

Neben dem Faktor Zeit sollte bei dem Einsatz elektronischer Medien immer auch an den Vorteil gedacht werden, der sich darin manifestiert, dass der Zugang zu relevanten Informationen und deren Verbreitung mit relativ geringer Infrastruktur möglich geworden ist – ein Novum und mittlerweile ein *Essential* wissenschaftlichen Arbeitens. Unsere Aufgabe, (wissenschaftliches) Wissen zu verbreiten, ist jedoch nur halb erfüllt, wenn die Information für die Gesellschaft nicht breit gestreut und leicht zugänglich ist. Es gilt daher, das Internet als Instrument für eine globale Basis wissenschaftlicher Erkenntnisse und geistiger Reflexion zu fördern. Diesem Grundverständnis entsprechend bekannte sich unsere Akademie bereits früh dazu, die neuen Formen der Wissensverbreitung nicht ausschließlich in der klassischen Form, sondern zunehmend auch nach dem Prinzip des „offenen Zugangs“ über das Internet zu fördern.

Konsequenterweise gehörte die Berlin-Brandenburgische Akademie der Wissenschaften daher 2003 – gemeinsam mit den großen deutschen Forschungsorganisationen, aber auch dem französischen Centre National de la Recherche Scientifique (CNRS) zu den Unterzeichnern der „*Berlin Declaration on Open Access to Knowledge in the Sciences and Humani-*

ties“ – ein Prinzip, das wir aktiv unterstützten.

Meine sehr verehrten Damen und Herren: Während es relativ leicht ist, den Nutzen des elektronischen Mediums darzustellen, sollte gleichwohl auch Raum gegeben werden darüber zu reflektieren, an welcher Stelle Nachteile oder – um in der Sprache des Mediziners zu bleiben –: „unerwünschte Wirkungen“ gleichzeitig ausgelöst werden. Und hier will ich – ebenfalls in Anlehnung an meine Erfahrungswelt – auf kurzfristig sichtbar werdende, aber vor allem auf sehr viel schwieriger sichtbar zu machende, langfristige unerwünschte Folgen verweisen: Gibt es solche? Und wie kann man eventuell deren Auftreten vermeiden?

Ich finde es richtig, dass unser Akademievorhaben sich dieser Frage mit einer Podiumsdiskussion unter Beteiligung ausgewiesener Expertinnen und Experten nähert, und ich danke Ihnen allen ganz herzlich, dass Sie sich die Zeit nehmen, diese Problematik mit uns heute Abend zu vertiefen. Christoph Cornelißen, Professor für Neuere und Neueste Geschichte an der Universität Kiel und gleichzeitig neuer wissenschaftlicher Projektleiter der *Jahresberichte für deutsche Geschichte*, danke ich insbesondere dafür, dass er die Moderation dieses wichtigen Podiumsgesprächs übernommen hat. Abschließend bleibt mir an dieser Stelle nur noch, dem Akademie Verlag, der uns in diesem Projekt sehr treu zur Seite steht, sowie dem Verlag Vandenhoeck & Ruprecht für die freundliche Unterstützung Dank zu sagen. Ihnen allen, meine sehr verehrten Damen und Herren, wünsche ich einen interessanten und spannenden Diskussionsabend.

Prof. Dr. Dr. h. c. Günter Stock ist Mitglied des Wissenschaftsrates und des Senats- und des Verwaltungsrates der Max-Planck-Gesellschaft, Vizepräsident der MPG, Senator der DFG, Senator der Deutschen Akademie der Naturforscher Leopoldina, Mitglied der Europäischen Akademie der Wissenschaften und Künste, Mitglied in verschiedenen Kuratorien, Mitglied mehrerer Aufsichtsräte und wissenschaftlicher Institutionen. Seit Januar 2006 ist er Präsident der Berlin-Brandenburgischen Akademie der Wissenschaften.

PODIUMSDISKUSSION: „VOM NUTZEN UND NACHTEIL
VIRTUELLER INFORMATIONSWELTEN FÜR DIE HISTORIE“

Zusammenfassung von Daniel Schlögl

*Podiumsdiskussion anlässlich der Tagung .hist 2006 am 24.02.2006 im
Leibniz-Saal der Berlin-Brandenburgischen Akademie der Wissenschaften,
Berlin*

Teilnehmer/innen:

Christoph Cornelißen, Universität Kiel (Moderation)

Friedrich Kittler, Humboldt-Universität zu Berlin

Claudia Lux, Zentral- und Landesbibliothek Berlin

Winfried Schulze, Universität München

Manfred Thaller, Universität Köln

Das Internet, stellt **Christoph Cornelißen** in seiner Einführung fest, ist inzwischen zu einem unverzichtbaren Bestandteil der Wissenschaftspraxis geworden. Jedoch ist die anfängliche Euphorie teilweise einer gewissen Skepsis gewichen. Die virtuellen Informationswelten mit ihren Chancen und Risiken stellen die Geschichtswissenschaft vor grundlegende Fragen: Kann die Wissenschaft mit ihren herkömmlichen, oft nur schwer verständlichen Diskursen der Geschwindigkeit und Omnipräsenz virtueller Angebote noch Stand halten? Bringt die Tendenz zur Visualisierung einen Wandel der Geschichtsbilder mit sich? Wie verändern sich Auffassung von Quellen und herkömmliche Fragen der Quellenkritik? Wer wird künftig historisch relevante Informationen im Netz bereitstellen, wer wird sie

wissenschaftlich kontrollieren? Wie steht es um die Kosten für die Nutzung der Informationen? Trotz der komplexen Thematik sind Pessimismus oder Beliebigkeitszynismus aber nicht angebracht. Mit Blick auf die vermeintlich neue Informationsflut hat Peter Weingart herausgearbeitet, dass sich die Wissenschaft bereits seit dem 18. Jahrhundert in einer Phase des exponentiellen Wachstums befinde. Auch das Phänomen der selektiven Wahrnehmung sei nicht erst mit dem Internet eingetreten; schon von jeher dürften sich circa 80 Prozent aller Zitate auf nur 20 Prozent aller gedruckten Artikel bezogen haben.¹ Nichtsdestoweniger sind starke Unterschiede in der Wahrnehmung der virtuellen Informationswelten festzustellen: Während Praktiker und Netzaktivisten vielfach noch ein ungebrochenes Verhältnis zu ihrer Tätigkeit haben, ist die Stimmung auf Seiten der herkömmlichen Wissenschaft auch durch Skepsis, Zurückhaltung und Unsicherheit gekennzeichnet.

Winfried Schulze weist darauf hin, dass auch die Diskussion um Vor- und Nachteile des Internets keineswegs neu ist, sondern bereits seit Jahren geführt wird. Die schnelle Zugänglichkeit von Rezensionen, aktuellen Zeitungsartikeln und digitalisierten Quellen haben die Techniken der Informationsbeschaffung für Historiker erheblich verändert. Insgesamt stehen jedoch noch zu wenig Inhalte frei im Netz; vielfach werden nur Meta-Informationen angeboten, während die eigentlichen Informationen im Druck oder über Umwege wie zum Beispiel JSTOR verfügbar sind. Eine Stärkung der Open Access-Bewegung ist demnach dringend erforderlich; die wenigen vorhandenen geschichtswissenschaftlichen Portale reichen für eine substantielle Anreicherung des Internets nicht aus. Hingegen sind schon jetzt die neuen Kommunikationswege von erheblichem Nutzen für die akademische Lehre, indem sie eine zeitsparende Begleitung von Lehrveranstaltungen und persönliche Betreuung ermöglichen. Das Problem, dass fremde Inhalte in Studienarbeiten hineinkopiert werden, hält sich durch die vorhandenen Kontrollmöglichkeiten in Grenzen. Was die Geschichtswissenschaft als Ganzes angeht, sieht Schulze eine Parallelität

¹ Weingart, Peter, Die Stunde der Wahrheit? Zum Verhältnis der Wissenschaft zu Politik, Wirtschaft und Medien in der Wissensgesellschaft, Weilerswist 2001.

zwischen dem Aufbau des Internets und der inzwischen beklagten thematischen Offenheit von Geschichtswissenschaft. Diese Tendenz hat zwar einerseits zu einer erfreulichen Themenvielfalt geführt. Andererseits erweist sich das Verschwinden kanonisierter Auffassungen und Themen als Problem; es fehlen die großen Geschichtsdebatten in der Öffentlichkeit, die früher für Bewegung gesorgt haben. Insgesamt jedoch ist die Synthese-Aufgabe guter Geschichtsschreibung noch nicht in Gefahr. Die Vorteile der Internetnutzung überwiegen eindeutig, so dass in Anlehnung an Kant gefordert werden kann: „Wage zu denken und auch im Netz zu schreiben“.

Friedrich Kittler sieht bei den Historikern/innen ein deutliches Defizit hinsichtlich der Beschäftigung mit dem Internet und seiner Geschichte. Imperien und Mächte sollten stärker von ihren technischen Medien und Netzen her gedacht werden. Wie im antiken Persien das Botensystem oder ab dem 19. Jahrhundert die Einrichtung von Tiefseekabeln und Funknetzen, prägen auch die neuen Kommunikationsmöglichkeiten in hohem Maße die Gegenwart. Der Ursprung des Internets liegt im militärischen Sektor: Ausgehend von der Vernetzung von 26 Kommandozentralen und Eliteuniversitäten in Nordamerika hat sich ein System entwickelt, in dem Rechneradressen miteinander kommunizieren und der Weg einer Nachricht nicht mehr vom Absender oder Empfänger, sondern von der Nachricht selbst bestimmt wird. Kittler plädiert dafür, die neuen Strukturen von Computern, Netzen und Daten für das Denken und Arbeiten aufzugreifen und auch bei der Analyse historischer Phänomene anzuwenden. Beispielsweise könnte damit die Chronologie der Vorsokratiker neu gefasst werden.

Manfred Thaller verweist zunächst auf größere Entwicklungstendenzen: Der DFG-Ausschuss für wissenschaftliche Bibliotheken und Informationssysteme, auf dessen Förderung circa 80 bis 90 Prozent aller bei .hist 2006 vorgestellten Projekte beruhen, geht in einem Perspektivpapier für die kommenden zehn Jahre davon aus, dass das digitale Informationsmedium zum Standard für die Informationsversorgung werden wird. Die Bedeutung des Buches als bibliothekarisches Sammelgut tritt damit zurück. Künftig werden als Standard digitale Angebote erstellt und aus diesen in vielen Fällen Bücher generiert – mit der Folge, dass diese Informationsangebote in allen Bibliotheken verfügbar sein werden. Ausgehend von den

vorhandenen Projekten ist die Digitalisierung historisch relevanter Buchbestände in großem Umfang vorgesehen, so dass die essentielle historische Überlieferung komplett digital zur Verfügung stehen wird. Aktuell wird die Digitalisierung der vom 16. bis zum 18. Jahrhundert erschienen Bücher, soweit sie in den Nationalbibliografien nachgewiesen sind, diskutiert. Der Forderungskatalog der DFG enthält darüber hinaus das Ziel, möglichst alle digitalen Angebote miteinander zu vernetzen. Vor diesem Hintergrund besteht bei den Infrastruktureinrichtungen ein erheblicher Nachholbedarf hinsichtlich der Ausweitung der IT-Kenntnisse. Für die Historiker bedeuten die technischen Möglichkeiten wichtige Chancen: 1. Nachdem die Entkanonisierung der Geschichtswissenschaft auch dazu geführt hat, dass die Editionsreihen ihre zentrale Rolle als gemeinsame Materialgrundlage der Forschung eingebüßt haben, bietet die massenhafte Bereitstellung digitalisierter Originalquellen die Chance, das entstandene Dilemma zwischen der Ausweitung der Gegenstandsbereiche und der bislang vielfach fehlenden Überprüfbarkeit der Aussagen zu überwinden. 2. Die bislang geltende Voraussetzung, dass die Beschreibung einer Quelle erheblich einfacher und günstiger ist als ihre Reproduktion, ist nicht mehr gültig. Die Geschichtswissenschaften stehen künftig vor der neuen Herausforderung von jederzeit und massenweise verfügbaren unedierten Quellen. 3. Die Informationswissenschaft ließ sich in den vergangenen Jahrzehnten von grundlegenden Annahmen leiten, wonach Wissen präzise und widerspruchsfrei ist. Das Internet hingegen ist durch Merkmale und Probleme gekennzeichnet, die dieser Annahme diametral entgegengesetzt sind. Aufgrund ihrer Fachkultur wären Historiker deutlich besser geeignet, mit den Herausforderungen des Internets umzugehen. Die Disziplin müsste aber zunächst lernen, mit den Entwicklungen konstruktiv umzugehen, um die neuen Medien mitbestimmen und dadurch gewinnen zu können.

Aus dem Bibliotheksbereich berichtet **Claudia Lux**, dass die Digitalisierung schon jetzt eine Schwerpunktaufgabe darstellt. Die Bibliotheken fordern aber eine deutlich darüber hinaus gehende nationale Digitalisierungsinitiative. Derzeit wird auf europäischer Ebene zwar in die Herausbildung von Strukturen und Standards investiert; für die historische Forschung käme es aber vielmehr auf die konkrete Bereitstellung von

Inhalten an. Immer noch beziehen sich Historiker hauptsächlich auf das Printmedium, während sich etwa Schüler bei Rechercheaufträgen auf CD-ROM, DVD und Video konzentrieren. Auch in der geschichtswissenschaftlichen Forschung muss es zu einer viel stärker multimedialen Entwicklung kommen, insbesondere da sich ein Teil der Historiker den neuen Medien noch weitgehend verschließt und diese nur am Rande benutzt. Angesichts der Beobachtung, dass viele Studenten erst gegen Ende ihres Studiums Bibliotheken aufsuchen, ist es unumgänglich, das Erlernen von Informationskompetenz und Recherchetechniken nicht nur als Pflichtleistung in das Studium, sondern auch als Unterrichtsfach an den Schulen zu verankern. Hinsichtlich der weiteren Beitragsmöglichkeit der Bibliotheken verweist Lux auf das Beispiel der von der Library of Congress gestarteten Initiative *World Digital Library*, entstanden aus dem Projekt *American Memory*.² Anders als bei Massendigitalisierungsinitiativen wie derjenigen von Google wird hier qualitativ auf den insbesondere für die Geschichte relevanten Bereich gesetzt. Die Größenordnung der finanziellen Förderung – 15 Millionen Dollar vom amerikanischen Kongress, weitere 45 Millionen Dollar von Stiftungen und Unternehmen – wäre beispielhaft für eine ähnlich gelagerte deutsche Initiative.

Kontrastiv zur grundsätzlich positiven Einschätzung der Diskutanten/innen weist **Cornelißen** auf mögliche Nachteile des Internets hin: Die erleichterte Zugänglichkeit von Informationen erschwert es, die Originalität von Studienleistungen zu überprüfen. Eine gewisse Tendenz zur Bequemlichkeit gefährdet traditionelle Arbeitsweisen der Historiker; wesentliche Erlebnisse bei der Archivarbeit wie etwa die Erfahrung von Authentizität und unvorhersehbare Funde drohen damit zu entfallen. Dem steht zwar eine neuartige Form von Entdeckungen im Internet gegenüber. Die Vielzahl und die Länge der ohne ausreichende qualitative Kontrolle ins Netz gestellten Texte, gerade auch im Bereich der Rezensionen, führen aber zur Übersättigung. Ist demnach trotz aller bisherigen Pluralisierung eine Einschränkung der Perspektiven zu befürchten?

Hinsichtlich der Kontrolle von Studienleistungen sieht **Schulze** keine

2 Siehe <<http://www.loc.gov/>>.

besondere Gefahr gegeben. Auch der Verlust des „goût des archives“ ist kaum zu befürchten – das Internet erleichtert zwar die Verfügbarkeit von Quellen, ersetzt aber nicht generell die Recherche vor Ort, zumal für die Auswertung von archivischem Massengut wie etwa Rechnungsbüchern. Die entscheidenden Fragen, an denen Historiker arbeiten sollten, werden weiterhin von der Gesellschaft und nicht vom Netz vorgegeben.

Dem Bedürfnis nach Kontrolle steht **Lux** grundsätzlich skeptisch gegenüber: Seit dem Einsetzen der Publikationsflut wurde an den Bibliotheken festgestellt, dass diese Möglichkeit nicht mehr gegeben ist. Der Zugang zu fehlerhaften Informationen ist auch im Internet offen.

Thaller benennt es als eine primäre Fähigkeit des Historikers, den Wert eines Dokuments beurteilen zu können. Ähnlich dem unterschiedlichen Renommee von Verlagen im Printbereich kann auch im Internet zwischen weniger ernst zu nehmenden und seriösen Informationsanbietern – etwa Universitäten und akademische Selbstorganisationen – unterschieden werden. Bezüglich der Befürchtung, die Archivnähe der Historiker könnte abnehmen, gibt Thaller zu bedenken, dass aufgrund der gestiegenen Studentenzahlen der Zugang zu Originalquellen aus konservatorischen Gründen stetig schwieriger geworden sei. Mit der Möglichkeit, digitalisierte Quellen ohne institutionelle Hemmnisse via Internet sehen zu können, könnten Studenten auch besser den Umgang mit den Quellen erlernen.

Bezüglich der Kontrollierbarkeit von Studienleistungen ist **Kittler** davon überzeugt, dass Betrug letztendlich primär den Betrügenden selbst schadet.

Schulze weist darauf hin, dass gegenüber dem früher sehr eingeschränkten Zitierhorizont in den letzten 20 Jahren eine dramatische Veränderung in der Geschichtsschreibung stattgefunden hat. Der bibliografische Erfahrungshorizont ist weltweit geworden. Es besteht nicht mehr die klare professionelle Dominanz einer bestimmten Denkhierarchie, sondern stattdessen eine aufgebrochene Szene ohne klare Positionen. Beispielsweise wird die Historiografie der Französischen Revolution entgegen der traditionellen Dominanz der Pariser Lehrstühle heute eher von amerikanischen Forschern geprägt. Das eigentliche Problem der aktuellen Geschichtsschreibung ist es demnach, diese weltweiten Entwicklungen wahrzunehmen.

men und in eine konsensfähige Rekonstruktion für die Gegenwart zu integrieren: Wo setzt man die Schwerpunkte, wenn man die Welt im Hintergrund hat?

Auf die Nachfrage von **Lux**, ob das „Früher“ eher im Sinne einer Beschränkung oder einer Beschränktheit zu analysieren sei, führt **Schulze** weiter aus, dass die ehemals fest gefügten Schulen der Interpretation historischer Ereignisse durchaus eine hilfreiche Beschränkung dargestellt hätten. Die heutige Situation trägt zur Problematik um die Wirkung der Geisteswissenschaften bei: Es gibt nicht mehr in gleicher Weise die kanonisierungs- und damit auch streitfähigen Gegenstände, an denen sich das Fach abarbeitet und die der – für die Wissenschaftsförderung entscheidenden – Politik vermittelt werden könnten. Anstelle der großen Meistererzählungen und Mandarinflugfiguren gibt es nun ein weites Feld, das integriert werden muss. Zwischen dem generellen Differenzierungs- und Pluralisierungsphänomen und dem, was konkret mit dem Internet geschieht, besteht eine Rückkopplung – es gilt, diesen Zusammenhang in den Griff zu bekommen.

Auch **Kittler** sieht es als notwendig, aber noch nicht hinreichend an, sich mit der Internetentwicklung und der Globalisierung der Historiografie vertraut zu machen. Er unterstreicht die Notwendigkeit, das Netz nicht nur als Kommunikationsmedium zu nutzen, sondern auch die Formen und Datenstrukturen zu adaptieren, die Computern und Datennetzen inhärent sind. Die Anwendung computerlogischer Strukturen zur Modellierung einfacher Sachverhalte, etwa Genealogien, würde die Geschichtswissenschaft weiterführen. Auch das Werk *Shadows and enlightenment* des Kunsthistorikers Michael Baxandall³, der mit Rendering-Techniken französische Gemälde auf ihre fotografische Treue hin analysiert hat, verdient in diesem Zusammenhang Beachtung.

Cornelißen greift den Gedanken der Enthierarchisierung auf: Gibt es in der neuen Informationswelt anstelle der Mandarine neue verborgene Hierarchien, etwa über das Google-Ranking? Besteht ein Wettbewerb zwischen den Informationen anbietenden Bibliotheken? Wie steht es um

3 Baxandall, Michael, *Shadows and enlightenment*, New Haven 1995.

die Konkurrenz zwischen europäischen und amerikanischen Angeboten, wie werden diese finanziert?

Lux führt aus, dass trotz der Förderung durch die DFG noch nicht genügend Mittel zur Verfügung stehen. Die Bibliotheken haben teilweise eigenständig Gelder requiriert oder auch auf Haushaltsmittel zurückgegriffen, wie etwa die ZLB für die Digitalisierung des Berliner Adressbuchs. Eine starke Zentralisierung im Bereich der Digitalisierung wäre nicht zielführend, vielmehr ist eine breit angelegte Kooperation unterschiedlicher Akteure – bis hin zu lediglich historisch Interessierten – notwendig. Sowohl Bibliotheken als auch Archive würden ihre Schätze in einem Großprojekt digitalisieren wollen; auch Museumsbestände sollten digitalisiert werden. Ideen und Initiativen sind zahlreich vorhanden, die Diskussion um Standards schreitet voran, doch es fehlen die Mittel zum Start groß angelegter Projekte. Lux plädiert für ein gemeinsames und energischeres Auftreten von Bibliotheken und Historikern; bestimmte Gelder wären für inhaltliche Projekte nachhaltiger angelegt als für wiederholte betriebswirtschaftliche Analysen. Trotz der zwischen den Bibliotheken bestehenden Konkurrenzen existieren gemeinsame Projekte, etwa die European Digital Library. Ein weltweiter Vergleich zeigt, dass in Asien bereits erhebliche Fortschritte erzielt worden sind: China digitalisiert in großem Umfang Zeitungen und hat mit der Digitalisierung seiner Bibliotheksbestände begonnen. Auf den Philippinen wurde in einem einjährigen Projekt das Nationalgut vollständig digitalisiert. Wo stehen wir?

Dazu merkt **Thaller** an, dass die Bundesrepublik im internationalen Vergleich sehr viel besser dasteht als gemeinhin angenommen wird. Gemessen an der Zahl der digitalen Objekte liegt Deutschland wohl vor Frankreich und etwa gleichauf mit England. Schwachstelle ist vielmehr die an den föderalen Strukturen scheiternde Vermittlung der digitalen Angebote in Form gemeinsamer Sucheinstiege. Bezüglich der Kostenfrage geht die DFG davon aus, dass die experimentelle Phase der Digitalisierung inzwischen überwunden ist. Für die Digitalisierung sämtlicher Inkunabeln in der Bundesrepublik – dies entspricht etwa 80 Prozent aller bis 1500 in Europa gedruckten Bücher – müsste ein Betrag zwischen fünf und zehn Millionen Euro aufgewendet werden, wobei davon auszugehen ist, dass die Digitali-

sierungskosten in den kommenden Jahren noch um jeweils zehn Prozent sinken werden.

Auf Fragen aus dem Publikum hin entwickelt sich zunächst eine Diskussion um das Nebeneinander von konventionellen und digitalen Medien. **Kittler** plädiert für eine Parallelität von Alt und Neu, zumal Digitalisate nicht alle Eigenschaften des Originals aufweisen können, dessen Realitätsdichte womöglich erst in Jahrhunderten erfasst werden kann. Bei allen Medienrevolutionen sind auch die alten Medien immer erhalten geblieben. Auf die Frage von **Lux**, wie sich Historiker den Medienmix wünschen, fordert **Schulze** eine Konzentration der Zeitschriften auf den digitalen Bereich. Gedruckte Zeitschriften kommen lediglich dem Gewinnstreben der Verlage entgegen und haben keine Berechtigung mehr. Bezüglich des Wandels in der Geschichtsschreibung führt er aus, dass historisch gesehen alle Pluralisierungsbemühungen immer wieder in neuen Autoritätsbildungen aufgefangen worden sind; Pluralisierung fasert niemals endlos aus, sondern resultiert in neuen Kanonisierungsversuchen, die dann erneut pluralisiert werden. Die Geschichtsschreibung hat dies längst erkannt: Anstelle der überkommenen gottgleichen, festschreibenden Erzählperspektive haben Historiker eine „moderierende Geschichtsschreibung“ entwickelt, die in die Darstellung bereits die unterschiedlichen Perspektiven integriert. Ranke wollte zeigen, wie es eigentlich gewesen ist. Heute wollen Historiker nicht nur zeigen, wie es gewesen ist: Die neue Geschichtsschreibung berücksichtigt idealerweise auch, was aus dem Beobachteten geworden ist, wie es unterschiedlich gesehen worden ist und welche Wirkung dies wiederum erzielt hat. Unterschiedliche Ansätze verlieren demnach an Bedeutung und werden integriert in ein komplexer gewordenes Bild, aber immer noch in der Absicht, die Darstellung mit einer überschaubaren Seitenzahl an die Rezipienten heranzubringen, ohne die die Historiografie nicht leben kann.

Auch **Thaller** will das Buch nicht generell abschaffen, für ihn ist jedoch die Frage entscheidend, in welchen Fällen man sich die Publikation in Buchform leisten kann. Die Geschäftspolitik bestimmter internationaler Verlage sieht er als Gewinnoptimierung und Ausplünderung des Steuerzahlers. Es ist zu hinterfragen, ob eine nicht mehr veränderliche oder

anpassbare Form in allen Fällen die ideale Form darstellt, oder ob zum Führen wissenschaftlicher Diskussionen eine Form sinnvoller ist, in der die Aussagen in sehr kurzen Abständen dynamisch fortgeschrieben und geändert werden können.

Cornelißen stellt abschließend fest, dass die Diskussion um Nutzen und Nachteil der neuen Medienlandschaft für die Geschichtsschreibung noch intensiver werden wird, nicht zuletzt aufgrund der elektronischen Medien selbst. Notwendig wären zunächst noch mehr propädeutische Hilfsmittel für den Umgang mit den Angeboten des virtuellen Informationsmarktes, nicht nur für Studierende, sondern auch für Dozierende. Die Tatsache, dass viele Lehrende die neuen Möglichkeiten noch nicht bzw. zurückhaltend nutzen, stellt angesichts der bereits erreichten Qualität vieler Angebote ein sehr bedauerliches Defizit dar. Sowohl Forscher als auch Internetspezialisten sollten sich demnach um Strategien und Hilfsmittel für die Überwindung dieser Diskrepanz zwischen Angebot und Nutzung bemühen.

Prof. Dr. Christoph Cornelißen hat den Lehrstuhl für Neuere und Neueste Geschichte an der Christian-Albrechts-Universität zu Kiel inne und ist Projektleiter des Akademienvorhabens Jahresberichte für deutsche Geschichte an der Berlin-Brandenburgischen Akademie der Wissenschaften.

Prof. Dr. Friedrich Kittler ist Lehrstuhlinhaber für Ästhetik und Geschichte der Medien an der Humboldt-Universität zu Berlin und stellvertretender Direktor des Helmholtz-Zentrums für Kulturtechnik.

Hon.-Prof. Dr. Claudia Lux leitet als Generaldirektorin die Stiftung Zentral- und Landesbibliothek Berlin. Sie ist designierte Präsidentin des Weltverbandes der Bibliotheken (IFLA).

Prof. Dr. Winfried Schulze hat den Lehrstuhl für Geschichte der Frühen Neuzeit an der Ludwig-Maximilians-Universität München inne.

Prof. Dr. Manfred Thaller ist Lehrstuhlinhaber für Historisch-Kulturwissenschaftliche Informationsverarbeitung an der Universität zu Köln.

von Stefan Gorjßen

.hist 2006, nach einer ersten Tagung im Jahr 2003¹ der zweite große Kongress, der sich mit „Geschichte im Netz“, Fragen zu Möglichkeiten und Grenzen der Nutzung von Kommunikationsnetzwerken durch die Geschichtswissenschaft beschäftigte, brachte eine große, für einzelne Besucher/innen kaum mehr zu überschauende Zahl von Anbietern und Benutzern von Online-Diensten für Historiker/innen zusammen. In zwölf Sektionen und Workshops bot die Konferenz am 23. und 24. Februar insgesamt 80 Vorträge und Präsentationen zu aktuellen Projekten und zwang die Besucher/innen damit schon beim Blick auf das Programm, sich auf einige wenige Angebote zu konzentrieren und eine Vielzahl nicht weniger interessanter Angebote an sich vorbei ziehen zu lassen. Auch der vorliegende Bericht kann nur Eindrücke über einen kleinen Ausschnitt bieten. Bleibt zu hoffen, dass der Tagungsband mit allen Beiträgen bald verfügbar sein wird.

Der unbestreitbar größte Nutzen des Internets – gemeint sind hier Angebote im „World Wide Web“ – besteht für Historiker/innen darin, dass es den Zugang zu konventionellen Materialien, insbesondere zu Quellen und

* Bericht zur Sektion: „Digitalisierung und langfristige Verfügbarkeit historischer Quellen – Strategien und kooperative Lösungsansätze“ sowie „Viele Wege führen nach Rom – Suchstrategien nach historischen Quellenbeständen im Internet“. Der Querschnittsbericht erschien am 19.06.2006 auf H-Soz-u-Kult: <<http://hsozkult.geschichte.hu-berlin.de/forum/id=726&type=diskussionen>>.

¹ Die Tagungsbeiträge finden sich publiziert in: Burckhardt, Daniel; Hohls, Rüdiger; Ziegeldorf, Vera (Hgg.), Geschichte und Neue Medien in Forschung, Archiven, Bibliotheken und Museen. Tagungsband .hist 2003 (Historisches Forum 7, 2 Bde.), Berlin 2005: <http://edoc.hu-berlin.de/e_histfor/> (24.04.2006).

Literatur, erheblich vereinfacht und erweitert hat und dies nicht nur durch die Bereitstellung von Bibliothekskatalogen und Findmitteln der Archive auf online zugänglichen Oberflächen, sondern zunehmend auch durch die Digitalisierung von Quellenbeständen (neudeutsch als „Digitalisate“ bezeichnet), die in wachsendem Umfang als Bilddateien direkt von Webservern abgerufen oder in einer multimedialen Umgebung bearbeitet werden können. Die „stoffliche Seite der Geschichte“, von der Alf Lüdtke beim Blick auf den Löschsand in alten Akten einst hoffte, sie könne zum „nützlichen ‚Sand im Getriebe‘ [...] (der) professionellen historischen Forschung“ werden², geht hierbei freilich verloren, ein Verlust, den die meisten angesichts eingesparter Zeit und Reisekosten aber gerne in Kauf nehmen werden.

Organisatorische und technische Probleme bei der Digitalisierung historischer Quellen standen im Mittelpunkt von zwei Sektionen, auf die sich der vorliegende Bericht konzentriert. Eine erste thematische Sektion widmete sich vor allem dem Problem der Verfügbarkeit und dauerhaften Sicherung bereits vorliegender oder in der Bearbeitung befindlicher Digitalisate sowie den in den letzten Jahren unternommenen Anstrengungen, Kooperation zwischen Einzelprojekten herzustellen und übergreifende Standards und Verfahren zu etablieren. Eines der größten, international ausgerichteten Projekte in diesem Feld, das Monika Hagedorn-Saupe vom Berliner Institut für Museumskunde der Staatlichen Museen vorstellte, bezeichnet das Akronym MICHAEL (Multilingual Inventory of Cultural Heritage in Europe). MICHAEL hat sich zum Ziel gesetzt, die vielfältigen Initiativen zur Digitalisierung in mehreren europäischen Staaten zu vernetzen und kulturwissenschaftliche digitale Sammlungen von Museen, Bibliotheken und Archiven über ein zentrales Portal in einem übergreifenden Inventar zugänglich zu machen. Das von der EU geförderte, auf Open-Source-Technologie basierende Projekt wird derzeit von den Kulturministerien Italiens und Frankreichs sowie durch das britische Museums-

2 Lüdtke, Alf, Stofflichkeit, Macht-Lust und Reiz der Oberflächen. Zu den Perspektiven von Alltagsgeschichte, in: Schulze, Winfried (Hg.), Sozialgeschichte, Alltagsgeschichte, Mikro-Historie. Eine Diskussion, Göttingen 1994, S. 65-80, Zitat S. 66.

Bibliotheks- und Archivkonsortium betrieben, ein Erweiterungsantrag um Kultureinrichtungen neun weiterer europäischer Staaten wird zurzeit bei der Europäischen Kommission beraten. Im kommenden Jahr soll das Portal online gehen, man darf gespannt sein, wie sich der konkrete Nutzen des Projekts in der Praxis darstellen wird.

Ein weithin ungelöstes, gleichwohl drängendes Problem des Umgangs mit digitalem Material in Archiven wird mit dem Begriff „Langzeitarchivierung“ umschrieben. Für das Problem einer dauerhaften Sicherung digital gespeicherter Überlieferung sind verlässliche, allgemein anerkannte Konzepte bis heute nicht in Sicht. Auf welchem schwankendem Boden sich die Archiv- und Bibliothekswissenschaften auf diesem Feld noch bewegen, belegte Max Vögler mit seinem Bericht über die Förderpraxis der Deutschen Forschungsgemeinschaft (DFG) in den letzten Jahren: Hier ist man offensichtlich der Ansicht, dass die parallele Förderung konkurrierender Projekte und Konzepte auf mittlere Sicht dazu führt, dass sich das beste Konzept durchsetzt. Die mit dem Schlagwort „Langzeitarchivierung“ verbundenen Probleme sind vielfältig, sie betreffen neben technischen (geeignete Datenträger, Datenformate, usw.) auch rechtliche Probleme etwa der Softwarelizenzierung oder Fragen der Zuständigkeit und Kompetenz der zahlreichen beteiligten Institutionen. Eines dieser Projekte zur „Langzeitarchivierung“ wird mit dem Akronym „KOPAL“ durch die Deutsche Bibliothek betrieben. Fluchtpunkt ist das nicht gerade bescheidene Ziel, das deutschsprachige Web zu archivieren. So wünschenswert ein solches Projekt auf den ersten Blick auch erscheinen mag, bleiben doch eine Vielzahl von technischen und konzeptionellen Hürden, für die noch keine befriedigende Lösung gefunden wurde. Thomas Wollschläger stellte in seinem Referat die Grundzüge des Systems vor, das auf einem offenen XML-basierten Datenformat basiert und in der in den vergangenen Wochen fertig gestellten Version sich vor allem auf die Archivierung von Online-Dissertationen, von CD-ROMs, Digitalisaten und digitalen Musikdaten des Deutschen Musikarchivs in Berlin beschränkt. Vor allem die Sicherstellung einer ausreichenden Performance, die auch ein Web Harvesting ermöglicht, zählt zu den im gegenwärtigen Stadium des Projekts noch ungelösten Problemen.

Solange man sich auf die Archivierung digitalisierter schriftlicher Quellen konzentriert, bleiben die Probleme überschaubar und scheinen auch – zumindest auf mittlere zeitliche Sicht – gut lösbar zu sein. Die im Umfeld der „Text Encoding Initiative“ bereits seit den 1980er Jahren entwickelten Verfahren geben hierfür einen praktikablen Rahmen ab, wie Stefan Cramme am Beispiel kleinerer Projekte, wie sie von der Bibliothek für Bildungsgeschichtliche Forschung in Berlin betrieben werden (Digitale Ausgabe der Werke Adolf Reichweins, digitale Briefedition Friedrich Fröbels), anschaulich vorführte. Die Vielzahl von Initiativen und Projekten zur Langzeitarchivierung zu bündeln und nach gemeinsamen, übergreifenden Lösungen zu suchen, diese Aufgabe verfolgt das am Bundesarchiv in Koblenz angesiedelte Projekt „NESTOR“ (Network of Expertise in long-term Storage Of digital Ressources), das Andrea Hänger und Karsten Huth vorstellten. NESTOR erarbeitet Richtlinien, veröffentlicht Empfehlungen und Kriterienkataloge und bietet Fortbildungsveranstaltungen an, um auf diese Weise Grundlagen für projektübergreifende Standards zu legen. Unsere französischen Nachbarn scheinen diese mühsame Arbeit des Vernetzens und des technisch-konzeptionellen Austauschs nicht zu kennen. Das französische Projekt PERSEE³ gibt allen Universitätsbibliotheken klare Vorgaben für Digitalisierungsprojekte. Entsprechend rasch wächst das Portal, das frei zugänglich ist und bereits nach drei Jahren Laufzeit einen bequemen transparenten Zugriff auf die Volltexte der Artikel in allen wichtigen kulturwissenschaftlichen Zeitschriften bietet. Copyright-Fragen, an denen die Realisierung entsprechender Projekte in Deutschland regelmäßig scheitert, lassen sich mit einem mächtigen Ministerium im Rücken offensichtlich leicht lösen: PERSEE konnte bislang jeden Autor und noch jede Autorin, wie Frédéric Blin vom französischen Bildungsministerium einem staunenden deutschen Publikum versicherte, dazu überreden, auf seine diesbezüglichen Ansprüche zu verzichten.

Die wichtigste Aufgabe von Portalen zu digitalen und digitalisierten Quellen ist, den Nutzern/innen einen einfachen und schnellen Zugriff auf die entsprechenden Bestände zu ermöglichen. Mit Such- und Recherche-

3 Vgl. <<http://www.persee.fr>>.

strategien beschäftigte sich eine weitere thematische Sektion der .hist 2006. Die Grenze zwischen Portalen, die Metakataloge und -verzeichnisse aufbauen, und Angeboten an spezialisierten Suchmaschinen sind dabei oftmals fließend. Die Erstellung von Online-Beständeübersichten im Bundesarchiv bedient sich, wie Anke Löbnitz berichtete, des 2003 auf XML-Technologie portierten Findbucheditors „MIDOSA“ (Mikrocomputergestütztes Informations- und DokumentationsSystem für Archive). Zielperspektive des großen Projekts „Netzwerk SED-Archivgut“, an dem außer dem Bundesarchiv auch die Landesarchive der neuen Bundesländer beteiligt sind, ist die Weiterentwicklung der MIDOSA-Software zu einer integrierten Suchmaschine über Beständeübersichten und Findbücher hinweg nach dem EAD-Standard („Encoded Archival Description“, 1998 in den USA standardisiert). Bis eine solche Suchmaschine durch die Forschung mit Gewinn zu nutzen ist, sind wohl noch einige Arbeitsjahre in die Erschließung und Retrodigitalisierung von Findmitteln zu investieren. Nicht zuletzt aus diesem Grund beteiligt sich das Bundesarchiv auch am Aufbau eines Findmittelkatalogs auf Clio-online, den Sebastian Bartheleit zusammen mit Robert Zepf von der Staatsbibliothek Berlin, Stiftung Preußischer Kulturbesitz, vorstellte. Das Angebot will den großen Bestand an Findmitteln der Staatsbibliothek online in einer Datenbank nutzbar und nach sachlichen Kriterien abfragbar machen. Gleichzeitig soll eine Schnittstelle bereitgestellt werden, über die Archive neue Findmittel direkt in die Datenbank einpflegen können. Das Projekt versteht sich dabei nicht als Konkurrenzprodukt, sondern als Ergänzung zum bereits seit mehreren Jahren existierenden BAM-Portal („Gemeinsames Portal für Bibliotheken, Archive und Museen“), das bereits auf der .hist 2003 vorgestellt wurde.⁴ Das BAM-Projekt greift hierbei auf bereits vorhandene digitale Bestandsverzeichnisse von Bibliotheken, Archiven und Museen zu und führt Meta-informationen über die dort gespeicherten Daten in einer zentralen Datenbank zusammen, in der dann eine Internetrecherche ermöglicht wird.

4 Maier, Gerald, Gemeinsames Internetportal für Bibliotheken, Archive und Museen. Das BAM-Portal, in: Burckhardt; Hohls; Ziegeldorf (Hgg.), Geschichte und Neue Medien (wie Anm. 1), S. 107-126, <http://edoc.hu-berlin.de/e_histfor/7_II#006013/> (24.04.2006).

Frank von Hagel vom Institut für Museumskunde in Berlin stellte aktuelle Bemühungen des BAM-Projekts vor, Objektdatenbanken der Museen zusammenzuführen und zugleich ein Erschließungsformat zu entwickeln, das kleineren Museen, die über kein eigenes Angebot verfügen, den Aufbau einer Objektdatenbank ermöglicht.

Das Institut für Museumskunde der Staatlichen Museen zu Berlin bemüht sich darüber hinaus, durch jährliche Befragungen einen Überblick über die in den deutschen Museen vorhandenen Objektdatenbanken zu gewinnen. Über Probleme bei der Erstellung einer Systematik kulturhistorischer Sachgüter und die unterschiedlichen Lösungsangebote im internationalen Maßstab berichtete Axel Ermert. Einen engeren thematischen Zugriff bezogen auf die Geschichte des Bergbaus verfolgt das Montanhistorische Dokumentationszentrum (montan.dok) beim Deutschen Bergbaumuseum. Primär- und Sekundärquellen schriftlicher, audio-visueller und sachlicher Natur werden hier in einer auf XML-Technologie basierenden Datenbank zusammengeführt.

Vergleicht man die Präsentationen auf der .hist 2006 zur Erschließung digitaler Quellen mit denen der .hist 2003 zeigt sich deutlich, dass die Diskussion sich weg von der (Retro-)Digitalisierung hin zur Vernetzung der inzwischen in großer Zahl realisierten Projekte und zum Aufbau übergeordneter Portale und Suchmaschinen verschoben hat. Hierfür gibt es kein Masterkonzept, sondern eine Vielzahl von Projekten auf unterschiedlichem Abstraktionsniveau. Viele der vorgestellten Projekte erfreuen sich einer oftmals leider nur sehr kurzfristigen Förderung durch die DFG, die damit auf Produktivitätsgewinn durch ein pluralisiertes Angebot zu hoffen scheint. Dieses durchaus sinnvolle Förderkonzept bedarf jedoch zwingend eines längeren Atems bei der Finanzierung, will man sicherstellen, dass die Projekte ihren Nutzen für die Geschichtswissenschaft und für die kulturelle Selbstvergewisserung der Gesellschaft entfalten können und nicht zu Datenfriedhöfen verkommen.

Dr. Stefan Gorißen ist Wissenschaftlicher Mitarbeiter an der Fakultät für Geschichtswissenschaft, Philosophie und Theologie der Universität Bielefeld. E-Mail: stefan.gorissen@uni-bielefeld.de

MITTELALTER IM WWW – EINE BESTANDSAUFNAHME*

von Harald Müller

Ein Querschnittsbericht zu .hist 2006 aus der Perspektive des Mittelalter-Historikers bzw. der Mittelalter-Historikerin bleibt eine künstliche Sache. Die meisten Themen der Veranstaltung wie „Suchstrategien“, „Fachportale“, „Elektronisches Lernen“ oder „Elektronisches Publizieren“ lassen sich nur bedingt epochenspezifisch aufsplitten. Sie fehlen daher im Folgenden, werden aber wohl an anderer Stelle auch für den häufig mit einer verhältnismäßig engen fachlichen Spezialisierung lebenden Mediävisten/innen angemessen aufbereitet. Zwei Sektionen hielten indes speziell für den am Mittelalter Interessierten viel versprechende Themen bereit. Deren Titel „Vernetztes Mittelalter. Entwicklungsstand und Perspektiven“ und „Quellen im Netz“ erweisen sich als charakteristisch für die Situation EDV-gestützten Arbeitens im Bereich der Mittelalterforschung. Im Mittelpunkt der einen stand die (möglichst) kooperative Verbesserung unterschiedlicher Informationsangebote im Netz, im Zentrum der anderen die Chancen digitaler Quellenaufbereitung. Beide Themenbereiche rühren damit an wichtige Felder historischen Arbeitens und dessen Veränderung durch die neuen Medien.

„Quellen im Netz sind anders!“ Diesen fundamentalen Satz schickte Patrick Sahle (Göttingen) der fast gleich lautenden Sektion „Quellen im Netz“ voraus, die von ihm und Georg Vogeler (München) verantwortet wurde. Anders in zweierlei Hinsicht: zum einen, weil das Internet als Werkzeug der Forschung mittlerweile selbstverständlich geworden ist,

* Bericht zur Sektion: „Quellen im Netz“ sowie „Vernetztes Mittelalter. Entwicklungsstand und Perspektiven“. Der Querschnittsbericht erschien am 23.05.2006 auf H-Soz-u-Kult: <<http://hsozkult.geschichte.hu-berlin.de/forum/id=707&type=diskussionen>>.

nicht aber als Ort, an dem Ergebnisse präsentiert werden; zum zweiten weil die Präsentation von Quellen in einem elektronischen Medium sich grundlegend von der Druckform unterscheidet – unterscheiden muss, will sie nicht leichtfertig die Chancen verschenken, die das Internet bietet. Quellen im Netz stellen also nicht nur eine neue Präsentationsform dar, sie erfordern eine eigene Herangehensweise an die Erschließung und Darstellung. Eine kohärente Methodik digitaler Quellenarbeit liegt aber bislang keineswegs vor. Mit sechs Vorträgen, davon vier, die sich mit dem Mittelalter und seinen Quellen befassten, sollten Wegmarken zur Erreichung dieses Ziels gesetzt werden.

Den Auftakt bildete eine ausführliche Vorstellung des so genannten *Monasterium*-Projekts durch Karl Heinz (Wien). Es handelt sich um eine Initiative zur Erschließung und Erforschung kirchlicher Quellen, die sich zunächst das Ziel gesetzt hat, die Urkunden der österreichischen Stiftsarchive zu erfassen und den Nutzern/innen mit vielfältigen Hilfestellungen zugänglich zu machen. Derzeit sind mithilfe von *Monasterium* rund 20.000 Urkunden verfügbar, was mehr oder weniger dem kompletten Bestand der österreichischen Stiftsarchive entspricht. Allerdings ist das Vorhaben weit ambitionierter geschnittenen, wie bereits das mehrsprachige Zugangportal dokumentiert.¹ Die Betreiber/innen planen eine Ausweitung ihrer Arbeit auf die Bestände der Nachbarländer, so dass im Idealfall der verfügbare Urkundenbestand verzehnfacht werden kann. Noch entscheidender aber ist, dass durch eine solche Kooperation die mittelalterliche Urkundenlandschaft wieder belebt werden könnte, die durch die heutigen politischen Grenzen willkürlich zerschnitten wird. *Monasterium* verwendet ältere Urkundenbücher und Regestenwerke in digitalisierter Form und ermöglicht den Benutzern/innen erstmals die vernetzte, übergreifende Recherche in disparaten Beständen, die zudem nicht immer leicht zugänglich sind. Abhilfe schafft hier die Einbindung digitaler Fotos der jeweiligen Urkunden. Für die Zukunft ist daran gedacht, nicht mehr nur konvertierend zu arbeiten, sondern mithilfe eines Redaktionssystems die Erschließung bislang nicht erfasster Bestände dezentral und unter Mithilfe der Benut-

1 Vgl. <<http://www.monasterium.net>> (20.03.2006).

zer/innen in Angriff zu nehmen. Fraglos eröffnet das Projekt mit seinen umfangreichen Abfragemöglichkeiten Heimat verbundenen Nutzern/innen wie professionellen Forschern/innen ganz neue Möglichkeiten. Vorsicht scheint allerdings bei dem Satz geboten, dass nunmehr in 95 Prozent der Fälle das Original nicht mehr konsultiert werden müsse. Der Einsparung von Zeit und Kosten auf der Nutzerseite steht als Risiko gegenüber, dass solche Aussagen die Archive ermuntern werden, ihrerseits Kosten für Benutzerbetreuung und wissenschaftliches Personal einzusparen.

Neben der Quellenbereitstellung, wie sie das *Monasterium*-Projekt wesentlich praktiziert, existieren auch primär analytische Vorhaben im Netz. Michael Gervers machte die Zuhörer mit dem DEEDS (Documents of Essex England Data Set)-Projekt der Universität Toronto vertraut, das sich einem äußerst verzwickten Problem der englischen Geschichte verschrieben hat. Von der Eroberung der Insel durch die Normannen 1066 bis ins 14. Jahrhundert sind die englischen Privaturkunden durchweg nicht mit einem Datum versehen; dies behindert eine stringente Auswertung erheblich. DEEDS versucht nun, mit elektronischen Mitteln Datierungshilfen für diese Quellengruppe bereitzustellen.² Dazu werden gedruckte Werke digitalisiert, aber auch Urkundenbestände systematisch neu bearbeitet. Gervers konnte eindrucksvoll vorführen, wie die Analyse von Wortverbindungen und formelhaften Wendungen mithilfe exakter statistischer Verfahren zu chronologischen Clusterbildungen führt, die immerhin eine Datierungsgenauigkeit von plus/minus fünf Jahren ermöglicht. Die grafische Darstellung von Buchstabenhäufigkeiten und der Vergleich dieser Muster erlaubt ferner die nahezu sichere Zuweisung von Urkunden zu identischen Schreibern. Eine Verfeinerung der Analysemethoden steht auf dem Programm der kommenden Jahre, wobei Gervers betonte, dass der Aufbau eines virtuellen Textarchivs keinesfalls das Ziel sei. Die acht Prozent der Dokumente, die ein Datum tragen, sollen weiterhin nur als Vergleichsmaßstab für die 92 Prozent der undatierten genutzt werden.

Mitten hinein in die Chancen digitaler Quellenedition führte Matthias Perstling (Graz). Er hat es übernommen, das „Steirische Marchfutterurbar

2 Vgl. <<http://www.utoronto.ca/deeds/research/research.html>> (20.03.2006).

von 1414/1426“ zu publizieren, ein detailliertes Verzeichnis von Futterabgaben, das über zwölf Jahre fortgeführt wurde. Die Veränderungen der Namen und Mengen sowie Streichungen und Ergänzungen, die über diesen Zeitraum hinweg vorgenommen wurden, haben die Liste zu einem so komplexen Dokument anwachsen lassen, dass eine traditionelle Edition in Buchform zum Scheitern verurteilt ist: Entweder reduziert sie die Komplexität soweit, dass die Quelle letztlich zerstört wird, oder sie bildet die verwirrende Vielfalt der Einträge und Streichungen sorgfältig ab, macht dadurch aber den Text unbenutzbar. Die viel versprechende Lösung heißt hier „dynamische Edition“. Dabei geht es nicht mehr um die Wiedergabe des Textes, die etwa bei der Edition einer mittelalterlichen Chronik im Mittelpunkt steht, sondern um die Aufbereitung der im Text – sofern es sich überhaupt noch um einen solchen handelt – bereitgestellten Informationen auf anderem Wege. Eine dynamische Edition stellt den Benutzern/innen die jeweils gewünschte Information zusammen, wobei der Inhalt von der Publikationsform getrennt wird. Die Nutzer/innen definieren also mit ihren Fragen die Textversion, die ihnen präsentiert wird. Wer nach einer Liste bestimmter Jahre fragt, erhält genau diese isoliert, wer bestimmte Namen sucht, kann diese herausziehen lassen, wen die einzelnen Schreiberhände interessieren, der bekommt die Beispiele vor Augen gestellt. Der „Text“ als Ganzes wird also elektronisch aufgelöst in eine Menge von Einzelinformationen, die mithilfe von Datenbanken und komplexen Suchabfragen jeweils bedarfsgerecht neu zusammengesetzt werden. Das ist unübersehbar ein Fortschritt, besonders für solche Quellen, die ihrerseits als dynamisch zu verstehen sind, Arbeitsdokumente, die nicht auf das Ideal eines fertigen, in Reinschrift vorliegenden Exemplars hin orientiert waren und die letztlich nie ein einheitliches Ganzes sein sollten. Die Ermöglichung ganz unterschiedlicher Fragestellungen, seien es namenkundliche, prosopografische, paläografische oder die wachsende Vielfalt statistischer Auswertungen, ist zu begrüßen; ein „lesbarer“ Text bleibt hier entbehrlich. Ob sich solche Verfahren für die Masse mittelalterlicher Quellen anbieten, scheint dem tief beeindruckten Betrachter dennoch eine genauere Abschätzung wert.

Das Spektrum der behandelten Quellen umfasste auch die mittelalter-

lichen Weltkarten. Günther Görz berichtete über die Tiefenerschließung mittelalterlicher *mappae mundi*, die in Erlangen zurzeit vorangetrieben wird und als dessen erstes Produkt hoffentlich noch in diesem Jahr mit der berühmten Ebstorfer Weltkarte eine Zimelie mittelalterlicher Kartografie auf ganz neue Art und Weise der Öffentlichkeit vorgestellt werden kann.³ Der besondere Reiz der digitalen Tiefenerschließung von Weltkarten liegt in der Verbindung dieser Quellen mit der zum späten Mittelalter hin zunehmenden Reiseliteratur, einer Wechselwirkung also aus Bild- und Schriftquellen. Die Erfordernisse des Reisens und die praktischen Erkenntnisse der Reisenden führten dazu, dass in die ursprünglich eher ein Weltbild im eschatologischen oder kosmologischen Sinne präsentierenden mittelalterlichen Karten neue Informationen eingetragen wurden, der Charakter der Karten dadurch vom Manifest der Weltanschauung zur geografischen Orientierungshilfe verändert wurde. Görz präsentierte im Detail das Herzstück des Projekts, die Datenbank, die Karten, Bilder und Metadaten verwaltet. Dabei gab er beeindruckende Einblicke in künftige Recherchemöglichkeiten. Sie lagen allerdings eher wie Inseln in einem Vortrag, der überwiegend von Problemen der Technik und der Aussagenlogik geprägt war. So ermüdend dieser Umstand auf manche Zuhörer/innen gewirkt haben mag, er verdeutlicht doch schlagend, dass die digitale Quellenwelt die Historiker/innen ebenso wie die Informatiker/innen fordert. Die bequeme Suche auf Knopfdruck ist insbesondere in solch komplexen Quellencorpora wohl nur um den Preis von Begrifflichkeiten wie „Konzepthierarchien“ oder „formale Ontologie der Datenbank“ und ihrer Inhalte zu realisieren. Die Sektion wurde durch zwei Beiträge zur Retro-Digitalisierung in der Zeitgeschichte und zur Aufbereitung von Filmen für das Internet abgerundet. Sie verschwanden erwartungsgemäß hinter dem begrenzten Interessenhorizont des Mediävisten, dokumentierten aber gleichwohl das erfolgreiche Bemühen der beiden Veranstalter, die Problematik digitaler Quellenedition mit einem wohl überlegten, breiten Fächer

3 Vgl. <http://univis.uni-erlangen.de/form?_s=2&dsc=anew/resrep_view&rprojs=phil2/IDSL/LGDP/kommen&anonymus=1&dir=phil2/IDSL/LGDP&ref=resrep&sem=2006s&_e=225> (20.03.2006).

von Zugangsweisen zu skizzieren.

Eine Präsentation einzelner Zeitalter und ihres aktuellen Standes in der digitalen Welt war von den Veranstaltern ausdrücklich nicht vorgesehen. Umso mehr überraschte es, dass ausgerechnet die Mediävistik sich mit einer epochenspezifischen Sektion vorstellte. Deren Titel „Vernetztes Mittelalter“ war durchaus programmatisch aufzufassen, wie die einleitenden Gedanken des Leiters, Arno Mentzel-Reuters, Mitarbeiter der *Monumenta Germaniae Historica* in München, schnell deutlich werden ließen. Er charakterisierte den gegenwärtigen Stand der Mediävistik auf diesem Gebiet als Scheidepunkt. Die bloße Bereitstellung von Informationen unterschiedlichster Art im Netz beschreibe die Arbeitsweise von gestern, die vernetzte Arbeit von Institutionen und Einzelforschern/innen markiere die Zukunft. Mit einigem rhetorischen Aufwand – der zuweilen den Eindruck der Selbstverliebtheit nicht vermeiden konnte – verdeutlichte Mentzel-Reuters diese Auffassung. Es dominiere heute das Internet-gestützte Arbeiten in der Form des Sterns: man baue Portale auf, von denen aus sich das Fach in alle möglichen Richtungen digital verzweige. Seine Vision aber sei der Kreis oder der Chor aufeinander bezogener autonomer Systeme, die Daten nicht nur bereitstellen, sondern sie weitergeben an den forschenden Nachbarn bzw. die forschende Nachbarin. Dies erhöhe die Effizienz und vermeide die Doppelerfassung, die heute noch gang und gäbe sei. Der praktische Einwand, dass die Sternförmigkeit des Internet-Angebotes wohl nicht zuletzt auch infolge der Förderlogik der DFG besteht, sei zunächst einmal zurückgestellt. Was Mentzel-Reuters nämlich provokant formulierte, hat auch viel Wahres an sich. Dem Sektionsleiter ist zuzustimmen, dass die Internet-Angebote bei allem Wert von einem enormen „Rauschen“ begleitet sind. Redundante Informationen, Mehrfachangebote, schlecht gepflegte Links und in ihrer Qualität kritikwürdige bis dubiose Seiten decken vielfach die wirklich nützlichen Dinge zu oder erschweren zumindest den selektiven Zugang. Jeder, der intensiv nach Informationen, Textausgaben etc. fahndet, wird diesem Befund zustimmen. Die Unbegrenztheit des Platzangebotes im Netz kann dazu verleiten, es als Müllhalde für Informationen aller Art zu missbrauchen; der Begriff der Hypertrophierung ist nur eine vornehme Umschreibung dessen.

Die Sektion versammelte in der Folge mit den *Monumenta Germaniae Historica*, den *Regesta Imperii* und der Handschriften-Datenbank *Manuscripta Mediaevalia* gleich drei zentrale Großprojekte der deutschen Mediävistik. Als viertes wäre das Internet-Angebot der Bayerischen Staatsbibliothek hinzugekommen, doch musste der Referent kurzfristig absagen. Allein diese Palette der großen Namen macht ein Verdienst der Veranstalter und gleichzeitig ein Problem der digitalen Mediävistik deutlich: Offenbar bedarf es des externen Anstoßes, damit diese Supertanker der Mittelalterforschung über Konvergenzen, Zusammenarbeit und die Vermeidung doppelter Arbeit nachdenken! Doch der Reihe nach.

Clemens Radl (München) präsentierte die digitalen *Monumenta Germaniae Historica*, die sich von der elektronischen MGH unter anderem dadurch unterscheiden, dass nicht nur die Texte bereitgestellt werden, sondern auch das gesamte Arbeitsinstrumentarium der Editionen. Nach der Realisierung des Projekts werden alle bis zum Jahr 2005 erschienenen *Monumenta*-Bände im Volltextzugriff verfügbar sein, die jüngeren Ausgaben jeweils sukzessive, sobald die mit dem Verlag ausgehandelte Schutzfrist für das gedruckte Buch abgelaufen ist.⁴ Mit den Texten werden auch die Register komplett über das Internet recherchierbar sein. Dies ist eine wichtige Entscheidung, da die Zusammenführung von Varianten, etwa disparater Ortsnamen, dort erfolgt, die Register also Zusatzinformationen bieten, welche die Volltextsuche in der Edition nicht unbedingt finden kann. Radl erläuterte anschließend technische Details zu möglichen Datentransfers. In der Diskussion war bei den Zuhörern/innen eine gewisse Enttäuschung darüber zu verspüren, dass die digitalen *Monumenta* offenbar nicht mit allen Such- und Verarbeitungsmöglichkeiten ausgestattet werden sollen, die heute denkbar sind.

Die *Regesta Imperii* teilen mit den *Monumenta Germaniae Historica* nicht nur einen Großteil der Quellen, auf die sie sich hauptsächlich beziehen, beide Unternehmen sind historisch und arbeitstechnisch in vielfältiger Weise miteinander verbunden. Dennoch erfordern die *Regesta* ein stärker dezentrales Arbeiten als die *Monumenta*, gilt es doch die gesamte Überlie-

4 Vgl. <<http://www.dmgh.de>> (20.03.2006).

ferung zu jeweils einem Herrscher flächendeckend zu erfassen und zu verarbeiten. Informationen und Korrekturen etwa aus einzelnen Archiven und Archivregionen sind hier mehr als willkommen, und die Anlage etwa der Regesten für Friedrich III. zeigt deutlich, dass in quellenreicher Zeit die Archivarbeit von einer Person allein nicht mehr zu leisten ist, sondern auf viele Schultern verteilt werden muss. In diesem Sinne sind die *Regesta Imperii Online*, wie Andreas Kuczera (Mainz) betonte, ein disloziertes Projekt.⁵ In erster Linie stellt das neue elektronische Angebot daher den Bearbeitern/innen praktische Hilfsmittel zur Verfügung. So dient der auch von externen Nutzern/innen gerne konsultierte digitale Bibliothekskatalog (RI-OPAC) vornehmlich dazu, allen Bearbeitungsstandorten eine einheitliche Kurztitelfindung zu gewährleisten. Als nächsten Schritt des Projektes stellte Kuczera die Suche in kumulierten Registern in Aussicht, womit die Abfrage in den immensen Datenbeständen des Unternehmens erheblich vielfältiger vorgenommen werden kann. Ebenfalls angestrebt ist die Vernetzung der Registerinträge mit digitalen Abbildungen der jeweiligen Textstücke.

In den Bereich der Handschriften-Erschließung führten die beiden weiteren Vorträge der Sektion. Robert Giel (Berlin) verzichtete darauf, die *Manuscripta Mediaevalia* nochmals eingehend vorzustellen. Stattdessen nahm er eine Einordnung in das digitale Umfeld vor, die bemerkenswerterweise statt auf ein PowerPoint-Gewitter auf die bloße Kraft des Wortes vertraute. Momentan sind in dieser Handschriften-Datenbank 60.000 in Deutschland liegende Manuskripte des Mittelalters erfasst. Zusätzlich werden sowohl alte Handschriftenkataloge digitalisiert als auch neue Erfassungen vorgenommen. Aus diesem zweigeteilten Vorgehen resultiert eine unterschiedliche Tiefe in den jeweiligen Handschriften-Beschreibungen; die Standards haben sich im Laufe der Jahre immer weiter verfeinert.⁶ Die Bearbeiter/innen verfolgen nunmehr das Ziel, die vorhandenen Informationen sachspezifisch zu erweitern. Das heißt konkret: Es ist geplant, die Nutzer/innen von der gesuchten Handschrift auf bestehende

5 Vgl. <<http://www.regesta-imperii.org>> (20.03.2006).

6 Vgl. <<http://www.manuscripta-mediaevalia.de>> (20.03.2006).

Editionen zu verweisen und umgekehrt aus Bibliotheks-OPACs heraus eine direkte Verbindung zur Handschriften-Datenbank herzustellen. Betreibt man dies konsequent, so zeichnet sich am Horizont auch schon die Vernetzung der Manuskript-Einträge mit der Forschungsliteratur ab. Den *Manuscripta Mediaevalia* würde dies sicherlich neue Nutzer/innen zuführen, und auch dem Fach wäre es zu wünschen, dass jedermann gewissermaßen automatisch von der Literatur auch wieder auf die Quellen zurückverwiesen wird.

Den Schlusspunkt setzte Torsten Schaßan (Wolfenbüttel) mit einem Beitrag, der den fordernden Titel „Desiderata der Handschriftenkatalogisierung und -digitalisierung“ trug. Dabei führte er vielfach am praktischen Beispiel aus, was sein Vorredner auf der analytischen Ebene angeschnitten hatte. Mit einer fulminanten Mischung aus grundlegenden Fragen und kleinteiligen Lösungen illustrierte Schaßan sein Credo, dass Handschriftenkataloge niemals fertig werden! Das war sicherlich anders gemeint, als Benutzer/innen gedruckter und digitaler Kataloge es auf den ersten Blick vermuten, die die Lücken der Erschließung oft genau dort finden, wo sie sie im Moment am wenigsten gebrauchen können. Der Referent zielte vielmehr darauf ab, dass die statische Präsentation von Erschließungsdaten nicht mehr zeitgemäß ist, dass – wie bei einigen Quellengattungen – einer dynamischen Aufbereitung der Daten der Vorzug zu geben sei. Nur in dieser Form könne man Daten nachtragen, Wissen vermehren, die beschriebene Handschrift sukzessive in all ihren Facetten erschließen.

So sehr diese Forderung einleuchtet, so sehr macht sie auch die Probleme deutlich, die mit den immensen Möglichkeiten digitaler Bereitstellung einhergehen. Unfreiwillig illustrierte der letzte Beitrag der Sektion auch zentrale Gedanken, die Arno Mentzel-Reuters zu Beginn formuliert hatte: die Übersättigung mit Information, das Rauschen des Beiwerks, das den Alarmton des eigentlich Gesuchten zu ersticken droht. Und wie ist es bei aller Dynamik mit der Stabilität der Information, die dauerhaft zitierfähig bleiben muss, solange sich die wissenschaftlichen Gepflogenheiten im Umgang mit Informationsquellen nicht nachhaltig ändern? Der Verfasser dieser Zeilen war sich jedenfalls nicht immer darüber im Klaren, ob er das euphorische „noch mehr“ an Details wirklich zu den Desiderata rechnen

möchte.

Bei genauerem Hinsehen entpuppt sich der Satz „Handschriftenkataloge sind nie fertig“ ohnehin als preiswerte Binsenweisheit. Jede Quellenedition und jede wissenschaftliche Arbeit überhaupt kann und muss dies für sich in Anspruch nehmen. Es ist allein die Form der gedruckten Publikation, die zu einem Einschnitt zwingt, weil ein Manuskript irgendwann abgeschlossen werden muss. Hier ist die digitale Welt im doppelten Vorteil, flexibler und kostengünstiger zu sein. Aber der Einschnitt, zu dem der Druck eines Buches nötigt, sollte nicht nur als Fessel betrachtet werden. Mit ihm geht ein heilsamer Zwang einher, sich und den Lesern/innen Rechenschaft abzulegen über Notwendiges und weniger Wichtiges. Darin liegt im Übrigen eine nicht zu unterschätzende Leistung der Wissenschaftler/innen.

Damit sind wir bereits bei der Gesamtbetrachtung dessen angelangt, was .hist 2006 an spezifisch Mittelalterlichem zu bieten hatte. Beide Sektionen zu diesem Thema lohnten den Besuch, weil über die Vorstellung konkreter Projekte und technischer Lösungsmöglichkeiten hinaus Schwerpunkte und Grenzen des digitalen Arbeitens am historischen Objekt erkennbar wurden. Quellenerschließung, Quellenpräsentation und selbst unmittelbare Quellenforschung sind mithilfe der elektronischen Datenverarbeitung und den dadurch gewaltig gestiegenen Möglichkeiten des Vergleichs längst in eine neue Dimension eingetreten. Das Staunen des Beobachters ist hier berechtigt, ebenso der Stolz der „Macher“ auf das Erreichte und ihre Vorfreude auf neue Entwicklungen. Dennoch muss man die Perspektiven klarer formulieren. Sie können nicht allein in einem quantitativen Wachstum digitaler Informationsangebote liegen. Es ist das Verdienst der Sektion „Vernetztes Mittelalter“, diese Frage zumindest gestellt zu haben. Sie betrifft die verbesserte Abstimmung der Angebote im Netz und die Strukturen der Zusammenarbeit von Forschungseinrichtungen untereinander. Damit verbunden sind indessen Themen, die weit über das Technische hinausgehen, etwa die Vision einer dezentralen digitalen „Forschungsgemeinschaft“, zu der jeder seine Bausteine beitragen kann. Vor der Debatte der Machbarkeit ist aber die Frage nach dem Sinn zu erörtern.

Auch scheint insgesamt ein wenig aus dem Blick geraten zu sein, wem man die Informationen bereitstellen will, welche Nutzer/innen mit den unterschiedlichen Angeboten eigentlich angesprochen werden sollen: Forscher/innen oder interessierte Laien? Oder beide? Für die Konzeption der digitalen Arbeit kann das nicht gleichgültig sein, denn was für den einen willkommene Instruktion ist, empfindet der andere möglicherweise als das genannte „Rauschen“. Die Zuhörer/innen konnten sich bisweilen des Eindrucks nicht erwehren, dass der Trend eher dahin geht, immer komplexere Datenbanken und Abfragemöglichkeiten zu schaffen, damit jede/r mit dem digital bevorrateten historischen Material nach Belieben verfahren kann. So faszinierend dies ist – auch weil es letztlich den umfassend ausgebildeten Historiker bzw. die Historikerin als Benutzer/in erfordert! –, so unwirtschaftlich erscheint diese Methode der grenzenlosen Aufbereitung für den eventuellen Fall, dass irgendjemand irgendwann irgendeine Frage stellt. Das Vorgehen, den Nutzern/innen endlose Informationsketten per Knopfdruck verfügbar zu machen, dürfte zudem die elektronische Datenverarbeitung, die sich mühsam zur angesehenen „historischen Fachinformatik“ gemausert hat, auf lange Sicht zur Hilfswissenschaft im engsten denkbaren Sinn machen. Ob sich dies mit dem Selbstverständnis der aktuellen Digital-Historiker/innen verträgt?

Dr. Harald Müller ist Privatdozent für Mittelalterliche Geschichte an der Humboldt-Universität zu Berlin. Seine Forschungsschwerpunkte liegen im Bereich der Kirchengeschichte und des Kirchenrechts sowie des Renaissance-Humanismus. Zurzeit vertritt er an der Universität Leipzig die Professur für Historische Hilfswissenschaften / Archivwissenschaft. Er ist seit 2001 Review Editor für das Mittelalter bei H-Soz-u-Kult. E-Mail: muellerh@geschichte.hu-berlin.de

THEMENPORTALE ZWISCHEN BIBLIOTHEK UND FACHWISSENSCHAFT*

von *Stephanie Marra*

Dass die personelle und finanzielle Verknappung von Ressourcen seit geraumer Zeit nicht nur den akademischen Betrieb in Deutschland zu Reformen drängt und zu einer Neupositionierung führt, sondern in einem verstärkten Maße auch die außeruniversitären Einrichtungen sowie vor allem Bibliotheken, Archive und Museen erfasst, machten auch die Sektionen der Fachtagung .hist 2006 – Geschichte im Netz deutlich.

„Wissensmanagement“, „Informationsvermittlung“, „Teaching Library“ – diese oder ähnlich lautende Schlagworte schreiben sich wissenschaftliche Bibliotheken gerne auf ihre Fahnen, wenn es darum geht, mit ihren vermeintlichen Imageproblemen aufzuräumen. Bibliotheken präsentieren sich seit jeher zu Recht als moderne, dienstleistungsorientierte Betriebe, die unmittelbar und nachhaltig auf technische Entwicklungen und neue Organisationsformen reagieren. Bibliothekare/innen entwickeln sich immer mehr hin zu „Wissensmanagern/innen“, die konventionelle und elektronische Medien, vielfältige Fachinformationen sowie umfangreiche Angebote aus den verschiedenen Fachdisziplinen überblicken, auswählen und zur Verfügung stellen. Angesichts rückläufiger Erwerbungssetats und zeitgleicher Preissteigerungen von Verlagserzeugnissen müssen jedoch zwangsläufig neue Wege in der Literatur- und Medienversorgung beschritten werden. Bibliothekarische Konsortien, die mit dem Gemeinschaftserwerb

* Bericht zur Sektion: „Die Sondersammelgebietsbibliotheken und ihre virtuellen Fachbibliotheken als Dienstleister für die historische Forschung“ sowie „Historische Fach- und Themenportale: Forschen, Lehren, Kommunizieren, Vermitteln, Publizieren“. Der Querschnittsbericht erschien am 22.05.2006 auf H-Soz-u-Kult: <<http://hsozkult.geschichte.hu-berlin.de/forum/id=719&type=diskussionen>>.

von elektronischen Fachzeitschriften und Fachdatenbanken zur Kostenreduktion in ihren jeweiligen Häusern beitragen, werden seit einigen Jahren ergänzt durch Bemühungen der wissenschaftlichen Bibliotheken, mit Hilfe von Nationallizenzen¹, Pay-Per-View-Modellen² und Pay-Per-Use-Angeboten³ die Verteuerungen von elektronischen Fachzeitschriften und Fachdatenbanken aufzufangen. Allein der Kommunikationsprozess mit den wissenschaftlichen Nutzergruppen scheint bei all diesen Aktivitäten etwas auf der Strecke zu bleiben. Während innovative Projekte und neue Entwicklungen in der Informationsversorgung offenbar nicht ausreichend in die Fachwelt (und erst Recht nicht in eine breite Öffentlichkeit) kommuniziert werden, scheinen dagegen kommerzielle Großprojekte, wie zum Beispiel Google Scholar⁴, eine breite Rezeption zu erfahren. Diese Diskrepanz ist keinesfalls eine neue Erscheinung, sondern begleitet den Prozess der Einführung von digitalen Medien seit seinem Beginn.

Das derzeitige und künftige Zusammenspiel von Bibliotheken und Wissenschaft stand im Mittelpunkt der Sektion „TS 3: Die Sondersammelgebietsbibliotheken und ihre virtuellen Fachbibliotheken als Dienstleister für die historische Forschung“. Innerhalb des durch die Deutsche Forschungsgemeinschaft (DFG) organisierten Systems der überregionalen Literaturversorgung kommt seit 1949 mehr als 30 Staats-, Universitäts-

1 Seit 2004 finanziert die DFG den Erwerb von Nationallizenzen, die den kostenlosen Zugang zu Datenbanken, digitalen Textsammlungen und elektronischen Zeitschriften für die deutschen Hochschulen ermöglicht. Bei den elektronischen Zeitschriften handelt es sich zumeist um so genannte Backfiles, also Jahrgänge, die mindestens zwei Jahre zurückliegen.

2 Das Pay-Per-View-Verfahren an Bibliotheken umfasst den kostenpflichtigen Bezug einzelner Volltextartikel aus wissenschaftlichen Zeitschriften.

3 Im Rahmen von Pay-Per-Use soll auch denjenigen Interessenten/innen der Zugang zu elektronischen Zeitschriften und Datenbanken ermöglicht werden, die keine eingetragenen Nutzer/innen einer zugangsberechtigten Bibliothek sind oder deren Bibliothek keine entsprechenden Lizenzen erworben hat. Individuellen Nutzern/innen wird für einen definierten Zeitraum der Zugang zu diesen Angeboten ermöglicht und anschließend ein Nutzungsentgelt in Rechnung gestellt.

4 Google Scholar, seit Mitte November 2004 als englischsprachige spezielle Suchmaschine für das Auffinden wissenschaftlicher Literatur online, findet sich unter <<http://scholar.google.com>> (23.03.2006).

und Spezialbibliotheken die Einrichtung fachspezifischer Sondersammelgebiete (SSG) zu. Ziel der jeweiligen Sammlungsschwerpunkte ist die weitestgehend vollständige Erwerbung und Verfügbarkeit der entsprechenden Fachliteratur des In- und Auslandes. Seit der Vereinigung der beiden deutschen Teilstaaten vor über 15 Jahren sind allerdings, bedingt durch die vergrößerte Wissenschafts- und Bibliothekslandschaft, auch Wandlungen innerhalb des SSG-Systems spürbar. Die Ausdifferenzierung der DFG-geförderten Sondersammelgebiete und die Fokussierung auf Spezialsammlungen bildeten sich ab Mitte der 1990er Jahre zeitgleich mit dem Aufbau so genannter Hybrider Bibliotheken aus, in denen konventionelle und digitale Medien gleichberechtigt nebeneinander stehen. Mit der Realisierung Virtueller Fachbibliotheken (ViFa), die von der DFG seit Ende der 1990er Jahre gesonderte Förderungen erhalten, kommt den Sondersammelgebietsbibliotheken damit zusätzlich die Aufgabe zu, disparate digitale Fachinformationen, erschlossen durch internationale bibliothekarische Standards, jeweils in einem Zentralportal zusammenzuführen. Neben der Einbindung genuin bibliothekarischer Kernmodule, wie relevanter OPACs, Datenbanken sowie Literatur-, Current Content- und Dokumentenlieferdiensten, finden sich in den weitestgehend gleichförmig aufgebauten ViFas auch wissenschaftliche Elemente, wie unter anderem die genannten qualitativ erschlossenen Fachinformationsführer, fachliche Suchmaschinen, Tutorials oder Volltexte.

In seiner Einführung umriss Klaus Kempf (Bayerische Staatsbibliothek [BSB] München) die Zielsetzung des Panels. Neben den Erfahrungsberichten exemplarisch vorgestellter ViFas sollten künftige Modelle und technische Entwicklungsmöglichkeiten in der geschichtswissenschaftlichen Literaturversorgung zur Sprache kommen. Diese neuen Anforderungen, denen sich allerdings nicht ausschließlich nur die fachzentrierten Sondersammelgebietsbibliotheken und deren ViFas zu stellen haben, thematisierte Ralf Goebel (DFG) in seinem Vortrag über allgemeine „Entwicklungen und Perspektiven der überregionalen Bibliotheksversorgung“.

Zu den Aufgaben, die künftig verstärkt an die Sondersammelgebietsbibliotheken und ihre ViFas heranzutragen sind, gehören neben einer schärferen Profilierung und wissenschaftlichen Vernetzung angesichts der

zunehmenden internationalen und kommerziellen Konkurrenz vor allem auch regelmäßig durchgeführte Evaluationen zur Messung der Nutzerzufriedenheit, Retrodigitalisierungen, der verstärkte Erwerb campusfreier Nationallizenzen und das Outsourcing spezieller Dienstleistungen.⁵ Gerade jedoch die letztgenannte Anforderung an die Sondersammelgebietsbibliotheken, sich künftig ebenso als überregionale Dienstleisterinnen zu profilieren, ist nicht frei von Problemen. Hier wird beispielsweise durch ein (kostenpflichtiges) Angebot der Auswahl, Beschaffung, Bearbeitung und Katalogisierung von Literatur bestimmter Fächer durch die SSG-Bibliotheken an den Kernkompetenzen aller übrigen wissenschaftlichen Bibliotheken, die nicht DFG- gefördert sind, gerüttelt. Ob es sich hierbei tatsächlich um eine von den Hochschulen und ihren Hochschulbibliotheken gewünschte „Optimierung des Personaleinsatzes“ handelt? Um den Bekanntheitsgrad der Sondersammelgebiete und ihres Leistungsspektrums innerhalb der Fachwissenschaften mehr als bisher zu steigern, erscheinen deshalb zunächst die von Goebel angeführten Aspekte der „Profilierung“, „Vernetzung“ und „Nutzungsanalyse“ dringlicher.

Im Anschluss an diesen programmatischen Ausblick wurden die Konzepte und inhaltlichen Ausrichtungen einzelner, bereits seit Jahren bestehender Virtueller Fachbibliotheken wie Vlib-AAC, Cibera und ViFa Ost⁶

5 Zur Aufgabenprofilierung der Sondersammelgebietsbibliotheken vgl. die Empfehlungen des Unterausschusses für Überregionale Literaturversorgung und des Bibliotheksausschusses von Juni 2004: Das DFG-System der überregionalen Sammelschwerpunkte im Wandel – Weitere Schritte zur Umsetzung des Memorandums zur Weiterentwicklung der überregionalen Literaturversorgung, S. 14, <http://www.dfg.de/forschungsfoerderung/wissenschaftliche_infrastruktur/lis/download/strategiepapier_ueberreg_lit_versorgung.pdf> (23.03.2006). Zur Analyse der Nutzerzufriedenheit des DFG-geförderten Informationsportals *vascoda* vgl. die im Jahr 2005 von der Universitäts- und Landesbibliothek Münster im Auftrag des BMBF durchgeführten „Evaluation von *vascoda* aus Nutzerperspektive“. Erste Teilergebnisse liegen vor: Gediga, Günther; Gildehorn, Antje; Cover, Britta: Evaluation von *vascoda* aus Benutzersicht. Ergebnisse der Nutzerbefragung 2005, Münster u.a. 2005, <http://www.dl-forum.de/dateien/Evaluation_vascoda_Ergebnisse_Befragung_2005.pdf> (23.03.2006). Siehe ergänzend dazu die Ergebnisse der Fokusgruppen-Interviews, <<http://www.ulb.uni-muenster.de/projekte/vascoda/vascoda-erfahrungen.pdf>> (23.03.2006).

6 Vlib-AAC. Virtual Library of Anglo-American Culture (SUB Göttingen)

vorgestellt. Im gelungenen Wechsel mit Vorträgen über spezialisierte Dienstleistungen für die künftige Informationsversorgung der BSB München als Sondersammelgebietsbibliothek für die Geschichtswissenschaften, zeigten die vorgestellten Fallbeispiele auch die Synergieeffekte, die sich aus einer gelungenen Zusammenarbeit von Bibliotheken und Wissenschaften ergeben können. Als eine der frühesten virtuellen Fachbibliotheken stellte Wilfried Enderle (Staats- und Universitätsbibliothek Göttingen) unter dem Titel „300.000 digitale Bücher für die Geschichtswissenschaft“ die Entwicklung der Vlib-AAC vor. Neben den besagten 300.000 Büchern, deren Summe sich aus der Anzahl der im Portal verfügbaren kommerziellen digitalen Textkorpora, thematischen Websites, subskribierten Zeitschriften und Mikroformen ergibt, wurden außerdem Nationallizenzen und Pay-Per-View-Angebote in das Fachportal aufgenommen. Angesichts der noch immer weit verbreiteten Haltung von Historikern/innen, Digitalisate und Online-Angebote für die Forschung und Lehre einzufordern und auch zu nutzen, weiterhin jedoch in möglichst renommierten Verlagshäusern zu publizieren⁷, plädierte Enderle abschließend für eine Vereinfachung der Ressourcenbereitstellung.

An diese Forderung knüpfte auch der Vortrag von Gregor Horstkemper (BSB München) an, der neue Konzepte der überregionalen Informationsversorgung für die Geschichtswissenschaften vorstellte. Neben dem Angebotsspektrum geschichtswissenschaftlicher Nationallizenzen präsentierte der Referent die im Herbst 2004 freigeschalteten Pay-Per-Use-

<<http://www.sub.uni-goettingen.de/vlib/history/>>(23.03.2006). Das interdisziplinäre Portal *cibera*: Virtuelle Fachbibliothek Ibero-Amerika / Spanien / Portugal ist seit Mitte November 2004 online, <<http://www.cibera.de/>> (23.03.2006). Zu den weiteren Inhalten von *cibera* siehe Karl, Annette; Mühlischlegel, Ulrike; Ullrich, Ralf: *cibera*: Virtuelle Fachbibliothek Ibero-Amerika / Spanien / Portugal, in: Bibliotheksdienst 39 (2005), S. 1588-1596, <http://www.zlb.de/aktivitaeten/bd_neu/heftinhalte2005/DigitaleBib1205.pdf> (23.03.2006). Auch die Virtuelle Fachbibliothek Osteuropa (ViFa Ost), die seit 2003 zugänglich ist, versteht sich als „fachübergreifendes Regionalportal“ zur Osteuropaforschung, <<http://www.vifaost.de/>> (23.03.2006).

7 Vgl. Blaschke, Olaf, Reputation durch Publikation. Wie finden deutsche Historiker ihre Verlage? Eine Umfrage, in: Geschichte in Wissenschaft und Unterricht 55 (2004), S. 598-620.

Angebote der BSB.⁸ Die technische Basis der Plattform bildet die Software „Hidden Automatic Navigator“ (HAN), die es registrierten Nutzern/innen ermöglicht, auch extern kostenfrei auf die Angebote der BSB zuzugreifen. Nicht registrierte Nutzer/innen nutzen den Pay-Per-Use-Zugang, der ihnen für ein geringes Entgelt die zeitlich befristete Nutzung der elektronischen Zeitschriften, Volltextsammlungen, Nachschlagewerke, CD-ROMs und Online-Datenbanken ermöglicht. Beide Angebotstypen sollen künftig in die Virtuellen Fachbibliotheken der BSB integriert werden. Eine bereits durchgeführte Erprobung des Pay-Per-Use-Modells innerhalb des geschichtswissenschaftlichen Portals Chronicon⁹ zeigte allerdings bereits, dass künftig individuelle Lösungsansätze notwendig sind, um den reibungslosen Ablauf des Pay-Per-Use-Zugangs zu gewährleisten. Nach einer fast einjährigen Erprobungszeit spiegelt die vergleichsweise geringe Nutzung der Nationallizenzen und des Pay-Per-Use-Modells in Chronicon die eingangs getroffene Feststellung eines möglichen Kommunikationsdefizits zwischen den Bibliotheken und den (Geschichts-)Wissenschaften, das es dringend aufzuheben gilt.

Den fachwissenschaftlichen Nutzern/innen stehen eventuell interdisziplinäre, zumeist nicht genuin bibliothekarisch durchstrukturierte und klassifizierte Module näher, wie beispielsweise das elektronische Pressearchiv „IberoDigital“ der ViFa Cibera, die Annette Karl (Ibero-Amerikanisches Institut Berlin) vorstellte. Retrodigitalisierte bzw. originäre wissenschaftliche Volltexte (Magisterarbeiten, Handbücher, Lexika und Jahrbücher zur Geschichte Osteuropas) und fachwissenschaftliche Informationsangebote (Forschungsdatenbank, Personenverzeichnis, Vorlesungsverzeichnisse, Veranstaltungskalender, usw.) werden dann auch künftig das Angebotsspektrum der seit 2003 verfügbaren ViFa Ost erweitern, die Alessandra Sorbello Staub (BSB München) in Vertretung für ihre erkrankte Kollegin Gudrun Wirtz (BSB München) und Olivia Griesche (Ludwig-Maximilian-Universität München) präsentierten. Im Fall von

8 Verfügbare Pay-Per-Use-Datenbanken der BSB München, <<http://www.bsb-muenchen.de/datenb/ppupreise.htm>> (23.03.2006).

9 Chronicon. Fachportal für Geschichtswissenschaften, <<http://www.chronicon.de>> (23.03.2006).

ViFa Ost haben eine kontinuierliche Öffentlichkeitsarbeit und das enge partnerschaftliche Verhältnis zwischen der Bibliothek und den beteiligten universitären und außeruniversitären Forschungseinrichtungen einen wesentlichen Anteil an dem Erfolg und der breiten Akzeptanz des Portals.

Als Eigenentwicklung der BSB stellte Alessandra Sorbello Staub das bereits zuvor erwähnte geschichtswissenschaftliche Informationssystem Chronicon vor, das unter einer Suchoberfläche alle bibliothekarischen und fachlichen Dienstleistungen, die an der BSB für die allgemeine und europäische Geschichtswissenschaft entwickelt wurden, vereinigt. Den Aufbau und die künftige Weiterführung von Compact Memory¹⁰, dem Fachportal für Jüdische Zeitschriften und Zeitungen bis 1938 im deutschsprachigen Raum, präsentierte Rachel Heuberger (Universitätsbibliothek [UB] Frankfurt am Main). Über eine ausdifferenzierte Rechercheoberfläche erschließen sich die zur Verfügung stehenden Digitalisate von derzeit über 100 deutschsprachigen Periodika, denen in einem nächsten Arbeitsschritt Periodika in hebräischer und jiddischer Sprache folgen sollen. Neben der langfristigen Bestandserhaltung und der Bereitstellung wichtiger Bestände für Forschung und Lehre trägt das vorgestellte Zeitschriftenarchiv dazu bei, die spezialisierte Literaturversorgung an der UB Frankfurt am Main als Schwerpunktbibliothek für die Fachgebiete Wissenschaft von Judentum und Israel zu optimieren. Die einzelnen Beiträge der Sektion haben in ihrer Gesamtheit deutlich gemacht, dass die Funktion der Bibliotheken auch in der Bereitstellung von Online-gestützten Informations- und Nutzerdiensten für die Fachwissenschaften ungebrochen wichtig ist.

Folgerichtig bietet sich ein Vergleich mit den Erfahrungswerten und

10 Das Angebot Compact Memory – Wissenschaftsportal für Jüdische Studien ist ein Digitalisierungsprojekt des Lehrgebiets Deutsch-jüdische Literaturgeschichte der RWTH Aachen, des SSG Judentum an der Universitätsbibliothek Johann Christian Senckenberg Frankfurt am Main und der Bibliothek Germania Judaica Köln, <<http://www.compactmemory.de>> (23.3.2006). Siehe hierzu auch Schicketanz, Till; Heiligenhaus, Kai: „Inseln im Meer des Beliebigen“. Architektur und Implementierung eines Internetportals Deutsch-Jüdischer Periodika, in: Jahrbuch für Computerphilologie 5 (2003), S. 67-96, <<http://computerphilologie.uni-muenchen.de/jg03/schicketanz-heiligenhaus.html>> (23.03.2006).

Perspektiven von geschichtswissenschaftlichen Fachportalen universitärer bzw. außeruniversitärer Provenienz, die weitestgehend mit den Sonder-sammelgebiets- und Spezialbibliotheken kooperieren, an. In der Sektion TS 8 „Historische Fach- und Themenportale: Forschen, Lehren, Kommunizieren, Vermitteln, Publizieren“ fanden sich dann neben den existierenden und geplanten Fachportalen zur überwiegend zeithistorischen Forschung auch „alte“ und „neue“ Bekannte wie die BSB, die ViFa Ost oder die ViFa Slavistik¹¹ ein, deren Vertreter/innen nun die fachwissenschaftlichen Anteile ihrer Virtuellen Bibliotheken präsentierten. Die eingangs von Sektionsleiter Jürgen Danyel (Zentrum für Zeithistorische Forschung, Potsdam) formulierten Erfahrungen und Perspektiven zeithistorischer Themenangebote im Internet machten allerdings schnell deutlich, dass Fachportale zur Neuere und Neuesten Geschichte wie das von Danyel vorgestellte Angebot Zeitgeschichte-online¹² mit nahezu denselben Problemen wie die zuvor thematisierten Virtuellen Fachbibliotheken zu kämpfen haben.

Trotz des ungebrochen starken öffentlichen Interesses an zeithistorischen Themen sowohl in konventionellen Veröffentlichungen wie in den Digitalen Medien konkurriert die Geschichtswissenschaft auf diesem Feld mit populärwissenschaftlichen und kommerziellen Anbietern um die Gunst der breiten Öffentlichkeit. Auch hier bietet eine verstärkte Vernetzung mit weiteren zeithistorischen Fachportalen, eine entsprechende Profilierung und Öffentlichkeitsarbeit, zu der auch die unmittelbare Umsetzung zeitaktueller Themenbereiche und „Forschungstrends“ gehört, ein probates Mittel, wissenschaftliche und außerwissenschaftliche Nutzergruppen langfristig zu binden. Eine explizit akademische Nutzerschicht wird hinge-

11 Die Virtuelle Fachbibliothek Slavistik (ViFa Slavistik) wird an der Staatsbibliothek zu Berlin in Kooperation mit der UB Bochum betreut. Das Slavistik-Portal, das künftig unter der Adresse <<http://www.slavistik-portal.de>> erreichbar sein wird, befindet sich derzeit noch im Aufbau.

12 Zeitgeschichte-online. Fachportal für die zeithistorische Forschung, ein Kooperationsprojekt des ZZf Potsdam und der Staatsbibliothek zu Berlin, ist als Modul der ViFa Geschichte, zu der gleichermaßen die Angebote Clio-online und Historicum.net gehören, konzipiert, <<http://zeitgeschichte-online.de>> (23.03.2006).

gen mit dem von Heidi Hein (Herder-Institut Marburg) vorgestellten Editionsprojekt „Dokumente und Materialien zur ostmitteleuropäischen Geschichte“¹³ anvisiert. Das noch im Aufbau begriffene Dokumentenportal wird thematisch konzipierte Module (Quellen, Fachinformationen, Unterrichtsmaterial) für Lehre und Studium beinhalten. Ob regelmäßig durchgeführte Evaluationen und die wissenschaftliche Betreuung der Module durch ein ausgewiesenes Fachkollegium die konstante Nutzerzufriedenheit gewährleisten können, wird sich erweisen. Aus der Werkstatt eines ebenfalls im Aufbau befindlichen Themenportals zur Europäischen Geschichte in der Moderne, das Bestandteil von Clio-online werden soll, berichtete Rüdiger Hohls (Humboldt-Universität zu Berlin). Die Erweiterungen schulischer und universitärer Curricula um Module europäischer Geschichte, die Einrichtung neuer Studiengänge und -abschlüsse sowie die Etablierung bzw. Umwidmung entsprechender Forschungseinrichtungen und Lehrstühle sind erste Indizien dafür, dass ein solches Themenportal nicht nur eine wichtige Anlaufstelle für Schüler/innen, Studierende, Lehrende und Wissenschaftler/innen werden wird, sondern sich auch dem Interesse der Öffentlichkeit gewiss sein kann. Aber auch dieses Themenportal wird nicht umhin kommen, sich mit Kooperationsprojekten und -partnern zu vernetzen, sich entsprechend thematisch zu zentrieren und die verschiedenen Nutzerschichten kontinuierlich anzusprechen. Michael Krölls (Leopold Franzens-Universität Innsbruck) abschließende, instruktive Analyse exemplarischer Webportale zur deutschsprachigen Zeitgeschichte brachte alle zuvor in beiden Sektionen immer wieder vorgebrachten Probleme und Perspektiven geschichtswissenschaftlicher Online-Angebote auf den Punkt. Kröll konnte nachvollziehbar darlegen, dass vor allem die Nutzer/innen von Vernetzungen der Portale untereinander, von inhaltlichen Abstimmungen und erweiterten fachlichen Kooperationen vermehrt profitieren können.

Allerdings bleibt abzuwarten, ob die in beiden Sektionen aufgezeigten

13 Als erster Teil des Editionsprojekts „Dokumente und Materialien zur ostmitteleuropäischen Geschichte“ steht bereits das Modul „Zweite Polnische Republik“ zur Verfügung, <<http://quellen.herder-institut.de/>> (23.03.2006).

Entwicklungen und Vorhaben in Zeiten knapper Finanzmittel langfristig realisierbar sind. Unabhängig von der Planungssicherheit digitaler Angebote sind hingegen die Aufgaben einer kommunikativen Vermittlung der Inhalte und einer stärkeren Öffentlichkeitsarbeit. Denn schließlich geht es bei allen institutionell angebundenen und von der öffentlichen Hand finanzierten Datenbanken, Virtuellen Fachbibliotheken, Themenportalen und Editionsprojekten um einen höchstmöglichen Nutzwert für die Zielgruppen.

Dr. Stephanie Marra ist Fachreferentin an der Universitätsbibliothek Dortmund und Lehrbeauftragte für die Geschichte der Frühen Neuzeit an den Historischen Instituten der Universitäten Dortmund und Bochum. Ihre Arbeitsschwerpunkte sind die Sozialgeschichte des frühneuzeitlichen Adels, die Sammlungsgeschichte von Bibliotheken und Museen sowie die Nutzung von online-gestützten Medien in Lehre und Forschung. Sie ist Redakteurin der Fachinformationsliste H-Museum und seit 1994 Herausgeberin von mehreren Online-Angeboten. E-Mail: stephanie.marra@ub.uni-dortmund.de

KOLLABORATIVES SCHREIBEN, LEHREN UND LERNEN*

von Daniel Burckhardt

Kein anderer Begriff wurde im Februar 2006 von deutschen Internet-Nutzern/innen bei Google so oft nachgefragt wie „Wikipedia“.¹ Eine breite Öffentlichkeit hat inzwischen die freie Online-Enzyklopädie als nützliche Anlaufstelle bei Recherchen im Netz entdeckt. Entsprechend präsent war die Wikipedia auch auf der Tagung .hist 2006. In verschiedenen Sektionen und Einzelgesprächen wurde wiederholt die Frage nach Verlässlichkeit und Zitierbarkeit der im ständigen Wandel befindlichen Wikipedia-Artikel gestellt. In der von Peter Haber geleiteten Sektion „Collaboratories. Das Schreiben der Geschichte im vernetzten Zeitalter“² standen aber weniger die Inhalte dieses Erfolgsprojekts kollektiven Schreibens, das in nur fünf Jahren auf rund 3,5 Millionen Artikel in über 200 Sprachen angewachsen ist, im Vordergrund, als vielmehr der Produktionsprozess sowie Theorie und Praxis des Hypertexts.

Entgegen dem generellen Trend im Webdesign, über eine weitgehend hierarchische Navigation die Zahl der Klicks zu nur schwach miteinander verknüpften Einzeltexten zu minimieren, feiert mit der Wikipedia der durch eine Vielzahl von Links rhizom-artig strukturierte Hypertext eine Renaissance.³ Ebenso wird die Frage nach Autorschaft erneut problemati-

* Bericht zur Sektion: „Collaboratories. Das Schreiben der Geschichte im vernetzten Zeitalter“ sowie „Entwicklung von E-Learning-Projekten als kollaborativer Forschungsprozess“. Der Querschnittsbericht erschien am 20.05.2006 auf H-Soz-u-Kult: <<http://hsozkult.geschichte.hu-berlin.de/forum/id=724&type=diskussionen>>.

1 Vgl. <<http://www.google.com/press/intl-zeitgeist.html#de>> (12.04.2006).

2 Zum Programm dieser Sektion siehe <http://www.clio-online.de/hist2006/userinfoviewsession.php4?session_id=21> (12.04.2006).

3 Siehe dazu auch Jakob Nielsens Alertbox-Kolumne vom 03.04.2006, <<http://www.useit.com/alertbox/hype.html>> (12.04.2006).

siert, die nicht nur für das Gesamtwerk, sondern auch für einzelne Artikel, die im Laufe der Zeit von verschiedenen Personen weiter geschrieben werden, kaum noch zu bestimmen ist. Wurden solche Entwicklungen – wie Jakob Krameritsch gegen Ende der Sektion ausführte – in der ersten Hochzeit des Hypertexts in den 1990er Jahren von der Geistes- und Kulturwissenschaft emphatisch diskutiert, kehrte mangels lange ausbleibender, erfolgreicher Umsetzungsversuche zunächst eine gewisse Ernüchterung ein. Möglicherweise erschien auch vielen durch Postmoderne und Dekonstruktivismus geprägten Geisteswissenschaftlern/innen das Wikipedia-Leitmotiv eines neutralen Standpunkts („neutral point of view“, kurz NPOV) als antiquiert. So steht die Wikipedia unter latentem Positivismus-Verdacht, was viele daran hinderte, in ihr eine technische und inhaltliche Realisierung der ein Jahrzehnt zuvor geäußerten Vorstellungen zu sehen.

Jakob Voß, in der deutschen Sektion der Wikipedia sehr aktiv, stellte die Gründe für den gegenwärtigen Erfolg der Wikipedia im günstigen Zusammenwirken verschiedener Faktoren dar: Der Wiki-Technologie und -Software, einer praktikablen Lizenz für die Inhalte, die die flexible Nachnutzung der Texte ermöglicht, einem einfachen gemeinsamen Projektziel, das im Aufbau einer möglichst umfangreichen Enzyklopädie gemäß dem NPOV besteht, sowie einer aktiven, selbstorganisierten *Community*, die Regeln für die Schlichtung von Meinungsverschiedenheiten und Konflikten entwickelt hat. Gemäß „Lotkas Gesetz“, wonach viele Autoren/innen nur einmal, ein paar wenige dagegen sehr oft publizieren⁴, sind auch bei der Wikipedia weniger als 10 Prozent der registrierten Nutzer/innen für über 90 Prozent der Bearbeitungen verantwortlich. Diese Gruppe der sehr aktiven Nutzer/innen ist überwiegend jung (20-30), gut gebildet und männlich (86 Prozent). Aus ihrem Kreis rekrutieren sich auch die im deutschen Sprachraum rund 200 Administratoren/innen, die für Sperrungen und Löschungen verantwortlich sind.

Zwei der soeben erwähnten Erfolgsfaktoren, die Wiki-Technologie sowie das Projektziel einer Enzyklopädie, in diesem Fall ein „Biogra-

4 „Die Anzahl der Personen, die n Aufsätze schreiben, ist proportional zu $1/n^2$ “, <<http://www.ib.hu-berlin.de/~wumsta/umdefa.html>> (12.04.2006).

phisch-bibliographisches Handbuch für das katholische Deutschland“, wurden von der Kommission für Zeitgeschichte in Bonn mit zwei Projektpartnern aus dem technischen Bereich erfolgreich in den Antrag für das WIKINGER-Projekt (Wiki Next Generation Enhanced Repository) im Rahmen der deutschen „E-Science“-Initiative aufgenommen.⁵ Wie Karl-Joseph Hummel ausführte, sollen die biografischen Einträge für ein „Who is Who“ des deutschen Katholizismus jedoch nicht von interessierten Nutzern/innen der Site stammen, sondern von ausgewählten Fachwissenschaftlern/innen erstellt werden. Als technische Innovation soll die Liste der im Handbuch berücksichtigten Personen durch linguistische Verfahren weitgehend automatisiert aus bestehenden Veröffentlichungen extrahiert werden. Es wird sich in den nächsten Jahren zeigen, ob diese *Top-Down*-Organisation im Rahmen eines Wissenschaftsprojektes – der Projektmitarbeiter stand das ganze Referat hindurch schweigend neben seinem Vorgesetzten – oder der *Bottom-Up*-Ansatz der Wikipedia bessere Ergebnisse produziert. Zu befürchten ist jedoch, dass die Lizenz der im Rahmen des WIKINGER-Projektes erstellten Biografien nicht ohne weiteres eine Weiternutzung in der Wikipedia erlauben wird.

Wie fremd für viele Fachwissenschaftler/innen die Vorstellung einer Arbeitsplattform ohne Zugangsbeschränkungen ist, zeigte sich auch in der Präsentation von Gisela Minn zum Forschungsnetzwerk und Datenbanksystem für den Trierer SFB 600 „Fremdheit und Armut – Wandel von Inklusions- und Exklusionsformen von der Antike bis zur Gegenwart“. Auf beeindruckende Weise wurden hier die Wünsche und Anforderungen einer fächerübergreifenden Forschergruppe analysiert und als netzbasiertes EDV-System implementiert. Die einzelnen Mitarbeiterinnen und Mitarbeiter stellen die von ihnen bearbeiteten Texte, Bilder, Töne und Objektinformationen, etwa zu einer antiken Münze, in eine gemeinsame Datenbank ein. Im Gegensatz zur Offenheit der Wikipedia herrscht hier aber das Grundprinzip, wonach derjenige, der die Daten ins System einträgt, allein über die Sichtbarkeit dieser Daten für andere Personen und Gruppen entscheiden kann. Neben dem Wunsch nach Exklusivität vor Ver-

5 Vgl. <<http://www.kfzg.de/Forschung/WIKINGER/wikinger.html>> (12.04.2006).

öffentlichung können auch Urheberrechte und – besonders bei neueren Akten – Persönlichkeitsrechte einer generellen Freigabe im Wege stehen.

Naturwissenschaftler/innen arbeiten in „Laboratories“ schon seit Jahrzehnten in größeren Gruppen zusammen. Dies spiegelt sich in einer Vielzahl von Autorennennungen bei den meisten Veröffentlichungen. In den Geschichtswissenschaften ist dagegen die mono-auktoriale Erzählung immer noch der Regelfall. Jan Hodels Skizze des „hist.collaboratory“⁶ stellte deshalb weniger ein konkretes Projekt vor, als eine Reihe von Überlegungen zu einem gemeinsam genutzten virtuellen Schreibtisch für die Geisteswissenschaften. Dabei machte er auf eine entscheidende Eigenschaft der Wiki-Systeme aufmerksam. In den Worten von Jim Wales, dem Gründer der Wikipedia: „The basic thing I think makes it work is turning from a model of permissions to a model of accountability. It isn't that you are allowed or not allowed to edit a certain thing; it's when you do it, that change is recorded, and if it's bad, people can see that.“⁷ Genau diese minutiöse Buchführung der Wiki-Systeme über sämtliche Veränderungen könnte ein wichtiger Faktor für die Akzeptanz von Kollektiv-Autorschaften in den Geisteswissenschaften werden. Wenn der Anteil eines jeden Autors und einer jeden Autorin an einem Text stets transparent ist, fällt vielleicht auch der Verzicht auf die alleinige Autorschaft nicht mehr so schwer.

Content Management Systeme (CMS) sind Team-Tools par excellence. Dieser Satz aus dem Eröffnungsreferat von Wolfgang Schmale passte besonders gut auf die von Madeleine Herren geleitete Sektion zur „Entwicklung von E-Learning-Projekten als kollaborativer Forschungsprozess“.⁸ Denn viele E-Learning-Plattformen sind nichts anderes, als auf den Unterricht spezialisierte CMS, allenfalls ergänzt um Kommunikationsfunktionen zwischen den Studierenden und den Lehrenden. Allerdings wurden nicht alle vorgestellten Projekte dem im Titel formulierten Anspruch gerecht. Oder wie einer der Tagungsteilnehmer in der Diskussion meinte:

6 Vgl. <<http://wiki.histnet.ch/index.php/HistCollaboratory>> (12.04.2006).

7 Vgl. <<http://news.com.com/2009-1025-5944453-3.html>> (12.04.2006).

8 Zum Programm dieser Sektion siehe <http://www.clio-online.de/hist2006/userinfoviewsession.php4?session_id=10> (12.04.2006).

Müsste man nicht von E-Teaching statt von E-Learning sprechen, wenn sich das Online-Angebot im Wesentlichen auf einen Seminarordner im Netz beschränkt?

Viele der nach dem ersten Boom der späten 1990er Jahre gestarteten E-Learning-Projekte sind aber weniger durch neue technische Möglichkeiten motiviert, bzw. proben offensiv neue Formen der Wissensvermittlung, sondern reagieren eher defensiv auf die Veränderungen in der universitären Lehre im Zuge des Bologna-Prozesses sowie auf weiter steigende Zahlen von Teilnehmern/innen in den Seminaren. Oder anders formuliert: Steht der Seminarordner im Netz, muss wenigstens keiner Schlange vor dem Kopierer stehen. Das Hosting in zugangsbeschränkten und meist nur über Campus-Lizenzen verfügbaren kommerziellen Systemen⁹ führt aber dazu, dass an verschiedenen Universitäten ähnliche Kursinhalte entwickelt werden. Mehr Kooperation zwischen den Lehrenden an den verschiedenen Hochschulen sowie der Einbezug der Universitätsbibliotheken als Informationsvermittler in die Angebote der Grundausbildung wurden gefordert, um den hohen Anfangs-Aufwand für den Aufbau einer netzgestützten Lehrveranstaltung zu reduzieren.

Mehr Erfolg als solche Appelle verspricht der Ansatz, der von Teilnehmern/innen des E-Learning-Projekts „Heidelberg 1934-Harvard 1940“ präsentiert wurde: Der Einbezug von Studierenden in die Konzeption und Realisierung eines Lehrangebots. Im Rahmen eines Oberseminars bildete das unveröffentlichte Manuskript „Nacht über Heidelberg“ von Dr. Barbara Schütz-Sevin, einer früheren Studentin und Doktorandin an der Universität Heidelberg, die Basis für einen netzgestützten Kurs zu den Themen Machtübernahme, Gleichschaltung, Antisemitismus und erzwungene Emigration in den 1930er und 1940er Jahren. Mit dem Einbezug von Bildern, Film- und Tonsequenzen, digitalisierten Zeitungen und Archivalien werden die medialen Vorteile des Webs gegenüber dem papiergebundenen Selbststudium ausgenutzt. Als technische Plattform wird *Moodle*¹⁰

9 In Innsbruck und Bochum Blackboard Academic Suite, <<http://www.blackboard.com>> (12.04.2006).

10 Vgl. <<http://moodle.org/>> (12.04.2006).

eingesetzt, das auch die Einbindung von Wiki-Komponenten und gegenseitigen Bewertungsmöglichkeiten unterstützt. Wie weit rechtliche Rahmenbedingungen, etwa Landeslehrverordnungen, einer Verlagerung von zuvor alleine von den Lehrenden übernommenen Aufgaben auf die Teilnehmer/innen der Veranstaltung selbst im Wege stehen, wird sich in Zukunft sicher zeigen. Urheberrechtliche Bedenken verhinderten auch die Öffnung der Veranstaltung für eine breite Netz-Öffentlichkeit außerhalb der Universität Heidelberg. Zu hoffen wäre, dass in Zukunft aber wenigstens Studierende von Harvard sich im Zuge einer transatlantischen Kooperation in diese Lehrveranstaltung einklinken können. Spätestens dann würde aus dem kollaborativen Forschungs- auch ein kollaborativer Lernprozess werden.

Wer die Programme der Tagung .hist 2006 und der entsprechenden Veranstaltung drei Jahre zuvor nur flüchtig vergleicht, könnte zum Schluss kommen, dass nach einem stürmischen Jahrzehnt nun eine Phase der Ruhe und Konsolidierung angebrochen ist. Bei der Retrodigitalisierung und in der Sektion zum elektronischen Publizieren traten vorwiegend bereits von der letzten Tagung bekannte Projekte auf, beim E-Learning hat sich die Zahl der Präsentationen sogar deutlich reduziert. Einzig die Sektion zu den „Collaboratories“ setzte einen ganz neuen Akzent. Der Erfolg der Wikipedia und die freie Verfügbarkeit der darunter liegenden Software-Plattform – das hat sich in der Sektion von Madeleine Herren zum E-Learning deutlich gezeigt – eröffnete in fast jeder Sektion ganz neue Perspektiven. Wie aus rohen Scans nach dem Wiki-Prinzip eine transkribierte und kommentierte Edition entstehen könnte, zeigt zurzeit die Zimmerische Chronik auf *Wikisource*.¹¹ Auch aus einer noch stark am gedruckten Artikel orientierten E-Publikation wird so vielleicht in wenigen Jahren ein schon vor der ersten Veröffentlichung von den Gutachtern/innen und danach von den Lesern/innen kommentiertes, berichtiges und ergänztes „Living Document“.¹² Gut möglich, dass diese kollaborativen Prozesse für das Fach in

11 Vgl. <http://de.wikisource.org/wiki/Zimmerische_Chronik> (12.04.2006), siehe auch eine entsprechende Mail von Klaus Graf, <<http://www.aedph.uni-bayreuth.de/2006/0004.html>> (12.04.2006).

12 Ein Beispiel für einen solchen Prozess bietet der Tagungsband Wikimania 2005,

den nächsten Jahren ähnlich bedeutsam werden, wie die Einführung von E-Mail und WWW im letzten Jahrzehnt.

Daniel Burckhardt ist Wissenschaftlicher Mitarbeiter am Institut für Geschichtswissenschaften der Humboldt-Universität zu Berlin und Web-Redakteur von H-Soz-u-Kult. E-Mail: BurckhardtD@geschichte.hu-berlin.de

<<http://meta.wikimedia.org/wiki/Transwiki:Wikimania05>> (12.04.2006).

ELEKTRONISCHES PUBLIZIEREN IN DEN GEISTESWISSENSCHAFTEN

ELEKTRONISCHES PUBLIZIEREN – EINLEITUNG

von Gudrun Gersmann

„Elektronisches Publizieren“? Vor einigen Jahren noch reagierten viele Historiker/innen geradezu verschreckt, wenn über Online-Publikationen gesprochen wurde. Zu exotisch schien damals diese Form der Veröffentlichung, zu weit entfernt von prestigeträchtigen klassischen Publikationsorten, zu fragil, um als wirkliche Alternative gelten zu dürfen. Selbst diejenigen, die grundsätzlich den „Neuen Medien“ gegenüber aufgeschlossen waren, legten oftmals Skepsis an den Tag: Würde es nicht so aussehen, als habe man keinen „vernünftigen“ Verlag für seine Dissertation finden können, wenn man das Manuskript einem Hochschulserver überantwortete? Würde man sich damit nicht in Fachkreisen disqualifizieren? Würde man sich denn überhaupt auf die Stabilität der Systeme verlassen können? Müsste man nicht von vornherein einkalkulieren, eines Morgens statt seines eigenen Textes auf dem Bildschirm eine schlichte Fehlermeldung („Server not found“) vorzufinden? Würde man mit der leicht kopierbaren elektronischen Fassung nicht zum Plagiierten geradezu einladen? So lauteten die Argumente, die „in jenen Zeiten“ immer wieder fast litaneihaft gegen das elektronische Publizieren vorgebracht wurden.

Wie gut, dass sich die Dinge geändert haben – oder zu ändern beginnen. Obwohl das eine oder andere Schreckgespenst („Langzeitarchivierung“) nach wie vor im Hintergrund herumgeistert und neue Formen der Wissenschaftskommunikation wie Weblogs oder Wikis vorerst nur zögerlich Eingang in die „Community“ finden, kann an der zunehmenden Bedeutung elektronischer Publikationsformen auch und gerade in den Geisteswissenschaften kein Zweifel mehr bestehen.

Auf welche Erfahrungen mit dem elektronischen Publizieren können wir Geisteswissenschaftler/innen mittlerweile zurückblicken? Wo existie-

ren nach wie vor Berührungsängste? Wie kann das elektronische Publizieren stärker als bisher in den Geisteswissenschaften verankert werden? Da der in den vergangenen Jahren erfolgte Umbruchs- und Aufbruchprozess hier nicht im Detail nachvollzogen werden kann, seien zum Auftakt nur einige wenige grundsätzliche Beobachtungen formuliert:

Erstens haben wir in der jüngsten Vergangenheit ein deutliches Wachstum der fachwissenschaftlichen Online-Journale, Repositorien und Plattformen erleben können, bedingt nicht zuletzt durch umfangreiche Fördermaßnahmen von Wissenschaftsorganisationen¹ und Ministerien: Im Rahmen der nordrhein-westfälischen Initiative „Digital Peer Publishing“² konnten seit 2004 beispielsweise zehn neue Online-Journale aus unterschiedlichen Fachbereichen ins Leben gerufen werden, weitere Zeitschriften werden folgen.

Zweitens hat die auf unterschiedlichen Ebenen geführte Debatte über die Realisierung von Open Access-Modellen im Bereich des geisteswissenschaftlichen Publizierens inzwischen erfreulich weite Kreise gezogen: Ungeachtet mancher Rückstände und hartnäckiger (Verlags)Widerstände konnten auf unterschiedlichen Ebenen wichtige Fortschritte erzielt werden. Nicht unterschätzen sollte man die Aufmerksamkeit, die die einschlägigen Fachgesellschaften inzwischen dem elektronischen Publizieren widmen: So hat etwa der Verband der Historiker und Historikerinnen Deutschlands auf dem Konstanzer Historikertag vom September 2006 einen eigenen Unterausschuss für den Bereich „Elektronisches Publizieren und Fachinformation in den Geschichtswissenschaften“ eingesetzt, der sich explizit um eine künftig stärkere Integration der neuen Publikationsformen in die Fachwissenschaft kümmern soll.

Drittens haben sich im Feld des elektronischen Publizierens fruchtbare neue Kooperationen zwischen Bibliotheken, Rechenzentren und Autoren/innen herausgebildet, die zu neuen Rollenverteilungen geführt haben: Wissenschaftler/innen agieren nunmehr auch schon als ihre eigenen Verleger, während auch auf Seiten der Verlage selbst teilweise zumindest ein

1 Wie zum Beispiel der DFG, vgl. <<http://www.dfg.de>>.

2 Vgl. <<http://www.dipp.nrw.de/journals/>>.

Prozess des Nachdenkens und des Reflektierens über die eigene Aufgabe in Gang gekommen ist.

Viertens hat die Einsicht, dass die Vorzüge elektronischer Publikationsformen über die bloße Faksimilierung einer gedruckten Vorlage oder die schnelle Veröffentlichung eines Beitrags weit hinausreichen, die Attraktivität des E-Publishing in den Augen vieler Autoren/innen maßgeblich erhöht. Ob es sich um die Anreicherung mit Bild- und Tonbeispielen handelt, um die Möglichkeit von Verlinkungen, die direkte „Anklickbarkeit“ in Bibliotheks-Opacs oder die Präsenz im „Google-Universum“ – jedes elektronisch vervielfältigte Manuskript zieht per se eine höhere „Visibilität“ nach sich als ein klassisch gedruckter und in Sammelbänden mit Kleinstauflagen auf ewig begrabener und verstaubender Artikel. Nicht das schlechteste Format also auch und gerade für Nachwuchswissenschaftler/innen, die mit ihren Forschungen weltweit bekannt werden – oder mit ihren Lesern/innen in Kontakt treten wollen.

Außerordentlich spannend dürfte es darüber hinaus sein, unter dem Stichwort Weblogs oder Wikis fünftens längerfristig die Entwicklung der neuen informellen Kommunikations- und Publikationsformen zu verfolgen, die experimentell etwa das kollaborative Schreiben praktizieren. Das Netzwerk „Knowledge Exchange“ belegt das internationale Interesse am Aufbau von „Virtual Research Environments“ (VRE), von „Umgebungen“ mithin, in denen Wissenschaftler/innen mit digitalisierten Quellen arbeiten, die im Netz verfügbaren Hilfsmittel nutzen und die Resultate ihrer Forschungen schließlich auch online veröffentlichen. Auch wenn auf Seiten der Wissenschaftler/innen häufig noch ein gewisser Hang zum gemütlichen Konsum vorherrscht und die Bereitschaft zum Engagement in Netzforen und Mailinglisten – aus welchen Gründen auch immer – durchaus noch gesteigert werden könnte, sind die Tendenzen in dieser Richtung doch unübersehbar. Es bleibt also spannend, und man darf hoffen, dass manche der noch existierenden Mängel und Kinderkrankheiten – angefangen beim unattraktiven „Gemischtwarencharakter“ vieler Hochschulschriftenserver bis hin zu ungelösten technischen Problemen – in absehbarer Zeit beseitigt werden können.

Viele der eingangs skizzierten Fragen konnten in der gut besuchten

Sektion „Elektronisches Publizieren“ der Berliner .hist 2006-Tagung wenn auch nicht erschöpfend beantwortet, so doch wenigstens angerissen und mit einem engagierten Publikum diskutiert werden. Der Bogen der Fragestellungen war weit gespannt: Aus unterschiedlichen Perspektiven versuchten Johannes Fournier (DFG) und Karsten Borgmann (HU Berlin) den derzeitigen Standort der „Neuen Medien“ in der Wissenschaftskommunikation zu bestimmen. Ausgehend von den Ergebnissen einer von der Deutschen Forschungsgemeinschaft (DFG) Ende 2004 in Auftrag gegebenen Studie über die Akzeptanz des elektronischen Publizierens in der Fachwissenschaft lieferte Johannes Fournier einen gesamteuropäisch angelegten Überblick über aktuelle Entwicklungen im Bereich des elektronischen Publizierens.³ Zu Recht wies er dabei auf die Langsamkeiten des Rezeptionsprozesses hin: Obwohl sich elektronische Zeitungen, Preprints, Wikis und Weblogs einer stetig wachsenden Lesergemeinde erfreuen, müssen diese Formen des Publizierens und Kommunizierens noch an Renommee gewinnen, um mit den etablierten Printproduktionen „mitziehen“ zu können. Diesen Befund bestätigte auch Karsten Borgmann vor dem Hintergrund einer im Januar 2006 abgeschlossenen, von den Projekten Clio-online⁴, H-Soz-u-Kult⁵ und Zeitgeschichte-online⁶ initiierten Umfrage über die Nutzung von Online-Angeboten in den Geschichtswissenschaften.⁷ So selektiv und vorläufig die Resultate solcher Umfragen sein mögen, liefern sie doch anregende Aufschlüsse über derzeitige Rezeptionsgewohnheiten und –erwartungen. Die Selbstverständlichkeit und Unbefangenheit, mit der Historiker/innen die fachwissenschaftlichen Netzangebote von Clio-online und H-Soz-u-Kult bis hin zu historicum.net⁸ in ihren Arbeitsalltag integriert haben, steht demnach – noch – in eklatantem Widerspruch zur Bereitschaft, eigene Texte online verfügbar zu machen: Zweifellos ist hier noch einiges an Arbeit zu tun.

3 Vgl. dazu den Beitrag in diesem Band.

4 <<http://www.clio-online.de/>>.

5 <<http://hsozkult.geschichte.hu-berlin.de/>>.

6 <<http://www.zeitgeschichte-online.de/>>.

7 Vgl. dazu den Beitrag in diesem Band.

8 <<http://www.historicum.net/>>.

Steckten die beiden erwähnten Beiträge den allgemeinen wissenschaftsimmanenten und wissenschaftspolitischen Rahmen des elektronischen Publizierens in den Geisteswissenschaften ab, so berichteten Michael Kaiser (Universität zu Köln)⁹ und Jan-Holger Kirsch (Zentrum für Zeithistorische Forschung, Potsdam)¹⁰ aus der Redakteurspraxis über ihre Erfahrungen mit dem Aufbau und der Gestaltung der Online-Fachjournale *sehepunkte*¹¹, *zeitenblicke*¹² und *Zeithistorische Forschungen*.¹³ Obwohl die drei Journale auf unterschiedlichen Publikationsgewohnheiten beruhen – bei den *sehepunkten* und *zeitenblicken* handelt es sich um reine Online-Journale, die *Zeithistorischen Forschungen* hingegen werden in Hybridform veröffentlicht –, ist das Fazit der Macher/innen letztlich ähnlich: Die hohen Zugriffszahlen sprechen einerseits für die hohe Popularität der Journale, die von der vielfach konservativen Wunschklientel mittlerweile offenbar ähnlich hochrangig eingeschätzt werden wie die klassischen Fachzeitschriften. Trotz der wachsenden Reputation kann keines der drei Journale freilich als gesichert gelten: Da der redaktionelle und personelle Aufwand bei der Einwerbung und Begutachtung von Artikeln bei einem Online-Journal ebenso hoch ist wie bei einem klassischen Journal, kann der Betrieb solcher Zeitschriften auf Dauer allein durch Anschubfinanzierungen nicht geleistet werden.

Solche Kosten sind nicht das einzige Problem, wie der Beitrag des Medienrechtlers Karl-Nikolaus Peifer (Universität zu Köln)¹⁴ anschaulich dokumentierte, der nicht nur höchst pointiert auf die Problematik des Urheberrechts und lizenzrechtlicher Klärungen verwies, sondern allen allzu euphorischen Publikationsvisionen in Bezug auf die Verwendung von Bild- und Tonquellen einen Dämpfer versetzte. Thomas Wolf-Klostermann (Bayerische Staatsbibliothek, München)¹⁵ ergänzte schließlich das Spektrum der Vorträge durch Bemerkungen zum Thema aus bibliothekarischer

9 Vgl. dazu den Beitrag in diesem Band.

10 Vgl. dazu den Beitrag in diesem Band.

11 <<http://www.sehepunkte.de/>>.

12 <<http://www.zeitenblicke.de/>>.

13 <<http://www.zeithistorische-forschungen.de/>>.

14 Vgl. dazu den Beitrag in diesem Band.

15 Vgl. dazu den Beitrag in diesem Band.

Sicht mit einem tröstlichen Fazit: Ungeachtet aller Kassandrarufer haben die Bibliotheken das digitale Zeitalter bisher nicht nur überlebt, sondern als Reaktion auf die tiefgreifenden Veränderungen im Kommunikations- und Publikationssystem neue dezentrale Organisationsformen und -strukturen entwickelt. Inwieweit es den Bibliotheken gelingen wird, als professionelle Informationsmanager zu agieren, wird die Zukunft zeigen.

Prof. Dr. Gudrun Gersmann ist Professorin für die Geschichte der Frühen Neuzeit an der Universität zu Köln, Begründerin des Internetportals [historicum.net](http://www.sehepunkte.de) und Mitherausgeberin der E-Journale

<<http://www.sehepunkte.de>>; <<http://www.zeitenblicke.de>>;

<<http://www.lesepunkte.de>>. E-Mail: gudrun.gersmann@uni-koeln.de

SEHEPUNKTE UND ZEITENBLICKE – ENTWICKLUNGSLINIEN NACH EINEM HALBEN JAHRZEHNT DES ELEKTRONISCHEN PUBLIZIERENS

von Michael Kaiser

*Mittlerweile erscheinen die *sehepunkte* im sechsten Jahrgang, die *zeitenblicke* in ihrem fünften Jahr. Absolut gesehen ist dies keine lange Zeitspanne, für Internetprojekte allerdings stellt es durchaus eine beachtliche Kontinuität dar.*

*Für beide Online-Formate lohnt es sich, einen Blick zurück zu werfen und zu rekapitulieren, von welchen Anfängen aus sie gestartet und wo sie heute in der Fachwelt zu verorten sind. Dies soll weniger chronologisch geschehen als vielmehr nach Sachpunkten geordnet, die helfen sollen, die Entwicklung der *sehepunkte* und der *zeitenblicke* einzuordnen und zu verstehen. Auf eine Kurzzvorstellung der Journale folgt eine Skizze der Startschwierigkeiten, danach werden die Punkte beleuchtet, die zur erfolgreichen Etablierung der Journale beitragen. Dabei spielen konzeptionelle Elemente der Online-Formate selbst eine Rolle, aber auch spezifische Eigenheiten des Mediums und das Verhältnis zu den Adressaten/innen. Abschließend werden vor allem mit den Rechtsfragen und den Finanzierungsproblemen Aspekte thematisiert, die die weitere Entwicklung der *sehepunkte* und der *zeitenblicke* mitprägen dürften.*

Eine Idee – zwei Projekte

Wie für viele andere Online-Publikationen stand auch für die *sehepunkte*¹ und *zeitenblicke*² am Anfang die Grundidee, die Möglichkeiten des elekt-

1 Vgl. <<http://www.sehepunkte.de>>.

2 Vgl. <<http://www.zeitenblicke.de>>.

ronischen Publizierens für die Geschichtswissenschaften nutzbar zu machen. Zwei Kriterien waren (und sind nach wie vor) handlungsleitend für die beiden Online-Projekte:

Zum einen ist dies die enorme Schnelligkeit des Mediums, die die Aktualität der publizierten wissenschaftlichen Inhalte merklich steigert. Dies gilt in ganz besonderem Maße für die Buchbesprechungen, die bedeutend schneller als in jedem anderen wissenschaftlichen Printmedium verfügbar sind und damit eine rasche Orientierung über aktuelle Neuerscheinungen ermöglichen. Zum anderen spielt die nahezu unbegrenzte Verfügbarkeit der Publikationen eine wichtige Rolle. Dass man bei Online-Publikationen nicht mehr von Bibliotheken und damit von Magazinaushubzeiten und Lesesaalöffnungszeiten abhängig ist, dass Online-Publikationen nie „ausgeliehen“ sind und damit nur einem Benutzer bzw. einer Benutzerin exklusiv, sondern für unbegrenzt viele Leser/innen gleichzeitig verfügbar sind, stellt den anderen unbestrittenen Vorteil dar. Beide Formate waren auch von Beginn an, das heißt auch schon lange vor der *Berliner Erklärung*, dem Prinzip des Open Access verpflichtet. Diese Vorteile des Online-Publizierens galt es nun zu verknüpfen mit den traditionellen wissenschaftlichen Standards. Fachwissenschaftliche Seriosität war die Voraussetzung, um die Gleichwertigkeit von Online-Formaten neben herkömmlichen Printjournalen zu erweisen. Mit diesem Anspruch sind die beiden Projekte *sehepunkte* und *zeitenblicke* angetreten.

Als erstes erblickte das Rezensionjournal *sehepunkte* das Licht der virtuellen Welt. Seine Geschichte begann im Jahr 2000 unter dem Namen PERFORM. Anfangs beschränkte sich PERFORM auf Publikationen aus dem Bereich der frühneuzeitlichen Geschichte. Doch rasch wurde hier die Öffnung zu allen anderen Teilepochen der Geschichte vollzogen, als Ende 2001 die *sehepunkte* das Erbe von PERFORM antraten. Erschien PERFORM anfangs nur alle zwei Monate, wurde mit dem Wechsel zu den *sehepunkten* auch ein einmonatiger Erscheinungszyklus eingeführt, der sich bis jetzt bewährt hat. Die *sehepunkte* bieten heute thematisch ein Komplettangebot an, indem die Epochengliederung das Altertum, das Mittelalter, die Frühe Neuzeit, das 19. Jahrhundert und die Zeitgeschichte umfasst – ergänzt durch eine allgemeine Rubrik für übergreifende Litera-

tur. KUNSTFORM, die kunstgeschichtliche Sektion der *sehepunkte*, besitzt zwar nicht diese explizite Substruktur, deckt aber prinzipiell ebenfalls den historischen Gesamtrahmen ab. Eine zusätzliche Möglichkeit, einen thematischen Schwerpunkt in einer Ausgabe zu setzen, bietet das Forum, das Publikationen zu einem bestimmten Themenkreis versammelt und mit einer knappen Einleitung versehen gebündelt präsentiert. Das mag mal die Stadtkultur in der Spätantike sein oder auch die kontroverse Diskussion über ein aktuelles Buch, wie sie sich in Mehrfach-Besprechungen durch Rezensenten unterschiedlicher Couleur und Standpunkte widerspiegelt – so etwa im Falle von Götz Alys Buch „Hitlers Volksstaat“.

Pro Ausgabe werden in den *sehepunkten* ungefähr 80 bis 100 Rezensionen publiziert. Bis dato sind über 3.000 Buchbesprechungen erschienen. Sie sind überwiegend in deutscher Sprache geschrieben, es gibt aber auch einige Rezensionen in Englisch, Italienisch, Französisch und Spanisch.

Etwas jünger sind die *zeitenblicke*, ein Journal für die Geschichtswissenschaften, das 2002 erstmalig online gestellt wurde. Formal ist hier offenkundig die größte Konstanz zu beobachten, da keine Umbenennungen und damit einhergehende Neukonzipierungen vorgenommen wurden. Eine thematische Schwerpunktsetzung zeichnet sich insofern ab, als in den aktuellen Ausgaben verstärkt frühneuzeitliche Themen aufgegriffen werden, während in früheren Nummern auch andere Epochen vertreten waren. Die positive Entwicklung der *zeitenblicke* ist ablesbar an der erhöhten Frequenz: Erschienen 2002 nur zwei Ausgaben, sind in allen folgenden Jahren stets drei Ausgaben realisiert worden. Momentan gibt es für jede Ausgabe ein festes historisches Thema. Hierbei kann es sich um Sujets handeln, die aktuelle Forschungstrends spiegeln wie es etwa die Ausgabe zu den Ego-Dokumenten „Das Ich in der Frühen Neuzeit“³ tut oder auch die Ausgabe zu „Raub und Wiedergutmachung“.⁴ Thematisiert wird aber auch die Situation des elektronischen Publizierens oder der aktuelle Stand der Diskussion zu den Studienreformen in Europa. In der Binnenstruktur gibt es mit den „Analysen und Beiträgen“ das typische Genre für ein

3 *zeitenblicke* 1 (2002), Nr. 2.

4 *zeitenblicke* 3 (2004), Nr. 3.

Fachjournal; Gleiches gilt für die Berichte aus der Forschung, während sich die Verlinkungen auf externe Netzangebote die Chancen des neuen Mediums zunutze machen. Neu und durchaus ungewöhnlich für ein wissenschaftliches Fachorgan ist das Interview, das an der Spitze einer jeden Ausgabe steht und meist herausragende Fachvertreter/innen und sonst thematisch ausgewiesene Gesprächspartner/innen zu Wort kommen lässt.

Die *zeitenblicke* veröffentlichen überwiegend Originalbeiträge; nur zu ungefähr zehn Prozent der Aufsätze gibt es Post- oder Pre-Prints. Die Aufsätze selbst erscheinen meist in deutscher Sprache, doch ist es möglich, dass Artikel auch in einer gängigen anderen Sprache publiziert werden. Bislang sind Beiträge in englischer und italienischer Sprache erschienen. Insgesamt sind seit 2002 über 200 Artikel veröffentlicht worden.

Auch wenn es sich um zwei separate Journale handelt, die jeweils eine eigene Geschichte haben und auch je für sich als Publikation existieren, bilden die *sehpunkte* und die *zeitenblicke* ein konzeptionelles Tandem, was in der Geschichtswissenschaft und in den angrenzenden Fachdisziplinen auch so wahrgenommen wird. Deutlich wird dies durch die komplementären Inhalte beider Formate, hinzu kommt die große strukturelle Ähnlichkeit. Beide Journale sind stark an traditionellen Lesegewohnheiten orientiert, verzichten etwa auf ein Day-By-Day-Publishing, wie es bei vielen anderen Online-Formaten praktiziert wird. Stattdessen gibt es eine feste Ausgabenstruktur mit festen Erscheinungsterminen. Überhaupt finden sich nur wenig dynamisierende Elemente, vielmehr handelt es sich um mehr oder weniger statische Online-Publikationen. Dabei stellt es durchaus einen bewussten Verzicht auf Möglichkeiten dar, die das Online-Publizieren ansonsten offeriert. Entscheidend aber ist hier gewesen, das Publikum mit seinen Gewohnheiten ernst zu nehmen und somit das neue Medium eines Online-Journals in weitgehend vertrauten Formen darzubieten.

Anfangsschwierigkeiten

Die Startphase der Projekte war von Problemen gekennzeichnet. Die typische Anfangseuphorie, die es auf Seiten der Projektinitiatoren gab, war durchaus notwendig, um der Skepsis und den Bedenken im Umfeld und bei

potenziellen Kooperationspartnern zu begegnen. Das Rezensionjournal musste in zwei Richtungen um Vertrauen werben: Da waren zunächst die Verlage zur Zusammenarbeit zu bewegen, dass sie den *sehpunkten* von ihren Verlagsprodukten Rezensionsexemplare zur Verfügung stellten. Es gab zu Beginn ein uneinheitliches Bild, da einige Verlagshäuser die *sehpunkte* von vornherein unterstützten, während andere Verlage zunächst deutlich distanziert reagierten und erst später ihre Bücher zur Rezension offerierten oder auf Anfrage der Redaktion zusandten. Die Kooperationsbereitschaft der Verlage hing verständlicherweise vom Erfolg der *sehpunkte* ab, der an den tatsächlich publizierten Besprechungen messbar war. Es ging also gerade in der kritischen Anfangsphase darum, auch Autoren/innen zu finden, die ihrerseits zuverlässig rezensierten, so dass damit Vorbehalte und Zweifel auf Seiten der Verlage zerstreut werden konnten. Bei den Wissenschaftlern/innen gab es ebenfalls Skepsis, ob die geschriebenen Besprechungen dann auch veröffentlicht und wahrgenommen, die Leistungen der Rezensenten/innen also auch in der Scientific Community sichtbar werden würden.

Die sehr kurze Frist von der Einreichung einer Besprechung bis zur Publikation bot kaum Anlass für die Bedenken der Rezensenten/innen. Der Vorteil der Schnelligkeit einer Online-Veröffentlichung, die eben verhinderte, dass der eingereichte Text über Monate, wenn nicht Jahre in der Redaktion lag, wurde hier sehr schnell deutlich. Auch die Reaktionen der besprochenen Autoren/innen, die sich durchweg sehr angetan zeigten, dass ihre Publikation derart schnell Resonanz gefunden hatte, bestätigten den positiven Eindruck und trugen ihren Teil zur guten Reputation des Rezensionsjournals bei. Der gute Name der *sehpunkte* bewirkte auch, dass diese in erstaunlich kurzer Zeit nicht nur ein Journal für junge Wissenschaftler/innen wurden, sondern ebenso von etablierten Fachvertretern/innen genutzt wurden, die ihrerseits keine Scheu zeigten, für die *sehpunkte* zu schreiben. Der Erfolg der *sehpunkte* half zweifelsohne, den Start der *zeitenblicke* zu befördern, die sowohl organisatorisch, als auch im Erscheinungsbild gleichsam als Schwesterprojekt der *sehpunkte* auftraten. Auch hier gab es Bedenken gegenüber dem Medium. Deutlich stärker als die *sehpunkte* waren die *zeitenblicke* mit Fragen nach der Nachhaltigkeit des

Formats konfrontiert – ein Punkt, der für die Buchbesprechungen, die doch eher als Tagesgeschäft wahrgenommen werden, sehr viel weniger relevant war. Ob aber ein Aufsatz auch in den kommenden Monaten und Jahren noch recherchierbar sein wird, ist wesentlich für die Bereitschaft der Wissenschaftler/innen, ihre Ergebnisse einer Online-Zeitschrift anzuvertrauen. Große Offenheit in dieser Frage seitens der Herausgeber/innen und der Redaktion sowie letztlich die Beständigkeit im Publikationsprozess selbst halfen hier ganz wesentlich, die Bedenken zu zerstreuen.

Gründe für den Erfolg

Beide Formate haben unbestritten großen Erfolg und können mittlerweile als im Fach etabliert gelten. Die Gründe dafür lassen sich auf ganz wenige Kernpunkte zurückführen. Beide Journale haben sich das Vertrauen der Autoren/innen und Leser/innen erarbeiten müssen. Wichtig war es deswegen von Anfang an, dass die Journale mit großer Zuverlässigkeit erschienen. Besonders die monatliche Erscheinungsweise der *sehepunkte* ist ein solches Kriterium, mit dem sich alle vier Wochen neu die Beständigkeit des Projekts bestätigt: Dass der 15. eines jeden Monats *sehepunkte*-Tag ist, ist gleichwohl nicht nur ein stets aktueller Erweis der Zuverlässigkeit, sondern umgekehrt auch Verpflichtung der *sehepunkte* selbst. Die *zeitenblicke*, die nicht in dem strengen Sinne auf ein Tagesdatum als Erscheinungstermin festgelegt sind, haben diesen Druck nicht, müssen gleichwohl in einer Taktung von jeweils ungefähr vier Monaten erscheinen.

Die Schnelligkeit des Mediums hat stets auch eine Schattenseite gehabt, mochte die Geschwindigkeit im Publikationsprozess doch auch implizieren, dass der Zeitvorteil mit einem Mangel an Qualität erkaufte sei. Um diesen Verdacht von vornherein zu entkräften, haben beide Journale von Beginn an auf die strikte Einhaltung von wissenschaftlichen Standards gesetzt. Bei den *sehepunkten* durchlaufen die Besprechungen ein elaboriertes Begutachtungsverfahren, das im Wesentlichen zweistufig funktioniert. Zunächst bearbeiten die Fachredakteure/innen den Text, wobei hier auch inhaltlich-konzeptionelle Fragen im Vordergrund stehen. Dieses spielt aber auch noch einmal im internen Begutachtungsverfahren durch die Redaktion und vor allem seitens der Herausgeber/innen eine wichtige Rolle, so

dass auf diese Weise ein hoher wissenschaftlicher Standard gewährleistet ist. Ebenso wird redaktionell auf die Beachtung der Formalia großer Wert gelegt, um das einheitliche Erscheinungsbild der *sehepunkte* zu wahren.

Um die fachwissenschaftliche Kompetenz in der redaktionellen Betreuung sicherzustellen, ist eine ganze Reihe von Wissenschaftlern/innen als Fachredakteure/innen tätig, die hier ihre eigenen besonderen Qualifikationen einbringen können und auch in diversen Spezialgebieten der historischen Forschung für hohe Standards bürgen. So ist auch die starke Verwurzelung der *sehepunkte* in der Wissenschaft selbst hergestellt, da die wissenschaftlichen Rezensenten/innen von Wissenschaftskollegen/innen betreut werden.

Das hier entstehende Vertrauen wird auch durch den Umstand befördert, dass beide Formate starke Partner-Institutionen haben, die ihrerseits Fachkompetenz für die Online-Formate einbringen. Dies ist das Herder-Institut in Marburg, das besonders für den weiten Bereich der osteuropäischen Geschichte zuständig ist, und auch das Institut für Zeitgeschichte in München, das weite Teile der Publikationen der Neuesten Geschichte mitbetreut. Zu erwähnen ist auch der wissenschaftliche Beirat, der aus renommierten Vertretern/innen der historischen Wissenschaften aller Teilepochen besteht, und auf seine Weise die Seriosität der Projekte gewährleistet. Gleiches gilt auch für die *zeitenblicke*, wobei das „nur“ dreimalige Erscheinen pro Jahr die Verwurzelung in der Fachwelt im Alltag nicht so augenfällig werden lässt wie bei den *sehepunkten*.

Neben diesen Faktoren spielt auch die bereits oben erwähnte Nachhaltigkeit der Online-Projekte eine große Rolle. Die Seriosität beider Journale schlägt sich auch im Umstand nieder, dass beide jeweils eine eigene ISSN (International Standard Serial Number) haben. Die Frage der Archivierung ist zweifelsohne ein Punkt, der immer noch einer wirklich abschließenden Antwort bedarf. Doch dürfen sich die *zeitenblicke* und *sehepunkte* etwa durch Kooperationen mit der Bayerischen Staatsbibliothek in München und dem Hochschulbibliothekszentrum für Nordrhein-Westfalen in Köln einigermaßen sicher fühlen, was die Dauerhaftigkeit im Netz angeht – ganz abgesehen vom Umstand, dass die Deutsche Bibliothek in Leipzig die Texte der *zeitenblicke* sichert und die Projekte ihrerseits nochmals eigene

Archiv- und Sicherungsfunktionen aufgebaut haben. Im Falle von technischen Schwierigkeiten ist daher auch gewährleistet, dass ein Rund-um-die-Uhr-Service für die permanente Sichtbarkeit der publizierten Inhalte im Netz sorgt. Neben der Nachhaltigkeit nicht minder wichtig ist das Prinzip der Unveränderlichkeit einer Online-Publikation. Hier gilt von Beginn an der Grundsatz, dass ein ins Netz gestellter Text als „gedruckt“ und damit nicht mehr korrigierbar galt – ein ganz entscheidender Satz, der in der Scientific Community als Ausweis von Seriosität wahrgenommen wird. Darin äußert sich weniger ein Konservatismus als vielmehr das Bedürfnis, Texte zu publizieren, die auch wissenschaftlich zitierfähig sind. Auf diese Weise ist auch für alle Wissenschaftler/innen, die für die *sehepunkte* und die *zeitenblicke* arbeiten, sichergestellt, dass diese Texte auch in die eigene Publikationsliste aufgenommen werden können. An dem Punkt kommen nicht zuletzt die durchaus starren Publikationsstrukturen ins Spiel, dass eben beide Journale in festen Ausgaben zu festen Terminen erscheinen. Hier wird die Zitierfähigkeit auch ohne Angabe der URL deutlich und unterscheidet sich praktisch nicht von den Zitierkonventionen, wie sie bei gewöhnlichen Printmedien angewandt werden. Die Wissenschaftler/innen, auch wenn sie zunächst dem neuen Medium mit Skepsis, Argwohn, vielleicht auch mit einigen Berührungsängsten begegneten, fanden hier Strukturen vor, die vertraut waren und damit Vertrauen schafften.

Neues Medium – neue Wissenschaft?

Nicht alles bei den Online-Projekten ist so vertraut wie die Publikationsweise. Das neue Medium wirkt auf seine Nutzer/innen in besonderer Art zurück – sowohl auf diejenigen, die für das Internet schreiben, als auch auf diejenigen, die die dort präsentierten Inhalte lesen. Dabei geht es weniger um Geschmacksfragen, die das Design der Homepages und das Layout einzelner Seiten prägen; vielmehr geht es um Formen des (wissenschaftlichen) Schreibens im und für das Netz. So besteht Konsens darüber, dass sich die kleine Form als adäquat für die Präsentation eben auch wissenschaftlicher Inhalte bewährt hat. Gerade Verknüpfungsmöglichkeiten, die einzelne, auch vom Umfang her überschaubare Texte in immer neue Zusammenhänge mit anderen Informationen stellen und sich damit dem

Begriff des Hypertextes nähern, sind vielfach der einleuchtende Beleg dafür. Dass das Lesen am Bildschirm von kürzeren, gut strukturierten Texten angenehmer ist, steht zudem außer Frage. Gegen diese Tendenz kleinerer Texte – wobei die Kleinheit ein relativer Begriff ist, zumal die *zeitenblicke* mitunter Aufsätze publizieren, die im Umfang traditionellen Veröffentlichungen in nichts nachstehen, ohne deswegen für das Netz ungeeignet zu sein – hat sich immer wieder Widerstand geregt. Die Forderung nach kurzen Texten gerät leicht in Verdacht, journalistische Gepflogenheiten zu adaptieren – für Autoren/innen mit wissenschaftlichem Anspruch ein tödliches Argument. Sicherlich ist die Gleichsetzung von kurzen Texten als unwissenschaftliche, ja sogar unseriöse Publikationsform auch ein „deutsches“ Problem, das sich umgekehrt immer wieder in Klagen über die Unlesbarkeit wissenschaftlicher Publikationen, speziell von wissenschaftlichen Qualifikationsarbeiten, niederschlägt. Dass aber im besten Sinne einfaches, prägnantes Schreiben vor allem eine stilistische Anforderung ist, die keineswegs mit Qualitätsverlusten und Einbußen an intellektueller Brillanz verbunden sein muss, ist eine Einsicht, die vielleicht weiter verbreitet ist als man denkt – zumal die konsequente Umsetzung dieser Einsicht eben einfach schwierig ist.

Ein anderes Argument, das besonders von Autoren/innen immer wieder für längere Texte ins Feld geführt wird, ist das der unbegrenzten Kapazität des Netzes: Anders als Bücher und Zeitschriften, die schon aus Kostengründen auf eine bestimmte Bogenzahl festgelegt sind, fällt diese Beschränkung für Online-Formate weg. Gerade auch das Kostenargument ist zweifelsohne richtig. Speziell bei den *sehepunkten* ist der Streit um den anzustrebenden Umfang einer Buchbesprechung die beständige Begleitmusik der alltäglichen Redaktionsroutine. Keineswegs Prinzipienreiterei, sondern vielmehr der konkrete und oft bestätigte Erfahrungswert, dass Besprechungen für die Masse der Leser/innen eine rasche Orientierung auf dem Buchmarkt bieten und daher auch im Umfang überschaubare Textgrößen sein sollten, lässt daher den Richtwert von 7.500 Zeichen für Rezension als angemessen erscheinen (was je nach Seitenlayout zwei bis drei Druckseiten entspricht). Fraglos sind hier kulturelle Entwicklungen im Gange, die keineswegs abgeschlossen sind, sondern vielmehr erst an ihrem

Anfang stehen. Klar dürfte lediglich die Tendenz sein, wobei die Neuen Medien stets bewiesen haben, dass sie immer gut sind, um überraschende Wendungen auszulösen. Bislang wird man aber festhalten können, dass vor dem Neuen Lesen zunächst das Neue Schreiben kommt. Den Herausgebern/innen von Online-Publikationen wächst somit nolens volens die Aufgabe zu, ihre Autorenschaft zu angemessenem Schreiben zu „erziehen“. Dies hört sich paternalistischer an, als es gemeint ist und als es sich in der Realität vollzieht. Vielmehr handelt es sich um einen interaktiven Prozess zwischen den Autoren/innen und Redakteuren/innen (die ja selbst auch als wissenschaftliche Autoren/innen tätig sind), der ein gemeinsames Lernen und Entwickeln von adäquaten Möglichkeiten des Schreibens für das Internet darstellt.

Es mag sein, dass andere Online-Projekte, die ihren Formaten eine andere, offenere Struktur gegeben haben, diese Entwicklung schneller vorantreiben. Doch auch die *sehepunkte* und die *zeitenblicke* partizipieren, wenn auch unmerklich, an diesen Tendenzen. Wichtiger als das Prädikat, Schrittmacher einer neuen Schreib- und Lesekultur zu sein, bleibt für beide Online-Journale daher der Ansatz, die hohe Akzeptanz in der Fachwissenschaft zu erhalten. Auf ihre eigene, ja durchaus erfolgreiche Weise sind auch die *sehepunkte* und die *zeitenblicke* Teil der sich neu entwickelnden Kommunikationsformen im Netz.

Rechtsprobleme

Die geradezu universelle Verfügbarkeit hat beiden Online-Projekten sehr geholfen, ihren Platz in der Fachwelt zu finden. Hier zeigte sich die große Berechtigung des Open Access-Prinzips. Allerdings zeigte sich auch, dass Open Access vielfach dahingehend missverstanden wurde, alles im Netz sei frei, alles sei möglich. Tatsächlich sind die technischen Spielräume immens, doch werden viele Ideen von rechtlichen Beschränkungen eingeholt – oder zumindest von juristischen Bedenken, die erkennen lassen, dass sich Betreiber/innen wie Nutzer/innen der Online-Projekte aufgrund einer unscharfen Rechtssituation auf juristisch unsicherem Terrain bewegen. Es war teilweise einer im besten Sinne naiven Euphorie des Anfangs geschuldet, mitunter aber auch eines schlichtweg fehlenden Bewusstseins in den

Fachwissenschaften selbst, dass diese Aspekte allzu lange ausgeblendet blieben. Nicht nur, dass einzelne Projekte dadurch Schaden nehmen könnten, auch eine Idee wie die des Open Access kann auf diese Weise ungerechtfertigterweise in Misskredit geraten.

Die Zeit der Unschuld ist jedenfalls längst vorbei; Herausgeber/innen und Redaktion beider Online-Formate sind sich dessen sehr bewusst. Damit sind zwei Konsequenzen verbunden. Zum einen müssen die Journale selbst auf eine juristisch sichere Basis gestellt werden. Dazu gehören klare Lizenzen, die die Rechte der Journalbetreiber/innen wie der Autoren/innen klar definieren. Für die *zeitenblicke* reguliert zum Beispiel die DiPP-Lizenz (Digital Peer Publishing) diese Fragen.⁵

Zum anderen geht es darum, auf Seiten der Autoren/innen das Bewusstsein für die juristische Problematik zu wecken – allerdings ohne eine Panik dahingehend auszulösen, dass das Online-Publizieren ein rechtlich fragwürdiges Unternehmen sei: mühsam erworbenes Vertrauen wäre hier schnell verspielt. Es gibt jedoch genügend Beispiele aus der Praxis der Redaktion, die die Relevanz der juristischen Problematik erkennen lassen. Gerade für die *zeitenblicke* spielt dies eine Rolle, da hier die Möglichkeit der Publikation von Abbildungen eine große Rolle und auch einen großen Anreiz für die Autoren/innen darstellt. Dass die Rechte von Quellenmaterial bei Online-Publikationen eigens zu klären sind, gehört somit zu den alltäglichen redaktionellen Routinen – ebenso dass im Extremfall Autoren/innen einsehen müssen, dass nicht alle multimedialen Möglichkeiten rechtlich abgedeckt sind. Man muss sich wohl mit dem Gedanken vertraut machen, dass die Rechteproblematik noch viel weniger als das Bemühen um Vertrauen für die Online-Journale und das Streben nach Seriosität kaum in eine Phase fallen, die in absehbarer Zeit vorbei sein wird. Die Klärung der jeweils rechtlich verbindlichen Normen für verschiedene Features wird auch angesichts der weiter voranschreitenden technischen Möglichkeiten höchstwahrscheinlich eine bleibende Aufgabe für die Betreiber/innen der Online-Journale bleiben. Allerdings werden sich auch die Autoren/innen, die online publizieren, zumindest teilweise diesem

5 Vgl. <http://www.dipp.nrw.de/lizenzen/dppl/dppl/DPPL_v2_de_06-2004.html>.

Problemkomplex zuwenden müssen.

Pecunia nervus rerum – auch im Internet

Geld ist der Maßstab aller Dinge. Und auch Online-Publikationen kosten Geld – die Banalität dieser Tatsache muss vor den Hoffnungen gesehen werden, die man hegte, als das Publizieren im Netz eine deutlich kostengünstigere Variante zu Hardcopy-Publikationen zu sein versprach. Heute wird immer deutlicher, dass auch im World Wide Web angesiedelte Projekte ihren Preis haben. Dies gilt umso mehr, wenn sie auf Qualität setzen, ja ein wissenschaftliches Niveau anstreben. Es gilt also, sowohl für das Personal als auch die technischen Standards, eine Finanzierung sicherzustellen.

Beide Projekte sind derzeit noch im Genuss einer öffentlichen Förderung, doch völlig unklar ist die zukünftige Finanzierung. Ob die *sehpunkte* und die *zeitenblicke* stets mit öffentlichen Geldern rechnen können, ist nicht sicher. Es gilt daher, neue Finanzierungsmöglichkeiten auszuloten, allerdings unter größtmöglicher Wahrung des Open Access-Prinzips. Festzuhalten ist, dass der freie Zugang zur Online-Version eine Gebührenerhebung für gegenständliche Publikationsformen nicht grundsätzlich ausschließt. Infrage kämen etwa CDs der *zeitenblicke*-Ausgaben, zu denen wäre auch an Print-On-Demand-Funktionen für beide Online-Formate. Selbst die PDF-Version, die derzeit zu jedem Text der *sehpunkte* und der *zeitenblicke* frei verfügbar ist, könnte hier einbezogen werden. Sehr viel weiter als die Finanzierung von Post-Prints und Hybridversionen gehen Gedankenspiele, das komplette Angebot über Campus- und/oder Nationallizenzen zu verkaufen, so dass die einzelnen Wissenschaftler/innen und Interessenten/innen immer noch einen freien Zugang zu den Online-Inhalten haben.

Bei allen Überlegungen in diese Richtung darf gleichwohl nicht außer Acht gelassen werden, dass solche Finanzierungsmodelle in der Scientific Community akzeptiert werden müssen. Es ist ja weder der Wissenschaft allgemein noch den Online-Projekten konkret geholfen, ein Finanzierungsschema zu entwickeln, das in keiner Weise von den potenziellen Nutzern/innen angenommen wird. Damit dürften auch für die Geschichts-

aber sicher auch für die Geisteswissenschaften insgesamt Überlegungen hinfällig sein, die Autoren/innen für das Publizieren in den Online-Formaten zur Kasse zu bitten. Derartige in den Naturwissenschaften übliche Verfahren sind in den Geisteswissenschaften nicht realistisch. Schließlich gebietet aber der Realitätssinn einzugestehen, dass momentan unabhängig von gängigen Finanzierungsmöglichkeiten kaum ein Geschäftsmodell vorhanden ist, das eine völlig eigenständige Kostendeckung für die Online-Projekte gewährleisten kann. Auch wenn zumindest ein Teil der Unkosten aufgefangen werden könnte, wird eine Grundförderung oder zumindest eine solide Grundausstattung der Projekte mit Stellen und Finanzmodellen aller Voraussicht nach unvermeidbar sein. Unbestritten ist jedenfalls, dass ähnlich wie die Rechteproblematik auch die Finanzierungsfragen schon in naher und mittelfristiger Zukunft an Bedeutung für die Online-Journale gewinnen werden.

Fazit

Beide Online-Journale können zum jetzigen Zeitpunkt eine durchweg positive Bilanz ziehen. Sie haben aus schwierigen Anfängen heraus ihren Weg gemacht und sind zu etablierten und renommierten Publikationsorganen ihrer Fachwissenschaft geworden. Entscheidend dafür war sicher die spezifische Konzeption beider Journale, die die Vorteile des neuen Mediums mit vertrauten Elementen der gewohnten Publikationskultur kombinierte. Abzuwarten ist, inwieweit das Online-Publizieren auch die Wissenschaftskultur, konkret das Lesen und vor allem das wissenschaftliche Schreiben, verändern wird. Daneben wird aber auch immer deutlicher, dass eine ganze Reihe von Fragen ungelöst ist und die zukünftige Arbeit mittragen wird: der weite juristische Bereich sowie die Finanzierungsproblematik stellen Herausforderungen dar, denen sich die Online-Projekte verstärkt zuwenden müssen. Die Kunst wird darin bestehen, die Rechte- und Finanzierungsfragen in einer Weise zu klären, die das Renommee der Online-Journale keinen Schaden nehmen lässt und das Online-Publizieren an sich nicht diskreditiert.

Dr. Michael Kaiser ist als Wissenschaftlicher Mitarbeiter am Historischen Seminar der Universität zu Köln, Mitherausgeber der zeitenblicke und Redakteur bei den sehepunkten tätig. Forschungsschwerpunkte: Politische Kultur der Landstände in der Frühen Neuzeit, der Dreißigjährige Krieg, Geschichte der Gewalt, Militärgeschichte. E-Mail: michael.kaiser@uni-koeln.de

« ZEITHISTORISCHE FORSCHUNGEN /
STUDIES IN CONTEMPORARY HISTORY » –
ERFAHRUNGEN MIT HYBRIDEN PUBLIKATIONSKONZEPTEN

von Jan-Holger Kirsch

Der Beitrag stellt das Konzept der 2004 gegründeten Zeitschrift „Zeithistorische Forschungen/Studies in Contemporary History“ vor, die dreimal jährlich sowohl online als auch gedruckt (im Verlag Vandenhoeck & Ruprecht) erscheint. Der Grundgedanke dieser „hybriden“ Publikationsform ist es, die spezifischen Vorteile beider Medien zu nutzen. Die (frei zugängliche) Online-Version ist dabei keine bloße Wiedergabe der (kostenpflichtigen) Druckausgabe im Internet, sondern nutzt die medialen Möglichkeiten, die gerade für die Zeitgeschichtsforschung attraktiv sind (Integration visueller und audiovisueller Quellen, interne und externe Verlinkung, Aktualisierung und Kommentierung). Neben den Vorteilen nennt der Beitrag auch die Probleme dieses Publikationsmodells: sehr geringe Abonnentenzahl der Druckausgabe, sehr hoher Redaktionsaufwand, erhebliche Kosten für Nutzungsrechte (audio)visuellen Materials. Vorerst ist offen, wie tragfähig ein solches Modell ist.

In einem „Zwischenbericht“ zur Nutzung der Neuen Medien in der Geschichtswissenschaft nannte Gudrun Gersmann 1999 drei Hauptprobleme: „Mangel an Beständigkeit, Mangel an inhaltlicher Qualität und schließlich Mangel an methodischem Know-how“. Sie plädierte für „ein verstärktes Bemühen um Qualitätssicherung“.¹ Noch vier Jahre später meinten Fabio

¹ Gersmann, Gudrun, Neue Medien und Geschichtswissenschaft. Ein Zwischenbericht, in: Geschichte in Wissenschaft und Unterricht 50 (1999), S.

Crivellari und Marcus Sandl: „Ein relevanter Fachdiskurs über die Relevanz, Qualität und Standards der Kommunikation wissenschaftlicher Inhalte mit und in den Neuen Medien ist derzeit nicht auszumachen.“² Inzwischen ist manches in Bewegung gekommen, wie auch die Konferenz .hist 2006 belegt. Weiterhin ist aber Skepsis angebracht, ob der fachspezifische professionelle Umgang mit elektronischen Medien schon jenes Maß an Selbstverständlichkeit erreicht hat, das möglich und wünschenswert wäre.

Dafür ist wohl weniger ein abstrakter „Fachdiskurs“ entscheidend als die überzeugende forschungspraktische Umsetzung. Neue Medien im allgemeinen und Internet-Zeitschriften im besonderen werden den Mainstream der Geschichtswissenschaft nur dann erreichen und mittelfristig verändern können, wenn auch die arrivierten Fachvertreter/innen erkennen, dass die kompetente Nutzung neuer medialer Möglichkeiten ihnen signifikante Vorteile bringt. Dabei dürfte es sinnvoll sein, bewährte Forschungs- und Publikationspraktiken nicht einfach durch neue zu ersetzen, sondern Altes und Neues zu kombinieren. Das E-Journal *zeitenblicke*³ und das elektronische Rezensionsorgan *sehepunkte*⁴, die Michael Kaiser in dieser Sektion vorgestellt hat⁵, sind zwei Aktivitäten in dieser Richtung; die hybride Zeitschrift *Zeithistorische Forschungen / Studies in Contemporary History*⁶, deren Konzept ich im Folgenden erläutern möchte, ist ein in ähnlicher Absicht, aber mit etwas anderen Mitteln und Schwerpunkten gestaltetes Projekt. Dem Untertitel der Sektion „Elektronisches Publizieren in den Geisteswissenschaften“ folgend, werde ich *Erfahrungen, Probleme und Perspektiven* skizzieren.

239-249, hier S. 245, 247.

2 Crivellari, Fabio; Sandl, Marcus, Die Medialität der Geschichte. Forschungsstand und Perspektiven einer interdisziplinären Zusammenarbeit von Geschichts- und Medienwissenschaften, in: Historische Zeitschrift 277 (2003), S. 619-654, hier S. 621 Anm. 5.

3 Vgl. <<http://www.zeitenblicke.de/>>.

4 Vgl. <<http://www.sehepunkte.de/>>.

5 Vgl. Kaiser, Michael, Shepunkte und Zeitenblicke – Entwicklungslinien nach einem halben Jahrzehnt des elektronischen Publizierens, in diesem Band.

6 Vgl. <<http://www.zeithistorische-forschungen.de/>>.



Abbildung 1: Startseite von *Zeithistorische Forschungen*

Erfahrungen

Ähnlich wie *zeitenblicke* und *sehepunkte* ist *Zeithistorische Forschungen/Studies in Contemporary History* (im Folgenden: ZF/SCH) eine Neugründung, bei der die Rahmenbedingungen von den Initiatoren/innen relativ eigenständig festgelegt werden konnten. (Mit welchen Strategien etablierte geschichtswissenschaftliche [Print-]Zeitschriften sich den Herausforderungen des elektronischen Publizierens stellen können, wäre ein eigenes, hier nicht zu untersuchendes Thema.) Natürlich erfolgte die Zeitschriftengründung nicht in einem Vakuum, sondern innerhalb des vorhandenen akademischen Feldes und auf der Basis einer Marktanalyse der vorhandenen Angebote.

Unabhängig von der Publikationsform sind Fachzeitschriften ein Mittel, bestimmten Vorverständnissen und Akzentsetzungen innerhalb der jeweiligen Disziplin – hier der Zeitgeschichte – Geltung zu verschaffen. Ähnlich wie wissenschaftliche Tagungen und Institute dienen sie nicht nur dazu, neue Forschungsergebnisse zu produzieren und zu distribuieren, sondern erfüllen auch den Zweck der Ein- und Abgrenzung dessen, was

man als „relevant“ und „innovativ“ verstanden wissen möchte.⁷ So erhofft sich das Zentrum für Zeithistorische Forschung (ZZF)⁸ von der Gründung einer eigenen Zeitschrift zweierlei (wobei dies eher intuitiv als explizit und intentional eine Rolle spielte): erstens die Bestimmung und Verbreitung eines erweiterten Verständnisses von Zeitgeschichte, das neben der herkömmlichen Forschung auch die Rolle von Zeitgeschichte in der Öffentlichkeit sowie ihre visuellen und audiovisuellen Dimensionen berücksichtigen soll; zweitens Mithilfe bei der Festigung des eigenen Renommées und der langfristigen institutionellen Verstetigung des Zentrums. Zeitschriften erzeugen eine regelmäßige und oftmals stärkere Aufmerksamkeit als Monografien – insofern sind sie ein Mittel, bestimmte Forschungsinteressen und -schwerpunkte in den Vordergrund zu rücken.

Es ist indes allgemein bekannt, dass an geschichtswissenschaftlichen Zeitschriften nicht gerade ein Mangel herrscht. Bibliotheken leiden unter der Last der stetig steigenden Anschaffungskosten, Forscher/innen unter der oft erdrückenden Literaturmenge. So bedarf die Neugründung einer Zeitschrift nicht nur verlegerischer Risikobereitschaft, sondern auch starker inhaltlicher Argumente. Bei Beginn der Planungen für ZF/SCH war unser Eindruck, dass auf dem Markt für geschichtswissenschaftliche Journale durchaus noch eine Leerstelle war. Neben den allgemeinhistorischen Zeitschriften wie *Historische Zeitschrift* oder *Geschichte in Wissenschaft und Unterricht* gab und gibt es in Deutschland zahllose Spezial- und Richtungszeitschriften, zum Teil auch mit ausdrücklich zeithistorischem Schwerpunkt. Es fehlte jedoch ein Organ, das das erwähnte breite Verständnis von Zeitgeschichte hinreichend repräsentieren würde. Dabei ist es keineswegs unsere Absicht, die traditionsreichen und verdienstvollen *Vierteljahrshefte für Zeitgeschichte* (VfZ) zu verdrängen; dies wäre abwe-

7 Vgl. Middell, Matthias, Vom allgemeinhistorischen Journal zur spezialisierten Liste im H-Net. Gedanken zur Geschichte der Zeitschriften als Elementen der Institutionalisierung moderner Geschichtswissenschaft, in: Ders. (Hg.), *Historische Zeitschriften im internationalen Vergleich* (Geschichtswissenschaft und Geschichtskultur im 20. Jahrhundert 2), Leipzig 1999, S. 7-31; Raphael, Lutz, *Geschichtswissenschaft im Zeitalter der Extreme. Theorien, Methoden, Tendenzen von 1900 bis zur Gegenwart*, München 2003, S. 36-40.

8 Vgl. <<http://www.zzf-pdm.de/>>.

gig und für die Zeitgeschichte als wissenschaftliche Disziplin schädlich. Eine wechselseitige Ergänzung und partielle Konkurrenz kann dagegen nur förderlich sein.

Im ursprünglichen DFG-Antrag für ZF/SCH war lediglich von einer „elektronischen Zeitschrift“ die Rede. Frühzeitig gab es beim Verlag Vandenhoeck & Ruprecht⁹ aber Interesse, parallel dazu eine gedruckte Version zu publizieren. Die Entscheidung zwischen einer reinen Online-Zeitschrift und einer „hybriden“ Erscheinungsweise wurde institutsintern und im Verbund Clio-online eingehend diskutiert. Letztlich haben wir uns für die hybride Option entschieden, um die Vorteile beider Formate nutzen und kombinieren zu können. Anzunehmen war auch – und dies hat sich in der Praxis bestätigt –, dass potentielle Leser/innen und Autoren/innen auf diesem doppelten Weg noch besser erreicht werden können.

Winfried Schulze hat mit Recht geschrieben: „Wer die Papierzeitschrift einmal kennen gelernt hat und mit ihr groß geworden ist, und wenn auch nur in der Form der preiswerten Kopie, ist vermutlich nicht zum sofortigen Totalumstieg auf die Online-Version zu bewegen. Hinzu kommt, dass die Materialität des gedruckten Buches und der gedruckten Zeitschrift einen eigentümlichen Reiz ausmacht, der nicht außer Acht gelassen werden darf.“¹⁰ Wenn Schulze allerdings hinzufügt, „dass wir gerade im Bereich der Geisteswissenschaften *noch* eine Epoche des doppelten Mediums werden genießen können oder ertragen müssen“¹¹, dann klingt es so, als handle es sich dabei um ein vorübergehendes, zeitlich begrenztes Phänomen, das zur reinen Online-Zeitschrift als eigentlichem Zukunftsmodell hinführe. Dies ist jedoch offen, zumal nicht nur die Rezeptionsgewohnheiten der Historiker/innen, sondern auch die Verlagsstrategien im Wandel begriffen sind. Aus mediengeschichtlicher Sicht ist zudem

9 Vgl. <<http://www.v-r.de/>>.

10 Schulze, Winfried, Zur Geschichte der Fachzeitschriften. Von der „Historischen Zeitschrift“ zu den „zeitenblicken“, in: *Historical Social Research/Historische Sozialforschung* 29 (2004), Sonderheft: Neue Medien in den Sozial- und Geisteswissenschaften. Elektronisches Publizieren und Open Access: Stand und Perspektiven, hg. von Katja Mruck und Gudrun Gersmann, S. 123-137, hier S. 135.

11 Ebd. (meine Hervorhebung).

seit langem bekannt, dass neue Medien die alten nicht einfach verdrängen, sondern sie ergänzen und mit ihnen neue Konstellationen eingehen.¹² Insofern könnte das hybride Konzept von ZF/SCH durchaus eine längere Zukunft haben.

Doch was ist mit „hybriden Publikationskonzepten“ eigentlich gemeint? Wie unterscheiden sich die gedruckte und die elektronische Version? Unter einer „elektronischen Zeitschrift“ im engeren Sinne soll hier ausdrücklich nicht verstanden werden, dass PDF-Dateien der gedruckten Hefte bzw. einzelner Artikel unverändert ins Internet gestellt werden und dort abrufbar sind, sei es kostenfrei oder gegen Gebühr. Dies ist bei den meisten internationalen Fachzeitschriften längst gängige Praxis (und wurde kürzlich beispielsweise auch bei den VfZ eingeführt¹³). Dabei handelt es sich lediglich um „digitale Replikat“¹⁴, während die medienspezifischen Möglichkeiten der Internet-Publikation verschenkt werden und der Primat der Druckausgabe erhalten bleibt. Mit ZF/SCH bemühen wir uns demgegenüber, die Druck- und die Internetausgabe als eigenständige und gleichwertige Angebote zu profilieren. Im Folgenden erläutere ich primär die Online-Version, weil im vorliegenden Kontext ja das *elektronische* Publizieren diskutiert werden soll.

Geht man auf die Startseite der Zeitschrift¹⁵, so findet man das Titelbild der aktuellen Ausgabe. Von dort aus gelangt man zum Inhaltsverzeichnis dieses Heftes oder aber zu den anderen Bereichen der Hauptnavigation: Archiv, Impressum, Suche, Autorenhinweise. Steuern wir hier einmal die Inhaltsübersicht von Heft 3/2005 zum Leitthema „Migration“ an.¹⁶ Nach der Einführung zum Heftthema gibt es vier Rubriken, die in unterschiedlichen Besetzungen in jedem Heft wiederkehren: Aufsätze, Debattenbeiträge, Quelleninterpretationen und Besprechungen. Die Auf-

12 Vgl. Schildt, Axel, Das Jahrhundert der Massenmedien. Ansichten zu einer künftigen Geschichte der Öffentlichkeit, in: Geschichte und Gesellschaft 27 (2001), S. 177-206, hier S. 200.

13 Vgl. <<http://www.atypon-link.com/OLD/loi/vfzg?cookieSet=1>>.

14 Vgl. Keller, Alice, Elektronische Zeitschriften. Eine Einführung (Bibliotheksarbeit 9), Wiesbaden 2001, S. 24, 51.

15 Vgl. <<http://www.zeithistorische-forschungen.de/>>.

16 Vgl. <<http://www.zeithistorische-forschungen.de/16126041-Inhalt-3-2005>>.

satzrubrik bedarf keiner Erläuterung; dort werden ausführlichere Beiträge mit neuen Forschungsergebnissen zur Zeitgeschichte veröffentlicht. Die Debattenrubrik ist für eher essayistische Formate bestimmt, die sich mit aktuellen Fragen der öffentlichen und/oder wissenschaftlichen Diskussion beschäftigen (zum Beispiel mit dem Streit um die RAF-Ausstellung in Heft 2/2004 oder mit dem Verhältnis von Zeitgeschichte und Geschichtsdidaktik in Heft 2/2005).

Die Rubrik „Quellen“ folgt der Überlegung, dass wir es gerade in der Zeitgeschichte mit einer enormen Menge und Bandbreite von Quellen zu tun haben, die je nach Gattung und Überlieferungssituation nähere methodische Reflexionen erfordern. Die Quellenrubrik präsentiert gleichsam Fingerübungen in zeithistorischer Dokumenteninterpretation, wobei „Dokumente“ je nach Fragestellung vieles umfassen können – Akten, Reden, Fotos und Fotoreportagen, Filme, Bau- und Kunstwerke usw. Diese Beiträge sollen sowohl interessante Themen vorstellen als auch anhand von Beispielen methodische Standards entwickeln. So hat ein Beitrag in Heft 1/2004 Robert Lebecks Fotoreportagen für den *stern* in die Geschichte der Fotografie und zugleich in die Geschichte des Magazinjournalismus eingeordnet.¹⁷

Die Rubrik „Besprechungen“ geht insofern über das Spektrum der meisten anderen geschichtswissenschaftlichen Zeitschriften hinaus, als hier bewusst keine neuen Bücher rezensiert, sondern andere Medien unter die Lupe genommen werden – aktuelle Ausstellungen, Websites, DVDs, CD-ROMs, Filme usw. Unser besonderes Steckenpferd ist die in jedem Heft auftauchende und anderswo inzwischen aufgegriffene Unterrubrik „Neu gelesen“: Ältere Bücher – seien es viel zitierte oder in Vergessenheit geratene – werden aus heutiger Sicht neu betrachtet, was oft zu überraschenden Entdeckungen führt.¹⁸ Den Abschluss jedes Hefts bildet ein redaktionell zusammengestellter Digest von Buchrezensionen, der auf

17 Siebeneicker, Arnulf, „Ich fixiere, was ist“. Robert Lebecks Fotoreportagen für den *stern*, in: ZF/SCH 1 (2004), S. 111-121. Vgl. unter <<http://www.zeithistorische-forschungen.de/16126041-Siebeneicker-1-2004>>.

18 Beispiel: <<http://www.zeithistorische-forschungen.de/16126041-Schloegel-3-2005>>.

Angebote von H-Soz-u-Kult verweist – zum einen zum jeweiligen Heftthema, zum anderen zu wichtigen Neuerscheinungen der letzten Monate.¹⁹

Ich könnte nun näher ausführen, mit welchen Themen, die in anderen Fachzeitschriften so nicht vertreten sind, wir das Profil von ZF/SCH zu prägen versuchen. Im vorliegenden Diskussionszusammenhang ist es allerdings entscheidender, die Besonderheiten, Vorteile und Probleme des „hybriden“ Modells zu erläutern. Die kurz beschriebenen Rubriken sind ja erst einmal unabhängig davon, ob die Beiträge nun gedruckt oder im Internet veröffentlicht werden. Dies erleichtert auch die Produktionsabläufe: Für die Druck- und die Internetausgabe werden exakt dieselben Texte verwendet. Wo liegt dann der „Mehrwert“ oder, neutraler formuliert, das Proprium der Internetversion? Zum einen ist die interne und externe Verlinkung zu nennen. Mit einfachem Mausclick gelangt man zu den im Text erwähnten weiterführenden Informationen, beispielsweise zu den Websites von Museen und Archiven²⁰ oder auch zu Bild- und Tondokumenten, die aus urheberrechtlichen Gründen nicht für die eigene Website übernommen werden dürfen.

Solche Links, unter anderem auch zu den persönlichen Websites der Autoren/innen, sind wichtig, aber weder spektakulär noch besonders neuartig. Gestalterisch und konzeptionell zentraler ist die Möglichkeit, inhaltlich bedeutsame visuelle und audiovisuelle Medienelemente direkt in die Texte zu integrieren. Zwei Beispiele seien genannt: Auszüge aus einer Rede Willy Brandts in einem Aufsatz über Brandts Bedeutung als Medienkanzler²¹; und eine Kamerafahrt durch die rezensierte Ausstellung „John F. Kennedy“ des Deutschen Historischen Museums von 2003.²² Durch solche erweiterten Darstellungsformen wird gerade für die Zeitgeschichte eine neue Ebene der Interpretation und der Überprüfbarkeit möglich. Wenn etwa Politikerreden nicht allein als geschriebene Texte verwendet, sondern

- 19 Beispiel: <http://www.zeithistorische-forschungen.de/portal/alias_zeithistorische-forschungen/lang_de/tabID__40208478/DesktopDefault.aspx>.
 20 Vgl. etwa die Links in <<http://www.zeithistorische-forschungen.de/16126041-Baur-3-2005>>.
 21 Vgl. <<http://www.zeithistorische-forschungen.de/16126041-Muenkel-2-2004>>, Abschnitt 8.
 22 Vgl. <<http://www.zeithistorische-forschungen.de/16126041-Patel-1-2004>>.

als performative Akte erschlossen und nachvollzogen werden können, lässt sich ihre Wirkung viel besser abschätzen. Die Internet-Präsentation ist also keine technische Spielerei, die man mitmachen oder auch lassen kann; sie kann vielmehr die Wissenschaftlichkeit der Zeitgeschichtsforschung maßgeblich unterstützen.²³

Zwei weitere Gestaltungselemente seien genannt, die die Internetausgabe bietet: Erstens ist es unabhängig vom periodischen Erscheinungsrhythmus (drei Ausgaben pro Jahr) jederzeit möglich, Aktualisierungen vorzunehmen und Kommentare einzufügen. Die Autorentexte bleiben unangetastet, doch können weiterführende Informationen nach Redaktionsschluss eines Hefts ohne weiteres nachgetragen werden – so etwa zusätzliche Rezensionen zu Band 4 von Hans-Ulrich Wehlers „Deutscher Gesellschaftsgeschichte“, der in Heft 1/2004 diskutiert wurde.²⁴ Auch kritische Stellungnahmen etwa zu den Essays im Debattenteil können mit vergleichsweise geringem Aufwand ergänzt werden.²⁵ Zweitens eignet sich die Website von ZF/SCH als Plattform, um bereits anderswo veröffentlichte Aufsätze zu den jeweiligen Heftthemen als digitale Reprints zugänglich zu machen.²⁶ Ein solcher Service dürfte gerade für Zwecke der Lehre nützlich sein. Viele Verlage sind in der Frage der Abdruckrechte (noch) entgegenkommend, zumal der Ort der Erstveröffentlichung natürlich genannt ist. Allerdings haben größere Verlage die Vermarktung einzelner Aufsätze inzwischen als Einnahmequelle entdeckt, so dass digitale Reprints vielleicht nicht mehr lange möglich sein werden.

Nach dem bisher Gesagten könnte man meinen, dass die Druckausgabe von ZF/SCH nur eine Schwundstufe und eine Konzession an die ver-

- 23 Vgl. dazu auch Lindenberger, Thomas, *Vergangenes Hören und Sehen. Zeitgeschichte und ihre Herausforderung durch die audiovisuellen Medien*, in: ZF/SCH 1 (2004), S. 72-85. Vgl. <<http://www.zeithistorische-forschungen.de/16126041-Lindenberger-1-2004>>.
 24 Vgl. <http://www.zeithistorische-forschungen.de/portal/alias_zeithistorische-forschungen/lang_de/tabid_40208402/Default.aspx>.
 25 Beispiel: <http://www.zeithistorische-forschungen.de/portal/alias_zeithistorische-forschungen/lang_de/tabid_40208485/Default.aspx>.
 26 Beispiel: <<http://www.zeithistorische-forschungen.de/16126041-Material-Debatte-1-2005>>.

mutlich sehr geringe Zahl von Historikern/innen ohne Internetanschluss sei; die Links und multimedialen Elemente fallen in der Papierversion ja weg. Dem ist entgegenzuhalten, dass die Druckausgabe eine eigene Funktion und Dignität besitzt – zum einen wegen der bisher nicht völlig verlässlichen Langzeitarchivierung von Internetpublikationen, zum anderen wegen der bequemen Möglichkeit, längere Texte und gerade Themenhefte unabhängig vom Rechnerzugang lesen zu können, ohne sich mit ausgedruckten Einzelblättern herumärgern zu müssen (bei unseren Diskussionen im Vorfeld der Zeitschriftengründung machte dies als „S-Bahn-Argument“ die Runde). Darüber hinaus ist die Kooperation mit dem Verlag Vandenhoeck & Ruprecht gegenüber Lesern/innen und Autoren/innen ein weiterer Seriositätsindikator. Verlag und Redaktion achten auch besonders auf eine ansprechende, mit den Inhalten korrespondierende Gestaltung der Druckausgabe – durch ein übersichtliches Layout, moderne Schriften, neugierig machende Coverfotos usw. soll sich die Zeitschrift vom oft monotonen Erscheinungsbild anderer Fachzeitschriften abheben.

„Hybrid“ ist ZF/SCH schließlich noch in einer weiteren Hinsicht, die die Zeitschrift von anderen Fachorganen unterscheidet: Die enge Verbindung mit Clio-online²⁷ und speziell mit Zeitgeschichte-online²⁸ hat sowohl für die redaktionellen Arbeitsabläufe als auch für die Nutzer/innen viele Vorteile. Der oben erwähnte Rezensionsdigest ist ein Beispiel; ein anderes Beispiel ist die Möglichkeit, in der Zeitschrift begonnene Diskussionen im Portal Zeitgeschichte-online fortzuführen oder umgekehrt. So hat sich aus zwei Zeitschriftenbeiträgen zur RAF-Ausstellung²⁹ später ein eigener Themenschwerpunkt bei Zeitgeschichte-online entwickelt³⁰; beide Angebote verweisen wechselseitig aufeinander. Dadurch, dass im größeren Zusammenhang von Clio-online mehrere Publikationswege zur Verfügung stehen (zu nennen ist hier auch die Reihe Historisches Forum³¹), ist trotz

27 Vgl. <<http://www.clio-online.de/>>.

28 Vgl. <<http://www.zeitgeschichte-online.de/>>.

29 Vgl. <<http://www.zeithistorische-forschungen.de/16126041-Kirsch-2-2004>> und <<http://www.zeithistorische-forschungen.de/16126041-Kraushaar-2-2004>>.

30 Vgl. <<http://www.zeitgeschichte-online.de/md=RAF-Inhalt>>.

31 Vgl. <<http://edoc.hu-berlin.de/histfor/impressum.php>>.

des begrenzten Umfangs der einzelnen Segmente auch eine größere Aktualität möglich als bei herkömmlichen Zeitschriften.

Probleme

Die praktischen Schwierigkeiten des geschilderten Publikationsmodells seien ebenso offen angesprochen wie die Vorteile. Das Hauptproblem, mit dem sich alle Redaktionen herumschlagen, ist ganz unabhängig vom gewählten Publikationskonzept: säumige Autoren/innen, inhaltlich und sprachlich mangelhafte Texte, überlastete Gutachter/innen. Der Produktionsdruck, unter dem Historiker/innen stehen oder zu stehen meinen, hat in den letzten Jahren weiter zugenommen, und die Zahl der Publikationen verhält sich mitunter umgekehrt proportional zu ihrer Qualität. Umso wichtiger werden Redaktionen als Vermittlungs- und Beratungsinstanzen, die bei der Textoptimierung helfen, manche von den Autoren/innen gewünschten Publikationen aber auch schlicht verhindern müssen, um die Scientific Community vor „Datenmüll und Infosmog“ zu schützen.³²

Ein Problem, das sich dagegen speziell bei unserer Hybridpublikation stellt, ist der geringe Kaufanreiz für die Druckausgabe. Es ist kein Geheimnis, dass die Abonnentenzahlen bislang weder befriedigend noch für den Verlag kostendeckend sind. Zwar haben nahezu alle einschlägigen Institute und auch etliche Privatleute ZF/SCH in gedruckter Form abonniert, aber solange sämtliche Inhalte kostenfrei im Netz abrufbar sind, sehen viele Bibliotheken angesichts knapper Etats keine wirkliche Notwendigkeit für ein Abonnement. Dies hatten wir zumindest für die ersten Jahre auch nicht anders vermutet, und man wird sehen müssen, wie sich die Nachfrage weiter entwickelt. Erfreulicherweise hat der Verlag Vandenhoeck & Ruprecht stets signalisiert, dass er von diesem Projekt keine kurzfristige Rendite erwartet, sondern es als Zukunftsinvestition und lohnendes Experiment betrachtet.

Schwierigkeiten bringt die Druckausgabe auch insofern mit sich, als

32 Vgl. Kirsch, Jan-Holger, Datenmüll und Infosmog. Über einige Hindernisse (kultur)wissenschaftlichen Erkenntnisgewinns heute, in: Potsdamer Bulletin für Zeithistorische Studien Nr. 32/33 (Nov./Dez. 2004), S. 9-18. Vgl. <http://www.zzf-pdm.de/bull/pdf/b3233/3233_kirsch.pdf>.

der Umfang aus verlegerischen Gründen auf 480 Seiten pro Jahrgang limitiert ist. Dies ist trotz aller Umfangskalkulationen im Vorfeld sehr schwer abzuschätzen, da einzelne Beiträge aus guten inhaltlichen Gründen länger geraten oder aber kurzfristig ganz ausfallen können. Für die Internetausgabe ist dies kein prinzipielles Hindernis, für die textidentische Druckausgabe dagegen sehr. Mit Hilfe von mehr oder weniger Bildmaterial muss versucht werden, den Umfang der Vorgabe anzupassen. Die Beschränkung ist andererseits ein Vorteil, da die Autoren/innen bestimmte Zeichengrenzen beachten müssen, was der Prägnanz der Texte durchaus zugute kommt. Heikler als die Umfangsbegrenzung ist es, dass die Vorbereitung der Druckausgabe – die Redaktion liefert eine komplette PDF-Datei des Heftes als Druckvorstufe an den Verlag – einen erheblichen Teil der Arbeitskraft bindet, so dass die Internetausgabe mitunter nicht so detailliert mit Zusatzangeboten bestückt werden kann, wie es inhaltlich sinnvoll und technisch möglich wäre.

Ein anderes, sowohl arbeitsökonomisches als auch finanzielles Problem sind die Nutzungsrechte für visuelles und audiovisuelles Material, das für das Konzept der Zeitschrift ja sehr zentral ist. Die Recherche nach geeigneten Dokumenten ist eigentlich Aufgabe der Autoren/innen, bleibt aus verschiedenen Gründen aber nicht selten an der Redaktion hängen. Hinzu kommt, dass die Kosten etwa für „Ikonen der Zeitgeschichte“, also für sehr bekannte und historisch aufschlussreiche Fotos, oft immens sind, zumal die Nutzung bei Bildagenturen sowohl für den Druck als auch für das Netz ausgehandelt werden muss. Wenn es gelingt, einen akzeptablen Paketpreis zu erzielen, stellt sich als nächstes Problem die zeitliche Befristung der Internetrechte. Etliche Dokumente müssen nach ein bis fünf Jahren von der Website entfernt oder aber nachvergütet werden – ein noch kaum abschätzbares finanzielles Risiko. Bei Quellen der Rundfunk- und Fernsehgeschichte haben die bisherigen Erfahrungen zudem gezeigt, dass die Sender bzw. Archive der Internetpublikation meist ablehnend gegenüberstehen, weil sie eine unkontrollierbare Weiterverteilung ihres auch kommerziell interessanten Materials befürchten. Hier ist noch viel Überzeugungsarbeit zu leisten und eine spezielle Vereinbarung für Forschungszwecke auf möglichst hoher Entscheidungsebene anzustreben.

Ein eher technisches Problem ist zumindest derzeit noch die Aktualität und Aktualisierung der externen Links. Bekanntlich ändern sich viele Internetadressen in sehr kurzen Rhythmen oder verschwinden ganz. In dem Maße, wie unser Archiv³³ weiter anwächst, erhöht sich auch die Gefahr, dass Links ins Leere bzw. zu Fehlermeldungen führen. Wir hoffen aber, dass sich dies in absehbarer Zeit durch einen automatisierten URL-Check vermeiden lässt. Inhaltlich bedeutsamer ist natürlich die Ausgestaltung der einzelnen Rubriken und Beitragsformate. Hier besteht das Problem, dass vielen Autoren/innen die hybride Publikationsweise noch nicht vertraut ist; die Erstfassungen der Manuskripte entsprechen meist den Konventionen, die die Autoren/innen aus dem Bereich der reinen Printmedien kennen. Die parallele Internetpublikation erfordert aber teilweise eine andere Textorganisation und vor allem das Mitbedenken der audiovisuellen Möglichkeiten. Auch wenn darauf in unseren Richtlinien zur Manuskriptgestaltung stets hingewiesen wird, bedarf es meist zusätzlicher Absprachen und redaktioneller Recherchen, um die Ausgangstexte für das Internet medienadäquat aufzubereiten. Zudem sind Beitragsformate wie Essays, Quelleninterpretationen, Ausstellungsbesprechungen etc. für viele Autoren/innen ungewohnt, selbst wenn sie Experten/innen für das jeweilige inhaltliche Thema sind. Um die Autoren/innen für die Pragmatik der genannten Textgattungen zu sensibilisieren, ist eine erhebliche Menge an Korrespondenz erforderlich, die über das sonst übliche Maß an redaktioneller Betreuung hinausgeht. Dies wird von den Autoren/innen aber durchaus honoriert, und auf lange Sicht wird sich die Forschung selbst durch die neuartigen Publikationsmöglichkeiten verändern – darauf hat Wolfgang Schmale zur Eröffnung der Konferenz .hist 2006 mit Recht hingewiesen.³⁴ Damit

33 Vgl. <<http://www.zeithistorische-forschungen.de/site/40208121/default.aspx>>.

34 Schmale, Wolfgang, Geschichte im Netz – Praxis, Chancen, Visionen, in Teilband 1: <http://edoc.hu-berlin.de/histfor/10_I/>. Ein Beispiel für veränderte Forschungsstrategien ist der Aufsatz von Paul, Gerhard, Die Geschichte hinter dem Foto. Authentizität, Ikonisierung und Überschreibung eines Bildes aus dem Vietnamkrieg, in: ZF/SCH 2 (2005), S. 224-245. Vgl. <<http://www.zeithistorische-forschungen.de/16126041-Paul-2-2005>>. Der Text beruht zu einem erheblichen Teil auf Bildrecherchen im Internet und macht diese durch Angabe der relevanten URLs auch nachvollziehbar.

komme ich abschließend zu den Perspektiven.

Perspektiven

„Die Wissenschaftler trauen der Online-Publikation nicht“, meinte der Wissenschaftsjournalist Dieter E. Zimmer im Jahr 2001. Er fügte hinzu, das Internet sei „der richtige Ort für die fluide Phase der wissenschaftlichen Information, nicht für die kristallisierte“.³⁵ Das Misstrauen ist zumindest unter Historikern/innen geringer geworden, nicht zuletzt deshalb, weil die Betreiber von Internetportalen und -journalen inzwischen überwiegend etablierte Forschungsinstitutionen sind, die die Bedeutung und den Nutzen des Internet für geschichtswissenschaftliche Zwecke erkannt haben. Zimmers zweite These ist ebenfalls überholt, und zwar in doppelter Hinsicht: Zum einen erweist sich das Internet durchaus als ein geeigneter Ort für „kristallisierte“, also fachlich begutachtete und sedimentierte Information. Zum anderen ist die Abgrenzung zwischen der „fluiden“ und der „kristallisierten“ Phase so nicht mehr haltbar; wie ich anhand der Aktualisierungen und Kommentare zu unserer Zeitschrift gezeigt habe, ist es gerade ein Vorteil des Internet, dass gefestigtes Wissen nicht starr bleiben muss, sondern fortgeschrieben und fluide gehalten werden kann – offen für Bestätigung, Korrektur oder Ergänzung. (Allerdings stellt sich eine solche Interaktivität nicht von selbst ein, sondern muss ebenso organisiert und moderiert werden wie die Ausgangsbeiträge.) Fragt man nun nach den Perspektiven speziell für das hybride Publikationskonzept von ZF/SCH – ein Konzept, für das es zumindest in der Geschichtswissenschaft keine unmittelbaren Vergleichsbeispiele gibt –, so sind idealtypisch drei Optionen zu nennen:

1. Verzicht auf die Internetausgabe, Beschränkung auf die Druckausgabe. Dies würde den Arbeitsaufwand deutlich reduzieren, ist aber keine ernsthaft zu erwägende Option, weil dann alle oben genannten Vorteile

35 Zimmer, Dieter E., Analoge und digitale Informationsmedien, in: Ansorge, Kathrin (Red.), Wissenschaftspublikation im digitalen Zeitalter. Verlage, Buchhandlungen und Bibliotheken in der Informationsgesellschaft. Vorträge eines Symposiums am 8. und 9. Februar 2001 in Berlin (Gesellschaft für das Buch 7), Wiesbaden 2001, S. 129-136, hier S. 133, 135.

der Online-Publikation und ihre Auswirkungen für ein zeitgemäßes Verständnis von Zeitgeschichte wegfallen würden. Selbst die Abonnentenzahl der Druckausgabe würde wahrscheinlich nicht signifikant steigen.

2. Verzicht auf die Druckausgabe, Beschränkung auf die Internetausgabe. Dies würde bedeuten, dem Publikationsmodell der *zeitenblicke* zu folgen und ganz konsequent medienadäquate Beiträge veröffentlichen zu können. Die Beachtung innerhalb der Zeitgeschichtsforschung wäre vermutlich aber geringer, und man würde die Vorteile der Druckausgabe verschenken.
3. Fortsetzung der Hybridpublikation. Das analoge und das digitale System haben einen je eigenen Mehrwert³⁶, und derzeit scheint es uns die beste Option zu sein, dieses Experiment fortzuführen. Dabei wird allerdings zu prüfen sein, ob die Devise des Open Access für die Internetausgabe weiter tragfähig ist. Gewiss sprechen gute Argumente für das Prinzip der freien Zugänglichkeit, aber es ist eine offene Frage, ob es sinnvoll und machbar ist, hochwertige Inhalte ohne Refinanzierung zu verbreiten, obwohl es bei den Nutzern/innen möglicherweise eine Zahlungsbereitschaft gibt. Würde man den Weg einer Teilkommerzialisierung beschreiten, ergeben sich daraus andererseits viele Folgeprobleme – besonders für öffentlich finanzierte Institute wie das ZZf. Dieser Diskussion werden sich die Geisteswissenschaften in den nächsten Jahren verstärkt stellen müssen. Publikationsmodelle der STM-Disziplinen (Science, Technology, Medicine) dürften dabei kaum übertragbar sein, weil die fachspezifischen Anforderungen und Möglichkeiten der Geisteswissenschaften deutlich anders gelagert sind.

Sicher ist indes, dass die personellen und technischen Fixkosten weder über Abonnements noch über Bezahlfunktionen im Internet, sondern nur durch institutionelle Förderung gedeckt werden können – dies verbindet fast alle geisteswissenschaftlichen Fachzeitschriften unabhängig von ihrer Veröffentlichungsform. Im Übrigen ist dies nicht Ausdruck einer „Krise“

36 So auch Schmale, *Geschichte im Netz* (wie Anm. 34).

des wissenschaftlichen Publizierens, sondern eher der Normalzustand. Nach einer Vielzahl von Zeitschriftenneugründungen der letzten Jahre und Jahrzehnte ist in Zukunft mit einem gewissen Konzentrationsprozess zu rechnen, zumal es sich als Irrtum erwiesen hat, dass qualitativ hochwertige Internetzeitschriften per se kostengünstiger seien als gedruckte Journale.

Dr. Jan-Holger Kirsch ist Wissenschaftlicher Mitarbeiter am Zentrum für Zeithistorische Forschung, verantwortlicher Redakteur von „Zeithistorische Forschungen/Studies in Contemporary History“ und in der Redaktion von „Zeitgeschichte-online“ tätig. Interessenschwerpunkte: Geschichte des 20./21. Jahrhunderts, Erinnerung an den Nationalsozialismus, Geschichte in der Öffentlichkeit, Geschichtstheorie und -didaktik. E-Mail: kirsch@zeitgeschichte-online.de

ZUR AKZEPTANZ DES ELEKTRONISCHEN PUBLIZIERENS IN DEN GEISTESWISSENSCHAFTEN – ERGEBNISSE UND FOLGEN DER DFG-STUDIE „PUBLIKATIONSSTRATEGIEN IM WANDEL“

von Johannes Fournier

Die an jegliche Veröffentlichung gebundene Funktion der Zuteilung wissenschaftlichen Renommées ist für elektronische Publikationen vor allem in den Geistes- und Sozialwissenschaften noch nicht selbstverständlich. Daten zur Akzeptanz elektronischer, insbesondere entgeltfrei über das Internet zugänglicher Publikationen (Open Access) hat die Deutsche Forschungsgemeinschaft (DFG) Ende 2004 in einer Umfrage erheben lassen, in der auch die von der DFG geförderten Fachgebiete der Geistes- und Sozialwissenschaften gut vertreten waren. Die Analyse dieser Daten lässt erkennen, welche Faktoren für Zuspruch oder Kritik an elektronischen Publikationen ausschlaggebend sind.

Zur Akzeptanz elektronischer Publikationen

Das Internet ist aus dem Alltag auch der Geisteswissenschaften heute nicht mehr wegzudenken. Die Suche nach einschlägiger Literatur, die Fachkommunikation mit den Kollegen/innen, die Lektüre von Tagesankündigungen, selbst die Rezeption von Literatur finden zunehmend häufiger in virtuellen Welten statt. Den Schritt von der Rezeption elektronischer Inhalte zum elektronischen Publizieren gehen viele Geisteswissenschaftler/innen hingegen noch zögernd, obwohl die Distribution von Forschungsergebnissen über das Internet einen weltweiten Leserkreis und damit höchste Sichtbarkeit für die Autoren/innen verspricht. Dass elektronische Publikationen derzeit nicht in allen Fächern unbeschränkte Aner-

kennung finden und dass manche Disziplinen heute noch eine gewisse Zurückhaltung üben, was die völlige Akzeptanz elektronischer Publikationen angeht, ist nur zu verstehen, wenn man sich vergegenwärtigt, dass jeder Publikation mindestens vier verschiedene Funktionen – Verbreitung von Information, Qualitätskontrolle, Kanonbildung und Zuweisung von Reputation – zukommen; darunter dürfte die Zuweisung von Renommee die für Autoren/innen wichtigste Funktion sein.¹ Nur ist die an jegliche Publikation geknüpfte Funktion der Zuteilung wissenschaftlichen Renommee für elektronische Publikationen noch nicht selbstverständlich. Andererseits lässt sich bereits heute absehen, dass die vielfältigen Transformationen im Prozess des wissenschaftlichen Arbeitens – die Kommunikation per E-Mail, der Rückgriff auf digitale Quellen, die Möglichkeit der unmittelbaren Einbindung auch multimedialer Komponenten in Publikationen – auf längere Sicht deutliche Auswirkungen auf die Publikationskultur haben werden: Dass kommunikative Prozesse, die im Vorfeld jeglicher wissenschaftlicher Publikation stattfinden, zunehmend stärker über elektronische Medien abgewickelt werden, führt zu einer zunehmend größeren Vertrautheit mit dem Internet, aus der wiederum eine zunehmende Akzeptanz auch elektronisch publizierter Forschungsergebnisse resultiert. Damit wird sich die elektronische Publikation à la longue auch in den Bereichen der Geisteswissenschaften, die Netz-Veröffentlichungen heute noch überwiegend kritisch sehen, völlig etablieren.

Als Organisation der Forschungsförderung unterstützt die Deutsche Forschungsgemeinschaft (DFG) satzungsgemäß auch den Austausch wissenschaftlicher Erkenntnisse. Daher liegt der DFG daran, den eingangs in aller Kürze charakterisierten Transformationsprozess zu unterstützen, zu begleiten und aktiv zu gestalten. Folglich hatte die DFG schon vor der Unterzeichnung der „Berliner Erklärung über offenen Zugang zu wissenschaftlichem Wissen“ im Herbst 2003 eine Reihe von Projekten gefördert²,

1 Vgl. Rowland, Fytton, *Print Journals: Fit for the Future?*, in: *Ariadne* 7 (1997); <<http://www.ariadne.ac.uk/issue7/fytton/>>: „The fourth function mentioned above (recognition of authors) should be seen by researchers as the most important function of academic journals“.

2 Vgl. den Text der Erklärung unter <<http://www.zim.mpg.de/openaccess-berlin/>

in denen sowohl Quellen als auch aktuelle Forschungsliteratur über das Internet für den freien Zugriff verfügbar gemacht wurden. Mit der Unterschrift unter die „Berliner Erklärung“ jedoch stand die Frage, in welcher Weise die DFG das Prinzip des Open Access unterstützen sollte, erneut zur Diskussion. Um sich Klarheit darüber zu verschaffen, wie auf dem Weg zu einer Kultur möglichst entgeltfrei zugänglicher Publikationen voranzuschreiten sei, hatte die DFG sich entschlossen, zunächst mit einer Umfrage unter DFG-geförderten Wissenschaftlern/innen zu ermitteln, welche Bedeutung elektronischen Publikationen, insbesondere im Open Access, in den verschiedenen Wissenschaftsfeldern zukommt.³ Bevor die Ergebnisse dieser Umfrage im Folgenden mit einem speziellen Blick auf den Rücklauf aus den Geisteswissenschaften vorgetragen werden, möchte ich jedoch zunächst einen Blick auf einige aktuelle Entwicklungen lenken, um den Kontext aufzuzeigen, in dem die Frage nach der – womöglich stärkeren – Akzeptanz elektronischer Publikationen meines Erachtens heute verortet werden muss. Abschließend sollen einige Maßnahmen skizziert werden, mit denen die DFG versucht, die Akzeptanz elektronischer Publikationen überhaupt und insbesondere in den Geisteswissenschaften aktiv zu fördern.

berlindeclaration.html>. – Mit ihrem Förderinstrument „Retrospektive Digitalisierung von Bibliotheksbeständen“ hat die DFG bereits seit 1997 in großem Stil dazu beigetragen, insbesondere für die Geistes- und Sozialwissenschaften relevante Quellen entgeltfrei über das Internet verfügbar zu machen (eine Liste der geförderten Projekte findet sich unter <<http://www.hki.uni-koeln.de/retrodig/index.html>>). Mit Projekten wie „Dissertationen Online“ (<<http://www.dissonline.de>>), „German Academic Publishers“ (<<http://www.gap-c.de>>) oder „SCOPE – Service Core for Open Publishing Environments“ (<http://edoc.hu-berlin.de/e_projekte/scope/kontakt.php>) wurden Infrastrukturen und Werkzeuge für das elektronische Publizieren aufgebaut. Schließlich wurden z.B. für „German Medical Science“ (<www.egms.de>), „Forum Qualitative Sozialforschung“ (<<http://www.qualitative-research.net/fqs/fqs.htm>>) oder „sehpunkte. Rezensionjournal für die Geschichtswissenschaft“ (<<http://www.sehpunkte.historicum.net/>>) Starthilfen zum Etablieren elektronischer Zeitschriften gewährt.

3 Deutsche Forschungsgemeinschaft, *Publikationsstrategien im Wandel? Ergebnisse einer Umfrage zum Publikations- und Rezeptionsverhalten unter besonderer Berücksichtigung von Open Access*, Bonn 2005. – Diese Publikation ist online zugänglich unter <www.dfg.de/lis/openaccess>.

Kommunikation und Publikation

Die Zeiten, in denen Netz-Publikationen als „wissenschaftliche Eintagsfliegen“ galten, haben wir inzwischen wohl hinter uns gelassen. Die Bedeutung elektronischer Publikationen für ein großes geisteswissenschaftliches Fachgebiet wie die Geschichte wird schon durch die schiere Zahl der Teilnehmer/innen an der Tagung .hist 2006 unterstrichen. Darüber hinaus scheint mir klar, dass die Bereitschaft, elektronische Publikationen für die eigene Forschung heranzuziehen oder auch selbst elektronisch zu publizieren, keine Frage allein der Generation ist. Doch ist man gut beraten, deutlich zwischen virtueller Kommunikation und elektronischer Publikation zu unterscheiden, da beide Phänomene sich mit unterschiedlicher Geschwindigkeit entwickeln: Während die Kommunikation über das Internet sich in rasantem Tempo ausbreitet, nimmt die Zahl wissenschaftlicher elektronischer Publikationen hingegen langsamer zu. Im Folgenden möchte ich mit wenigen Beispielen illustrieren, wie selbstverständlich der Umgang mit dem Medium Internet vielfach bereits wurde; dabei soll der Weg von der Diskussion zur Publikation nachgezeichnet werden, um die Beispiele beiden Sphären der „elektronischen Wissenschaft“ – der Kommunikation und der Publikation – zu entnehmen.

Vielfach dienen Seminare dazu, wissenschaftliche Arbeiten zu einem bestimmten Thema zu vertiefen und eigene Thesen vor der Veröffentlichung zu verifizieren. Heute bieten zudem Wikis und vor allem Weblogs als eine Art „virtueller Teeküchen“ die Möglichkeit, Ideen gleich weltweit hinterfragen und diskutieren zu lassen.⁴ Daher ist es nur konsequent, wenn Weblogs als neue Form der Kommunikation Einzug in die Wissenschaft halten.⁵ Als Autoren/innen von Weblogs nutzen Wissenschaftler/innen konsequent die (Rück-)Entwicklung des Internet zu einem „social web“

4 Vgl. Butler, Declan, Science in the web age: Joint Efforts, in: Nature 438 (1.12.2005), S. 548f. – Wie auch andere Formen der elektronischen Publikation können Weblogs zudem zu einer gesteigerten Sichtbarkeit ihrer Autoren/innen und auf diese Weise zu deren größerem Renommee beitragen.

5 Die dem Seminar ähnliche Funktion scheint mir dadurch bestätigt, dass, soweit ich sehe, Weblogs ebenso wie (geschichtswissenschaftliche) Wikis in erster Linie zur Unterstützung der akademischen Lehre eingesetzt werden, vgl. etwa die unter „Earthshine“ (<<http://www.earthshine.org/>>) zugänglichen Seiten.

und adaptieren neue Technologien, um ihre Thesen weltweit diskutieren zu lassen. Im deutschsprachigen Raum scheint diese Entwicklung allerdings noch am Anfang zu stehen. Eine über das Wissenschaftsportal vascoda ausgeführte Suche nach dem Suchwort „Blog“ im Fach „Geschichte“ führt zu dem über Clio-online nachgewiesenen, von Klaus Graf gepflegten Weblog „Archivalia“; bei zwölf weiteren, über chronicon ermittelten Treffern war mir meist nicht unmittelbar klar, ob und wie sie das Fach betreffen.⁶ Und offenbar misstraut die Zunft der als wenig seriös empfundenen Kommunikation im Blog: In Jan Hodels „Hist.net.blog“ ist nachzulesen, wie Fachvertreter/innen, als er unter anderem Weblogs und Wikis im Rahmen einer fachdidaktischen Tagung als Instrumente kollaborativen geschichtswissenschaftlichen Arbeitens präsentierte, den Autor mahnten, wieder ‚mehr Bodenhaftung‘ zu erlangen.⁷ Ich gehe dennoch davon aus, dass Weblogs, obwohl sie bislang keine signifikante Verbreitung in der deutschen akademischen Welt gefunden haben, in Zukunft verstärkt zum Einsatz kommen werden, da sie kollaboratives Arbeiten optimal unterstützen können. Dabei dürften sich genauere Funktionsbereiche herauschälen, die diesen Instrumenten ihre Rolle neben anderen, längst etablierten Formen der Wissenschaftskommunikation zuweisen.

Während Beiträge in Wikis und Weblogs dem Bereich der informellen akademischen Kommunikation zugerechnet werden müssen, der insbesondere unter dem wichtigen Aspekt des Erwerbs von Renommee solange keine große Rolle spielen wird, bis auch hier gewisse Standards für die Qualitätssicherung etabliert werden, kann die Vorabveröffentlichung wissenschaftlicher Forschungsergebnisse in Form von – wiederum nicht qualitätsgeprüften – Preprints bereits als unmittelbare Publikationsform gelten. Wie inzwischen weit bekannt ist, spielt diese Publikationsform insbesondere in der Physik und Mathematik eine große, aber auch in den

6 Vgl. <<http://www.vascoda.de/>>.

7 Vgl. <<http://hodel-histnet.blogspot.com/2006/01/hok-lesenschreiben-einstieg.html>>. „Zurecht mahnte mich die Runde [i.e. die Teilnehmer/innen einer Tagung „Geschichte lehren an der Hochschule“, Universität Paderborn, 8./9.09.2005] an die Praxistauglichkeit der Vorschläge“ (<<http://hodel-histnet.blogspot.com/2005/09/bloggen-und-geschichte-oder-warum.html>>).

Wirtschaftswissenschaften (in der Form sogenannter „working papers“) eine nicht unbedeutende Rolle. Demgegenüber wird die Praxis, Forschungsergebnisse vor der eigentlichen Publikation über das Datennetz zu verbreiten, in den Geisteswissenschaften nur sehr zurückhaltend geübt. Doch mag hier ein Prozess des Umdenkens in Gang sein. Ende des Jahres 2005 annoncierten die Universitäten Princeton und Stanford, dass sie unter dem Namen „Princeton/Stanford Working Papers in Classics“ ein gemeinsames Preprint-Archiv für die Altertumswissenschaftler/innen dieser Universitäten eröffnet haben⁸; in diesem Archiv werden Publikationen gezielt veröffentlicht, um einen Diskussionsprozess in Gang zu bringen, dessen Ergebnisse nach Möglichkeit vor der endgültigen Publikation altertumswissenschaftlicher Arbeiten berücksichtigt werden sollen. Es scheint, als ob die Publikationskultur insbesondere der Physiker/innen in diesem Fall die der Geisteswissenschaftler/innen beeinflusst hat.

Für die „eigentliche“ Publikation, also die unmittelbare Veröffentlichung von Forschungsergebnissen in referierten Zeitschriften, spielen elektronische Open Access Zeitschriften wohl gerade in den Geisteswissenschaften eine Rolle: „Among the humanities, history is at the forefront when it comes to electronic full text journals, based in particular in the US and the UK“, heißt es bei Michael Nentwich.⁹ Das inzwischen wohl etwas bekanntere, an der Universität Lund gepflegte *Directory of Open Access Journals* (DOAJ) weist unter der Rubrik „Geschichts- und Altertumswissenschaften“ elf Open Access Zeitschriften aus der Archäologie, 76 aus den Geisteswissenschaften und einen Titel aus der Diplomatie im engeren Sinne nach.¹⁰ Lediglich acht der dort verzeichneten geschichtswissenschaftlichen Zeitschriften sind deutschsprachig. Dieses Bild ändert sich allerdings beträchtlich, wenn zur Analyse geschichtswissenschaftlicher Open Access Zeitschriften nicht das DOAJ, sondern die *Elektronische Zeitschriftenbibliothek* (EZB) der Universität Regensburg

8 Princeton/Stanford Working Papers in Classics, vgl. <<http://www.princeton.edu/~pswpc/>>.

9 Nentwich, Michael, *Cyberscience. Research in the Age of the Internet*, Wien 2003, S. 115.

10 Vgl. <<http://www.doaj.org>>.

herangezogen wird.¹¹ Insgesamt weist die EZB ein Vielfaches der über DOAJ verfügbaren Open Access Zeitschriften nach. Circa 40 Prozent der in der EZB verzeichneten 25.302 Titel sind über das Internet entgeltfrei zugänglich. Von den 1.048 für die Geschichtswissenschaft nachgewiesenen Zeitschriften sind 479 entgeltfrei über das Netz verfügbar, darunter 78 deutschsprachige Titel.

Eine gewisse Zurückhaltung dem Internet gegenüber deutet sich schließlich hinsichtlich so genannter Sekundärpublikationen an, also wenn untersucht wird, wie viele der bereits auf herkömmliche Weise publizierten Arbeiten von ihren Autoren/innen zusätzlich zur Verlagsveröffentlichung im Internet bereitgestellt werden. Eine vergleichende Untersuchung zu so genannten „institutionellen Repositorien“ hat zu dem interessanten Ergebnis geführt, dass die Repositorien australischer oder italienischer Universitäten fast zur Hälfte mit geistes- und sozialwissenschaftlichen Arbeiten bestückt sind, während vergleichbare Inhalte in Repositorien deutscher Universitäten einen verschwindend geringen Prozentsatz ausmachen.¹²

Nimmt man die Indizien, die sich aus dem Umgang mit Weblogs, Preprints, Open Access Zeitschriften und Repositorien ergeben, zusammen, liegt die Vermutung nahe, dass elektronisches Publizieren in den Geisteswissenschaften gerade in Deutschland noch an einem Anfang steht. Die damit einhergehende Vermutung, dass das Internet für die fachliche Arbeit bislang als wenig relevant angesehen wird, bestätigt sich nicht nur durch die offenbar geringe Akzeptanz elektronischer Veröffentlichungen, sie zeigt sich nicht nur in der aktiven Publikationstätigkeit, sondern wird auch im Rezeptionsverhalten der Geisteswissenschaftler/innen gespiegelt. Im Rahmen des DFG-Förderinstruments *Kulturelle Überlieferung* wurden zahlreiche für Geisteswissenschaftler/innen relevante Quellen in (retro-) digitalisierter Form zugänglich gemacht¹³; ebenso konnte der orts- und

11 Vgl. <<http://rzblx1.uni-regensburg.de/ezeit/>>.

12 Vgl. van Westrienen, Gerard, *Completed Questionnaires. Country Update on Academic Institutional Repositories*, <<http://www.surf.nl/download/country-update2005.pdf>>, S. 14f. – Hier wäre jedoch weiter zu fragen, inwieweit neben wissenschaftlichen Publikationen in engerem Sinne auch Lehrmaterialien usw. in den Repositorien verfügbar gemacht werden.

13 Vgl. <<http://www.hki.uni-koeln.de/retrodig/index.html>>.

zeitunabhängige Zugriff auf für die Geisteswissenschaften relevante digitale Text- und Werkausgaben in den Jahren 2004 und 2005 durch den Erwerb von Nationallizenzen begünstigt werden.¹⁴ Die faktische Nutzung dieser im Internet frei verfügbaren Materialien lässt jedoch Rückschlüsse auf ihren derzeit noch geringen Bekanntheitsgrad zu.¹⁵ Hier liegt deshalb der Schluss nahe, dass elektronisches Publizieren wohl erst dann zur geläufigen Praxis im Wissenschaftsfeld werden kann, wenn es auch für die dem Publizieren vorausliegenden Arbeitsschritte der Quellenlektüre, des Bibliografierens, der Rezeption von Sekundärliteratur viel alltäglicher geworden ist, auf die im Internet nachgewiesene, leicht zu findende und verlässlich qualitätsgesicherte, fachspezifische Information zurückzugreifen.

DFG-Umfrage zum Publikationsverhalten

Da Konzeption, Methodik und Durchführung der Umfrage bereits anderweitig publiziert wurden, sollen diese Aspekte hier in aller Kürze behandelt werden.¹⁶ Befragt wurden 1.023 DFG-geförderte Wissenschaftler/innen aus den vier großen Wissenschaftsbereichen (Geistes- und Sozialwissenschaften, Lebenswissenschaften, Naturwissenschaften und Ingenieurwissenschaften), die je zu circa einem Viertel vertreten waren. Innerhalb jedes Wissenschaftsbereichs wurden zudem circa zehn Prozent Nachwuchswissenschaftler/innen befragt. Um den Rücklauf fachspezifisch auswerten zu können, wurde schließlich die Zuordnung zu

14 Eine Liste der erworbenen Datenbanken findet sich unter <http://www.nationallizenzen.de>.

15 Mit dem Aufbau des Portals digitalisierter Drucke (<http://www.zvdd.de>) und einer Seite für alle über Nationallizenzen verfügbare Text- und Datensammlungen (<http://www.nationallizenzen.de>) wird der Zugriff auf die jeweils zentral nachgewiesenen Materialien bequem möglich sein; daher bleibt zu hoffen, dass die Angebote auch bekannter und besser genutzt werden.

16 Ausführlicher dazu Fournier, Johannes, Zur Bedeutung von Open Access für das Publikationsverhalten DFG-geförderter Wissenschaftler. Bericht über die Ergebnisse einer Umfrage im Auftrag der Deutschen Forschungsgemeinschaft, in: Zeitschrift für Bibliographie und Bibliothekswesen 52 (2005), S. 235-244; Ders., In die Zukunft publizieren. Positionen und Reflexionen, in: Stempfhuber, Maximilian (Hg.), In die Zukunft publizieren. Herausforderungen an das Publizieren und an die Informationsversorgung in den Wissenschaften, Bonn 2006, S. 279-293.

einem Wissenschaftsfeld nach der Systematik der DFG-Fachkollegien abgefragt. Das Hauptergebnis der Auswertung lässt sich rasch zusammenfassen: Quer durch alle Wissenschaftsbereiche haben sich die Befragten mit deutlicher Mehrheit dafür ausgesprochen, dass die DFG entgeltfrei zugängliche elektronische Publikationen stärker fördern sollte. Auffällig war, dass – wenn auch mit nur kleinen prozentualen Verschiebungen – sich der Nachwuchs gerade der Geisteswissenschaften weniger stark als die etablierten Professoren/innen für deren verstärkte Förderung ausgesprochen hat. Dies kann nur daraus erklärt werden, dass – anders als etwa in den Naturwissenschaften – die Zahl der Angebote, in denen junge aufstrebende Akademiker/innen elektronisch publizieren und so ihre Reputation beachtlich steigern können, in den Geisteswissenschaften noch relativ gering ist. Schließlich ist zu bedenken, dass die Habilitation bzw. das so genannte „Zweite Buch“, dessen Online-Publikation als eher nachrangig betrachtet wird, nach wie vor eine bedeutsame Rolle spielt.

Wenn ich im Folgenden den Versuch unternehme, einzelne Fragen gezielt im Hinblick auf die Antworten vorzustellen, die von Vertretern/innen geisteswissenschaftlicher Fächer vorgebracht wurden, ist höchste Vorsicht angebracht. Die Zahl der Antworten, die den DFG-Fachkollegien „Alte Kulturen“, „Geschichtswissenschaften“, „Kunstwissenschaften“, „Sprachwissenschaften“, „Literatur-, Theater-, Medienwissenschaften“, „Ethnologie, Außereuropäische Kulturen, Religionswissenschaften“, „Theologie“ und „Philosophie“ zuzuordnen sind, war vielfach so klein, dass sie nicht signifikant und damit statistisch eigentlich nicht verwertbar sind. Die folgenden Ausführungen können daher nur als Momentaufnahmen betrachtet werden, die allenfalls Tendenzen erkennen lassen, die erst auf dem Hintergrund der auf Ebene der Fachgebiete bzw. der Wissenschaftsbereiche erzielten Gesamtergebnisse der DFG-Umfrage bzw. auf der Basis auch anderer Analysen und Studien zum Thema¹⁷ zu werten sind. Zudem muss an dieser Stelle einschränkend darauf hingewiesen werden, dass sich die Fragen gezielt auf *entgeltfrei* zugängliche elektronische Publikationen (Open Access) bezogen haben und das Spektrum der

17 Zum Beispiel Nentwich, Cyberscience (wie Anm. 9).

Umfrage damit von vornherein nicht elektronischen Publikationen generell galt. Allerdings beziehen sich insbesondere die Themen, die mit Frage Nr. 20 („Hier haben wir eine Reihe von Aussagen zum Thema Open Access zusammengestellt, die in der gegenwärtigen Diskussion wiederholt genannt werden. Welche halten Sie für zutreffend, welche nicht?“) abgedeckt werden sollten, auf Problembereiche, die nicht allein Open Access Publikationen, sondern elektronische Veröffentlichungen im Allgemeinen betreffen. Daher sollen zunächst Auswertungen zu diesem Fragekomplex vorgestellt werden.

Die Langfristverfügbarkeit von Publikationen ist bereits ein wichtiges Kriterium, wenn Geisteswissenschaftler/innen entscheiden, welchen Zeitschriften sie ihre Arbeiten zur Veröffentlichung anbieten sollen (so zeigen es auch die Antworten auf Frage Nr. 6 „Wenn Sie selbst eine Arbeit in einer wissenschaftlichen Zeitschrift veröffentlichen wollen: Wie wichtig sind Ihnen die folgenden Aspekte?“).¹⁸ Auch wenn sich circa ein Drittel der befragten Geisteswissenschaftler/innen außer Stande sieht, die Frage zu beantworten, ob die langfristige Verfügbarkeit elektronischer Publikationen gewährleistet werden kann, überwiegt auf Seite derjenigen, die die Frage dezidiert beantwortet haben, eine deutliche Skepsis.¹⁹ Ein durchaus vorsichtiges, etwas größeres Zutrauen, elektronische Publikationen auch längerfristig verfügbar zu halten, deutet sich allenfalls in der Geschichts- und in der Sprachwissenschaft an.

Geschichte und Sprachwissenschaft: Dies ist eine Konstellation, die noch häufiger anzutreffen ist und die in gewissem Sinne bestätigt, was auch die Auswertung quer durch die Fächer bestätigt hat: Der – auch in

18 Vgl. den so genannten „Tabellenband“ zur Umfrage unter <http://www.dfg.de/dfg_im_profil/zahlen_und_fakten/statistisches_berichtswesen/open_access/download/oa_tabband.pdf>, S. 38.

19 Die Auffassung, dass die langfristige Verfügbarkeit elektronischer Publikationen nicht ebenso gesichert werden kann wie die gedruckter Veröffentlichungen vertreten 60% (9 von 15) der Befragten aus den „Alten Kulturen“, 42% (13 von 31) aus den Geschichtswissenschaften, 88% (7 von 8) aus den Kunstwissenschaften, 33% (7 von 21) aus den Sprachwissenschaften, 58% (11 von 19) aus den Literaturwissenschaften, 50% (3 von 6) aus der Ethnologie, 56% (5 von 9) aus der Theologie und 63% (5 von 8) aus der Philosophie.

diesen beiden Fächern – intensivere Umgang mit elektronischen Publikationen²⁰ führt offenbar dazu, dass das Vertrauen in die Langfristverfügbarkeit, in die Qualitätssicherung auch elektronischer Veröffentlichungen, in deren gesicherten Nachweis oder die häufigere Zitation steigt.²¹

Die Qualität der Begutachtung (Frage Nr. 6), die ebenfalls überaus wichtig für die Wahl eines Publikationsortes ist, wird im Falle elektronischer Publikationen fast ebenso oft bezweifelt wie bestätigt. Besonders häufig werden Zweifel an der Qualitätssicherung in der Theologie und Philosophie artikuliert. Neben der hohen Zahl der Befragten, die sich nicht im Stande sahen, die Frage „Ist die Qualitätssicherung bei elektronischen Publikationen ebenso sichergestellt wie bei konventionellen?“ verlässlich zu beantworten, muss auf diejenigen Wissenschaftler/innen hingewiesen werden, die offenbar sowohl bei gedruckten als auch bei elektronischen Publikationen die unterschiedlichsten Erfahrungen im Begutachtungsprozess machen konnten. Dabei fällt ins Auge, dass wiederum in den Geschichts- und in den Sprachwissenschaften fast ebenso viele Befragte die Qualitätssicherung elektronischer Publikationen bestätigen wie bezweifeln.²²

Ein weiterer, wichtiger Aspekt war die Frage nach der Zitationshäufigkeit. Die Sichtbarkeit von Publikationen spielt ja insbesondere für den wissenschaftlichen Nachwuchs eine große Rolle. Quer durch alle geisteswissenschaftlichen Fächer allerdings wird die Frage, ob elektronische seltener als gedruckte Publikationen zitiert werden, eher bestätigt – auch das gewiss ein Grund dafür, dass eigene Arbeiten eher in konventionellen Druckzeitschriften veröffentlicht werden. Falls diese Annahme zutrifft,

20 Vgl. dazu Nentwich, Cyberscience (wie Anm. 9), S. 112 (Sprachwissenschaften) und S. 114-116 (Geschichtswissenschaften).

21 Vgl. DFG, Publikationsstrategien im Wandel (wie Anm. 3), S. 48-52.

22 22% der Geschichts- und 19% der Sprachwissenschaftler/innen gehen davon aus, dass die Qualitätssicherung elektronischer Publikationen ebenso sichergestellt ist wie die gedruckter; das Gegenteil nehmen 26% der Geschichts- und wiederum 19% der Sprachwissenschaftler/innen an. 16% (Geschichte) bzw. 9% (Sprachwissenschaft) der Befragten sind unentschieden in ihrer Bewertung; 35% (Geschichte) bzw. 52% (Sprachwissenschaft) geben an, die Frage nicht beantworten zu können.

wäre die einleuchtende Erklärung wohl darin zu suchen, dass elektronische Publikationen bibliografisch seltener nachgewiesen werden als konventionelle. Die Einschätzung der Befragten hierzu scheint dies zu bestätigen²³; zugleich wäre das eine Bestätigung dafür, dass die Literaturrecherche in den geisteswissenschaftlichen Disziplinen noch oft auf traditionelle bibliografische Hilfsmittel zurückgreift, während der Zugriff auf Publikationen über Suchmaschinen demgegenüber weniger prominent ist.

Erfreulich für die DFG war der Umstand, dass bei der Frage nach den wichtigsten wissenschaftlichen Open Access Zeitschriften im eigenen Forschungsfeld nicht nur H-Soz-u-Kult (7x) und Zeithistorische Forschungen. Studies in Contemporary History (4x), sondern auch die sehepunkte (5x), zeitenblicke (3x) und Forum Qualitative Sozialforschung (4x) genannt wurden; auch die übrigen „häufigeren“ Nennungen – öfter als zweimal wurde kein anderer Titel angeführt –, kommen eher aus dem Bereich der Geschichts- und Altertumswissenschaften als aus anderen Disziplinen: The Bryn Mawr Classical Review (3x), The British Museum Studies in Ancient Egypt and Sudan (2x), Göttinger Forum für Altertumswissenschaft (2x) und Forum historiae iuris (2x).²⁴

Rückschlüsse auf die Akzeptanz des elektronischen Publizierens lassen sich nicht zuletzt aus den Angaben ziehen, mit der die eigene Publikationspraxis der Befragten untersucht wurde. Im Folgenden ist daher noch kurz darauf einzugehen, wie die Disziplinen sich – immer unter Hinweis auf die kaum gegebene statistische Signifikanz – zu wissenschaftlichen Erstveröffentlichungen in Open Access Zeitschriften, zur Vorabveröffentlichung von Forschungsergebnissen oder auch zum so genannten self-archiving verhalten.

Die Publikation in Open Access Zeitschriften spielt, den Angaben der

23 Dass elektronische Publikationen bibliografisch seltener nachgewiesen werden als konventionelle, nehmen 53% (8 von 15) der Befragten aus den „Alten Kulturen“, 68% (21 von 31) aus den Geschichtswissenschaften, 63% (5 von 8) aus den Kunstwissenschaften, 43% (9 von 21) aus den Sprachwissenschaften, 74% (14 von 19) aus den Literaturwissenschaften, 33% (2 von 6) aus der Ethnologie, 33% (3 von 9) aus der Theologie und 75% (6 von 8) aus der Philosophie an.

24 Vgl. DFG, Publikationsstrategien im Wandel (wie Anm. 3), S. 42.

Befragten zufolge, offenbar keine Rolle.²⁵ Dieses Bild entspricht der Tendenz, die sich generell erwiesen hat.²⁶ Auffällig ist im Bereich der Geisteswissenschaften höchstens – wobei wiederum die geringe Fallzahl gerade der befragten Nachwuchswissenschaftler/innen zu bedenken ist –, dass, mit Ausnahme des Fachgebiets „Alte Kulturen“, nur etablierte Wissenschaftler/innen bekundet haben, bereits in einer Open Access Zeitschrift veröffentlicht zu haben, während der wissenschaftliche Nachwuchs dies bislang nicht getan hat.

Gänzlich unüblich scheint die – aus dem Bereich besonders der Physik bekannte – Praxis zu sein, Vorabveröffentlichungen wissenschaftlicher Arbeiten über das Internet bereitzustellen (Frage Nr. 19). Die verschwindend geringe Zahl derjenigen Befragten, die hier überhaupt Auskunft gaben, könnte zugleich darauf hindeuten, dass die Frage im geisteswissenschaftlichen Feld nicht richtig zu verstehen war, da die entsprechende Publikationspraxis nicht geübt wird.²⁷

Demgegenüber ist es offenbar in manchen Fällen durchaus gängig, dass Wissenschaftler/innen ihre bereits publizierten Arbeiten selbst im Internet bereitstellen (Frage Nr. 16), um so den kostenfreien Zugriff auf diese zu ermöglichen. Auf diese Weise bereitgestellt werden naturgemäß neben Zeitschriftenaufsätzen vor allem Beiträge aus Tagungsbänden.²⁸ In

25 Nach einer Berechnung des Mittelwerts wurden 0,08 Publikationen aus den „Alten Kulturen“, 0,22 aus den Geschichtswissenschaften, 0,13 aus den Kunstwissenschaften, 0,4 aus den Sprachwissenschaften, 0,22 aus den Literaturwissenschaften, 0,17 aus der Ethnologie und keine aus der Theologie und der Philosophie in Open Access Journals veröffentlicht.

26 Vgl. DFG, Tabellenband (wie Anm. 18), S. 46.

27 Vgl. DFG, Tabellenband (wie Anm. 18), S. 51f.

28 Der Mittelwert für Postpublikationen von Zeitschriftenaufsätzen beträgt in den „Alten Kulturen“ 1, in den Geschichtswissenschaften 1,47 in den Kunstwissenschaften 0,37, in den Sprachwissenschaften 7, in den Literaturwissenschaften 4,01, in der Ethnologie 5,24, in der Theologie 3,17 und in der Philosophie 3,57. Die Mittelwerte für zusätzlich veröffentlichte Beiträge aus Tagungsbänden sind in den „Alten Kulturen“ 5,56, in den Geschichtswissenschaften 5,42, in den Kunstwissenschaften 3,33, in den Sprachwissenschaften 14,03, in den Literaturwissenschaften 5, in der Ethnologie 3,7, in der Theologie 0 und in der Philosophie 10,71. In den meisten Fächern ist es somit eher üblich, Beiträge aus Tagungsbänden entgeltfrei verfügbar zu machen, als Zeitschriftenbeiträge sekundär zu publizieren.

den Sprachwissenschaften, in denen Sekundärpublikationen offenbar gängiger sind als in anderen Forschungsfeldern, finden sich sogar zusätzlich elektronisch publizierte Monografien wieder. Die Tendenz, in erster Linie bereits anderweitig veröffentlichte Beiträge über das Internet entgeltfrei verfügbar zu machen, entspricht dem Verhalten auch anderer Wissenschaftsbereiche.²⁹

Die Praxis der Sekundärpublikation bietet den Vorteil, dass Arbeiten im Internet publiziert werden können, die bereits das „Gütesiegel“ einer anerkannten Fachzeitschrift tragen. Insofern bietet die Unterstützung des so genannten self-archiving derzeit den vielleicht vielversprechendsten Ansatz, zur Akzeptanz elektronischer Publikationen beitragen zu können. Allerdings muss auch klar sein, dass Wissenschaftler/innen sich durch erst nachträgliche elektronische Publikation ihrer Forschungsergebnisse mancher medienspezifischer Vorteile, die die direkte Publikation in Open Access Zeitschriften bieten könnte – etwa die unmittelbare Verknüpfung zu Quellen, Karten, Statistiken – begeben bzw. mit hohem Aufwand rechnen müssten, um diese nachträglich zu integrieren.

Im Folgenden soll dargestellt werden, mit welchen Aktivitäten, die teilweise auf Basis der Umfrageergebnisse, die zuvor nur in einem sehr begrenzten Ausschnitt referiert werden konnten, die DFG versuchen möchte, zur Akzeptanz von elektronischen Publikationen beizutragen.

Förderaktivitäten zur Unterstützung der Akzeptanz von E-Journals

Nur in aller Kürze sei darauf hingewiesen, dass im Rahmen des Förderinstruments *Elektronisches Publizieren* sowohl die Entwicklung von Publikationstools als auch die Gründung innovativer elektronischer Zeitschriften unterstützt werden können; diese Aktivitäten der DFG sind weithin bekannt.³⁰ Etwas ausführlicher soll beschrieben werden, welche Aktivitäten derzeit geplant sind, um verstärkt zur Akzeptanz elektronischer Publikationen beizutragen.

29 Vgl. DFG, Tabellenband (wie Anm. 18), S. 48f.

30 Vgl. auch <<http://www.dfg.de/lis>>, insbesondere unter <<http://www.dfg.de/lis/openaccess>>.

1. Im Januar 2006 haben sich Senat und Hauptausschuss der DFG dafür ausgesprochen, alle DFG-geförderten Wissenschaftler/innen aufzufordern, ihre Publikationen nach Möglichkeit im Open Access verfügbar zu machen. Der zu diesem Zweck eigens in die Verwendungsrichtlinien eingearbeitete Text weist explizit darauf hin, dass damit nicht allein die zusätzliche Veröffentlichung bereits anderweitig publizierter Forschungsergebnisse gemeint ist, sondern dass Wissenschaftler/innen in Open Access Journals publizieren sollten, sofern für das jeweilige Fach entsprechend renommierte Zeitschriften existieren.
2. Zugleich arbeitet die Geschäftsstelle derzeit intensiv an einer Neugestaltung des Förderinstruments *Publikationsbeihilfe für Zeitschriften*. In diesem Programm können, sofern ein Zuschussbedarf besteht, die Herstellungskosten wissenschaftlich exzellenter Zeitschriften unter bestimmten Voraussetzungen mitfinanziert werden. Im Rahmen des bestehenden Programms werden 36 geistes- und sozialwissenschaftliche Printzeitschriften, doch keine elektronische Zeitschrift gefördert. Ein neues Programm zur Zeitschriftenförderung soll dagegen auch das Ziel verfolgen, den Prozess hin zur zunehmenden Akzeptanz elektronischer Publikationen zu unterstützen, indem die Konditionen der Förderung bewusst im Hinblick auf die bei Redaktion und Herstellung elektronischer Zeitschriften anfallenden Arbeiten und Kosten ausgerichtet werden. Insbesondere die Möglichkeit einer Beihilfe, mit der die Transformation gedruckter in elektronische Zeitschriften unterstützt werden soll, dürfte attraktiv sein, um den Prozess zur Umstellung einer Druckkultur auf eine Kultur elektronischer Zeitschriften zu fördern. Die Vorschläge mit den Grundzügen für das neue Förderprogramm lagen im Frühjahr 2006 dem Ausschuss für wissenschaftliche Bibliotheken und Informationssysteme sowie den geistes- und sozialwissenschaftlichen Fachkollegien vor, um deren Rückmeldungen zum neuen Konzept berücksichtigen und einarbeiten zu können, bevor die Entscheidungsgremien der DFG mit dem neuen Vorschlag konfrontiert werden sollen.
3. Schließlich soll die Struktur für wissenschaftliche Sekundärpublikationen möglichst attraktiv gestaltet werden. Hier müssen vor allem As-

pekte der Qualitätssicherung in Repositorien in den Blick genommen und umgesetzt werden, um das Vertrauen in die Qualität der über Repositorien verfügbaren Inhalte als Voraussetzung ihrer Akzeptanz zu steigern. Gleichzeitig müssen Dienstleistungen entwickelt werden, die es Wissenschaftlern/innen attraktiv erscheinen lassen, ihre besten Arbeiten auch über qualitativ hochwertige Hochschulschriftenserver anzubieten. Hierzu wurden Mitte Februar 2006 auf einem zusammen mit der Deutschen Initiative für Netzwerkinformation (DINI e.V.) ausgerichteten Symposium Vorschläge entwickelt, wie Repositorien durch Etablierung von Statistik- und Zitationsdiensten für Autoren/innen attraktiver gestaltet werden können. Die unseres Erachtens vordringlich zu lösende Aufgabe wird darin bestehen, Hochschulschriftenserver so zu vernetzen, dass sie auch über Ländergrenzen hinweg interoperabel sind und dass deren Inhalte über alle Suchinstrumente – vom elektronischen Bibliothekskatalog über Virtuelle Fachbibliotheken bis hin zu allgemeinen Suchmaschinen wie *Google* – nachgewiesen und für das wissenschaftliche Arbeiten rezipiert werden können.

4. Einen wesentlichen Schub könnte die Akzeptanz elektronischer Publikationen auch in den Geisteswissenschaften erhalten, wenn renommierte Wissenschaftler/innen mit einer langen Liste reputierlicher Veröffentlichungen dafür gewonnen werden könnten, diese in einer konzertierten Aktion entgeltfrei verfügbar zu machen. Die Idealvorstellung wäre die Auflage eines nach dem niederländischen Vorbild des *Cream of Science*³¹ gestrickten Projekts, mit dem sämtliche Publikationen deutscher Spitzenwissenschaftler/innen digital nachgewiesen und möglichst vollständig entgeltfrei verfügbar gemacht würden. Dies hätte Vorbildfunktion für viele andere Forscher/innen. Zur Umsetzung entsprechender Pläne müssen jedoch eine stattliche Zahl organisatorischer, administrativer, finanzieller und insbesondere rechtlicher Fragen und Probleme geklärt werden. Als langfristiges Ziel bleibt eine Art deutsches *Cream of Science*-Vorhaben jedoch Aufgabe der DFG.

Gerade für die Geisteswissenschaften stellt die langfristige Verfügbarkeit

31 Vgl. <<http://www.creamofscience.org/en/page/language.view/keur.page>>.

ihrer Texte eine besondere Notwendigkeit dar. Vielfach ist die noch geringe Akzeptanz elektronischer Publikationen darauf zurückzuführen, dass die langfristige Archivierung wissenschaftlicher Dokumente letztlich nicht geklärt ist. Die Vielzahl der hier zu lösenden Probleme wird insbesondere über die vom Bundesministerium für Bildung und Forschung geförderten Projekte *kopal – Kooperativer Aufbau eines Langzeitarchivs digitaler Informationen* – und *nestor – Kompetenznetzwerk Langzeitarchivierung* – gebündelt³², wo Problemlösungsstrategien entwickelt und umgesetzt werden. Gezielte Einzelprojekte, die dem Problem der Archivierung von Netzpublikationen gewidmet sind, werden auch von der DFG gefördert.

Fazit

Das Spannungsfeld zwischen der zunehmenden Akzeptanz elektronischer Publikationen einerseits und einer gewissen Zurückhaltung diesen gegenüber andererseits tritt insbesondere im Verhältnis sekundär publizierter wissenschaftlicher Artikel zu Originalbeiträgen in Open Access Zeitschriften zutage. Das für geisteswissenschaftliche Autoren/innen erforderliche Renommee ist in vielen Fällen vorerst nur durch Publikationen in Druckzeitschriften und Buchreihen, die als besondere Gütesiegel angesehen werden, zu erwerben. Man darf jedoch zurecht davon ausgehen, dass bestehende und auch neue elektronische Zeitschriften sich im Laufe der Zeit als diesen gleichberechtigte ‚Marken‘ etablieren, sofern die durch Wissenschaftler/innen, Verlage und – insbesondere, doch nicht nur im Falle von Sekundärpublikationen – Bibliotheken vorgenommene Qualitätskontrolle auch elektronischer Publikationen höchsten Ansprüchen genügt.

Dr. Johannes Fournier ist bei der Deutschen Forschungsgemeinschaft als Programmdirektor für den Förderbereich „Elektronisches Publizieren“ verantwortlich. E-Mail: johannes.fournier@dfg.de

32 Vgl. <<http://kopal.langzeitarchivierung.de/>>; <<http://www.langzeitarchivierung.de/index.php>>.

ELEKTRONISCHES PUBLIZIEREN – DIE NUTZERPERSPEKTIVE

von Karsten Borgmann

Der Beitrag untersucht die Charakteristik einer geschichtswissenschaftlichen Publikationskultur vor dem Hintergrund einer zunehmenden Verfügbarkeit elektronischer Fachpublikationen. Bisherige Untersuchungen haben sich in diesem Zusammenhang vor allem mit den Ansprüchen und Bedürfnissen von geschichtswissenschaftlichen Autoren/innen beschäftigt, deren Skepsis gegenüber elektronischen Veröffentlichungsorten nach wie vor groß ist. Der Beitrag stellt diesen, aus der Autorenperspektive verständlichen Einstellungen, dezidiert die Perspektive von Wissenschaftlern/innen als geschichtswissenschaftlichen Nutzern/innen und Rezipienten/innen von elektronischen Fachpublikationen gegenüber. Daten für diese Analyse liefert die erste Sichtung einer Befragung über die Nutzung geschichtswissenschaftlicher Online-Angebote, die Ende 2005 und Anfang 2006 durchgeführt wurde. Eine genauere Analyse der Umfragedaten wird im Zuge der Publikation der Umfragergebnisse erfolgen.

Aktuelle Untersuchungen zur Nutzung des Internets als Medium wissenschaftlicher Kommunikation zeigen eine offenbar quer durch alle Wissenschaftsbereiche gehende Diskrepanz bei der Nutzung des Internets, einerseits als Publikationsort eigener Veröffentlichungen, andererseits als Medium, über das auf Informationen und Publikationen zugegriffen werden kann.¹ So gaben in einer Umfrage der Deutschen Forschungsgemein-

1 Hier ist vor allem die von der Deutschen Forschungsgemeinschaft im Jahr 2004 in Auftrag gegebene Studie: Publikationsstrategien im Wandel? Ergebnisse einer Umfrage zum Publikations- und Rezeptionsverhalten unter besonderer Berücksichtigung von Open Access, Bonn 2005 (im Weiteren zitiert als „DFG [2005]“), vgl. <http://www.dfg.de/dfg_im_profil/zahlen_und_fakten/statistisches_berichtswesen/open_access/index.html> (28.02.2006) zu nennen, die ausführlicher im Beitrag von Johannes Fournier in diesem Band vorgestellt wird.

schaft (DFG) 83 Prozent der befragten Wissenschaftler/innen in Natur-, Lebens- und Geisteswissenschaften an, mehr oder weniger häufig Fachpublikationen zu nutzen, die im Open Access² zur Verfügung stehen.³ Über 80 Prozent sind zudem der Meinung, dass die freie Zugänglichkeit zu elektronischen Publikationen öffentlich gefördert werden sollte. Demgegenüber wurde das Internet als Ort zur Veröffentlichung eigener Open Access-Publikationen bislang nur von rund zwölf Prozent der befragten Wissenschaftler/innen genutzt.⁴ Eine deutliche Mehrheit der Befragten äußerte sich zudem skeptisch, was Qualität, Renommee, langfristige Zitierfähigkeit und, ganz allgemein, die fachliche Bedeutung von Open Access-Publikationen angeht.⁵

Vieles spricht dafür, dass diese Diskrepanz sich im Blick auf die Geschichtswissenschaften noch verschärft. Zwar sind Open Access-

2 Mit dem Prädikat Open Access werden im Folgenden Publikationen und Informationsangebote bezeichnet, die Nutzern/innen im Internet weitgehend barrierefrei und kostenlos zur Verfügung stehen und deren Autoren/innen bzw. weitere Inhaber/innen von Nutzungsrechten dieser Nutzung zugestimmt haben. Diese Definition enthält noch keine Aussage über Mechanismen der Qualitätssicherung einer Publikation, wie sie in anderen Vorschlägen zu Open Access ausgeführt wird, vgl.: Harnad, Stevan; Brody, Tim; Vallières, François u.a. (Hgg.), The Access/Impact Problem and the Green and Gold Road to Open Access, Wizards of OS 3 Konferenz 9.-14. Juni 2004, vgl. <<http://wizards-of-os.org/index.php?id=1773>> (20.04.2006) sowie DFG (2005), S. 40f.

3 DFG (2005), S. 51f.

4 DFG (2005), S. 43f.

5 DFG (2005), S. 48-50, zusammengefasst in: Mugabushaka, Alexis-Michel; Fournier, Johannes, Offener Zugang zu wissenschaftlichem Wissen: Erfahrungen und Ansichten DFG-geförderter Wissenschaftler, DFG infobrief (2005), 1, S. 3., vgl. <http://www.dfg.de/dfg_im_profil/zahlen_und_fakten/statistisches_berichtswesen/ib/download/ib01_2005.pdf> (24.03.2006); auch in einer 2004 mit 145 Teilnehmern/innen durchgeführten Umfrage der Online-Rezensionszeitschrift Querelles-Net, schätzte ein bemerkenswert deutlicher Prozentsatz der Befragten Prestige und Vertrauenswürdigkeit von elektronischen Publikationen negativ ein, vgl.: Elektronisches Publizieren. Ergebnisse unserer Umfrage vom Dezember 2004. Zusammengestellt von Rochus Wolff, in: <<http://www.querelles-net.de/umfrage/index.shtml>> (20.04.2006).

Zeitschriften in den Geisteswissenschaften allgemein wenig etabliert, was auch mit der bleibenden Bedeutung von Monografien und Sammelbänden als typischen Publikationsformaten dieser Fächergruppe zu tun hat.⁶ In anderer Hinsicht lässt sich jedoch vermuten, dass innerhalb der Geisteswissenschaften besonders Historiker/innen als potenzielle „Power User“ von Online-Informationen aller Art in Frage kommen. So zeigte eine bereits 2003 ebenfalls von der DFG in Auftrag gegebene Studie zur Nutzung des Systems der überregionalen Literaturversorgung in Deutschland, dass insbesondere die Geschichtswissenschaften sich in ihrer Arbeit besonders vielfältiger Informationsmöglichkeiten bedienen.⁷ So nehmen Historiker/innen eher als Vertreter/innen anderer Disziplinen Informationen auch aus anderen Bereichen als ihrem eigenen Fachgebiet wahr, sie interessieren sich mehr als andere Fächer für Forschungsprojekte, also für Informationen rund um den Forschungsbetrieb. Ferner kalkulieren sie einen verhältnismäßig hohen Zeitaufwand für Informationsbeschaffung ein und sie sichten Informationen lieber selber, als sich auf Bewertungen und Klassifikationen anderer zu verlassen.⁸ So könnte man annehmen, dass mit Blick auf dieses wenig selektive Rezeptionsverhalten Geschichtswissenschaft und Internet wie für einander geschaffen sind. Allerdings steht dieser vermeintlich harmonischen Beziehung eine ebenso signifikant hohe Ablehnung des Internets als geeignetem Ort für eigene Veröffentlichungen gegenüber.

Gründe für diese Skepsis können in der geringen Bekanntheit und der fehlenden Vertrautheit mit den gegebenen Möglichkeiten für eigene Open Access-Veröffentlichungen vermutet werden, dem durch geeignete Öffentlichkeitsarbeit und gezielte Fördermaßnahmen gegenzusteuern wäre.⁹

6 DFG (2005), S. 22.

7 Universitäts- und Landesbibliothek Münster; Infas (Hgg.), Nutzungsanalyse des Systems der überregionalen Literatur- und Informationsversorgung: Teil I: Informationsverhalten und Informationsbedarf der Wissenschaft, o.O., 2003 (im Weiteren zitiert als: „ULM / Infas [2003]“). Befragt wurden in dieser Studie 1.581 Personen, davon 142 aus den Geschichtswissenschaften, vgl. <http://www.dfg.de/forschungsfoerderung/wissenschaftliche_infrastruktur/lis/download/ssg_bericht_teil_1.pdf> (17.03.2006).

8 ULM / Infas (2003), S. 57-74, besonders S. 59, 62.

9 Vgl. Fournier, Johannes, Wege zum Wissen. Aktionsfelder zur Förderung des Open Access durch die DFG, o.O., o.J., <http://www.dfg.de/dfg_

Zugleich wird jedoch auch deutlich, dass solche Einstellungen als Teil einer fachspezifischen Publikationskultur offenbar langfristig recht stabil sind¹⁰, eine schnelle Änderung grundlegender Einstellungen lediglich durch verstärkte Werbung für elektronische Veröffentlichungswege also nicht unbedingt zu erwarten ist. Da in den Geschichtswissenschaften, im Verhältnis zu anderen Wissenschaftsbereichen, sehr lange an Publikationen gearbeitet wird, ist es verständlich, wenn Autoren/innen für ihre Werke auch eine „aufwändige“ Veröffentlichungsform wünschen. Die Publikation soll herausragen aus der Masse anderer, sie soll an einem renommierten Ort erscheinen, durch Aufnahme in eine Bibliothek langfristig verfügbar sein und vor allem wissenschaftliche Aufmerksamkeit finden. Einer der deutlichsten Kritikpunkte an der Präsenz von Werken im Internet, den wissenschaftliche Autoren/innen in diesem Zusammenhang äußern, ist, dass gegenüber konventionellen Publikationswesen im Netz keine ausreichenden Qualitätsprüfungen existieren.¹¹ Demgegenüber werden die aufwändigeren Produktionsverfahren des „klassischen“ Publikationsbetriebs als natürlicher „Filter“ angesehen, der nur in Inhalt und Form hochwertige Publikationen passieren lässt. An dieser Stelle wird der Konflikt zwischen dem Anliegen nach singulärer Veröffentlichung und der Praxis einer breiten, umfassenden Nutzung besonders offenkundig.

Folgt man nun in diesem Dilemma, das, wie ich behaupten möchte, zur Gesamtcharakteristik der spezifisch geschichtswissenschaftlichen Publikationskultur dazugehört, vor allem der Perspektive der Autoren/innen, hängt die Entscheidung für Open Access mit dem Vorhandensein von hochwertigen Veröffentlichungsorten im Netz zusammen. Durch ein vermehrtes Angebot von redaktionell selektiven und aufwändig produzierten Open Access-Zeitschriften soll den Autoren/innen der Wechsel zur Publikation in einem neuen Medium erleichtert werden. Der Schwerpunkt dieses Ansatzes besteht in der Verbesserung der Präsentation sowie in einer größeren Selektivität bei den Inhalten.

im_profil/zahlen_und_fakten/statistisches_berichtswesen/open_access/download/oa_stellungnahme.pdf>.

10 Mugabushaka; Fournier, Offener Zugang (wie Anm. 5), S. 2.

11 DFG (2005), S. 50.

Eine andere Herangehensweise könnte jedoch auch darin bestehen, die Nutzerperspektive einzunehmen und zu fragen, wie im Prozess der Nutzung von Online-Publikationen die Qualität einer Arbeit erkannt werden kann. Welche qualitativen Auszeichnungsmechanismen hält das Internet bereit, bzw. welche Möglichkeiten beginnen sich herauszubilden, die den Nutzern/innen erlauben, die Spreu vom Weizen zu trennen? Die Kenntnis dieser Zusammenhänge könnte Befürchtungen von Seiten der Autoren/innen abbauen, mitsamt ihren aufwändig produzierten Werken in der Masse des Internets unterzugehen. Was lässt sich aus der genaueren Analyse der Nutzung lernen, das hilft, Internet-Publikationen so zu optimieren, dass Autoren- und Nutzerinteressen zumindest in gewissem Maße versöhnt werden können?

Ich möchte im Folgenden zu diesem Problem einige recht vorläufige Überlegungen anführen, die auf der ersten Sichtung von Ergebnissen einer Online-Umfrage beruhen, die von den Projekten Clio-online, H-Soz-u-Kult und Zeitgeschichte-online Ende 2005 und Anfang 2006 durchgeführt wurde.¹² Die Befragung hatte nicht in erster Linie das Ziel, das Akzeptanzproblem elektronischer Fachpublikationen zu klären. Sie sollte vor allem einen Eindruck von der Nutzung verschiedener existierender Fachangebote erlauben, also neben elektronischen Publikationen auch von Recherche-techniken, der Verwendung von Suchmaschinen und dem Bedarf an weiteren Informationsdiensten. Mit der Umfrage sollte den Nutzern/innen eine Möglichkeit zum Feedback eingeräumt werden, das wiederum als Ansatzpunkt zur Verbesserung der Projektangebote dienen sollte. Einige Fragen und Antworten lieferten jedoch auch einen Eindruck von der Nutzung des

12 An der Online-Umfrage beteiligten sich 2.215 Personen, die den 18-seitigen Fragebogen beantworteten. 1.581-mal wurde er komplett ausgefüllt. Fast 60 Prozent aller Befragten arbeiten an einer Universität oder Forschungseinrichtung, ein beachtliches Drittel ist jedoch an nicht-akademischen Institutionen tätig oder arbeitet freiberuflich. Insgesamt dominiert unter den Teilnehmern/innen der Umfrage das wissenschaftliche Interesse an Geschichte im universitären Kontext, andere berufliche Arbeitsfelder, z.B. als Dozent/in oder Lehrer/in, im Museum, als Bibliothekar/in, Archivar/in, Autor/in oder Publizist/in treten dagegen zurück. Die Ergebnisse werden in Kürze unter der Adresse <<http://www.clio-online.de/umfrage>> publiziert.

Internets als Teil des geschichtswissenschaftlichen Publikationswesens, der zeigt, wo aus Nutzerperspektive Stärken und Schwächen gesehen werden. Im Kreis der geschichtswissenschaftlichen Internet-Nutzer/innen, der mit der Umfrage angesprochen wurde, waren zunächst keine generellen Vorbehalte gegenüber dem Internet als Plattform für Fachpublikationen festzustellen. Dies ist zum einen sicherlich der Internet-Affinität der Teilnehmer/innen einer Online-Umfrage zuzuschreiben, zum anderen aber vermutlich auch auf die in den letzten Jahren stark fortgeschrittene Integration von Netztechnik in den geschichtswissenschaftlichen Arbeitsplatz zurückzuführen. Nicht nur die offensichtlichen technischen Vorzüge wie Schnelligkeit und Zugänglichkeit werden an Internet-Angeboten geschätzt, auch die Qualität und Zuverlässigkeit von Internet-Informationen allgemein beurteilt eine sehr deutliche Mehrheit der Befragten positiv.¹³

Dieser erste Eindruck wird bestätigt durch eine Frage, die auf verschiedene, alternative Formen möglicher Qualitätssicherung von Angeboten zielte. So entscheidet sich nur eine Minderheit für aufwändigere inhaltliche Prüfverfahren (*Peer Review*). Eine deutliche Mehrheit akzeptiert auch Veröffentlichungen, die lediglich formalen Kriterien (wie zum Beispiel Vollständigkeit, klare Urheberschaft usw.) genügen. Diese Form der Veröffentlichung wird normalerweise von Fachportalen oder auch Dokumentenservern der Hochschulen gepflegt (Abbildung 1).

Es bestätigt sich die bereits genannte Beobachtung, dass das Internet aus Nutzersicht eher wegen der darüber zugänglichen Breite von Informationen und nicht aufgrund der qualitativen Vorauswahl der Angebote geschätzt wird. Die genannten Veröffentlichungsformen stellen in der Realität natürlich keine Alternativen dar, sondern ergänzen sich, wie weiter unten noch ausgeführt wird. Dass die Frage diese Optionen alternativ darstellte erhöhte jedoch die Aussagekraft, da die Probanden/innen an dieser Stelle eine deutlichere Präferenz zeigen mussten.

13 Eine ausführliche Analyse der Antworten, aufgeschlüsselt nach Alters-, Qualifikations- und Berufsgruppen, wird im Zuge der Publikation der Umfrageergebnisse erfolgen.

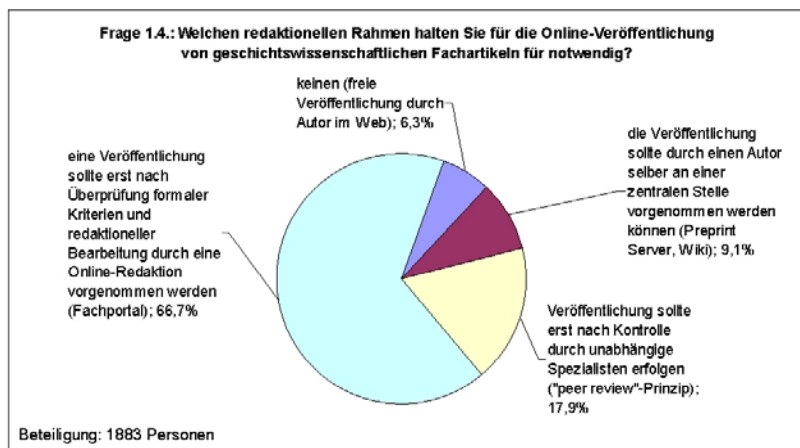


Abbildung 1: Notwendige Qualitätskontrolle für die Veröffentlichung von Fachartikeln

Die Umfrage berührte die Frage nach stärkerer inhaltlicher Vorauswahl von Angeboten auch noch in einem anderen Zusammenhang. Gefragt nach Möglichkeiten der Vor-Selektion im Rahmen eines E-Mail-Abonnements entscheiden sich deutlich weniger Nutzer/innen für eine inhaltliche Vorauswahl gegenüber lediglich formalen Ordnungsprinzipien der Angebote. Vor-Selektionen, sei es durch Fachredakteure/innen, Fachbibliothekare/innen oder eine automatische Klassifikation nach Stichwörtern rufen die Skepsis der Nutzer/innen hervor. In der befragten Gruppe lässt sich eine Präferenz beobachten, E-Mail-Abonnements nach funktionalen Eigenschaften (Rubriken) zu ordnen, auch wenn dabei unvermeidlich auch Mailings im Postfach landen, die außerhalb des eigenen inhaltlichen Interessengebiets liegen (Abbildung 2).

Fasst man diese Beobachtungen zusammen, ließe sich hypothetisch äußern, dass aus der Nutzerperspektive nicht unbedingt die unter hohem redaktionellen Aufwand erstellte E-Zeitschrift, sondern das nach formalen Kriterien arbeitende Fachportal, bzw. im Falle der Veröffentlichung von Forschungsarbeiten der universitäre Dokumentenserver die effizienteste

Variante der Online-Publikation ist. Aber bestätigen sich damit nicht alle Befürchtungen, die Autoren/innen in Bezug auf das Verschwinden ihrer Werke in der Masse formal ähnlicher, aber möglicherweise qualitativ schlechterer Veröffentlichungen hegen? Zwei Antworten lassen sich aus der Perspektive der Nutzer/innen geben.

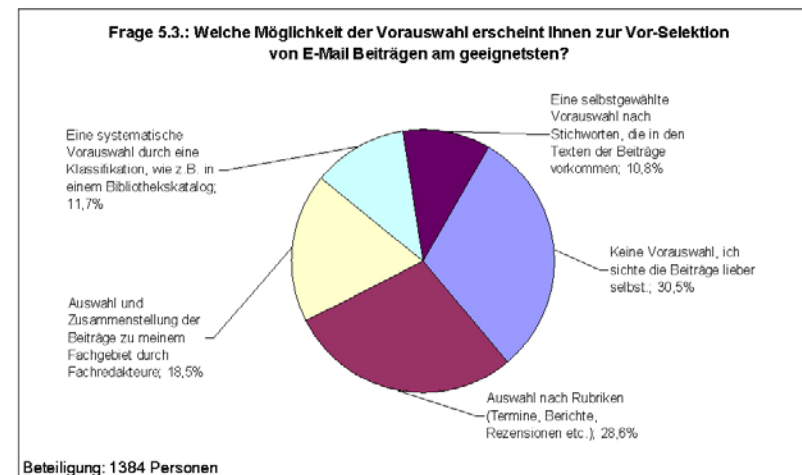


Abbildung 2: Präferenz bei der Vor-Selektion von E-Mail-Abonnements

Zum einen ist nicht anzunehmen, dass sich die große Bandbreite von Publikationsformen und -orten unterschiedlicher Qualität, die das „klassische“ Publikationswesen hervorgebracht hat, in Zukunft reduziert. Im Gegenteil, große Digitalisierungsinitiativen, die von privater wie öffentlicher Seite vorangetrieben werden, sorgen dafür, dass vergangene wie zukünftige Publikationen, ob in Buchform oder als Zeitschrift, elektronisch abrufbar sein werden.¹⁴ Damit wird die gesamte redaktionelle Kompetenz

¹⁴ Vgl. Lewandowski, Dirk, Google Scholar. Aufbau und strategische Ausrichtung des Angebots sowie Auswirkungen auf andere Angebote im Bereich der wissenschaftlichen Suchmaschinen. Expertise im Auftrag des Hochschulbibliothekszentrums Nordrhein-Westfalen, Düsseldorf 2005, vgl. <http://www.dl-forum.de/dateien/Expertise_Google-Scholar.pdf>.

etablierter Titel, Reihen, Verlage und Herausgeberkollektive ins digitale Zeitalter transferiert – und damit auch alle Höhen und Tiefen des „klassischen“ Publikationswesens, deren Kenntnis ein wichtiger Teil der *working knowledge* von Historikern/innen ist. Experimente mit ausschließlich elektronisch herausgegebenen Fachzeitschriften haben angesichts dieses in großem Umfang laufenden Prozesses der Digitalisierung nie eine große Rolle gespielt.¹⁵ Auch die Monopolisierung elektronischer Vertriebswege durch einige Großverlage, wie sie international in Bereichen der Natur- und Lebenswissenschaften zu beobachten ist¹⁶, ist in einem durch viele kleine und mittelgroße Unternehmen geprägten, deutschen Verlagswesen auch mittelfristig eher unwahrscheinlich.¹⁷ Wissenschaftlichen Autoren/innen bleibt somit weiterhin eine große Auswahl an distinktiven Offline-Veröffentlichungsorten erhalten, allerdings ist mit digitalen Parallelpublikationen und -verwertungen durch die Verlage zu rechnen, sofern sich Autoren/innen nicht eigene Nutzungsrechte vorbehalten.

Zum anderen werden über das Netz in vielfacher Weise zusätzliche Informationen und Bewertungen über Publikationen verfügbar, die den Nutzern/innen eine schnelle Einschätzung der Qualität eines Werks erlauben. Über einen Dienst wie die Rezensionssuche „Historische Rezensionen Online“ werden im einfachen Zugriff alle wichtigen Anbieter elektronischer Buchbesprechungen erschlossen, die zuvor an unterschiedlicher Stelle im Netz aufgesucht werden mussten.¹⁸ Es spricht vieles dafür, dass neben der Erschließung eines wachsenden Bestandes digitalisierter Publikationen selbst, das Internet vor allem als Speicher solcher „Meta-Informationen“ über den primären wissenschaftlichen *Content* genutzt

15 Keller, Alice, Elektronische Zeitschriften. Entwicklungen in den verschiedenen Wissenschaftszweigen, in: *zeitenblicke* 2 (2003), 2, Absatz 28, vgl. <<http://www.zeitenblicke.historicum.net/2003/02/keller.html>> (24.03.2006) weist auf den sinkenden Anteil reiner Online-Zeitschriften gegenüber dem bei über 80 Prozent liegenden Anteil digitalisierter Print-Zeitschriften hin.

16 Zum Einstieg in diese Problematik vgl. den Artikel Zeitschriftenkrise, in: Wikipedia, <<http://de.wikipedia.org/wiki/Zeitschriftenkrise>> (20.04.2006).

17 Vgl. den interessanten Diskussionsbeitrag: Steinhauer, Eric, Wissenschaftliche Verlage und open access, in: *InetBib*, 09.03.2006, vgl. <<http://www.ub.uni-dortmund.de/listen/inetbib/msg29889.html>> (20.04.2006).

18 Vgl. <<http://www.clio-online.de/rezensionen>> (24.03.2006).

wird. So geben Befragte an, am häufigsten Rezensionen über wissenschaftliche Neuerscheinungen im Internet zu suchen. Die zweithäufigste Kategorie stellen Literaturübersichten und Bibliografien dar, also ebenfalls ein Genre, in dem bibliografische Information fachlich evaluiert und zusammengestellt wurde (Abbildung 3).¹⁹

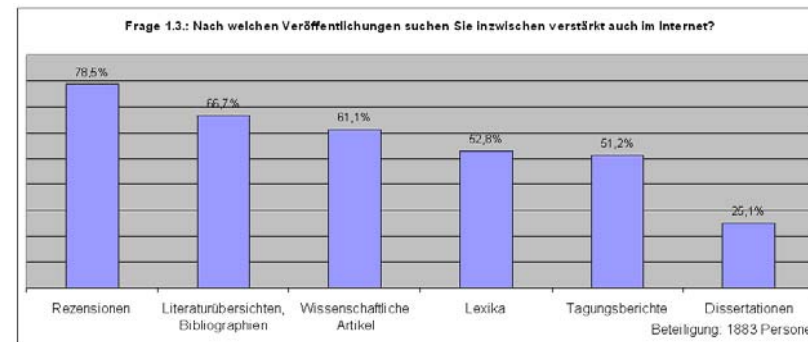


Abbildung 3: Wissenschaftliche Publikationen, nach denen im Internet gesucht wird

Wie in einem enorm erweiterten sozialen Netz²⁰ werden so über die bislang verfügbaren Kontakte der Wissenschaftler/innen hinaus, Wissen und Kenntnisse anderer Wissenschaftler/innen sichtbar, die eine schnelle

19 Dieses Ergebnis wird durch die Analyse von Zugriffen auf verschiedene Rubriken des H-Soz-u-Kult Webservers gestützt. Die am häufigsten aufgerufenen Rubriken sind die Rezensionen, die mit deutlichem Abstand führen. Es folgen Berichte über Konferenzen, Stellenausschreibungen, Konferenzankündigungen und Zeitschrifteninhaltsverzeichnisse. Die von der Redaktion verantworteten Primärtexte, Artikel, Themenschwerpunkte und Diskussionen liegen erst im hinteren Teil der Zugriffe. Vgl. H-Soz-u-Kult „Mediadaten“ jeweils aktuell unter: <<http://hsozkult.geschichte.hu-berlin.de/index.asp?pn=mediadaten>>.

20 ULM / Infas (2003), S. 2-3 beschreibt das informelle Netz von Kollegen/innen, mit denen man beruflich verkehrt, als wichtiges Informationsreservoir der professionellen Arbeit. Anschaulich wird die elektronische Form des Networkings anhand von rund 10.000 H-Soz-u-Kult-Beiträgen, die in circa drei Jahren von Nutzern/innen an andere Nutzer/innen über die Funktion „Beitrag versenden“ der Website verschickt wurden.

Einschätzung der fachlichen Relevanz von Informationen verschiedenster Art ermöglichen. Zum zentralen Instrument des Zugriffs auf diese verstreuten Informationseinheiten entwickeln sich die großen Suchmaschinen. Ihre Bedeutung als universeller Zugang zu Inhalten wächst.²¹ Nutzer/innen erhalten hier nicht nur Informationen über den Ort einer Veröffentlichung im Netz, sondern auch eine automatisierte Bewertung des Dokuments selbst, die, ähnlich dem Ranking in den etablierten „Citation Indexes“, durch die Anzahl von Referenzen auf den Text gebildet wird. Texte und Dokumente, die in dieser Hinsicht eine große Zahl von Bezügen aufweisen können, erlangen durch bessere Platzierung leichter Bekanntheit und Aufmerksamkeit. Dass die Suchmaschine Google inzwischen auch bei der Frage nach den bevorzugten Wegen speziell bibliografischer Recherchen unmittelbar hinter den primären Bibliothekskatalogen rangiert, zeigt Abbildung 4.

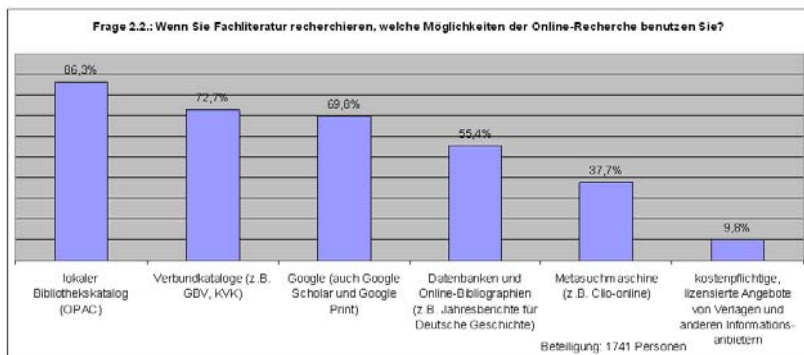


Abbildung 4: Suchmaschinennutzung bei bibliografischen Recherchen

Damit relativiert sich zunehmend die Bedeutung des Veröffentlichungsor-

21 So erfolgen über 60 Prozent der Zugriffe auf einzelne Inhaltsseiten der H-Soz-u-Kult-Homepage direkt über die Suchmaschine Google. Da H-Soz-u-Kult per Mail Adressen zum Aufruf von Beiträgen auf dem Server verschickt, dürfte dieser Prozentsatz im Vergleich zu anderen Angeboten noch relativ niedrig sein, da viele Beiträge direkt aufgerufen werden.

tes zugunsten einer durch die Menge der im Internet zur Verfügung stehenden Referenzen erzeugten Prominenz des einzelnen Textes. So entscheidet nicht mehr allein die Veröffentlichung in einer renommierten Schriftenreihe, sondern auch das Vorhandensein von ausreichenden Meta-Informationen und Rückbezügen im Netz über die Sichtbarkeit einer Veröffentlichung. Wissenschaftliche Autoren/innen sind also gut beraten, auf leichte Zugänglichkeit ihres veröffentlichten Werkes zu achten, was eine Veröffentlichung im Open Access voraussetzt. Sie sollten zudem dafür Sorge tragen, dass weitere Referenzen, Hinweise und Erwähnungen über Autor/in und Werk im Internet zu finden sind. Dies geschieht fast automatisch, sobald Autoren/innen eine aktive Veröffentlichungstätigkeit im Netz pflegen, zum Beispiel als Beiträger/innen für Mailinglisten, als Mitarbeiter/innen an Artikeln von Online-Wörterbüchern wie Wikipedia oder indem sie ihr professionelles Profil, inklusive der zugehörigen Publikationen und Referenzen, veröffentlichen.²² Die eingangs festgestellte Diskrepanz zwischen Autorschaft und Nutzung von Web-Publikationen wird in dieser Form des „gebenden und nehmenden“ Internet-Gebrauchs um einiges relativiert.

Aber, und das wird deutlich, wenn man sich noch einmal Abbildung 1 vor Augen führt: Stabilität und transparente Kontextinformationen, die die „Zitierfähigkeit“ von elektronischen Publikationen für den wissenschaftlichen Gebrauch herstellen, sind über die Sicherstellung von Open Access hinaus zentrale Voraussetzungen für die Nutzung von Online-Publikationen. Nicht institutionell gesicherte Inhalte (Wikis) oder formal nicht standardisierte Eigenveröffentlichungen stoßen auf Vorbehalte im wissenschaftlichen Gebrauch. Allein wenn Autorschaft, Entstehungszeitpunkt und Entstehungskontext eines Textes gesichert sind, lässt sich dieser in wissenschaftlichen Argumentationen verwenden.

Als ein vorläufiges Fazit ließe sich festhalten: Zur Einschätzung der gesamten Publikationskultur in einem Fach gehört die Beachtung sowohl

22 Möglich ist dies beispielsweise an zentraler Stelle im Clio-online Forscher/innen Verzeichnis für die Geschichtswissenschaften, vgl. <<http://www.clio-online.de/forscherinnen>> (25.03.2006).

der Produktion wie auch der Nutzung elektronischer Veröffentlichungen. Aus den vorhandenen Bedenken von Autoren/innen hinsichtlich einer Publikation im Netz sollte nicht zu schnell auf die tatsächliche Bedeutung von elektronischen Publikationen in der Praxis wissenschaftlicher Nutzung geschlossen werden. Auch wenn Veröffentlichungen standardisiert in Fachportalen oder Dokumentenservern erfolgen, führt ihre erleichterte Nutzung zu neuen qualitativen Auszeichnungen in Form von Referenzen. Trotz standardisierter Präsentationsverfahren, die auf so manchem Universitäts-Dokumentenserver noch sehr schlicht wirken, können Veröffentlichungen sich so leicht von der Masse anderer Inhalte abheben. Neben einer weiter bestehenden Qualifizierung eines Werks durch etablierte redaktionelle Verfahren im Sinne des „klassischen“ Publikationswesens entstehen im Netz neue Informationsbestände, die den Nutzern/innen die schnelle Einschätzung des Gebrauchswerts von Publikationen erlauben. Freier Zugang zum Werk und eine aktive Nutzung des Internets als Publikationsmedium unterstützen die Entstehung dieser „Meta“-Informationen und fördern damit die Sichtbarkeit und Bekanntheit eigener elektronischer Veröffentlichungen

Systematische Untersuchungen zur Nutzung von fachwissenschaftlichen Internet-Angeboten stehen noch weitgehend am Anfang, auch die hier vorgetragenen Überlegungen stellen noch nicht den letzten Stand der Auswertung der genannten Umfragedaten dar. Aber für die Gesamtbewertung einer fachlichen Publikationskultur und die Begründung förderpolitischer Maßnahmen sollte das Rezeptionsverhalten von Wissenschaftlern/innen genauso viel Beachtung finden, wie ihre Präferenzen bei der Veröffentlichung ihrer Werke. Der Blick auf die Geschichte von Marketing und Marktforschung zeigt, dass Studien des tatsächlichen Nachfrageverhaltens von Kunden/innen überwiegend der von der Produktionslogik bestimmten Unternehmenspraxis nachfolgten.²³ Viele Anbieter einer expandierenden New Economy am Ende des letzten Jahrhunderts

23 Zum Einstieg: Engel, Alexander, Tagungsbericht: Grundzüge der Marketinggeschichte. Vom betrieblichen Absatzinstrument zur universellen Sozialtechnik, in: H-Soz-u-Kult, 24.03.2006, vgl. <<http://hsozkult.geschichte.hu-berlin.de/tagungsberichte/id=1084>> (20.04.2006).

überschätzten die tatsächliche Nachfrage für ihre Angebote. Angesichts der extremen Knappheit von öffentlichen Ressourcen für die Informationsinfrastruktur sollten auch im Rahmen von steuerfinanzierten Angeboten Nutzungsstudien vermehrt zur Steuerung von Fördermaßnahmen eingesetzt werden, die dann umso näher am wissenschaftlichen Bedarf liegen.

Karsten Borgmann ist Wissenschaftlicher Mitarbeiter im Projekt Clio-online an der Humboldt-Universität zu Berlin sowie langjähriger Redakteur von H-Soz-u-Kult. E-Mail: borgmannk@geschichte.hu-berlin.de

ZUR RECHTLICHEN PROBLEMATIK DES ELEKTRONISCHEN PUBLIZIERENS

von Karl-Nikolaus Peifer

Die technischen Möglichkeiten zur elektronischen Publikation durch Universitäten sind heute gegeben. Auch zur Umsetzung des Prinzips des offenen Zugangs nach der „Berliner Erklärung“ benötigt man das Urheberrecht, denn der Raum des urheberrechtlich Ungeschützten ist mittlerweile sehr klein geworden; auch Open Access-Publikationen sind letztlich Lizenzproduktionen. Das Hauptproblem wird darin bestehen, nicht verlagsgebundene Wissenschaftler/innen zu Open Access-Publikationen zu bewegen. Gelingt dies, so ist ein Lizenzvertrag mit den Autoren/innen zu schließen, für die die Creative Commons-Lizenz eine erste Annäherung erlaubt. Sie ist allerdings auf das US-amerikanische Urheberrecht zugeschnitten und daher nur eingeschränkt importfähig. Bei verlagsgebundenen Autoren/innen wird man die Verlage zur Kooperation bewegen müssen. Die Bereitschaft hierzu besteht eventuell, wenn ansonsten eine Abwanderung insbesondere jüngerer Autoren/innen in die Open Access-Szene droht. Die dienst- und urhebervertragsrechtlichen Bedingungen für Open Access sind dürftig. Weder können Universitäten derzeit ihre Wissenschaftler/innen zwingen, die Formulierung von Forschungsergebnissen in Universitätsforen einzustellen, noch können Verlage rechtlich gezwungen werden, den Autoren/innen „Repository Rights“ zu belassen. Beides erforderte Rechtsänderungen, deren Realisierbarkeit gering ist.

Einleitung

Die Verteuerung von Zeitschriftenabonnements, die elektronische Absi-

cherung elektronischer Verlagsprodukte und die immer noch massive Bezuschussung der Produktion, Distribution und Anschaffung wissenschaftlicher Literatur haben in den Natur- und Geisteswissenschaften den Ruf nach neuen Publikationsformen laut werden lassen. „Offener Zugang zu Wissen“ lautet die Forderung der Stunde. Die elektronische Publikation und das elektronische Universitätsarchiv werden aus vielen öffentlichen Töpfen gefördert. Braucht das elektronische Publizieren rechtliche Begleitung oder ist das Urheberrecht ein Hindernis auf dem Weg zu modernen Formen der Wissensverbreitung? Die *Berliner Erklärung über offenen Zugang zu wissenschaftlichem Wissen*¹ geht davon aus, dass sich die praktischen und wirtschaftlichen Bedingungen für die Verbreitung von wissenschaftlichen Inhalten „grundlegend verändert“ haben. In der Tat haben sich Publikationswege verbessert und verbilligt, sofern eine Publikation nicht auf dem Papier-, sondern auf elektronischem Wege erfolgt. Elektronische Publikationen lassen sich zwar durch kommerziell tätige Unternehmen, aber auch im Selbstverlag realisieren. Die bloße Veröffentlichung eines Inhaltes ist jedermann möglich, der über einen technischen Zugang zu einem Server und ausreichend Speicherraum verfügt. Die Knappheiten der Papierwelt spielen dabei kaum noch eine Rolle. Ebenso sieht es auf der Nutzerseite aus. Der Traum vom jederzeitigen schnellen Zugang zum Wissen der Welt ist technisch realisierbar. Die an den Juristen gestellte Frage lautet, ob der Zugang auch rechtlich ohne weiteres realisierbar ist. Urheberrechtler/innen müssen sich der Frage stellen, ob es einen Paradigmenwechsel im Recht des Geistigen Eigentums vom Eigentums- zum Freiheitsideal gibt oder ob man nach wie vor das Recht des Geistigen Eigentums benötigt, um die Ziele der Berliner Erklärung umzusetzen? Die Antwort darauf kann kurz ausfallen: Auch für die Umsetzung der Berliner Erklärung benötigt man das Urheberrecht. Allerdings ändert

¹ Die Berliner Erklärung über offenen Zugang zu wissenschaftlichem Wissen vom 23.10.2003, abrufbar unter <http://www.zim.mpg.de/openaccess-berlin/berlin_declaration.pdf> (07.02.2006); weitere Nachweise zum Fragenkreis bei Dreier, Thomas, Creative Commons, Science Commons. Ein Paradigmenwechsel im Urheberrecht?, in: Ohly, Ansgar; Bodewig, Theo; Dreier, Thomas u.a. (Hgg.), Perspektiven des Geistigen Eigentums und Wettbewerbsrechts (Festschrift für Gerhard Schricker), München 2005, S. 283-298, 291.

dieses Instrumentarium seine Funktion, wenn und solange es in der Hand der Kreativen verbleibt und von diesen im Sinne der Berliner Erklärung eingesetzt wird. Es ist dann nicht mehr ein Instrument der Zugangskontrolle, sondern ein Mittel, um die an der Quelle gewährten Zugangserleichterungen auch in der Verwertungskette einer Publikation zu sichern. Seine klassische Bedeutung wird das Urheberrecht behalten, wenn es darum geht sicherzustellen, dass das offen zugängliche Werk weiter mit dem Namen seines Autors / seiner Autorin verbunden und unverändert in der Form verbleibt, in der es vom Schöpfer / von der Schöpferin autorisiert wurde.

Elektronisches Publizieren als Rechtsproblem

Wer seine Gedanken in eine individuelle, also durch seinen persönlichen Stil geprägte sprachliche, grafische oder bildliche Form bringt, um sie der Öffentlichkeit oder den Fachkollegen/innen mitzuteilen, schafft regelmäßig ein urheberrechtlich geschütztes Werk, sei es ein Sprachwerk, ein Werk der Bildenden Kunst, des angewandten Grafikdesign oder aber eine wissenschaftliche Darstellung in Form eines Modells, einer Zeichnung oder eines komplexen Diagramms.² Der urheberrechtliche Schutz hieran entsteht allein durch den Akt der Schöpfung, also eine Realhandlung, die unabhängig vom Willen der Schöpfer/innen ist. Ist das Werk einmal geschaffen, so ist der urheberrechtliche Schutz hieran unverzichtbar und unveräußerlich (Paragraf 29 Absatz 1 Urheberrechtsgesetz – UrhG). Das ist eine Besonderheit des deutschen Rechts, die allerdings in vielen weiteren kontinentaleuropäischen Urheberrechtssystemen geteilt wird.

Soll das Werk auf einem elektronischen Server gespeichert und zum Abruf bereitgestellt werden, so erfordert dies Nutzungshandlungen derjenigen, die die elektronische Publikation verantworten. Diese Nutzungshandlungen können Urheber/innen zwar dulden; um aber Herausgebern/innen oder Betreibern/innen einer elektronischen Plattform Rechtssicherheit zu verschaffen, müssen Urheber/innen sie jedoch lizenzie-

² Nach verbreiteter Auffassung genügt dazu auch eine geringe Kreativität, vgl. Heermann, Peter W., Der Schutzzumfang von Sprachwerken der Wissenschaft und die urheberrechtliche Stellung von Hochschulangehörigen, Gewerblicher Rechtsschutz und Urheberrecht (GRUR), 1999, S. 468-476, 469.

ren, sie müssen also Nutzungsrechte bezüglich all derjenigen Handlungen einräumen, die erforderlich sind, um die Publikation zu fixieren, sie gegebenenfalls als elektronische Datei im PDF- oder HTML-Format zu versenden oder aber auf einem Server so bereit zu halten, dass sie durch Nutzer/innen zu Zeiten und von Orten ihrer Wahl abgerufen werden können. Mit den drei erwähnten Handlungen sind Verwertungsrechte der Urheber/innen genannt, die als Vervielfältigungs-, Verbreitungs- und als Recht der öffentlichen Zugänglichmachung den Urhebern/innen des Werkes als ausschließliche Befugnisse gesetzlich zugewiesen sind (Paragrafen 16, 17, 19a UrhG). Damit ist der Prozess der Verbreitung von Werken, die urheberrechtlich geschützt sind, lückenlos erfasst. Das Urheberrecht als gesetzliches Instrumentarium hat seit den 1990er Jahren sukzessive sämtliche Handlungen einbezogen, die als Besonderheiten nur in der elektronischen, nicht aber in der gedruckten Welt anfallen. Dazu gehört insbesondere die Befugnis, temporäre Vervielfältigungen den Urhebern/innen zuzuweisen (Paragraf 16 UrhG), soweit diese nicht allein den Zweck verfolgen, eine Übertragung von Daten in elektronischen Netzen oder die rechtmäßige Nutzung von Werken zu ermöglichen (Paragraf 44a UrhG). Zwar darf das Browsen im Internet zustimmungsfrei von jedermann vorgenommen werden, das Lesen eines elektronischen Buches und das Laden eines Computerprogramms fallen jedoch unter die Ausschließlichkeitsrechte der Urheber/innen oder der sonstigen Rechtsinhaber/innen (zum Beispiel Verleger/innen). Wer eine elektronische Publikation verantwortet, benötigt mithin eine Reihe von Nutzungsrechten, die er sich von den Urhebern/innen als ersten und originären Inhabern/innen dieser Befugnisse einräumen lassen muss.

Das Urheberrecht als Zugangs- und Verbreitungskontrollinstrument

In der Welt der „alten Medien“ werden die Ausschließlichkeitsrechte nicht von den Werkschöpfern/innen selbst, sondern von den Zugangsvermittlern, also Verlagen und Medienunternehmen geltend gemacht. Sie sind zwar auch darauf angewiesen, sich diese Rechte von den Urhebern/innen einräumen zu lassen, doch erfolgt dies üblicherweise bei Verlagsproduktionen über die Einräumung umfassender Nutzungsrechte, die sowohl die Print-

als auch die elektronische Publikation umfassen (*Buy-Out*). Urheber/innen können sich in deutschen und in vielen kontinentaleuropäischen Rechtsordnungen ihrer Verwertungs- und Persönlichkeitsrechte an dem von ihnen geschaffenen Werk gleichwohl nicht gänzlich entäußern. Stets behalten sie das Recht, gegen Plagiate oder Entstellungen ihrer Werke vorzugehen (Paragrafen 13, 14 UrhG). Allerdings können sie nicht mehr allein eine elektronische Publikation autorisieren, sofern sie die Rechte an ihrem Werk einem Verleger eingeräumt haben. Die Verleger/innen behalten dann eine Kontrollposition, die sie selbständig, also ohne Abstimmung mit den Kreativen, ausüben können. In dieser Funktion kann das durch das Urheberrecht gewährte Ausschließlichkeitsrecht auch dazu verwendet werden, das Werk und seine Bestandteile umfassend zu kontrollieren. Das umfasst die Befugnis, Nutzer/innen von den meisten Nutzungsarten auszuschließen bzw. die Gestattung von Nutzungshandlungen von der Zahlung einer angemessenen Lizenzgebühr abhängig zu machen. Das Instrumentarium wurde in der Vergangenheit stetig rechtlich fortentwickelt, wenn die technischen Möglichkeiten zur einfachen Vervielfältigung und Verbreitung von Inhalten faktische Kontrollbefugnisse behinderten. Interessanterweise fanden diese Fortentwicklungen statt, wann immer die technischen Möglichkeiten zur Verbreitung von Werken derart erleichtert wurden, dass die Kontrollbefugnisse desjenigen, der in die Verwertung des Manuskripts investiert hatte, zurückgingen (*Napsterproblematik*). Die Entwicklung des Urheberrechts ist daher stets maßgeblich auf die Initiative der Werkverwerter/innen, nicht aber auf die Initiative der Kreativen zurückzuführen. Das lässt sich anhand der Entwicklung vom Buchdruck bis zur Möglichkeit der Rundfunksendung, Digitalisierung und Internetverbreitung von Werken nachvollziehen.³ Um das Werk zur handelsfähigen Ware zu machen, bedarf es nämlich eines starken und umfassenden Urheberrechts.

³ Peifer, Karl-Nikolaus, Individualität im Zivilrecht, Tübingen 2001, S. 61; Woodmansee, Martha, The Genius and the Copyright. Economic and Legal Conditions of the Emergence of the 'Author', in: 17 Eighteenth Century Studies. (1983/84), S. 425-448. Daher ist für die ökonomische Analyse des Verbreitungsprozesses allein die Angebotsfunktion der Verleger maßgeblich, vgl. Giesecke, Michael, Der Buchdruck in der frühen Neuzeit, Frankfurt am Main 1998, S. 642.

In der Hand der Verleger/innen, die ihre Nutzungsbefugnisse von den Urhebern/innen ableiten, wird das Urheberrecht zum Zugangs- und Nutzerkontrollinstrument. Es kontrolliert den Weg des Werkes über die Einräumung von Lizenzen (Nutzungsrechten, Paragraph 31 UrhG). Lizenzen können dazu verwendet werden, Nutzerhandlungen in zeitlicher, räumlicher und inhaltlicher Hinsicht einzuschränken (Paragraph 31 Absatz 1 Satz 2 UrhG). Die Lizenzen können ihrerseits weiter übertragen (Paragraph 34 UrhG) oder durch abhängige Sublizenzen weiter aufgefächert werden (Paragraph 35 UrhG). Dieser Prozess der Aufspaltung und Weitergabe von Nutzungsrechten ist grundsätzlich nur mit Zustimmung der Urheber/innen zulässig (Paragrafen 34 Absatz 1 Satz 1, 35 Absatz 1 Satz 1 UrhG). In der Logik der Lizenzeinräumung liegt es, dass im Zweifel, das heißt bei Fehlen von ausdrücklichen Vereinbarungen, eine Befugnis als nicht eingeräumt gilt, also bei den Rechtsinhabern/innen zurückbleibt. Doch werden Verleger/innen ihren Vermarktungsbeitrag in der Regel davon abhängig machen, dass die Urheber/innen ihnen die benötigten Befugnisse umfassend einräumen. Zweifel über den Umfang der eingeräumten Befugnisse werden sie nicht aufkommen lassen. Das Lizenzrecht ist daher maßgebliches Kontrollinstrument auch bei der kommerziellen Verbreitung elektronischer Publikationen im Verlagsgewerbe. Es sorgt auf der Nutzerseite dafür, dass Zugangsrechte typischerweise an Lizenzzahlungen geknüpft werden. Für Verleger/innen stellt die Lizenzmöglichkeit das nachvollziehbare Mittel dar, ihre Investitionen in die Produktion und Distribution der Werkstücke amortisierbar zu machen.

Open Access und Creative Commons als Grundlage freien und erweiterten Zugangs bei elektronischen Publikationen

Open Access beschreibt schlagwortartig das Interesse daran, dass im Bereich kreativer Schaffensprozesse Nutzer/innen in weitgehendem Maße Zugang zu Werken erhalten. Gedacht ist vor allem, aber nicht nur, an den Zugang durch Nutzer/innen aus den Bereichen Wissenschaft und Forschung. Die Umsetzung der Idee setzt voraus, dass diejenigen Personen, die Urheberrechte an den Werken halten, der Nutzung zustimmen. Bei verlagsgebundenen Produktionen erfordert dies die Zustimmung nicht nur

der Kreativen, also der Urheber/innen, sondern auch derjenigen, die als Verleger/innen ausschließliche Nutzungsrechte an den betreffenden Werken haben.

Die in den USA und in Großbritannien praktizierte Lösung eines wissenschaftlichen Gemeinschaftseigentums (Creative Commons) ist hierzulande rechtlich nicht umsetzbar. Creative Commons (Science Commons) geht davon aus, dass jedenfalls die durch staatliche Finanzierung entstandenen Forschungsergebnisse in einen Pool der Gemeinfreiheit Eingang finden können, aus dem jede/r frei schöpfen kann.⁴ Creative Commons hat im Softwarebereich eine Parallele im Bereich der Public Domain-Software.⁵ Im deutschen Urheberrecht fehlt es an einer rechtlichen Grundlage für Creative Commons-Überlegungen, weil der urheberrechtliche Schutz – anders als im US-amerikanischen Recht – unverzichtbar ist.⁶ Creative Commons im Sinne von Gemeinfreiheit entsteht daher frühestens nach Ablauf der Schutzfrist eines Werkes (in der Regel 70 Jahre nach dem Tod des Urhebers / der Urheberin, Paragraph 64 UrhG). Open Access-Modelle müssen daher im Rahmen von Lizenzvereinbarungen realisiert werden, sofern urheberrechtlich geschützte Inhalte dabei Verwendung finden.

4 Vgl. <<http://creativecommons.org/>; <http://creativecommons.org/>> (07.02.2006); zu den dahinter stehenden Ideen vgl. aus urheberrechtlicher Sicht Lessig, Lawrence, *Free Culture*, New York 2004, S. 282. In Großbritannien: <<http://creativearchive.bbc.co.uk/>> (07.02.2006).

5 Der Begriff ist mit „gemeinfreie Software“ zu übersetzen, im Einzelnen kann Public Domain-Software allerdings auch für Software verwendet werden, die lediglich eine gewisse Zeit frei benutzt werden kann (Shareware), vgl. im Einzelnen Jaeger, Till; Metzger, Axel, *Open Source Software. Rechtliche Rahmenbedingungen der Freien Software*, München 2002, S. 1.

6 Der Grundsatz wird aus § 29 Abs. 1 UrhG abgeleitet, wonach das Urheberrecht als Ganzes oder in seinen Bestandteilen nicht übertragen werden kann, vgl. Schrickler, Gerhard, *Die Einwilligung des Urhebers in entstellende Änderungen des Werkes*, in: Forkel, Hans; Kraft, Alfons (Hgg.), *Beiträge zum Schutz der Persönlichkeit und ihrer schöpferischen Leistung*, Frankfurt am Main 1985, S. 409-419, 413; zur früheren Rechtslage ebenso Gierke, Otto von, *Deutsches Privatrecht Band I*, Leipzig 1895, S. 702ff.

Ausschließlichkeit und Freiheit: Was ist geschützt?

Grundsätzlich schützt das Urheberrecht nur verkörperte und persönlich individuell geformte Gedanken. Es gilt der Grundsatz, dass der (ungeformte) Inhalt selbst frei bleibt und nur die konkrete Form der Gedankenäußerung Schutz genießt (Idea-Expression). Gesetze und Gerichtsurteile, abstrakte Ideen, Theorien, Methoden, Entdeckungen, Stilrichtungen und Alltagsäußerungen, wie Briefe und gewöhnliche Kommunikationsbeiträge stellen keine persönlich-geistigen Schöpfungen im Sinne des Urheberrechts dar und dürfen daher grundsätzlich von jedermann frei verwertet werden. Das betrifft auch einige der in der *Berliner Erklärung* erwähnten Gegenstände, wie Rohmaterialien und Metadaten. Auch Quellenmaterial ist in der Regel urheberrechtsfrei, weil vielfach die Schutzfristen abgelaufen sein werden. Die Verwertungsfreiheit endet allerdings, wenn die Gedanken ihrerseits in eine organisierte Informationssammlung (Datenbank, Paragraph 87a UrhG) eingebettet sind und der Zugang zu dieser Informationssammlung faktisch kontrollierbar ist. Ebenso ist es, wenn Quellenmaterial „remonetisiert“ wird, etwa dadurch, dass Archiv- oder Museumsfotografen/innen die Bestände ablichten. An solchen Ablichtungen entsteht wiederum ein eigener Lichtbildschutz (Paragraph 72 UrhG), der zwar eine kürzere Schutzfrist von „nur“ 50 Jahren ab erstem Erscheinen hat (Paragraph 72 Abs. 2 UrhG), allerdings für diesen Zeitraum dafür sorgt, dass Vervielfältigung, Verbreitung und öffentliche Zugänglichmachung erlaubnispflichtig werden.⁷ Die Freiheit von Verwertungshandlungen ist daher in der elektronischen Welt der Ausnahmefall. In der Regel wird man die Handlung lizenzieren lassen müssen.

Ausgenommen von der Zustimmungspflicht sind Handlungen, die unter urheberrechtliche Schrankenbestimmungen fallen (Paragraphen 44a bis 63a UrhG). Im Bereich der wissenschaftlichen Nutzung zählen hierunter die Befugnisse zum wissenschaftlichen Zitat (Paragraph 51 UrhG), zur Vervielfältigung von Werken zu Zwecken des persönlichen (auch wissenschaftlichen) oder Archivgebrauchs (Paragraph 53 UrhG), zur Versendung

7 Vgl. Bundesgerichtshof (BGH), in: *Gewerblicher Rechtsschutz und Urheberrecht (GRUR)* 1990, S. 669, 673 – Bibelreproduktion.

von Kopien durch öffentliche Bibliotheken (künftig Paragraf 53a UrhG-Entwurfassung), zur Anzeige von Inhalten an Leseplätzen in Bibliotheken, Museen und Archiven (künftig: Paragraf 52b UrhG-Entwurfassung), zur öffentlichen Wiedergabe im Hörsaal (Paragraf 52 UrhG) und zur öffentlichen Zugänglichmachung in internen Forschernetzen (Paragraf 52a UrhG). Sämtliche genannten Befugnisse – mit Ausnahme der Zitatschranke – sind zwar zustimmungs-, nicht aber vergütungsfrei. Daran zeigt sich, dass Zugangsfreiheit nicht identisch ist mit Gratiszugang.

Hinzu kommt, dass die urheberrechtlichen Schranken nur die Rezipienten/innen privilegieren, nicht jedoch auch die Plattformbetreiber/innen, also die Verwerter/innen, zum Beispiel die Herausgeber/innen einer elektronischen Zeitschrift. Das lässt fragen, unter welchen Umständen die Wissensproduzenten/innen dafür sorgen können, dass die von ihnen öffentlich zugänglich gemachten Inhalte auch frei zugänglich bleiben. Dies ist wiederum die Aufgabe des Lizenzrechts, also der Ausgestaltung von Nutzungsbedingungen. Bevor auf deren Inhalt eingegangen wird (Kapitel 7), ist zu fragen, wer befugt ist, Lizenzen zu erteilen.

Lizenzierungsbefugnisse: Wer darf Lizenzen erteilen?

Forschungsergebnisse werden überwiegend im Rahmen von arbeits- oder beamtenrechtlichen Beschäftigungsverhältnissen erbracht. Der allgemeine arbeitsrechtliche Grundsatz, wonach das Arbeitsergebnis den Arbeitgebern/innen zusteht, gilt nicht einschränkungslos für urheberrechtlich geschützte Werke, jedenfalls nicht im Hinblick auf Hochschullehrer/innen.⁸ Als Urheber/in gilt stets der (menschliche) Schöpfer. Arbeitge-

8 Zentral hierzu ist der 112. Band, S. 243 der amtlichen Sammlung des Bundesgerichtshofs vom 27.09.1990. Dort ging es um die Frage, ob die Erben eines verstorbenen Hochschullehrers Anspruch auf Herausgabe von Aufzeichnungen über archäologische Grabungsaufzeichnungen und wissenschaftliche Korrespondenz haben. Der BGH nahm an, dass die Arbeitsergebnisse originär dem Wissenschaftler selbst zustünden und er allenfalls eine Pflicht zur Andienung gegenüber dem Dienstherrn habe. Vgl. im Übrigen OLG Karlsruhe, in: Gewerblicher Rechtsschutz und Urheberrecht (GRUR) (1988), S. 536, 540; GRUR (1991), S. 523, 525 – Grabungsmaterialien; Ullmann, Eike, Das urheberrechtlich geschützte Arbeitsergebnis. Verwertungsrecht und Vergütungspflicht, in: GRUR (1987), S. 6, 8; Ulmer, Eugen, Urheber- und

ber erhalten je nach Ausgestaltung des Arbeitsverhältnisses nur Nutzungsrechte an den „in Erfüllung der Dienstpflicht“ geschaffenen Werken. Das setzt voraus, dass die Publikation Gegenstand des Arbeitsverhältnisses ist. Das wiederum wird man bei Wissenschaftlern/innen nur sehr eingeschränkt annehmen können, da zur Wissenschaftsfreiheit auch die Freiheit gehört, Erkenntnisse unveröffentlicht zu lassen. Nicht einmal eine Pflicht zur Publikation ist daher ohne weiteres vertraglich oder dienstrechtlich geschuldet.

In der rechtswissenschaftlichen Literatur wird überwiegend angenommen, dass auch Forscher/innen, die sich zur Publikation entschließen, keine Pflicht aus dem Anstellungs- oder Beamtenverhältnis haben, ihre Publikation der Hochschule zur Verwertung anzudienen.⁹ Stets haben Forscher/innen das Recht, selbst über das „Ob“ und „Wie“ einer Erstveröffentlichung ihrer Ergebnisse zu entscheiden.¹⁰ Die Frage wird nur im Arbeitnehmererfindungsrecht nach Abschaffung des so genannten „Hochschullehrerprivilegs“ anders gesehen.¹¹ Im Urheberrecht ist eine Lösung, welche eine Andienungspflicht der Forscher/innen gegenüber ihrer Universität oder einem Projektförderer vorsieht, konzeptionell zwar nicht vollständig ausgeschlossen, sie wird aber erheblichen Widerstand der Betroffenen auslösen, sofern es keine Kompensationen für die Andienung (wie im Arbeitnehmererfindungsrecht) gibt. Zudem wird sie eine wissenschaftspolitische Debatte darüber auslösen, ob und inwieweit die Forschungsfrei-

Verlagsrecht, 3. Aufl. Berlin 1980, S. 403; a.A. Verwaltungsgericht Berlin, in: Neue Juristische Wochenschrift (1978), S. 848.

9 Leuze, Dieter, Kritische Anmerkungen zu § 42 Arbeitnehmererfindungsgesetz, in: Gewerblicher Rechtsschutz und Urheberrecht (2005), S. 27-33, 28; Schack, Haimo, Urheber- und Erbeervertragsrecht, Tübingen 2005, Randnummer 982 mit weiteren Nachweisen; kritisch neuerdings aber Pflüger, Thomas; Ertmann, Dieter, E-Publishing und Open Access – Konsequenzen für das Urheberrecht im Hochschulbereich, in: Zeitschrift für Urheber- und Medienrecht (ZUM), Heft 6 (2004), S. 436-443, 441.

10 Vgl. hierzu Hansen, Gerd, Zugang zu wissenschaftlicher Information – alternative urheberrechtliche Ansätze, in: Gewerblicher Rechtsschutz und Urheberrecht – Internationaler Teil (2005), S. 378, 380.

11 Zu der Neuregelung Bartenbach, Kurt; Volz, Franz-Eugen, Erfindungen an Hochschulen. Zur Neufassung des § 42 Arbeitnehmererfindungsgesetz, in: Gewerblicher Rechtsschutz und Urheberrecht (2002), S. 743-758.

heit Menschen oder den sie anstellenden Institutionen zusteht.¹² Im Ergebnis würde zumindest eine Gesetzesänderung erforderlich, um dem öffentlichen Dienstherrn die Möglichkeiten zu verschaffen, auf die Publikationsbefugnis der selbständig in der Universität tätigen Wissenschaftler/innen einzuwirken.¹³ Angesichts der erbitterten Kämpfe, die seit 2003 um die Anpassung des Urheberrechts an die Erfordernisse der Informationsgesellschaft geführt werden, erscheint die Durchsetzung einer solchen Änderung nicht sehr wahrscheinlich.

In der Regel wird es also dabei bleiben, dass Wissenschaftler/innen selbst zunächst originär berechtigt sind, ihre Werke zu verwerten. Im Hinblick auf neu geschaffene Werke sind die Ziele der Berliner Erklärung realistischer, weil diese Werke den Urhebern/innen selbst zustehen und sie diesbezüglich lizenzbefugt und möglicherweise auch lizenzbereit sind. Die Situation ändert sich, sobald Wissenschaftler/innen in Verlagsprojekte eingebunden werden. Sofern Urheber/innen Verlagsverträge abschließen, werden sie den Verlegern/innen genau bezeichnete Nutzungsrechte einräumen, die im Ergebnis die umfassende wirtschaftliche Verwertung durch die Verleger/innen ermöglichen. Dann verschiebt sich die Lizenzbefugnis auf die Verwerter/innen. Open Access wird es hier nur geben, wenn die Verleger/innen kooperieren.

Eine Lücke für Altvorgänge gilt in Grenzen für diejenigen Nutzungsarten, die zum Zeitpunkt der Einräumung von Verlagsrechten noch nicht als wirtschaftlich eigenständig und technisch realisierbar bekannt waren, so etwa die Rechte zur elektronischen Nutzung im Internet.¹⁴ Solche unbe-

12 Das Grundrecht der Wissenschaftsfreiheit nach Art. 5 Abs. 3 GG steht sowohl den in der Universität tätigen Wissenschaftlern/innen als Individuen als auch der Universität als Einrichtung selbst zu, so Bundesverfassungsgericht in der amtlichen Entscheidungssammlung des Gerichts Band 68, S. 193, 267 und Band 75, S. 192, 196.

13 Ein Gesetzgebungsvorschlag hierzu findet sich bei Pflüger; Ertmann, ZUM, Heft 6 (2004), S. 436, 441 (wie Anm. 9).

14 Das Internet mit seinen Möglichkeiten des elektronischen Abrufs gilt als bekannte Nutzungsart frühestens seit Mitte der 1990er Jahre, OLG Hamburg, in: Zeitschrift für Urheber- und Medienrecht (ZUM) (2000), S. 870, 872. Die Digitalisierung von Werken war im Bereich der CD-Produktion seit 1979 möglich und ab 1983 marktreif, vgl. OLG Düsseldorf, in: ZUM (2001), S. 164, 165; Kammergericht

kannten Nutzungsarten können nach noch (7/2006) geltender Rechtslage nicht wirksam eingeräumt werden (Paragraf 31 Absatz 4 UrhG), sie verbleiben also bei den Urhebern/innen. Haben Urheber/innen in Altverträgen ausschließliche Vervielfältigungs-, Verbreitungs- und Wiedergaberechte einem Verlag oder Medienunternehmen eingeräumt, so fällt die Nutzung des Werkes im Internet nicht darunter. Sofern Verleger/innen diese Verwertungsform nutzen möchten, müssen sie die Rechte hierfür nachlizenzieren. Das gilt übrigens auch für Verwertungsgesellschaften, die mit den betreffenden Urhebern/innen Wahrnehmungsverträge geschlossen haben.¹⁵ In der zurzeit anhängigen Urheberrechtsreform (so genannter *Zweiter Korb*) soll die hier erörterte Regelung allerdings aufgehoben werden.¹⁶ Für Altverträge sieht das Gesetz vor, dass Verwerter/innen, denen umfassende ausschließliche Rechte eingeräumt wurden, künftig auch unbekannte Nutzungsarten ausüben dürfen, sofern die Urheber/innen dem nicht binnen eines Jahres widersprechen.¹⁷ Diese Regelung wird im Ergebnis dazu führen, dass auch voll amortisierte Repertoires, die in den Archiven der Verwerter/innen lagern, von diesen zunächst allein ausgewertet werden. Die Vision, die Archive für die Öffentlichkeit frei nutzbar zu gestalten, bleibt insoweit ein Traum.

Die Möglichkeit, Rechte für elektronische Plattformen zurückzuhalten, sehen Verlagsverträge typischerweise nicht vor. Sie müsste gesondert ausgehandelt werden. Das wird nur in Kooperation mit den Wissenschaftsverlagen und mit Unterstützung durch Wissenschaftsorganisationen gelingen. Zwar könnte man sich auch vorstellen, im Urhebervertragsrecht eine

Berlin (KG), in: MultiMedia und Recht (MMR) (1999), S. 727, 725.

15 Dreier, Thomas; Schulze, Gernot, Urheberrechtsgesetz. Kommentar, München 2006, Kommentierung Nr. 83 zu § 31 UrhG; OLG Hamburg, ZUM (Fußn. 4) 2002, S. 480, 481 (Handy-Klingeltöne). Die Einzelheiten sind verzwickelt. So gilt der vorbeschriebene Grundsatz erst seit Inkrafttreten des Urheberrechtsgesetzes 1965, nicht aber für Altverträge, die vorher geschlossen wurden, Dreier; Schulze, Kommentierung Nr. 86 zu § 31 UrhG (wie oben).

16 Art. 1 Nr. 3 des Entwurfs eines Zweiten Gesetzes zur Regelung des Urheberrechts in der Informationsgesellschaft vom 26. 02 2006, abrufbar unter <<http://www.bmj.bund.de/media/archive/1122.pdf>> (07.02.2006).

17 § 1371 des Entwurfs vom 26.01.2006 (wie Anm. 16). Die Jahresfrist beginnt voraussichtlich mit dem Inkrafttreten des Gesetzes zu laufen.

Bestimmung vorzusehen, wonach Urheber/innen stets das Recht behalten, ihr verlagsgebundenes Werk etwa nach einer gewissen Karenzzeit auf elektronischen Datenbanken zu hinterlegen (Repository Right)¹⁸, doch würde auch eine solche Regelung eine Gesetzesänderung erfordern, für die im Moment der politische Handlungsspielraum fehlt. Jedenfalls würde ein zwingendes Repository Right einen erheblichen Widerstand der Verlage erzeugen. Dass dieser Widerstand erfolgreich praktiziert werden kann, zeigte die Debatte um die Einführung einer Vorschrift, welche die Zugänglichkeit von Werken in Intranets von Universitäten beabsichtigte. Die nach dem erbitterten Widerstand der Verlage Gesetz gewordene Fassung (Paragraf 52a UrhG) wird nach den derzeitigen Plänen das Jahr 2006 nicht überleben.¹⁹

Um Open Access-Modelle zu ermöglichen, wird man bezüglich verlagsgebundener Werke mithin die Kooperation der Verlage künftig auch in Bezug auf Altproduktionen benötigen. Man wird mit den Verlagen gemeinsame Vermarktungsmodelle entwickeln müssen, die es ermöglichen, den Zielen der *Berliner Erklärung* nahe zu kommen. Das kann etwa bedeuten, dass nur Auszüge aus Verlagsproduktionen elektronisch zugänglich gemacht werden. Volltexte könnten in einer Weise zugänglich gemacht werden, welche die Seitenzahlen des Printprodukts unterschlagen, das heißt in nur eingeschränkt dauerhaft zitierfähiger Form erscheinen. Auch eine Verlinkung mit Verlagsservern dürfte eine Option sein. Die Verlage müssen an dieser Stelle sehen, dass Aufmerksamkeit durch Blick- und Abrufkontakte auch die Bereitschaft erhöht, das Printprodukt zu erwerben, elektronische Server können so zu Marketing-Tools für die Verlage werden. Letztlich muss den Verlagen klar sein, dass bei mangelnder Kooperationsbereitschaft möglicherweise ein Teil ihres angestammten Geschäfts wegbriecht. Diese Gefahr besteht, wenn Wissenschaftlicher/innen freiwillig

18 Ein entsprechender Regelungsvorschlag findet sich bei Hansen, Gerd, Zugang zu wissenschaftlicher Information. Alternative urheberrechtliche Ansätze, in: Gewerblicher Rechtsschutz und Urheberrecht (GRUR) – Internationales (2005), S. 378, 386 (wie Anm. 10).

19 Die durchaus enge Vorschrift des § 52a UrhG ist nach § 137k UrhG „mit Ablauf des 31. Dezember 2006 nicht mehr anzuwenden“. Eine Streichung der Norm sieht der Regierungsentwurf vom 26.01.2006 (vgl. Anm. 16) nicht vor.

für künftige Produktionen elektronisch offen zugängliche Server benutzen.

Lizenzierungspraxis bei Open Access-Modellen

Offener Zugang bedeutet, dass mit öffentlichen Mitteln elektronische Plattformen errichtet werden, auf denen Wissenschaftler/innen publizieren können. Der Preis für die Einstellung der Publikation liegt darin, dass die dort zur Verfügung gestellten Werke zur freien Nutzung durch andere offen stehen müssen. Das entlastet die Etats der Bibliotheken und Nutzer/innen. Bei intelligenten Suchmaschinen findet man die Inhalte schnell, kostengünstig und von jedem Ort aus. Ein entscheidender Vorteil liegt darin, dass Inhalte sofort abrufbar sind, das Wissen also nicht bereits in der Publikationsphase veraltet. Nicht unwichtig ist, dass die digital eingestellten Publikationen auf einfache Weise zählbar sind. Derzeit zögern Wissenschaftler/innen allerdings noch, diesen Weg zu gehen, weil die klassische Veröffentlichung mehr Reputation und nicht zuletzt auch einen Zusatzverdienst erbringt.²⁰

Sofern sich Wissenschaftler/innen dazu entschließen, Open Access-Server zu nutzen, sollen die dort platzierten Beiträge nach der *Berliner Erklärung* das „freie, unwiderrufliche und weltweite Zugangsrecht und die Erlaubnis, die Veröffentlichung für jeden verantwortlichen Zweck zu kopieren, zu benutzen, zu verteilen, zu übertragen und abzubilden unter der Bedingung der korrekten Nennung der Urheberschaft“ nach sich ziehen. Urheberrechtlich übersetzt bedeutet dies, dass Wissenschaftler/innen den Betreibern/innen der Plattform ein einfaches Nutzungsrecht dahingehend einräumen, dass das Werk öffentlich zugänglich gemacht („übertragen“, Paragraf 19a UrhG) und für Zwecke der Zurverfügungstellung vervielfältigt werden darf („kopieren“, Paragraf 16 UrhG). Sofern über den Server auch elektronische Kopien versendet werden sollen („verteilen“), ist ein Verbreitungsrecht (Paragraf 17 UrhG) einzuräumen. Nicht ganz klar sind

20 In diesem Zusammenhang kommt es auf die Äußerungen zur Unterstützung des Prinzips des offenen Zugangs in der Berliner Erklärung an. Vgl. auch den Beitrag von Fournier, Johannes, Zur Akzeptanz des elektronischen Publizierens in den Geisteswissenschaften. Ergebnisse und Folgen der DFG-Studie „Publikationsstrategien im Wandel“, in diesem Band.

die Formulierungen „abbilden“ und „benutzen“. Die Abbildung erfordert nicht mehr als das Vervielfältigungs- und elektronische Wiedergaberecht („Zugänglichmachen“, Paragraph 19a UrhG), ist also an sich bereits durch die vorhergehenden Formulierungen umfasst. Ein denkbarer, darüber hinausgehender Sinn kann darin gesehen werden, dass Grafiken oder Abbildungen in einem wissenschaftlichen Beitrag mittels dieser Lizenz auch in einen anderen Zusammenhang gestellt, also „zitiert“ werden dürfen. An sich ist die Zitatbefugnis allerdings bereits per urheberrechtlicher Schrankenbestimmung erlaubt (Paragraph 51 UrhG), also ohnehin zustimmungsfrei. Doch erlaubt das Zitatrecht nicht die Verwendung einer Abbildung zur bloßen Illustration eines Zusammenhangs. Zitieren dürfen Nutzer/innen nur in dem Umfang, in dem das Zitat erforderlich ist, um einen eigenen Gedanken wissenschaftlich zu erläutern. In den meisten Fällen geht dies auch ohne Zitat. Doch sagt ein Bild nicht nur sprichwörtlich mehr als tausend Worte, es erklärt einen Vorgang auch eingängig und somit in einer auch für die Lehre im Hörsaal vor einem Massenpublikum zeitgemäßen Weise.²¹ Die Illustration der Lehre im Präsenzünterricht oder in Systemen des „Blended Learning“ durch vorlesungsbegleitende Materialien geht häufig über das hinaus, was das Zitatrecht erlaubt.²² Gemeint sein kann also ein „Illustrationsrecht“.

Problematisch kann ein freies Benutzungsrecht sein. Die im US-amerikanischen Recht gebräuchliche Creative Commons-Musterlizenz²³ spricht hier klarer von einem Bearbeitungsrecht. Gegen die generelle Wirksamkeit solcher Bearbeitungs- oder Umarbeitungsrechte sprechen

21 Vgl. Weidenmann, Bernd (Hg.), Wissenserwerb mit Bildern. Instruktionale Bilder in Printmedien, Film/Video und Computerprogrammen, Bern 1994; speziell zur Anwendung in der juristischen Kommunikation Ulbrich, Stefan, Visuelle Rechtskommunikation, Diss. Bochum 2004, S. 46.

22 Peifer, Karl-Nikolaus, Medienrecht, Urheberrecht und die Wissensgesellschaft. Stimmen die rechtlichen Rahmenregeln für die Zukunft von Forschung und Lehre?, demnächst in: Verein zur Förderung der Rechtswissenschaft e.V., Fakultätsspiegel Wintersemester 2005/2006 (Antrittsvorlesung an der Universität zu Köln vom 16.12.2006, als PDF-Datei vom Verfasser zu beziehen über medienrecht@uni-koeln.de.

23 Auch in deutscher Übersetzung und für den deutschen Lizenzmarkt bestimmt, abrufbar unter <http://creativecommons.org> (07.02.2006).

nicht nur wissenschaftsethische Bedenken, sondern vor allem die Unverzichtbarkeit der urheberrechtlichen Befugnis auf Schutz der Integrität des Werkes in der von den Urhebern/innen geschaffenen Form (Paragraph 14 UrhG). Das Benutzungsrecht ist also entweder zu weitgehend oder aber nutzlos, weil es bereits in der Befugnis zur Vervielfältigung, Verbreitung und Wiedergabe enthalten ist.

Rechtliches Instrument für die Verwirklichung von Open Access-Publikationen ist das Lizenzrecht. Open Access-Modelle setzen die Existenz des Urheberrechts voraus, weil nur ein rechtlicher Schutz die Lizenzierung von Einzelnutzungen zulässt. Sie finden eine Parallele in der Free Software- oder Open Source-Bewegung bei Computerprogrammen.²⁴ In beiden Anwendungsfeldern kehrt man die Logik des Lizenzrechts um. Während Lizenzen von Verwertern/innen klassischerweise dazu benutzt werden, Nutzerbefugnisse eng und passgenau zu begrenzen, wird im Free Software-Bereich die Lizenz dazu genutzt, umfassend Befugnisse einzuräumen, die das Kopieren und die Weiterverbreitung von Software gestatten.

Damit der offen zugängliche Inhalt auch weiterhin offen bleibt, also nicht in proprietäre Systeme eingebunden wird, ferner die Befugnisse der Urheber/innen auf Anerkennung ihrer Urheberschaft beachtet werden, schließen die Lizenzgeber/innen (die Kreativen) mit den Nutzern/innen einen Lizenzvertrag, der durch die Vornahme einer Nutzungshandlung zustande kommt. Voraussetzung für die wirksame Einbeziehung der Lizenzbedingungen ist, dass die Geltung der Lizenz deutlich sichtbar für die Nutzer/innen angekündigt wird, der Lizenztext selbst auf einfache Weise abrufbar ist (Paragraph 305 Absatz 2 BGB). Die Lizenzerteilung kann inhaltlich beschränkt eingeräumt werden (zum Beispiel auf Vervielfältigung, Verbreitung, Bearbeitung, kommerzielle/nicht kommerzielle Nutzung). Sie kann zudem unter die Bedingung gestellt werden, dass die Weitergabe des Inhalts zu den gleichen Konditionen erfolgen muss wie die Ursprungsnutzung, also der Inhalt offen bleiben muss. Im Ergebnis führt

24 Dazu ausführlich Jaeger; Metzger, Open Source (wie Anm. 5); Spindler, Gerald (Hg.), Rechtsfragen bei Open Source, Köln 2004, S. 9.

eine Verletzung dieser Bedingungen dazu, dass die Lizenz erlischt, die Nutzungshandlung mithin zur Urheberrechtsverletzung wird (Paragraf 158 Absatz 2 BGB). Auf diese Weise wird das ursprünglich Ausschließlichkeit ermöglichende Lizenzrecht zur wertvollen Waffe bei der Verteidigung des freien Zugangs auch über mehrere Nutzergruppen hinweg.

Die Erteilung entgeltfreier Jedermann-Lizenzen ist mittlerweile im Urhebervertragsrecht ausdrücklich vorgesehen (Paragraf 32 Absatz 3 Satz 2 UrhG). Dies war nötig, weil Urheber/innen grundsätzlich nicht pauschal auf jede Form der Vergütung verzichten können.

Ergebnisse

Open Access-Modelle sollten für die Wissensproduzenten/innen so attraktiv sein, dass deren Mitwirkung freiwillig erfolgen kann. Insbesondere langjährig tätige Wissenschaftler/innen, die in Verlagsprodukte eingebunden sind, können auf diese Weise gewonnen werden. Eine dienstrechtliche Verpflichtung zur Open Access-Publikation ist auch bei Projektförderungen fragwürdig im Hinblick auf ihre Vereinbarkeit mit der (negativen) Wissenschaftsfreiheit der Forscher/innen. Bei noch „freien Werken“ ist dagegen eine (freiwillige) Open Access-Mitwirkung von Wissenschaftlern/innen weitergehend unproblematisch. Das Urheberrecht kann mit seinen Lizenzregelungen dafür sorgen, dass frei zur Verfügung gestellte Werke nicht ohne Genehmigung verändert werden oder die Urheberschaft an ihnen unterdrückt wird. Eine schnelle Vertragshilfe sind hierzu die im Netz erhältlichen Creative-Commons-Lizenzen, sofern sie auf das deutsche Urheberrecht abgestimmt wurden.

Open Access-Modelle bei gebundenen Autoren/innen müssen auf die Kooperation der Wissenschaftsverlage setzen. Eine Verpflichtung von Wissenschaftlern/innen zur Veröffentlichung von Forschungsergebnissen auf einem Online-Server setzt voraus, dass diese Befugnis den Wissenschaftlern/innen noch zusteht. Den Vorbehalt an eine Projektförderung zu knüpfen, bringt die Betroffenen in ein Dilemma, wenn sie die betreffenden Rechte schon Verlagen eingeräumt haben. Wissenschaftlern/innen aufzugeben, Rechte für Online-Server im Rahmen von Open Access-Projekten vorzubehalten, setzt voraus, dass diese eine Unterstützung bei

der Vertragsgestaltung erhalten. Die gängigen Creative Commons-Lizenzen, die im Netz erhältlich sind, sind dafür keine Hilfe, denn es geht weniger darum, die richtige Formulierung zu finden, als darum, den Rechteinhaber gegenüber dem strukturell stärkeren Vertragspartner durchzusetzen.

Viele Open Access-Modelle sind vor dem Hintergrund eines verwerterfreundlichen Urheberrechtssystems konzipiert, das weitgehende Verzichtsmöglichkeiten der Urheber/innen vorsieht. Diese Modelle können nicht vollständig auf das kontinentaleuropäische Urheberrecht übertragen werden. Auch international-privatrechtlich eignen sich vor dem Hintergrund einer ausländischen Rechtsordnung konzipierte Modelle nicht notwendig für einen weltweiten Einsatz.

Schluss

Die Zukunft von Open Access liegt entweder in der Kooperation zwischen Verlegern/innen und Wissenschaftlern/innen oder der Ablösung bestimmter Publikationsformen durch Wissenschaftler/innen, die zunehmend selbst Verleger/innen werden. Wissenschaftler/innen als Autoren/innen müssen dazu bewegt werden, Open Access als seriöse und reputationsträchtige Publikationsform zu akzeptieren. Dann werden sie auch die Bereitschaft finden, diesen Weg zu gehen. Die Zugangsfreiheit kann über Lizenzregelungen auch in der Verwerterkette abgesichert werden. Kooperation ist auch im Verhältnis zu den Verlagen erforderlich. Kooperationsunwillige Verlage riskieren, den Wissenschaftsmarkt zu verlieren. Kooperationsunwillige Herausgeber/innen von Open Access-Publikationen riskieren, zum Urheberrechtsverletzer zu werden, wenn nämlich die von ihnen publizierten Inhalte bereits in Verlagsproduktionen eingebunden sind und die publizierenden Urheber/innen über die betreffenden Rechte selbst nicht mehr verfügen. Das Urheberrecht kann und muss Open Access begleiten. Allerdings befindet es sich damit historisch gesehen noch in einer durchaus ungewohnten Rolle, die mancherorts noch etwas Starrsinn erzeugt.

Prof. Dr. Karl-Nikolaus Peifer hat einen Lehrstuhl an der Universität zu Köln inne und ist dort Direktor des Instituts für Medienrecht und Kommunikationsrecht sowie Direktor des Instituts für Rundfunkrecht. Im Nebenamt ist er Richter am Oberlandesgericht Hamm. Er befasst sich unter anderem mit Fragen des Persönlichkeits- und Unternehmensschutzes in den Medien sowie mit Fragen des Geistigen Eigentums. E-Mail: medienrecht@uni-koeln.de

LANGZEITARCHIVIERUNG ELEKTRONISCHER PUBLIKATIONEN
IM DIENSTE DER WISSENSCHAFT –
EINE NEUE AUFGABE DER BIBLIOTHEKEN

von Thomas Wolf-Klostermann

Die Bayerische Staatsbibliothek (BSB) trägt als wissenschaftliche Universalbibliothek die Verantwortung für einen steigenden Anteil von elektronischen Publikationen. Sie kauft, sammelt oder digitalisiert, erschließt, stellt bereit und archiviert unterschiedlichste Dokumente. Mit rund zehn Millionen elektronisch vorliegenden Seiten allein aus Retrodigitalisaten verwaltet sie bereits heute eine der größten und am schnellsten wachsenden digitalen Sammlungen im deutschsprachigen Raum. Monatlich kommen mehrere zehntausend neue Seiten hinzu. Neben der Erschließung und Bereitstellung gehört es zu den Aufgaben der Bibliothek, diese Daten auch langfristig zu erhalten. Sie ist außerdem dabei, das elektronische Archiv um wissenschaftlich relevante Netzpublikationen der Sondersammelgebiete und amtliche sowie kommerzielle Publikationen auszuweiten. Dieser Artikel berichtet aus der Praxis der Langzeitarchivierung an der BSB und stellt aktuelle Zukunftskonzepte vor.

Elektronische Publikationen an der Bayerischen Staatsbibliothek

Die BSB ist eine der größten wissenschaftlichen Universalbibliotheken in Deutschland und verfügt über einen bedeutenden Altbestand und eine Reihe von umfangreichen Sondersammlungen. Neben ihrer Rolle als lokale und regionale Informationsversorgerin fungiert die BSB als zentrale bayerische Landes- und Archivbibliothek mit Pflichtexemplarrecht für Bayern. Zusätzlich nimmt die BSB mit einer Reihe von geisteswissen-

schaftlichen Fächern am deutschlandweiten Sondersammelgebietsplan der Deutschen Forschungsgemeinschaft (DFG) teil. Weil elektronische Publikationen im Wissenschaftsbetrieb und im öffentlichen Leben einen immer größeren Stellenwert einnehmen, sieht sich auch die BSB mit einer jährlich wachsenden Zahl von digitalen Objekten konfrontiert, für die sich, analog zum konventionellen Bestand, ein bibliothekarischer Archivierungsauftrag ableiten lässt. Doch anders als bei gedruckten Medien, die einfach katalogisiert und ins Regal gestellt werden können, sind bei elektronischen Dokumenten ganz andere Anstrengungen nötig um die langfristige Verfügbarkeit zu erhalten.

Die elektronischen Dokumente in der BSB lassen sich in vier Gruppen einteilen: Den größten Anteil stellen derzeit die (1) eigenen Retrodigitalisate. Die zweitgrößte Gruppe bilden die (2) lizenzpflichtigen E-Medien. Einen verhältnismäßig geringen Anteil haben im Augenblick noch die (3) freiwilligen Ablieferungen von Verlagen, Ämtern und anderen Institutionen. Schließlich (4) gibt es noch die wissenschaftlich relevanten Internet Quellen in den Sondersammelgebieten, die zwar in der Bibliothek erschlossen und präsentiert werden, jedoch derzeit dort noch nicht gehostet bzw. als Kopie aufbewahrt werden.

Eigene Retrodigitalisate

Seit 1997, als mit Unterstützung der DFG das *Münchener Digitalisierungszentrum* gegründet wurde, wird in der BSB auch digitalisiert. Heute ist das Referat Digitale Bibliothek integraler Bestandteil der BSB. Technische Ausstattung und Know-how wurden seit 1997 sukzessive erweitert, zuletzt 2004-2005, als im Zuge der Konzeption des „retrodigitalen Bestandsaufbaus“ mit Hilfe der Förderung des Bayerischen Staatsministeriums für Wissenschaft, Forschung und Kunst eine moderne „Digitalisierungsstraße“ eingerichtet wurde. Die aktuelle Ausstattung mit acht Scannern für unterschiedliche Anwendungszwecke ermöglicht eine professionelle und qualitativ hochwertige Digitalisierung von allen gängigen Medientypen und Formaten. Mittelalterliche Handschriften, Alte Drucke und historische Landkarten bilden den Schwerpunkt eines breiten Spektrums an Medien. Zusätzlich wird ein Teil der Digitalisierungsaufträge durch externe

Dienstleister abgewickelt. Die Digitalisierung erfolgt auch im Rahmen von Kooperationsprojekten. Durch den Service *Digitization On Demand* steht es den Benutzern/innen offen, digitale Kopien aus nahezu dem gesamten Bestand der Bibliothek zu erwerben. Seit ihrem Bestehen hat die Digitale Bibliothek rund zehn Millionen Digitalisate mit einer Gesamtmenge von mehr als zehn Terabyte erstellt.

Lizenzpflichtige E-Medien

Die Bibliothek bietet Zugriffsmöglichkeiten auf eine große Zahl von lizenz- und gebührenpflichtigen elektronischen Informationsquellen. Dies geschieht im Rahmen des Erwerbungs- und Sammelprofils der Bibliothek und auch im Auftrag der bayerischen Hochschulbibliotheken, für die Konsortialabkommen über elektronische Zeitschriften und Datenbanken abgeschlossen werden. Bei Netzpublikationen werden für die lokal lizenzierten Medien und für die im Rahmen des Bayern-Konsortiums erworbenen digitalen Materialien nach Möglichkeit Vereinbarungen über die Langzeitverfügbarkeit der lizenzierten Periodika und Datenbanken getroffen.

Pflichtablieferung und Amtliche Veröffentlichungen

Im Rahmen des gültigen Landesgesetzes über die Pflichtablieferung erhält die BSB als regionale Pflichtexemplarbibliothek für Bayern im Rahmen bereits heute von den Produzenten freiwillig bereitgestellte elektronische Medien auf physischen Datenträgern, was einen Archivierungsauftrag impliziert. Für die Pflichtablieferung elektronischer Publikationen sind gesetzliche Regelungen auf nationaler Ebene in Vorbereitung.

Über die Pflichtablieferung im engeren Sinne hinaus hat die BSB auch bei amtlichen Veröffentlichungen von Regional- und Bundesbehörden sowie Gebietskörperschaften einen Sammelauftrag. Gerade Ämter und Behörden verlegen aus Kostengründen einen immer größeren Anteil ihrer Publikationen ausschließlich in elektronischer Form. Eine Verlinkung auf die Seiten des Amtes ermöglicht zwar kurzfristig den Zugriff auf diese Dateien, die Langzeitverfügbarkeit, auch von veralteten Publikationen, ist aber nur dann gewährleistet, wenn es der BSB möglich ist, eine Kopie dauerhaft einzulagern.

Elektronische Medien in den Sondersammelgebieten

Im Rahmen der Sondersammelgebiete (SSG) ist die BSB für die Fächer Geschichte, Osteuropa, Klassische Altertumswissenschaften und Musikwissenschaft verantwortlich. Dies gilt auch für die zahlreichen Internet-Ressourcen zu diesen Fachgebieten. Wissenschaftlich relevante Internetquellen zu den Sondersammelgebieten werden systematisch gesammelt und erschlossen, bis dato aber nicht an der BSB gespeichert. Zumindest für einen Teil dieser Quellen, deren dauerhafte Zugänglichkeit auf den Seiten der Urheber/innen nicht gesichert ist, sieht sich die BSB als zuständige SSG-Bibliothek im Einvernehmen mit den Rechteinhabern/innen in der Archivierungspflicht.

Kernaufgaben der Langzeitarchivierung

Der Begriff Langzeitarchivierung (LZA) steht für alle Aktivitäten, die dem Erhalt der Verfügbarkeit der Dokumente über die Lebensdauer der Trägermedien und die Grenzen des technologischen Wandels hinweg dienen. Er entstammt dem anglo-amerikanischen Ausdruck „*Long-Term Digital Preservation*“. Dort ist mit *Preservation* eher „beschützen“ und „bewahren“ gemeint, weniger „archivieren“. „Langzeit-“ bezieht sich dabei auf eine unbegrenzt in die Zukunft reichende Zeitspanne mit tiefgreifenden technologischen Veränderungen und einem sich wandelnden Nutzungsverhalten. Die LZA-Aktivitäten lassen sich in zwei Gruppen einteilen: Die physische und die logische Langzeitarchivierung. Die *physische LZA* (*Bitstream Preservation*) umfasst Aktivitäten, die dem rein physischen Erhalt der Daten dienen, also zum Beispiel der Betrieb geeigneter Speichersysteme samt aller Maßnahmen der Systemsicherheit und Prävention von Datenverlusten, zum Beispiel Notfallvorkehrungen. Zur physischen LZA gehört auch das in der Fachsprache *Refreshing* genannte regelmäßige Umkopieren der Daten auf neue Speichersysteme.

Der rein physische Erhalt der Daten ist zwecklos, wenn diese zu einem späteren Zeitpunkt nicht mehr gelesen werden können. Das zu verhindern, ist Aufgabe der *logischen LZA*. Sie stellt mit einer Reihe von Maßnahmen die spätere Nutzbarkeit der erhaltenen Datenströme sicher.

Dazu gehört die rechtzeitige und verlustfreie Migration in neue zeitgemäße Dateiformate. Zusätzlich kann die Emulation der alten Betriebsumgebung auf neuen Computersystemen angestrebt werden. Der Erhalt von zeitgenössischer Hard- und Software kann dagegen lediglich als Zwischenlösung gelten, bietet aber keine ausreichende Sicherheit.

Durch bibliografische, strukturelle, technisch-administrative und rechtliche Metadaten lassen sich die zu archivierenden Daten beschreiben, in Ordnungssysteme integrieren und Erhaltungsmaßnahmen zuführen. Die bibliografischen Daten ermöglichen die Identifizierung des Werkes. Strukturdaten vermitteln die ursprüngliche Einteilung in Bände und Kapitel. Technisch-administrative Daten wie zum Beispiel Angaben zu Dateiformaten und -größen erleichtern die physische Langzeitarchivierung und dokumentieren Erhaltungsmaßnahmen, zum Beispiel vergangene Migrationschritte. Durch rechtliche Informationen werden zum Beispiel Copyright-Informationen und Zugriffsberechtigungen definiert.

Erfolgreich in der Praxis – Digitalisierung und Langzeitarchivierung an der BSB

An der BSB ist die Langzeitarchivierung Aufgabe des Referats *Digitale Bibliothek*. Hier wurde man sich frühzeitig der Problematik einer wachsenden Menge erhaltenswerter digitaler Dokumente bewusst. Bereits 1999-2001 wurden deshalb erste strategische Konzepte entwickelt und im Rahmen eines ersten DFG-Projektes zum Thema Langzeitarchivierung zusammen mit dem Institut für Softwaretechnologie der Universität der Bundeswehr eine Bestandsaufnahme gemacht und die Migration von erhaltenswerten Altdaten durchgeführt. Seit 2003 engagiert sich die BSB im nationalen Kompetenznetzwerk nestor.¹ Neben der Beteiligung an allen

¹ Das Gemeinschaftsprojekt nestor (Network of Expertise in long-term Storage and availability of digital Resources in Germany), das die Bayerische Staatsbibliothek zusammen mit anderen Gedächtnisorganisationen initiiert hat, ist aus der Überzeugung heraus entstanden, dass der Erhalt des kulturellen Erbes nur gemeinschaftlich gelingen kann und soll, und diente dem Aufbau eines Kompetenznetzwerkes zur Langzeitarchivierung. Nestor ist es gelungen, das Bewusstsein für die Probleme der LZA in Deutschland zu wecken, das Wissen und die Kompetenzen in diesem Bereich zu mehren und zu verbreiten, eine

gemeinschaftlichen Aufgaben liegt der inhaltliche Schwerpunkt der Mitarbeit der BSB auf der Entwicklung eines Kriterienkatalogs für vertrauenswürdige digitale Langzeitarchive. Um die Vertrauenswürdigkeit der Archive nachzuweisen, sollen jene Aspekte ihres organisatorischen Aufbaus und der verwendeten technischen Verfahren geprüft und bewertet werden, die für die Langzeitarchivierung relevant sind. Die Kriterien und deren Bewertungen dienen als Grundlage für die Entwicklung eines Zertifizierungsverfahrens digitaler Langzeitarchive. Zusammen mit der Humboldt-Universität zu Berlin hat die BSB die Federführung der nestor-Arbeitsgruppe „Vertrauenswürdige Archive – Zertifizierung“ übernommen, die sich dieses Themas annimmt und Experten/innen aus allen Lebenszyklen digitaler Objekte von der Produktion über die Bereitstellung und Nutzung bis zur Archivierung einbezieht. Dabei baut sie auf Vorarbeiten der international besetzten RLG/OCLC Working Group on Digital Archive Attributes² auf, an der auch die BSB beteiligt war.

Das Sachgebiet *Digitalisierung* hat sich seit dem ersten Retrodigitalisierungsprojekt, den Reichstagsprotokollen (1867 bis 1895), an Verfahren guter Praxis orientiert und diese zum Teil mitgestaltet.³ In der konsequen-

kooperative Lösung für Deutschland vorzubereiten und diese in den internationalen Kontext einzubinden. Eine ausführliche Beschreibung des Projektes nestor und des Arbeitsschwerpunktes der BSB findet sich auf den Webseiten des Projektes, vgl. <<http://www.langzeitarchivierung.de/>> (11.10.2006).

- 2 RLG/OCLC Working Group on Digital Archive Attributes, Trusted Digital Repositories: Attributes and Responsibilities. An RLG-OCLC Report, Mountain View, CA: RLG, 2002, <<http://www.rlg.org/en/pdfs/repositories.pdf>> (11.10.2006).
- 3 Vgl. Deutsche Forschungsgemeinschaft, Bericht der Arbeitsgruppe Technik zur Vorbereitung des Programms Retrospektive Digitalisierung von Bibliotheksbeständen im Förderbereich Verteilte Digitale Forschungsbibliothek, <http://www.sub.uni-goettingen.de/ebene_2/vdf/endfas1.htm> (11.10.2006); Deutsche Forschungsgemeinschaft, Die Erschließung und Bereitstellung digitalisierter Drucke. Vorschläge des Unterausschusses für Kulturelle Überlieferung, <http://www.dfg.de/forschungsfoerderung/wissenschaftliche_infrastruktur/lis/download/konzept_digitale_drucke.pdf> (11.10.2006) sowie Deutsche Forschungsgemeinschaft, Praxisregeln in Förderprogramm Kulturelle Überlieferung, <http://www.dfg.de/forschungsfoerderung/formulare/download/12_151.pdf> (11.10.2006). Siehe auch: Boudrez, Filip; Dekeyser, Hannelore, DAVID – Digitaal archiveren in de praktijk. Handboek,

ten Anwendung offener und offen gelegter Standards für die Bilddigitalisierung und Erschließung liegt die Voraussetzung für eine langfristige Archivierung des Materials.

Ursprünglich wurden die bei der Digitalisierung entstandenen Daten auf archivtauglichen CD-ROMs gespeichert und im Magazin abgestellt. Schon nach kurzer Zeit erwies sich diese Methode aufgrund steigender Datenmengen und der Unsicherheit des Speichermediums als nicht mehr praktikabel. Deshalb wurde 2004 – die gesamte Datenmenge betrug damals circa 0,8 Terabyte – eine Zusammenarbeit mit dem Leibniz-Rechenzentrum (LRZ) begonnen. Das CD-ROM-Archiv wurde aufgelöst, sämtliche Daten wurden ausgelesen und konnten verlustfrei in das Archiv- und Backupsystem des LRZ überführt werden. Die Digitale Bibliothek baute von Anfang an auf selbst entwickelte Lösungen auf Open Source-Basis. Rückgrat der Digitalisierung ist seit 2004 die „Zentrale Erfassungs- und Nachweisdatenbank für Digitalisate“ (ZEND), die den gesamten Workflow von der Vorbereitung über die Digitalisierung bis zur Bereitstellung und Archivierung mit maßgeschneiderten Funktionen unterstützt und die Verarbeitung der Daten weitgehend automatisiert, andererseits die bibliografischen, strukturellen, technischen und administrativen Metadaten verwaltet.

Die Einrichtung der Digitalisierungsstraße mit mehreren neuen Großscannern 2004-2005 ließ stark steigende Datenmengen erwarten, so dass eine grundlegende Umstellung der Workflows erforderlich wurde. Eine manuelle Archivierung der Daten wäre auf Dauer nicht praktikabel gewesen, so dass bei der Weiterentwicklung der ZEND verstärkt auf automatisierte Verarbeitungsschritte gesetzt wurde. Es wurde geplant, den Archivierungsprozess vollständig in den Produktionsprozess zu integrieren und die Überspielung der Dateien ins Archiv weitgehend zu automatisieren. Dieses Ziel wurde 2005 erreicht. Die einzelnen Etappen dieses Workflow seien im Folgenden kurz beschrieben:

<<http://www.antwerpen.be/david/index.htm>> und Dörr, Marianne, Retrodigitalisierung in Bibliotheken und wissenschaftlichen Einrichtungen, <<http://computerphilologie.uni-muenchen.de/jg01/doerr.html>> (11.10.2006).

Vor Beginn der eigentlichen Produktion steht die Auswahl des Werkes. Motivation für ein Digitalisierungsprojekt ist entweder der Bestandschutz (Schonung bzw. Sicherheitsverfilmung eines wertvollen Werkes), der retrodigitale Bestandsaufbau (Bereitstellung von häufig nachgefragten Werken aus dem Altbestand der BSB) oder die Nachfrage von Bibliotheksnutzern/innen nach einem einzelnen Werk, sei es aus wissenschaftlichen oder kommerziellen Erwägungen (Faksimile-Produktion).

Bevor ein Werk auf den Scanner kommt, werden zunächst alle wesentlichen Parameter definiert und die nachfolgenden Arbeitsschritte festgelegt. Ausgangsbasis für alle Verarbeitungsschritte ist eine von der ZEND erzeugte unveränderliche numerische ID. Sie dient als Namensbasis für alle zugehörigen Images und für das Server-Verzeichnis, in dem diese abgelegt werden. Außerdem bildet sie den buchspezifischen Kern der National Bibliography Number (NBN), eines persistenten Identifikators, der über die Deutsche Bibliothek (DDB) in Frankfurt am Main verwaltet wird und die dauerhafte Ansprechbarkeit des Titels im Internet garantiert. Über die ID werden gleichzeitig die Digitalisierungsaufträge und die technischen Metadaten verwaltet. Über eine Z39.50-Schnittstelle werden die bibliografischen Daten aus dem lokalen Katalogsystem der BSB in die ZEND übernommen.

Mit der Erfassung eines Titels in der ZEND sind unmittelbar alle endgültigen und unveränderlichen Identifikatoren vergeben. Über ein elektronisches Auftragsformular wird der Digitalisierungsauftrag erzeugt. Beim Absenden wird überprüft, ob die eingegebenen Parameter plausibel sind, die technischen und administrativen Metadaten werden abgespeichert. Zu diesen gehören etwa Angaben zur Objektart (zum Beispiel Karte, Zeitschrift), zur geforderten Bildauflösung und Farbtiefe oder besondere Hinweise für die Einstellung der Scanner, wie sie beispielsweise beim Digitalisieren von gold- oder silberfarbigen Abbildungen notwendig sind. Mitgeführt werden darüber hinaus das Bestelldatum, der Name und die Abteilung des Auftraggebers / der Auftraggeberin sowie der Ausführenden bzw. die Anschrift des externen Dienstleisters.

Die Digitalisierung wird hausintern oder vom externen Dienstleister nach den im Auftrag vorgegebenen Parametern durchgeführt. Der digitale

Master entsteht immer im TIFF-Format (Tagged Image File Format), welches höchste Qualität und verlustfreie Speicherung der Bilddaten garantiert. Daraus werden je nach Bedarf Derivate für die Web-Präsentation (JPGs, GIFs, PDFs), Druckvorstufen für die Faksimileproduktion, aber auch analoge Mikroformen zum Zwecke der Sicherheitsverfilmung erstellt. Nach dem Digitalisieren werden die Images – ohne Bildung von Zwischenverzeichnissen – in einem Sammelverzeichnis auf dem FTP-Server der BSB abgelegt. Dieses Sammelverzeichnis wird mit Hilfe von zeitgesteuerten Verfahren überwacht. Immer dann, wenn Bilddaten zur Weiterverarbeitung dort vorhanden sind, wird das Bereitstellungsverfahren gestartet. Zunächst wird die Datenbank anhand der Dateinamenbezeichnung abgefragt, ob Titeldaten erfasst sind. Ist das der Fall, wird automatisch im Bereitstellungsbereich auf dem öffentlich zugänglichen Server ein Verzeichnis unter dem Namen der ID des Digitalisats eingerichtet. In dieses Verzeichnis werden alle zusammengehörenden Images kopiert. Die Dateien werden geordnet und ihre Namen in der richtigen Reihenfolge in die XML-Strukturdatei geschrieben. Die BSB bezieht sich beim Erstellen der XML-Dateien auf die Richtlinien und Standards der Text Encoding Initiative (TEI).⁴ Die XML-Datei enthält im so genannten TEI-Header die bibliografischen Daten des Objekts. Bei Bedarf kann die XML-Datei die tiefere Erschließung des Buches, etwa das Inhaltsverzeichnis, und gegebenenfalls den Volltext in maschinenlesbarer Form aufnehmen. In der einfachsten Form enthält sie neben den bibliografischen Angaben nur die Bezeichnung der Images in der Digitalisierungsreihenfolge. Dies ist die Voraussetzung dafür, dass die Digitalisate zu einem elektronischen Buch zusammengefügt werden können und eine Navigation auf dem Bildschirm möglich wird. Durch die Wahl des XML-Formates und des TEI-Standards sind diese Dateien ebenfalls für die Archivierung und Migration geeignet. Sie werden mit den Images gemeinsam und gleichzeitig archiviert.

Über die Datenbankabfrage werden die Titelangaben ausgelesen und im TEI-Header-Format ebenfalls in die XML-Datei eingefügt. Die ZEND steuert darüber hinaus die Programme, die die automatisierte Anfertigung

4 Zur Text Encoding Initiative vgl. <<http://www.tei-c.org/>> (11.10.2006).

von browserkompatiblen Derivaten (JPG, GIF und PDF) durchführen, sowie die Tools für die Strukturdatenbearbeitung und die Bildnachbearbeitung. Aus den TIFF-Dateien werden Derivate in zwei Vergrößerungsstufen hergestellt. Dafür wird der TIFF-Header mit den Angaben zu Größe, Auflösung, Ausrichtung usw. ausgelesen. Nach dem Schreiben der XML-Strukturdatei und der Bildherstellung für das Web ist eine einfach durchzublätternde digitale Ausgabe des Buches fertig und benutzbar. Vom Ablegen der digitalen Master im Sammelverzeichnis bis zur Fertigstellung dieser einfachen Webversion sind keine manuellen Tätigkeiten erforderlich. Für die Produktion eines 200 Seiten umfassenden Buches benötigt ein leistungsstarker Rechner zurzeit etwa eine Stunde. Die fertig gestellte digitale Ausgabe kann jederzeit tiefer erschlossen und gegebenenfalls auch korrigiert werden. Dafür stehen eine ganze Reihe von Tools zur Verfügung, wovon als wichtigster der XML-Editor zu nennen ist, der von jedem Ort innerhalb der Staatsbibliothek webgestützt aufrufbar ist und es erlaubt, Kapitelüberschriften zuzuweisen und Seitenzählungen (arabisch, römisch, blatt-, spalten- und seitenweise) halbautomatisch einzufügen. Komplexere Zählungen können auch per Hand bei jeder Seite eingefügt werden. Vom XML-Editor aus lässt sich auch eine Bildkorrektur – etwa durch Drehen einzelner Seiten – anstoßen. Eine tiefere Erschließung des Inhalts wird insbesondere bei den großen Digitalisierungsprojekten, wie der Digitalisierung ganzer Zeitschriften oder vielbändiger Werke, durchgeführt. Hier werden Aufsatztitel, Register, Lemmata usw. erfasst und vielfach auch der Volltext über OCR gewonnen.

Mit der Freigabe für die Benutzung im Internet werden NBN und momentane URL über eine OAI-Schnittstelle der ZEND dem Frankfurter NBN-Management übermittelt, das dafür sorgt, dass die dem Digitalisat zugewiesene NBN im Internet aufrufbar wird. Die NBN wird auf Verbundebene und im lokalen Katalogsystem der BSB nachgewiesen, so dass Benutzer/innen, die den Titel recherchieren, direkt auf das Digitalisat hingewiesen werden. Die Archivierung wird nach Beendigung aller Arbeitsschritte automatisch gestartet. Der Produktionsserver der Digitalen Bibliothek ist über einen Client des Archiv- und Backup-Systems Tivoli Storage Manager von IBM mit dem Leibniz-Rechenzentrum verbunden.

Im Archivsystem des Rechenzentrums werden die Daten in eine Verzeichnisstruktur gelegt, die nach der ID des Digitalisats benannt wird. Anfänglich wurden die Daten nur zur Nachtzeit transferiert, um das Münchner Wissenschaftsnetz nicht unnötig zu belasten. Seit dem Ausbau der Datenleitung auf Gigabit-Geschwindigkeit im Herbst 2005 kann der Datentransfer unterbrechungsfrei rund um die Uhr laufen.

Im Archivsystem werden alle zum Objekt gehörenden Daten abgespeichert, also sämtliche Images im TIFF-Format und die zugehörigen XML-Strukturdaten. Die zusätzliche Archivierung der Bildderivate erfüllt eine Backup-Funktion für den Fall eines Datenverlustes auf dem Bereitstellungsserver. Für die Langzeitarchivierung kommt dabei den XML- und den TIFF-Dateien die größte Bedeutung zu. Das Leibniz-Rechenzentrum selbst speichert die zu archivierenden Daten mehrfach auf verschiedenen Speichermedien. Nach abgeschlossenem Archivierungsvorgang erfolgt eine Rückmeldung an die ZEND, in der dieser Arbeitsschritt – wie alle anderen auch – in den administrativen Metadaten dokumentiert wird. Bei Bedarf können die Daten innerhalb kürzester Zeit aus dem Langzeitarchiv zurückgeholt werden. Dieser Prozess wird durch eine E-Mail-Funktion an den Archivserver angestoßen. Aufgrund der ID wird das gesamte Verzeichnis eines Objektes zurückgeholt. Ein manueller Zugriff auf einzelne Dateien ist ebenfalls möglich.

Insgesamt handelt es sich bei der ZEND um ein mit der Zeit und an den Aufgaben gewachsenes Instrument, das mit maßgeschneiderten Funktionen die Aufgaben erfüllt. Die Verwendung von offenen und offenen Standards (Datenformate TIFF, XML/TEI) stellt einerseits die Langlebigkeit der durch Digitalisierung erzeugten Daten sicher, ermöglicht andererseits die Migration zu neueren Bereitstellungssystemen. Die Automatisierung aller Arbeitsschritte von der Auftragsvergabe bis zur Archivierung sichert zudem die hohe Qualität der Objekte. Die Vorhaltung bibliografischer, struktureller, technischer sowie administrativer Metadaten ermöglicht zukünftig die gezielte Anwendung von Langzeiterhaltungsstrategien. Die leistungsfähige Infrastruktur des Leibniz-Rechenzentrums garantiert den langfristigen Erhalt der Daten. Damit ist die BSB für die Anforderungen der Zukunft gerüstet.

Weitere Schritte – das Pilotprojekt zur Langzeitarchivierung von Netzpublikationen

Die Aufgabe der Archivierung der Retrodigitalisate konnte durch Aufbau der ZEND zufrieden stellend gelöst werden. Wie bereits zu Beginn dargestellt, trägt die BSB als Universal- und Landesbibliothek Verantwortung für ein viel breiteres Spektrum an digitalen Objekten, für die es bislang noch kein konkretes Verfahren der Langzeitarchivierung gab. Deshalb wurde an der BSB im Juli 2005 ein von der DFG gefördertes zweijähriges Pilotprojekt begonnen, das die Langzeitarchivierung dieses breiten Spektrums an elektronischen Medien vorbereiten soll. Der Fokus richtet sich sowohl auf die Sammlung, Verwaltung und Speicherung von elektronischen Ressourcen im Langzeitarchiv, als auch auf deren Bereitstellung zur Nutzung im Kontext der Bibliothek. Kooperationspartner in diesem Projekt ist das Leibniz-Rechenzentrum. Ziel ist der Aufbau eines Bibliothekarischen Archivierungs- und Bereitstellungssystems (BABS), die Ausarbeitung von Workflows und Kooperationsmodellen und die Archivierung und (partielle) Bereitstellung von wissenschaftlichen Netzpublikationen aus unterschiedlichen Quellen. Um die Medienvielfalt der BSB möglichst umfassend zu repräsentieren, sollen Medien verschiedener Provenienz in das Projekt einbezogen werden. Darunter fallen:

- Veröffentlichungen von potentiellen Pflichtverlagen,
- im Bayernkonsortium lizenzierte E-Zeitschriften und Datenbanken,
- amtliche Veröffentlichungen von Regional- und Bundesbehörden sowie Gebietskörperschaften,
- wissenschaftlich relevante Literatur aus der nationalen und internationalen Produktion in den SSG-Fächern Geschichte, Altertumswissenschaften, Osteuropa und Musikwissenschaft,
- in den Fachportalen zusammengestellte und erschlossene Netzpublikationen,
- eigene und in Kooperationsprojekten erstellte Retrodigitalisate.

Ähnlich wie schon bei der ZEND soll das Datenmanagement von der Sammlung über die Erschließung bis zur Nutzung in der Bibliothek angesiedelt sein und in die bestehenden Bibliothekssysteme integriert werden.

Die Aufbewahrung der Daten erfolgt, wie schon bei den Digitalisaten, am Leibniz-Rechenzentrum.

Solange grundlegende rechtliche Fragen bei der Langzeitarchivierung (zum Beispiel die Pflichtablieferung, das Anfertigen mehrfacher Kopien, verändernde Eingriffe bei der Konvertierung bzw. Migration, Bereitstellung über Netze) noch nicht allgemein geregelt sind, können die einzelnen Arbeitsschritte bei der Langzeitarchivierung nur in Absprache mit den Eigentümern/innen bzw. Produzenten/innen der Daten modelliert und realisiert werden. In den Gesprächen mit den Produzenten/innen – Verlage, Regional- und Bundesbehörden, Wissenschaftler/innen, Bibliotheken – liegt einer der Schwerpunkte des Projektes. Speziell in Bezug auf urheberrechtlich geschützte digitale Medien ist mit den Anbietern/innen bzw. Produzenten/innen weiterführend die Frage nach möglichen Geschäfts- und Nutzungsmodellen zu diskutieren. Die BSB hat für lizenzpflichtige E-Medien ein dreistufiges Zugriffskonzept entwickelt, das von der lokalen Nutzung in den Räumen der Bibliothek, über den externen Zugriff für registrierte BSB-Benutzer/innen bis hin zu speziell vereinbarten Pay-Per-Use-Modellen für die überregionale Literaturversorgung reicht. In welcher Weise sich diese Modelle etwa auf den Bereich der Pflichtablieferung digitaler Medien übertragen lassen bzw. welche Art von Modifikation erforderlich ist, wird im Projekt geprüft.

Ein weiterer Schwerpunkt des Projektes ist die Entwicklung neuer bzw. die weitere Ausgestaltung existierender Workflows in der Bibliothek und im Rechenzentrum. Die enge Verzahnung mit den bewährten Arbeitsabläufen garantiert den nahtlosen Übergang zur Routine und damit die Nachhaltigkeit der hier zu entwickelnden Modelle. Vorgesehen ist ferner die Erprobung von Methoden der Langzeitarchivierung, die sowohl den physischen Erhalt der Daten als auch den Erhalt der Verfügbarkeit betreffen. Eine durch den Umzug des Leibniz-Rechenzentrums an einen neuen Standort und der damit verbundenen Einführung neuer Hard- und Software erforderlich gewordene Migration des kompletten Archivspeichersystems wurde bereits erfolgreich durchgeführt. Die technische Umsetzung erfolgt weitgehend unter Nachnutzung der vorhandenen Infrastruktur, die überwiegend aus Eigenentwicklungen auf Open Source-Basis besteht. Neben

der bereits bewährten ZEND kommen die verschiedenen vorhandenen Präsentationslösungen ebenso zum Tragen wie die bereits im Umfeld der elektronischen Zeitschriften und Online-Datenbanken der BSB erfolgreich eingesetzten Authentifizierungsmodule. Ein wichtiges Fundament für die Organisation der Langzeitarchivierung wird die kommerzielle Repository-Software DigiTool bilden, die vom Bibliotheksverbund Bayern beschafft wurde und derzeit in Betrieb genommen wird. DigiTool ist bedarfsgerecht konfigurierbar, fügt sich optimal in die vorhandene bibliothekarische Systemumgebung ein und bringt wesentliche technische Voraussetzungen für die Langzeitarchivierung mit, so zum Beispiel Möglichkeiten einer Objektmodellierung mit feiner Granularität bis auf Dateiebene, die flexible Verwaltung von Metadaten und eine Web-Präsentation des archivierten Materials. Die Anbindung des bibliothekarischen Bereitstellungssystems an das Archivsystem des Rechenzentrums wird, in Erweiterung der aktuellen Lösung, über eine leistungsfähige File- System-Schnittstelle realisiert werden.

Zusammenfassung

Mit ihrem 1997 gegründeten Digitalisierungszentrum gehört die BSB zu den Pionieren der Digitalisierung in Deutschland. Entsprechend früh setzte man sich hier mit der Problematik der Langzeitarchivierung auseinander und begann, am eigenen Bedarf orientierte Lösungen zu erarbeiten. Die Digitalisierung hat sich als Geschäftsgang etabliert und ist mit dem eigenständigen Referat *Digitale Bibliothek* integraler Bestandteil der BSB. Dank der massiven Erweiterung der Scanner-Ausstattung konnte die Digitalisierungsproduktion seit Mitte 2005 signifikant gesteigert werden. Mit der Zentralen Erfassungs- und Nachweisdatenbank für Digitalisate ist seit 2004 ein leistungsfähiges Arbeitsinstrument erfolgreich in Betrieb, das auch die Langzeitarchivierungsaufgaben weitgehend automatisiert. Bis 2007 wird im Rahmen des vorgestellten DFG-Projektes die vorhandene Infrastruktur um Module und Workflows zur Langzeitarchivierung von Netzpublikationen erweitert. Auch zukünftig wird durch die Mitarbeit bei nestor auf nationaler und internationaler Ebene der Austausch von Know-how und Erfahrungen mit anderen Gedächtnisorganisationen weiter vertieft werden.

Durch die Zusammenarbeit mit dem Leibniz-Rechenzentrum als kompetentem und erfahrenem Partner steht ein zuverlässiges, sicheres Speichersystem mit hoher Kapazität zur Verfügung. Unter diesen Voraussetzungen sieht sich die BSB in einer günstigen Startposition für kommende Aufgaben auf dem Gebiet der Langzeitarchivierung.

Dr. Thomas Wolf-Klostermann ist im Referat „Digitale Bibliothek“ der Bayerischen Staatsbibliothek tätig. Aufgabenschwerpunkte: Erschließung und Web-Bereitstellung von Digitalisaten, technische Betreuung der Produktion, Langzeitarchivierung. E-Mail: thomas.wolf-klostermann@bsb-muenchen.de

DIGITALISIERUNG UND STRATEGIEN
DER LANGZEITARCHIVIERUNG

DIE DEUTSCHE FORSCHUNGSGEMEINSCHAFT UND
DEUTSCHE INITIATIVEN ZUR LANGZEITARCHIVIERUNG
FÜR DIE WISSENSCHAFT

von Max Vögler

Die Erschließung und Bereitstellung historischer Dokumente ist ein etablierter Förderschwerpunkt der Deutschen Forschungsgemeinschaft. Aber es ist ein Förderschwerpunkt im Umbruch: die neuen Möglichkeiten im Bereich des digitalen Publizierens haben die Arbeitsweise von Historikern/innen radikal verändert und stellen auch neue Anforderungen an geschichtswissenschaftliche Infrastrukturen. Dieser Beitrag befasst sich vor allem mit den strukturellen Aspekten dieser Entwicklung auf nationaler Ebene und fragt hier insbesondere nach den Projekten, Perspektiven und Problemen im Bereich der Langzeitarchivierung digitaler Dokumente.

Als im Jahre 1086 Wilhelm der Eroberer das *Domesday-Projekt* ins Leben rief, wollte er eine für alle Zeit nachvollziehbare und *endgültige* Darstellung der damaligen Grundbesitz-Verhältnisse in England schaffen. Dies ist ihm gelungen: Fast ein Jahrtausend später gilt das so genannte *Domesday-Buch* immer noch als erster umfassender Überblick über Grundbesitz in England.¹ 1986 hat die BBC zum 900. Jubiläum des Buches einen ähnlichen Ansatz verfolgt. Das *BBC Domesday-Projekt* sollte einen Querschnitt der „images and sounds of Britain 1986“ – insgesamt 200.000 digitale Dokumente (Ton, Film, Text, Bild, usw.) – als kulturelle Überlieferung für das nächste Jahrtausend sichern. 2,5 Millionen britische Pfund hat das Projekt gekostet und die Ergebnisse wurden auf besonders unempfindliche

1 Vgl. <<http://www.domesdaybook.co.uk/>> (11.08.2006).

Laserdisketten gebrannt. Allerdings konnten die Laserdisketten nur auf einem Acorn special BBC Microcomputer gelesen werden und würden demnach nur so lange verfügbar sein, wie das letzte Exemplar des Acorns noch funktionsfähig blieb. Schon wenige Jahre später war die Technik obsolet, die Inhalte unlesbar. Erst 2002, nach dreijähriger Arbeit, ist es Wissenschaftlern/innen an der University of Michigan im so genannten *Camelion-Projekt* gelungen, eine Emulationsumgebung für das Acorn Betriebssystem zu programmieren, die die Inhalte des *BBC Domesday-Projektes* wieder zugänglich macht. Der Projektmanager bemerkte 2002 zu Recht, „*BBC Domesday* has become a classic example of the dangers facing our digital heritage.“²

Die zunehmende Verwendung von digitalen Inhalten – ob *born digital*, wie zum Beispiel E-Mails oder Webseiten, oder nachträglich digitalisiert – im wissenschaftlichen Alltag ist für Bibliotheken, Archive, Rechenzentren, Behörden, Verlage und auch für die Wissenschaftler/innen selbst eine enorme Herausforderung. Die ständig sinkenden Kosten der Digitalisierung und Optical Character Recognition-Kodierung, die neuen Möglichkeiten der semantischen Erschließung, Tagging und anderen Entwicklungen sind dabei, auch die traditionelle Arbeitsweise der Historiker/innen – in Archiven und Bibliotheken handgeschriebene oder gedruckte Quellen zu lesen und darüber Aufsätze und Monografien zu schreiben – radikal zu verändern. In den Sprach- und Literaturwissenschaften sowie in der Philosophie gibt es schon die ersten Ansätze einer neuen, kollaborativen Arbeitsweise.³

Dieser Beitrag bietet eine kurze Einführung in die Probleme und Perspektiven der Langzeitarchivierung. Es werden anfänglich allgemeine Betrachtungen zur Langzeitarchivierung gemacht, anschließend einige Entwicklungen in Deutschland vorgestellt und letztlich die momentanen Überlegungen, Fragestellungen und Förderperspektiven der DFG dargestellt.

2 Zum Domesday Projekt vgl. <<http://news.bbc.co.uk/1/hi/technology/2534391.stm>>; <<http://www.si.umich.edu/CAMILEON/>> (11.08.2006).

3 Hierzu vgl. z.B. das Projekt Hyper-Nietzsche: <<http://www.hypernietzsche.org/>> (11.08.2006).

Warum Langzeitarchivierung?

Das Ziel der Langzeitarchivierung ist schnell formuliert, aber tückisch in der Umsetzung. Im Gegensatz zur „herkömmlichen“ Langzeitarchivierung – die Aufbewahrung eines Buches oder Schriftstücks in einer Bibliothek oder einem Archiv – muss bei der Langzeitarchivierung digitaler Objekte nicht nur auf den Erhalt des Inhaltes geachtet werden, also auf die Integrität der digitalen Information, sondern auch darauf, dass diese Inhalte verwendbar bleiben. Digitale Medien langfristig zu archivieren, bedeutet also nicht in erster Linie das Speichern, sondern die dauerhafte Nutzung zu ermöglichen, so dass Texte, Bilder, E-Learning-Objekte (Leselisten, virtuelle Semesterapparate, Projektergebnisse der Studierenden), digitale Reproduktionen von Museumsstücken, Datenbanken, sowie Forschungsdaten und jegliche andere Art digitaler Information über viele Jahre und Softwareversionen hinweg verwendbar bleiben.

Grundsätzlich gibt es zwei fast entgegengesetzte Strategien, die jeweils die dauerhafte Nutzung digitaler Objekte zum Ziel haben. Die erste ist eine Migrationstrategie, bei der die regelmäßige Überspielung von digitalen Inhalten auf das jeweils gängige Format im Mittelpunkt steht. Die zweite ist eine Emulationsstrategie, bei der das Daten- oder Softwareformat erhalten bleibt und die Nutzung durch eine in regelmäßigen Abständen angepasste Emulationsumgebung dauerhaft ermöglicht wird. Dahinter stehen zwei konträre Philosophien: Was soll erhalten bleiben, Inhalt oder Objekt? Marilyn Geller hat dieses Dilemma mit all seinen Konsequenzen gut beschrieben:

„If it is the object we want to preserve, then we must choose emulation, which requires the building of backward compatible software. If instead, we want only to preserve the intellectual content, then the more important issue is to normalize files in some way that will allow us to store and manipulate them to assure readability in a future software and hardware environment.“⁴

Zu diesen technischen Herausforderungen kommen weitere. Die naht-

4 Vgl. Geller, Marilyn, *Models for E-Journal Archives. Future Pathways into the Past*, *The E-Resources Management Handbook*, Bd. 1 (2006), S. 52-53.

lose Einbindung *rechtlicher* Regelungen bei automatisierten Zugriffssystemen ist noch weitgehend unerprobt. Hierzu müssen Verlage aktiv beteiligt werden, wenn es um die Speicherung von lizenzierten Inhalten geht. Archive, in denen der Zugang zu (digitalisierten) Beständen zunehmend über das Internet gewährleistet wird, werden robuste und automatisierte Datenschutzregelungen implementieren müssen. Weiterhin kann eine Langzeitarchivierungsstrategie nur im Netzwerk und in Zusammenarbeit mit vielen Partnereinrichtungen vollzogen werden. Das heißt, dass *organisatorische* Modelle – Workflows der Datenab- und -weitergabe, Bearbeitung, Aufbau von Speicherungsnetzwerken durch Datenspiegelung usw. – auch wichtige Bestandteile einer Archivierungsstrategie sind. Und verbunden mit dieser organisatorischen Herausforderung gibt es die *finanzielle* Komponente: Wie können Geschäftsmodelle erarbeitet werden, die die Kosten der Langzeitarchivierung in einer handhabbaren Weise im System verteilen?

Diese vier Problemkomplexe der Langzeitarchivierung – Technik, Organisation, Recht und Finanzierung – müssen auf lokaler, regionaler sowie nationaler und internationaler Ebene angegangen werden. Sie müssen auch für verschiedene Arten von digitalen Objekten – E-Journals, E-Books, Forschungsprimärdaten aus verschiedensten Disziplinen, Digitalisate, usw. – umgesetzt werden. Nur so kann ein wirklich „vertrauenswürdiges Netzwerk“ für die Langzeitarchivierung digitaler Objekte geschaffen werden.

Lösungsansätze in Deutschland

In Deutschland bilden zwei Projekte den Kern einer bundesweiten Langzeitarchivierungsstrategie. Vor allem aufbauend auf die Vorarbeiten der Kollegen/innen an den Nationalbibliotheken in den Niederlanden und in Australien⁵ verfolgt das deutsche Projekt *kopal* die Entwicklung einer auf die hiesigen Bedürfnisse angepassten technischen Lösung für die Langzeitarchivierung.⁶ Gleichzeitig zielt das Projekt *nestor – Kompetenznetzwerk*

5 Vgl. <<http://www.kb.nl/dnp/e-depot/e-depot-en.html>> sowie <<http://www.nla.gov.au/preserve/>> (11.08.2006).

6 Das Projekt ist an der Deutschen Nationalbibliothek in Frankfurt sowie der Gesellschaft für wissenschaftliche Datenverarbeitung mbH und der Staats- und

Langzeitarchivierung darauf ab, Konsens über eine dauerhafte Organisationsform sowie die Abstimmung über die Übernahme von Daueraufgaben herzustellen.⁷

Weitere Maßnahmen sind jedoch auf verschiedenen Ebenen und mit unterschiedlichen Ansätzen notwendig, um ein wirklich funktionierendes System der digitalen Langzeitarchivierung zu schaffen. Die Deutsche Forschungsgemeinschaft (DFG) hat als zentrale Selbstverwaltungseinrichtung der Wissenschaft ein besonderes Interesse an dem Erhalt, der Pflege und der Gewährleistung des langfristigen Zugriffs auf wissenschaftlich relevante digitale Inhalte. Die Förderung von Langzeitarchivierungsstrategien ist der DFG bei den nachfolgend aufgeführten drei Segmenten wissenschaftlicher Informations-Infrastrukturen besonders wichtig:

1. Das System der überregionalen Literaturversorgung: Sondersammelgebietsbibliotheken (SSG-Bibliotheken) sammeln zunehmend elektronische Dokumente in Form von E-Books, E-Journals, Webseiten, pdf-Dokumenten usw. Für diese Materialien sollen sie „die über die Beschaffung hinausgehenden Funktionen der professionellen bibliothekarischen Pflege der Sammlungen von der Erschließung über die überörtliche Bereitstellung der Dokumente bis zur Sicherung der langfristigen Verfügbarkeit“ gewährleisten.⁸
2. Digitalisierungsprojekte: Im Rahmen von (Retro-)Digitalisierungsprojekten wurden seit 1996 mit erheblichem Einsatz von Fördermitteln „Primärdaten der geisteswissenschaftlichen Forschung“ online verfügbar gemacht.⁹

Universitätsbibliothek in Göttingen (SUB) angesiedelt und wird durch das Bundesministerium für Bildung und Forschung (BMBF) gefördert. Vgl. <<http://kopal.langzeitarchivierung.de/>> (11.08.2006).

7 Nestor ist auch in Göttingen an der SUB angesiedelt und wird durch das BMBF finanziert. Vgl. <<http://www.langzeitarchivierung.de/>> (11.08.2006).

8 „Richtlinien zur überregionaler Literaturversorgung der Sondersammelgebiete und virtuellen Fachbibliotheken“ (Stand 17.07.2006), S. 6. Vgl. <http://www.dfg.de/forschungsfoerderung/wissenschaftliche_infrastruktur/lis/download/richtlinien_lit_versorgung_ssg_0607.pdf>. Zum System der SSG-Bibliotheken vgl. <<http://webis.sub.uni-hamburg.de/>> (11.08.2006).

9 Hierzu gehören die DFG-geförderten Digitalisierungsprojekte (für eine Übersicht vgl. <<http://www.zvdd.de/sammlungen.html>>); die zunehmende Zahl der archi-

3. Forschungsprimärdaten: DFG-geförderte Forschungsprojekte in allen Disziplinen produzieren große Mengen an Primärdaten. Hier sind Beiträge zu einer kohärenten Strategie zu leisten, um die Speicherung, das Auffinden und den dauerhaften Zugang von Forschungsdaten zu ermöglichen.

Im Rahmen dieser Handlungsfelder hat die DFG in den vergangenen Jahren eine Reihe von Projekten gefördert, die weniger den Gesamtanspruch von *nestor* oder *kopal* verfolgen, sondern vielmehr einen Beitrag zu Teilaspekten und -fragen leisten. Hier sind als Projekte hervorzuheben:

- *Datenbankgestützte Langzeitarchivierung digitaler Objekte*, Universität der Bundeswehr München: allgemeine Konzeptentwicklung für die Überführung digitaler Publikationen in Standarddatenbanksysteme.¹⁰
- *BABS: Langzeitarchivierung von Netzpublikationen aus dem Spektrum der Bayerischen Staatsbibliothek (BSB)*, BSB München: organisatorische und rechtliche Musterlösungen für elektronische Medien aus Verlagen und Behörden.¹¹
- Erfassung, Erschließung und Sicherung von Websites politischer Parteien der Bundesrepublik Deutschland sowie ihrer Fraktionen in den Parlamenten, Archiv der Friedrich-Ebert-Stiftung Bonn: Entwick-

varischen Findmittel (und z.T. auch Digitalisate), die online auffindbar sind (vgl. <http://www.bundesarchiv.de/bestaende_findmittel/findmittel_online/index.html>; <<http://www.archive.nrw.de/>>; <<http://www.landesarchiv-bw.de/>> sowie <<http://kalliope.staatsbibliothek-berlin.de/>> für Nachlässe und Autografen); die DFG-geförderten Nationallizenzen für die Wissenschaft, in denen auch viele für die Geisteswissenschaften relevanten Quelleneditionen enthalten waren (z.B. „Early English Books“, Comintern Classified Archives, oder das Patrologia graeco-latina; hierzu vgl. <<http://www.nationallizenzen.de/>>); die weiteren Digitalisierungsprojekte in der EU und der restlichen Welt (hier z.B. die vielen Digitalisierungsprojekte des Joint Information Systems Committee (JISC), aufgefasst unter: <http://www.jisc.ac.uk/index.cfm?name=coll_subject_a> sowie das gigantische „American Memory“ Projekt der Library of Congress (LOC), vgl. <<http://memory.loc.gov/ammem/index.html>>); und noch die nicht-staatlichen not-for-profit Digitalisierungsprojekte, vor allem im Bereich der Zeitschriften (vgl. <<http://www.jstor.org/>> und <<http://www.digizeitschriften.de/>>) (11.08.2006).

10 Vgl. Borghoff, Uwe M.; Rödiger, Peter; Scheffczyk, Jan; Schmitz, Lothar, *Langzeitarchivierung. Methoden zur Erhaltung digitaler Dokumente*, Heidelberg 2003.

11 Vgl. <<http://www.babs-muenchen.de/>> (11.08.2006).

lung und Erprobung von Verfahren zur automatisierten Spiegelung von Websites.¹²

- *Wissenschaftliche Primärdaten in der Meteorologie*, Technische Informationsbibliothek (TIB) Hannover: Bibliothek als Vermittlungsdienst – eine DOI (Digital Object Identifier) Vermittlungsstelle – zu wissenschaftlichen Primärdaten.¹³
- *Dokumentation und Archivierung von Rohdatensätzen aus der psychologischen Forschung*, Zentrum für Psychologische Information und Dokumentation (ZPID) Trier: gesicherte Archivierung durch Metadaten genau beschriebener Rohdatensätze zwecks Re- und Sekundäranalysen durch Dritte.¹⁴
- *Pressearchive des Hamburger Weltwirtschaftsarchivs (HWWA)*, HWWA Hamburg: sukzessive Freischaltung von urheberrechtlich geschützten Digitalisaten.¹⁵

Als nächsten Schritt plant die DFG die gezielte Förderung von verteilten Speichersystemen für digitale Inhalte. So können die Integrität der gespeicherten Daten systemseitig permanent überprüft und gegebenenfalls auftretende Verluste selbsttätig kompensiert werden. Eine solche verteilte Speicherung bietet sich zum Beispiel für die Inhalte von Open Access-Zeitschriften, Materialien aus den Projekten zur retrospektiven Digitalisierung von Bibliotheksbeständen sowie die Inhalte der Fachinformationsführer an.

Darüber hinaus gibt es aber noch weitere Ansätze und Fragestellungen, die wiederholt in verschiedenen Kontexten der wissenschaftlichen Datenhaltung erprobt bzw. beantwortet werden müssen:

1. *Technik*: Welche bereits verfügbaren technischen Lösungen können für die Langfristarchivierung im jeweiligen Handlungsfeld (SSG-Bibliotheken, Digitalisierungsprojekte, Forschungsprimärdaten) eingesetzt werden? Welche Vor- und Nachteile resultieren aus dem Einsatz verschiedener Lösungen?

12 Vgl. <<http://www.fes.de/archiv/spiegelung/default.htm>> (11.08.2006).

13 Vgl. <http://www.std-doi.de/front_content.php> (11.08.2006).

14 Vgl. <<http://www.zpid.de/>> (11.08.2006).

15 Vgl. <<http://webopac.hwwa.de/digiview/>> (11.08.2006).

2. *Organisation*: Wie kann es gelingen, Netzwerke von vertrauenswürdigen Partneereinrichtungen aufzubauen? Unter welchen Voraussetzungen und aufgrund welcher organisatorischen Vorbedingungen können die im System der überregionalen Literaturversorgung engagierten Bibliotheken ein derartiges Netzwerk bilden? Nach welchen Modellen wären die langfristige Verfügbarkeit und der dauerhafte Zugang zu den Ergebnissen von Digitalisierungsprojekten verlässlich zu garantieren? Mit welchen Modellen könnte das Auffinden von und der Zugriff auf wissenschaftliche(n) Forschungsprimärdaten langfristig und gegebenenfalls über die Grenzen einzelner Disziplinen hinweg ermöglicht werden?
3. *Recht*: Welche rechtlichen Regelungen sind bei der Langzeitarchivierung lizenzpflichtiger und urheberrechtlich geschützter Inhalte zu beachten? Inwieweit sind auch datenschutzrechtliche Aspekte zu bedenken? Wie können etwa maschinell lesbare Rechtssysteme, oder die Zusammenarbeit mit Verlagen in Public-Private-Partnership-Modellen in diesem Bereich zu Problemlösungen beitragen? Und, last but certainly not least,
4. *Kosten*: Mit welchen Kosten ist beim Aufbau eines Langzeitarchivierungs-Netzwerks in verschiedenen Handlungsfeldern zu rechnen? Wie skalieren sich diese Kosten beim nachhaltigen und dauerhaften Betrieb dieser Systeme mit mehreren Partnern? Welche Geschäftsmodelle können dazu beitragen, vertrauenswürdige und tragfähige Netzwerke für die Archivierung digitaler Objekte aufzubauen?

Nur indem solche Fragen für verschiedene Arten von digitalen Objekten und in verschiedenen Kontexten beantwortet – und diese Antworten dann in konkrete Maßnahmen umgesetzt – werden, kann in der Tat von einem funktionierenden Langzeitarchivierungssystem gesprochen werden.

Dr. Max Vögler ist Referent bei der Deutschen Forschungsgemeinschaft in der Gruppe Wissenschaftliche Literaturversorgungs- und Informations-

systeme. Er ist Historiker und hat 2005 mit einem Thema zur Religionsgeschichte in der Habsburgmonarchie an der Columbia Universität in New York promoviert. Von 2002 bis 2005 war er Projektkoordinator bei Clio-online. E-Mail: Max.Voegler@dfg.de

DIE EUBAM-GRUPPE UND IHRE MITWIRKUNG IN DEN
EUROPÄISCHEN KOOPERATIONSPROJEKTEN
MINERVA UND MICHAEL

von Monika Hagedorn-Saupe

In diesem Beitrag wird die EUBAM-Gruppe vorgestellt, die als beratendes Gremium zu einer verstärkten Präsenz deutscher Kultureinrichtungen in Europa beiträgt und den Dialog zwischen den Kultureinrichtungen und europäischen Stellen fördert. Mit dieser Zielsetzung wirkt EUBAM auch mit an europäischen Kooperations- und Koordinierungsprojekten.

EUBAM steht für „Europäische Angelegenheiten für Bibliotheken, Archive, Museen und Denkmalpflege“ in Deutschland und beschreibt eine Arbeitsgruppe, die sich aus Vertretern der Kultusministerkonferenz (KMK), der Bundes- und Länderministerien, der Deutschen Forschungsgemeinschaft (DFG) sowie von Experten/innen der Sparten Bibliothek, Archiv, Museum und Denkmalpflege zusammensetzt. EUBAM wurde im Jahr 2001 ins Leben gerufen, um — wie im Leitbild (siehe Abbildung) im April 2005 formuliert — in fachlicher und strategischer Hinsicht Maßnahmen zur Verbesserung des Zugangs zu und der Erhaltung von beweglichem und unbeweglichem Kulturgut zu entwickeln und zur Koordinierung von Digitalisierungsaktivitäten im europäischen Kontext Beiträge zu leisten. EUBAM nimmt eine nationale Focus- und Transferfunktion für entsprechende Förderprogramme der EU wahr, indem sie

- die Entwicklung einer nationalen (Bund und Länder) Digitalisierungsstrategie – unter dem Gesichtspunkt des Zugangs zu kulturellen und wissenschaftlichen Inhalten für Bibliotheken, Archive, Museen und

- der Denkmalpflege – initiiert und begleitet,
- entsprechende Aktivitäten in Deutschland in moderierender Weise anregt und verbindet,
- den Aufbau nationaler, spartenübergreifender Kooperationsstrukturen fördert,
- den Informationstransfer zwischen den EU-Gremien und den oben genannten nationalen Kultursparten organisiert (unter anderem mit Internet-Angebot, Mailingliste, Informationsveranstaltungen und Erstberatung von Antragstellern),
- nationale Interessen im Vorfeld von Planungen zu EU-Förderprogrammen bündelt und an die politischen Instanzen weiter leitet sowie
- im Rahmen ihrer Mitgliedschaft in koordinierenden EU-Projekten, wie MINERVA und MINERVA-PLUS sowie MICHAEL und MICHAEL-PLUS, nationale Positionen in den entsprechenden EU-Gremien vertritt.

Vorsitzender von EUBAM ist derzeit Prof. Dr. h.c. Klaus-Dieter Lehmann, Präsident der Stiftung Preußischer Kulturbesitz in Berlin. Das Sekretariat von EUBAM wird kooperativ geführt durch die Staatsbibliothek zu Berlin Preußischer Kulturbesitz (für den Bereich Bibliotheken), das Institut für Museumskunde der Staatlichen Museen zu Berlin Preußischer Kulturbesitz (für den Bereich der Museen), das Bundesarchiv (für den Bereich der Archive) und die Oberste Denkmalschutzbehörde der Senatsverwaltung für Stadtentwicklung in Berlin (für den Bereich der Denkmalpflege).

Leitbild für EUBAM

Die interministerielle Bund-Länder-Arbeitsgruppe EUBAM (EUropäische Angelegenheiten für Bibliotheken, Archive und Museen) ist ein Zusammenschluss von Vertretern der Kultusministerkonferenz (KMK), der Bundes- und Länderministerien, der Deutschen Forschungsgemeinschaft (DFG) sowie von Experten der Sparten Bibliothek, Archiv, Museum und Denkmalpflege. Ihr vorrangiges Ziel ist, in fachlicher und strategischer Hinsicht deren Aktivitäten und Interessen vor allem für Maßnahmen zur Verbesserung des Zugangs zu und der Erhaltung von beweglichem und

unbeweglichem Kulturgut zu entwickeln und zur Koordinierung Beiträge zu leisten. Die EUBAM nimmt damit insbesondere für Maßnahmen zur Erhaltung des wissenschaftlichen und kulturellen Erbes eine nationale Focus- und Transferfunktion für entsprechende Förderprogramme der EU wahr, indem sie

- die Entwicklung einer nationalen (Bund und Länder) Digitalisierungsstrategie – unter dem Gesichtspunkt des Zugangs zu kulturellen und wissenschaftlichen Inhalten – für Bibliotheken, Archive, Museen und der Denkmalpflege initiiert und begleitet,
- entsprechende Aktivitäten in Deutschland in moderierender Weise anregt und verbindet,
- den Aufbau nationaler, spartenübergreifender Kooperationsstrukturen fördert,
- den Informationstransfer zwischen den EU-Gremien und den oben genannten nationalen Kultursparten organisiert (unter anderem mit Internet-Angebot, Mailingliste,
- Informationsveranstaltungen und Erstberatung von Antragstellern),
- nationale Interessen im Vorfeld von Planungen zu EU-Förderprogrammen bündelt und an die politischen Instanzen weiter leitet sowie
- im Rahmen ihrer Mitgliedschaft in MINERVA und MINERVA-PLUS nationale Positionen in den entsprechenden EU-Gremien vertritt.

Die EUBAM unterstützt die interessierten Institutionen bei der Vertretung ihrer Interessen gegenüber amtlichen Stellen der Europäischen Union und unterstützt staatliche Stellen in Deutschland bei der Formulierung von Politikzielen im nationalen und europäischen Kontext sowie bei deren Implementierung. Durch Information, Beratung und Koordination leistet die EUBAM einen unverzichtbaren Beitrag zur Schaffung der für die Wettbewerbsfähigkeit notwendigen Transparenz und Motivation für die oben genannten Kultureinrichtungen in Deutschland nach innen und nach außen und unterstützt so im Sinne eines kooperativen Föderalismus die Bemühungen der Länder und des Bundes auf dem Weg des „Fit-Machens

für Europa“.¹

EUBAM nimmt damit zunächst strategische Aufgaben wahr. Um diese auch effektiv umzusetzen und zum einen die für europäische Netzwerke notwendigen Stellungnahmen zu erarbeiten sowie zum anderen die Kultursparten auch intensiv zu beteiligen, wurde im Jahr 2004 die EUBAM-Arbeitsgruppe „Digitalisierung im Kulturbereich in Deutschland“ (auch als EUBAM-MINERVA-Spiegelgruppe bezeichnet) eingerichtet. Ernannt wurden sechzehn Experten/innen aus Bibliotheken, Archiven, Museen und dem Bereich der Denkmalpflege, deren Aufgabe es ist, Arbeitspapiere und Stellungnahmen zu den auf europäischer Ebene diskutierten Themen vorzubereiten und Vorschläge zu erarbeiten, wie deutsche Museen, Bibliotheken, Archive und Einrichtungen der Denkmalpflege im europäischen Kontext sichtbar gemacht werden können. Auf Bitten von EUBAM hat die Arbeitsgruppe ein Konzept für ein Internet-Portal „Kulturerbe-digital“ entwickelt. Mit dem Aufbau dieses Internet-Portals soll die deutsche und internationale Öffentlichkeit in umfassendem Maße und gebündelt über die zahlreichen deutschen Aktivitäten im Bereich der Digitalisierung von Kulturgut informiert werden. Das Portal soll die vielfältigen vorhandenen Ressourcen miteinander vernetzen. Durch die Zusammenarbeit mit dem „Projektträger Neue Medien in der Bildung + Fachinformation (PT-NMB+F) im DLR (Deutsches Zentrum für Luft- und Raumfahrt e.V.)“ ist auch die Internet-Plattform **DL-Forum**² beteiligt, die Informationen zu Projekten, Förderprogrammen, Förderbekanntmachungen, Veranstaltungen und Anlaufstellen im Umfeld des Förderschwerpunkts „Digitale Bibliothek“ insbesondere des Bundesministeriums für Bildung und Forschung (BMBF) und der Deutschen Forschungsgemeinschaft (DFG) zusammenführt.

In Kooperation mit dem DL-Forum erarbeiten die Sekretariate gemeinsam mit der EUBAM-Arbeitsgruppe „Digitalisierung im Kulturbereich

1 Vorsitzender der EUBAM ist derzeit Prof. Dr. h.c. Klaus-Dieter Lehmann, Präsident der Stiftung Preußischer Kulturbesitz, Berlin. Website: <<http://www.eubam.de>> (Am 19.04.2005 verabschiedete Fassung). Alle Links wurden zuletzt am 06.11.2006 überprüft.

2 Vgl. <<http://www.dl-forum.de>>.

reich in Deutschland“ eine Übersicht über Digitalisierungsprojekte im Kulturbereich. Diese Informationen, zunächst als Liste vorgelegt und auf der Homepage von EUBAM³ zum Herunterladen bereitgestellt, werden nun in eine Datenbank eingepflegt, so dass sie in Zukunft besser recherchierbar allen Interessierten zur Verfügung stehen.

MINERVA

Die EUBAM-Arbeitsgruppe „Digitalisierung im Kulturbereich in Deutschland“ wird auch als „EUBAM-MINERVA-Spiegelgruppe“ bezeichnet, da sie die Schnittstelle bilden soll zwischen den deutschen und europäischen Diskussionen zum Themenkomplex „Digitalisierung im Kulturbereich“. MINERVA (Ministerial NETwork for Valorising Activities in digitisation) begann als im Rahmen des IST-Programms gefördertes Projekt (Laufzeit: 2002-2005) von überwiegend Kultusministerien mit dem Ziel, eine europäische Plattform für Digitalisierungsaktivitäten im kulturellen Bereich zu schaffen, die nationalen Digitalisierungsstrategien und -aktivitäten stärker aufeinander abzustimmen und gemeinsam mit der Europäischen Kommission die im eEurope-Aktionsplan festgelegten Ziele zu erreichen.⁴

Die Europäische Kommission hatte im Dezember 1999 die Initiative eEurope gestartet, um Europa „ans Netz zu bringen“. Hierfür wurde ein eEurope-Aktionsplan erstellt. Im März 2000 in Lissabon hat die Kommission auf dem Treffen des Europäischen Rates einen Bericht vorgelegt. Am 19./20. Juni 2000 in Feira (Portugal) hat der Europäische Rat dem eEurope-Aktionsplan 2002 zugestimmt. Dieser Aktionsplan umfasste auch den Bereich „Europäische digitale Inhalte für globale Netze“, wobei explizit auch der kulturelle Bereich genannt wurde. Dem eEurope-Aktionsplan 2002 folgte der eEurope-Aktionsplan 2005, diesem wiederum die Initiative i2010, die von der Kommission im Juni 2005 verabschiedet wurde.⁵ In i2010 wird formuliert, dass eine Informationsgesellschaft gefördert werden soll, die **alle Menschen einbezieht**. Um diese zu erreichen, sollen eine

3 Vgl. <<http://www.eubam.de>>.

4 Vgl. <<http://www.minervaeurope.org>>.

5 Vgl. <http://ec.europa.eu/information_society/ceurope/i2010/index_en.htm>.

Reihe von Initiativen angeregt und umgesetzt werden. Eine davon ist der Aufbau von digitalen Bibliotheken, über die alle Zugang zu Multimedia und zur multilingualen europäischen Kultur haben sollen.

Für Deutschland erhob sich die Frage, wer Partner in MINERVA werden könnte. Bedingt durch die Kulturhoheit der Länder, liegt die Hauptverantwortung für Kultur bei den sechzehn Kultusministerien der Länder, daneben nimmt der/die Bundesbeauftragte für die Kultur und die Medien, angesiedelt am Bundeskanzleramt, eine Reihe von Aufgaben wahr. Aufgrund dieser verteilten Verantwortlichkeiten für Kultur in Deutschland wurde die Stiftung Preußischer Kulturbesitz für EUBAM Partner in MINERVA. Im Februar 2004 wurde MINERVA durch MINERVA Plus erweitert. MINERVA Plus ist die Erweiterung und Fortführung des bis 2005 befristeten Projektes MINERVA. Das Projekt umfasst die Kooperation zwischen den Ländern Österreich, Belgien, Tschechien, Dänemark, Estland, Finnland, Frankreich, Deutschland, Griechenland, Großbritannien, Ungarn, Niederlande, Irland, Italien, Malta, Polen, Portugal, Spanien, Schweden, und Slowenien (Israel und Russland sind als assoziierte Partner beteiligt).

Das MINERVA-Projekt hatte zum Ziel, zum einen koordinierende Maßnahmen zur Entwicklung abgestimmter Digitalisierungsstrategien im Kulturbereich zu unterstützen und zu befördern; zum anderen sollten europäische Arbeitsgruppen abgestimmte Empfehlungen und Arbeitsmaterialien erstellen, die die Digitalisierung im Kulturbereich und das Zugänglichmachen der Digitalisate unterstützen. Eine Arbeitsgruppe mit dem Titel „Identification of user needs, content and quality framework for common access points“ behandelt die Fragen der Zugänglichmachung der Daten im Internet. Gemeinsam wurde ein Handbuch entwickelt („Handbook für Quality in Cultural Websites: Improving Quality for Citizens“), in dem Rahmenbedingungen und Grundprinzipien formuliert wurden, die von Relevanz sind, wenn man eine qualitativ hochwertige Website aufbauen möchte. Ausgehend von diesen allgemeinen Grundsätzen wurden zehn Anforderungen an eine gute Webseite formuliert:

„Eine qualitativ hochwertige Webseite erfüllt folgende Kriterien:

- **Transparenz:** sie soll deutlich strukturiert sein und den Zweck sowie

die Verantwortlichen für die Seite benennen;

- **Effektivität:** sie soll ihre Inhalte sinnvoll auswählen, digitalisieren, präsentieren und validieren, um eine größtmögliche Nutzbarkeit zu gewährleisten;
- **Pflege:** sie soll Richtlinien für eine Qualitätssicherung beachten und damit sicher stellen, dass die Webseite gepflegt und in angemessenem Zeitraum aktualisiert wird;
- **Zugänglichkeit:** sie soll für alle Benutzer zugänglich sein, unabhängig von der verwendeten Technologie und von persönlichen Behinderungen; dies schließt Navigation, Inhalt und alle interaktiven Elemente ein;
- **Benutzerorientierung:** sie soll die Nutzerbedürfnisse berücksichtigen, Relevanz und leichte Bedienbarkeit gewährleisten sowie Evaluationsergebnisse und Feedback einbeziehen;
- **Interaktivität:** sie soll den Nutzern die Möglichkeit bieten, mit den Verantwortlichen Kontakt aufzunehmen und eine zweckdienliche Antwort zu erhalten. Sie soll – wo es angemessen ist – die Nutzenden ermutigen, Anfragen zu stellen und Informationen gemeinsam zu nutzen sowie Diskussionen mit und unter den Nutzern anregen;
- **Mehrsprachigkeit:** sie soll sich der Wichtigkeit der Mehrsprachigkeit bewusst sein, indem sie als Mindeststandard einen Zugang in mehr als einer Sprache anbietet;
- **Interoperabilität:** sie soll in kulturelle Netzwerke integriert sein und es den Nutzenden ermöglichen, Inhalte und Services, die ihren Bedürfnissen entsprechen, leicht aufzufinden;
- **verantwortliches Management:** sie soll den rechtlichen Bestimmungen wie Urheberrecht und Datenschutz entsprechen. Die Bedingungen der Nutzung der Webseite und ihrer Inhalte müssen deutlich gemacht werden;
- **dauerhafte Bewahrung:** sie soll solche Richtlinien und Standards übernehmen, die sicherstellen, dass die Webseite und ihr Inhalt langfristig erhalten werden können.⁶

6 Vgl. <<http://www.minervaeurope.org>>.

Für jeden dieser Punkte wurde eine Checkliste aufgestellt sowie einige Fragen formuliert, die die Verantwortlichen für die Webpräsentation in Kultureinrichtungen darin unterstützen sollen, die eigene Webseite auf Qualität zu prüfen und nutzergerechte Webseiten aufzubauen. Diese zehn Prinzipien und die ihnen zugeordneten Checklisten und Fragen wurden in mehreren europäischen Sprachen publiziert (weitere sind in Vorbereitung). Während große Kultureinrichtungen oft bereits einen ausgebauten Internetauftritt haben, ist dies für kleinere Museen, Bibliotheken und Archive oft noch nicht der Fall. Um diese Einrichtungen beim Aufbau einer eigenen Webseite zu unterstützen, wurde ein weiteres Instrument entwickelt: „Museo & Web“. Hierbei handelt es sich um eine webbasierte Anleitung für den Aufbau einer Webseite, die den zehn Qualitätskriterien genügt. Eine weitere Arbeitsgruppe erarbeitete ein Handbuch, das die Kultureinrichtungen bei der Durchführung von Digitalisierungsprojekten unterstützen soll, das in zehn europäischen Sprachen verbreitet wird. Der deutsche Titel lautet: „Good Practice – Handbuch für Digitalisierungsprojekte“. Eine Arbeitsgruppe unter britischer Leitung erarbeitete technische Richtlinien, die bei Digitalisierung berücksichtigt werden sollten. Auch diese Empfehlung liegt in vier Sprachen (Deutsch, Englisch, Französisch und Griechisch) vor. Alle diese Publikationen stehen auch auf der Homepage von MINERVA/MINERVAPLUS⁷ zur Verfügung und viele Interessierte haben inzwischen davon Gebrauch gemacht. Ende Januar 2006 ist die finanzielle Förderung der EU von MINERVA/MINERVA-PLUS ausgelaufen. Insgesamt haben sich alle beteiligten Partner weit über das vertraglich festgelegte engagiert und führen auch nach Auslaufen der Projektförderung gemeinsame Aktivitäten weiter.

Aus den gemeinsamen Aktivitäten zur Förderung von Digitalisierung im Kulturbereich heraus entstanden auch die Überlegungen zum Aufbau eines europäischen Portals, in dem die im Kulturbereich vorhandenen digitalen Sammlungen aus den verschiedenen europäischen Ländern nachgewiesen werden. Die Kultusministerien Frankreichs und Italiens und die dem britischen Kultusministerium nachgeordnete Einrichtung MLA

7 Vgl. <<http://www.minervaeurope.org>>.

(Museums, Libraries and Archives Council), die in ihren Ländern Digitalisierungsvorhaben und die Veröffentlichung von digitalen Daten in großem Umfang fördern, beantragten eine Förderung bei der EU, um die nationalen Aktivitäten europäisch zu vernetzen. Dies ist das Projekt MICHAEL.

MICHAEL

Im Rahmen des eTEN-Programms wird inzwischen das Vorhaben MICHAEL (Multilingual Inventory of Cultural Heritage in Europe)⁸ gefördert. Zielsetzung des Projektes ist es, europäisches kulturelles Erbe weltweit zugänglich zu machen. MICHAEL ist darauf ausgerichtet, die vielfältigen europäischen digitalen Sammlungen von Museen, Bibliotheken und Archiven zu vernetzen und diese über ein zentrales Zugangportal zu erschließen. MICHAEL soll damit die Recherche schwer auffindbarer Informationen zu digitalen Sammlungen von Europas Kultureinrichtungen erleichtern. Unter Nutzung der bereits vom französischen Kultusministerium genutzten Open Source-Technologie für den Nachweis von digitalen Sammlungen im Kulturbereich in Frankreich⁹ wird ein mehrsprachiges Portal aufgebaut. Das für die Präsentation der digitalen Sammlungen formulierte Datenmodell basiert auf den Empfehlungen von MINERVA zur Interoperabilität und orientiert sich an den W3C-Empfehlungen für die Beschreibung von digitalen Sammlungen. Die Software-Plattform umfasst zwei Module: Ein Produktionsmodul, mit dessen Hilfe webbasierte Inhalte in das jeweilige Portal eingepflegt werden können und ein Publikationsmodul. Letzteres bietet ein intuitives, einfach nutzbares Interface, damit interessierte Nutzer/innen mit ihrem Webbrowser digitale kulturelle Inhalte leicht auffinden können. Mit Hilfe dieses Publikationsmoduls werden nationale MICHAEL-Portale aufgebaut, die über eine OAI-PMH-Schnittstelle Daten an den europäischen Server liefern.

Großbritannien, Italien und Frankreich haben bereits einen solchen nationalen Server aufgesetzt. Inzwischen sind auch diesem Vorhaben elf weitere Länder (Belgien, Deutschland, Finnland, Griechenland, Malta,

⁸ Vgl. <<http://www.michael-culture.org>>.

⁹ Vgl. <<http://www.culture.fr>>.

Niederlande, Polen, Schweden, Spanien, Tschechische Republik, Ungarn) beigetreten (Projekt MICHAEL-PLUS), die bis 2007 weitere nationale Server einrichten werden. Eine erste Version des europäischen Servers wird voraussichtlich im November 2006 online sein. Eine besondere Herausforderung für den Aufbau des europäischen Servers ist die Sprachenvielfalt in Europa. Das Portal soll den interessierten Nutzern/innen in ihrer eigenen Sprache zur Verfügung stehen, die Navigation und allgemeine Beschreibungen müssen in allen Sprachen der beteiligten Partner vorliegen. Damit die Nutzer/innen über alle Sammlungen in ihrer eigenen Sprache recherchieren können, wird zunächst auf ausgewählte Begriffe des UNESCO-Thesaurus aufbauend eine Liste von Begriffen mit ihren jeweiligen Entsprechungen in den Sprachen der Partner im Portal implementiert, über die die Suche erfolgen wird. In der Regel sind auch in MICHAEL/MICHAEL-PLUS wie auch in MINERVA/MINERVA-PLUS die jeweiligen Kultusministerien Projektpartner. Für Deutschland hat EUBAM wegen der verteilten Verantwortlichkeiten für Kultur einen anderen Weg vorgeschlagen und es konnten sieben große Kultureinrichtungen als Partner für das Projekt gewonnen werden: das Bundesarchiv und das Landesarchiv in Baden-Württemberg, Die Deutsche Bibliothek und die Bayerische Staatsbibliothek, das Deutsche Museum in München und das Naturmuseum Senckenberg in Frankfurt am Main sowie die Stiftung Preußischer Kulturbesitz, die mit der Staatsbibliothek zu Berlin, den Staatlichen Museen zu Berlin und dem Geheimen Staatsarchiv alle drei Kultursparten umfasst. Diese sieben Partner werden gemeinsam das deutsche Portal aufbauen und dafür werben, dass möglichst viele deutsche Kultureinrichtungen Informationen über ihre digitalen Sammlungen für das deutsche und europäische Portal bereitstellen.

Auch EUBAM und die EUBAM-MINERVA-Spiegelgruppe unterstützen den Auf- und Ausbau des MICHAEL-Portals, weil damit die Vielfalt und der Reichtum der Kultur in Deutschland besser sichtbar werden.

Auf europäischer Ebene werden derzeit große Anstrengungen unternommen, digitalisiertes bzw. digitales Kulturerbe möglichst schnell und in möglichst großem Umfang einem weltweiten Publi-

kum zugänglich zu machen. Man spricht von der Europäischen Digitalen Bibliothek, die Informationen zu Beständen aus allen Kultursparten umfassen soll. Um mit dieser rasanten Entwicklung Schritt halten zu können, muss schnell reagiert werden können. Hier ist EUBAM als koordinierendes Gremium eine geeignete Instanz.

Monika Hagedorn-Saupe ist Leiterin des Referats „Besucherbezogene Museumsforschung und Kulturstatistik“ am Institut für Museumskunde der Staatlichen Museen zu Berlin Preußischer Kulturbesitz und stellvertretende Leiterin des Instituts Lehraufträge in den Bereichen Besucherforschung, Museumspädagogik und Internationale Medieninformatik an der Fachhochschule für Technik und Wirtschaft in Berlin. E-Mail: m.hagedorn@smb.spk-berlin.de

DIE DIGITALISIERUNGSPROJEKTE DER
UNIVERSITÄTSBIBLIOTHEKEN IN FRANKREICH –
PERSEE UND ANDERE ANGEBOTE

von *Frédéric Blin und Valérie Néouze*

Digitalisierungsprogramme werden in den französischen Bibliotheken seit ungefähr 20 Jahren geleitet. Heutzutage führen die meisten Einrichtungen, wie etwa öffentliche Bibliotheken (Ministère de la Culture) oder akademische Bibliotheken (Ministère de l'Education nationale, de l'enseignement supérieur et de la recherche = MENESR, Ministerium für Bildung und Forschung), Digitalisierungsprojekte ihrer Sammlungen durch. Gallica, die digitale Bibliothek der Bibliothèque nationale de France mit ihren circa 70.000 Werken im Volltext und 80.000 Bildern, ist das bekannteste französische Beispiel solcher Programme auf internationaler Ebene. Darüber hinaus werden zurzeit weitere wichtige Projekte in den akademischen Bibliotheken entwickelt. Diese Vorhaben werden vom MENESR finanziell unterstützt. Dabei entstehen auch nationale Programme, wie PERSEE, das französische Portal für retrospektive Sammlungen wissenschaftlicher Zeitschriften in den Geistes- und Sozialwissenschaften in französischer Sprache.

Die Digitalisierungsprojekte akademischer Bibliotheken in Frankreich

Nationalen und internationalen Beispielen folgend, sind in den letzten Jahren zahlreiche lokale Digitalisierungsprojekte in den französischen Universitätsbibliotheken initiiert worden. Gleichzeitig reagieren sie damit ebenfalls auf eine neue Realität des wissenschaftlichen Veröffentlichungswesens. Diese Projekte, welche die auf französischem Niveau angestellten

Überlegungen über die Schaffung einer europäischen digitalen Bibliothek vorangetrieben haben¹, deuten die Grundlagen eines digitalen dokumentarischen Netzes an, das jenes der CADIST-Bibliotheken (*Centres d'Acquisition et de Diffusion de l'Information Scientifique et Technique* – Zentren des Erwerbs und der Verbreitung der wissenschaftlichen und technischen Information²) vervollständigt.

Unter diesem Gesichtspunkt haben einige Bibliotheken seit mehreren Jahren Portale zu digitalisierten Dokumenten, von nationalem oder sogar internationalem Interesse, entwickelt. Unter den repräsentativsten Verwirklichungen in diesem Bereich kann man folgende erwähnen:

- *Medic@*³, die digitale Bibliothek für Medizin der *Bibliothèque interuniversitaire de Médecine de Paris* (BIUM), präsentiert aus ihren Sammlungen mehr als 3.200 Werke im Volltext, sowie eine Datenbank von Bildern und Arztporträts von fast 6.000 Dokumenten. Die BIUM beteiligt sich darüber hinaus mit ihren Sammlungen auch an anderen Digitalisierungsprojekten, so etwa im Portal *Criminocorpus*, welches zum Beispiel die Zeitschrift *Archive de l'anthropologie criminelle* von 1886 bis 1914 im Volltext zugänglich macht.⁴ Damit hat sich das Internetprojekt der BIUM mit seinem föderalen Charakter letztendlich zu einer maßgeblichen, virtuellen Medizinbibliothek entwickelt.⁵
- Das *Conservatoire numérique des Arts et Métiers* (CNUM⁶) ist eine virtuelle Bibliothek, die sich der Wissenschafts- und Technikgeschichte widmet. Sie besteht aus den Sammlungen des *Conservatoire natio-*

1 Jeanneney, Jean-Noël, *Quand Google défie l'Europe*, Paris 2005.

2 Diese CADIST-Bibliotheken sind, wie die Sondersammelgebietsbibliotheken in Deutschland, spezialisierte Forschungsbibliotheken. Sie wenden sich hauptsächlich an die Gesamtheit der französischen Wissenschaftler/innen in ihren spezifischen Bereichen.

3 *MEDIC@*, vgl. <<http://www.bium.univ-paris5.fr/histmed/medica.htm>> (Alle Links wurden zuletzt überprüft am 20.09.2006).

4 *Criminocorpus*, vgl. <<http://www.hstl.crhst.cnrs.fr/criminocorpus>>.

5 Unter den angebotenen Diensten findet man Verbindungen in Richtung Bücher und elektronischer Zeitschriften, spezialisierter Internet-Adressen, virtueller Ausstellungen, Bilddatenbanken, Namensverzeichnissen und Zeitplänen usw.

6 *Conservatoire numérique des arts et métiers*, vgl. <<http://cnum.cnam.fr/>>.

nal des arts et métiers (CNAM⁷). Das Portal, das mehr als 400 digitalisierte Werke vom 16. bis zum 20. Jahrhundert (circa 120.000 Seiten), darunter auch die nationalen Ausstellungen und Weltausstellungen des 19. Jahrhunderts, anbietet, registriert etwa 25.000 Besucher/innen im Monat.

- Die Manuskripte der Rheinmystik der *Bibliothèque nationale et universitaire de Strasbourg*: Es handelt sich um 25 religiöse Bücher des 14. und 15. Jahrhunderts. Diese wurden als Bild und nicht als Text digitalisiert.⁸
- Die Online-Ausgaben der *Ecole nationale des chartes*: Eine große Bildungs- und Forschungseinrichtung, die auf historische Wissenschaften spezialisiert ist. Sie verfolgt eine aktive Politik der Digitalisierung und Publikation historischer Quellen, wie etwa Chroniken, Chartulari, Prägestempel, Manuskripte usw.⁹ Diese Dokumente werden mit Kommentaren und kritischen Anmerkungen von den Lehrenden und Studierenden der Schule ergänzt. Die Bibliothek begleitet dieses Programm, indem sie ihre Sammlungen zur Verfügung stellt.
- Der Nachlass des Wissenschaftlers Jean-Martin Charcot wurde gemeinsam mit dem Dokumentationsdienst der Universität Paris VI. „Pierre und Marie Curie“ digitalisiert.¹⁰ Benutzer/innen können dort Unterrichtsmanuskripte, Thesen, Zeitschriften und seltene Werke sowie Alben der Assistenzärzte/innen des Krankenhauses *de la Salpêtrière* finden.
- Zahlreiche Universitäten haben ebenfalls ihre Dissertationsbestände digitalisiert. In einigen Fällen stellen sie sogar ihre wissenschaftlichen Gutachten zur Verfügung.

Weitere Projekte mit einer national koordinierten Digitalisierungspolitik sind zum Beispiel die Projekte der Bibliothek Cujas in Paris, französische

7 *Conservatoire national des arts et métiers*, vgl. <<http://www.cnam.fr/>>.

8 Digitalisate der *Bibliothèque nationale et universitaire de Strasbourg*, vgl. <<http://www-bnus.u-strasbg.fr/BNU/FR/>>.

9 Die *Editions en ligne de l'Ecole nationale des chartes*, vgl. <<http://elec.enc.sorbonne.fr/>>.

10 Charcot Nachlass: <<http://jubil.upmc.fr/repons/portal/bookmark?MainTab=CMS>ShowChannelDoc&ShowDocChannel=presentation/SCDM/charcot/historique>>.

Hauptbibliothek für Rechtswissenschaften¹¹, und die *Bibliothèque de documentation internationale contemporaine* (BDIC), Hauptbibliothek für Zeitgeschichte¹² mit ihren alten und wertvollen Sammlungen. Da diese Projekte noch sehr jung sind, konnten bisher nur eine geringe Zahl von Digitalisaten, beispielsweise historische, rechtswissenschaftliche Werke der Cujas-Bibliothek, online zur Verfügung gestellt werden. Alle diese Projekte wurden von den lokalen Einrichtungen initiiert und unter anderem vom Ministerium für Bildung und Forschung gefördert. In der Tat werden die erwähnten Vorhaben – entweder im Rahmen vierjähriger, staatlicher Verträge, oder im Rahmen außergewöhnlicher Aktionen – durch die Ausschüsse des Ministeriums (Unterabteilung der Bibliotheken) geprüft, die dadurch den Bibliotheken helfen können, ihre wissenschaftlichen und technischen Projekte zu entwickeln. Die Bewilligung einer finanziellen Unterstützung erfolgt mittels mehrerer Kriterien:

- Wissenschaftliche Qualität der Inhalte sowie des kritischen Apparats, das die digitalisierten Dokumente begleitet,
- die Wahl der technischen Lösungsansätze (Qualität der Digitalisierung, Textformat bezüglich des Bildformates, Integration in die Informationssysteme der Universitäten),
- Zugang zum Produkt der Digitalisierung (Suchmaschinen, Sudoc-Portal¹³, OAI-Harvesting...). Die Frage des Zugangs, bisher zu wenig berücksichtigt, gilt nun als essentiell und wird deshalb besonders vom Ministerium geprüft.

Durch diese Begutachtungs- und Unterstützungsrolle will das Ministerium die Koordinierungspolitik der Digitalisierungsprogramme auf eine nationale Ebene stellen. Um besser zu identifizieren, welche Sammlungen schon digitalisiert worden sind oder wo bereits Digitalisierungsprogramme angelaufen sind, aber auch um eine Annäherung lokaler Projekte zu för-

11 Bibliothek Cujas, vgl. <<http://biu-cujas.univ-paris1.fr/>>.

12 Bibliothèque de documentation internationale contemporaine (BDIC), vgl. <<http://www.bdic.fr/index.php3>>.

13 SUDOC Portal, vgl. <<http://www.portail-sudoc.abes.fr/>>. Das Sudoc-Portal ist das förderierende Portal zu Online-Ressourcen der französischen Hochschulen und Universitäten.

dern, hat das Ministerium Ende 2005 beschlossen, eine Datenbank zu den schon digitalisierten Beständen in den französischen Hochschulen einzurichten. Um dieses Ziel zu erreichen, haben sich das Ministerium für Bildung und das Ministerium für Kultur zu einer Zusammenarbeit entschlossen. Das Ministerium für Kultur hat im Rahmen des europäischen Programms MICHAEL (Multicultural Inventory of Cultural Heritage in Europe¹⁴) sein eigenes Portal zu den digitalisierten kulturellen Beständen seiner Institutionen, wie öffentliche Bibliotheken, Archive und Museen, gebaut. Das MENESR hat also das von MICHAEL benutzte Datenmodell überprüft, um es an die Anforderungen der Bestände der Hochschulen, mit Hinblick auf eine technische Kompatibilität mit dem Portal MICHAEL, anzupassen. Das besagte Portal zu den digitalisierten Beständen der französischen Hochschulen soll Ende des Jahres 2006 zugänglich gemacht werden.

Parallel zu dieser Unterstützung lokaler Projekte kann das Ministerium ebenfalls die Initiative ergreifen, nationale Projekte einzuführen. In den letzten Jahren sind zwei Hauptprojekte entstanden; Liber Floridus, eine Datenbank von Buchmalereien mittelalterlicher Manuskripte¹⁵, die im Besitz der Hochschulen sind, und PERSEE¹⁶, das nun etwas ausführlicher vorgestellt werden soll.

PERSEE

Der Kontext

Das Programm PERSEE erfolgt in einem schwierigen wirtschaftlichen Kontext für das Verlagswesen in den Geistes- und Sozialwissenschaften. Während die Forscher/innen in den Naturwissenschaften im Zeitschriftenbereich – Hauptvektor der wissenschaftlichen Information – seit langem die neuen Technologien nutzen, um Methoden für die elektronische Erfassung und Verbreitung der wissenschaftlichen Ergebnisse zu entwickeln und dadurch den schnellen Umlauf, die internationale Verbreitung und die

14 MICHAEL Portal, vgl. <http://www.michael-culture.org/index_f.html>.

15 Liber Floridus, vgl. <<http://liberfloridus.cines.fr/>>.

16 PERSEE, vgl. <<http://www.persee.fr/>>.

optimale Nutzung dieser Ergebnisse für ihre Gemeinschaft favorisieren, bleibt der Bereich der Zeitschriften in den Geistes- und Sozialwissenschaften dagegen aus mehreren Gründen noch abseits dieser Entwicklung.

Charakterisiert wird der geistes- und sozialwissenschaftliche Bereich durch eine Aufsplitterung der Akteure, unter denen die Hochschulen und die gelehrten Gesellschaften einen ausschlaggebenden Platz einnehmen. Die zahlreichen Zeitschriften befinden sich im Allgemeinen in einer sehr heiklen wirtschaftlichen Lage, die mit zwei Hauptfaktoren zusammenhängt.

- Einerseits leiden sie unter einer schwachen internationalen Sichtbarkeit, was hauptsächlich auf eine vorrangig französisch sprechende Leserschaft zurückzuführen ist, aber auch, weil diese Zeitschriften auf einem Modell der Herstellung und Verbreitung beruhen, das auf dem Papier begründet wurde, und welches sich infolgedessen stark von den neuen Praktiken der elektronischen Herausgabe von Forschungsergebnissen unterscheidet.
- Andererseits sind die Redaktionen, die einen anderen Umgang mit Informatiktechnologien pflegen als die naturwissenschaftlichen Publikationsorgane, an eine Kultur des Verlagswesens gebunden.

Trotzdem stellen die Zeitschriften in den Geistes- und Sozialwissenschaften, insbesondere in Frankreich, ein wissenschaftliches Kulturgut sehr hoher Qualität dar. Einer technischen Angleichung sehen sowohl die Forscher/innen als Erzeuger/innen und Verbraucher/innen dieser Information, als auch Redakteure/innen der Zeitschriften mit großer Ungeduld entgegen, da dies die nationale und internationale Sichtbarkeit ihrer Produktion fördern und die Dauerhaftigkeit einiger extrem spezialisierter Titel gewährleisten könnte.

Der Wert dieser Zeitschriften und der potentielle wirtschaftliche Mehrwert, den sie darstellen, sind den angelsächsischen Hauptakteuren nicht verborgen geblieben. Seit mehreren Jahren unternehmen sie eine systematische Akquisition französischer Publikationen aus dem Bereich der Geistes- und Sozialwissenschaften, mit den anspruchsvollsten Titeln beginnend. Sie haben ihren französischen Kollegen/innen vorgeschlagen, die digitalisierten Zeitschriften in ihre eigenen virtuellen und kostenpflich-

tigen Sammlungsportale zu integrieren. Die Angebote dieser Datenverteiler bzw. Akteure sind hinsichtlich der Strategien zur Verbreitung der französischen Wissenschaftsforschung sehr aufschlussreich. In der Tat hätte im Falle mangelnder Alternativen die Wahl dieser kommerziellen Logik in großem Ausmaße verhängnisvolle Folgen gehabt:

- Höhere Kosten, die die wissenschaftliche Gemeinschaft hätte übernehmen müssen, um Zugang zu diesen Zeitschriften zu haben. Das heißt paradoxerweise, dass sie dafür bezahlen müsste, ihre eigenen Forschungsergebnisse zu befragen, obwohl diese schon weitgehend durch öffentliche Mittel subventioniert wurden.
- Das Risiko für die Zeitschriften in den Geistes- und Sozialwissenschaften, einer fast monopolistischen elektronischen Verbreitung zu unterliegen, besteht und kann im Bereich der Naturwissenschaften bereits beobachtet werden.
- Eine weitere Folge wäre eine zusätzliche Schwächung eines Teiles der wissenschaftlichen Produktion, die sehr spezialisiert ist und in Zeitschriften veröffentlicht wird, die in einer strikten Handelslogik für nicht rentabel gehalten würden. Es könnte sogar das Ende einiger Titel wegen einer erzwungenen Koordinierung bedeuten, welche die Konkurrenz zwischen den eigenen Titeln eines Datenverteilers vermeiden soll.

In den Geistes- und Sozialwissenschaften scheint die Mehrzahl der Herausgeber/innen unter bestimmten Bedingungen nicht gegen einen freien Zugang zu den digitalisierten Auflagen der Zeitschriften zu sein. Die Bewilligung einer öffentlichen finanziellen Beihilfe für die retrospektive Digitalisierung der Sammlungen als Gegenleistung zu ihrer kostenlosen Bereitstellung im Netz könnte sich ihrer Meinung nach wie folgt gestalten:

- Ein kostenloses Onlinestellen der Zeitschriften, im Durchschnitt drei bis fünf Jahre nach ihrer ersten Veröffentlichung, bedeutet in der Regel keinen Verdienstausschlag, da wenig Artikel oder thematische Faszikel eine langfristige wirtschaftliche Rentabilität haben.
- Die Verbreitung kostenloser Artikel auf dem Internet stellt ein wirksames Abrufprodukt dar, welches eine neue Leserschaft anziehen könnte, die möglicherweise für ein Abonnement der laufenden Veröffentlichungen in Frage käme.

- Die Handelsverbreitung der letzten Lieferungen (mit Varianten bezüglich der Dauer der *Moving Wall*, die im Durchschnitt für einen Zeitraum von drei bis fünf Jahren zu bewahren ist, um den Verkauf der Zeitschriften zu gewährleisten) soll aufrechterhalten werden.
- Die gedruckte Ausgabe, parallel oder in Anlehnung an die elektronische Version ihrer Veröffentlichungen, soll beibehalten werden.

Die Richtlinien von PERSEE

Aus diesem Protokoll heraus ist das Projekt PERSEE als öffentliches Portal für die elektronische Verbreitung der retrospektiven Sammlungen französischsprachiger Zeitschriften der Geistes- und Sozialwissenschaften entstanden. Es gibt drei Hauptzielsetzungen:

- dieses wissenschaftliche Kulturgut durch seine Verbreitung im Internet zu bewahren und zu valorisieren,
- an der Wirkung der Forschung in französischer Sprache durch die Verbreitung eines ausgedehnten Zeitschriftencorpus teilzunehmen, um dieses in das internationale wissenschaftliche Angebot zu integrieren und so eine linguistische Isolierung zu vermeiden,
- durch die Suchmethoden, die nach der Digitalisierung ermöglicht werden, erweiterte Arbeits- und Nutzungsmöglichkeiten dieser Zeitschriften anzubieten, die jetzt schon in anderen Bereichen üblich geworden sind (automatisierte Recherche in ausgedehnten Sammlungen, Downloads, Anmerkungen usw.).

Unter Leitung der *Direction de l'enseignement supérieur* hat das Projekt PERSEE eine weitgehende Zusammenarbeit mit den anderen Institutionen (unter anderem mit der Direktion für Forschung und Direktion für Technologie) erlaubt, die das wissenschaftliche Veröffentlichungswesen aktiv unterstützen. Drei allgemeine Konzepte sind von diesen verschiedenen Partnern im Rahmen von konkreten Aktionen entwickelt worden:

Ein pragmatisches Konzept:

- Die Notwendigkeit eines kurzfristigen Aktionsplans: In einem wettbewerbsfähigen, internationalen, wissenschaftlichen Kontext und in einer Situation, in der in Frankreich noch keine Alternative vorgeschlagen wurde, um die Veröffentlichung der Zeitschriften online zu erlauben,

- haben einige Herausgeber bilaterale Verhandlungen mit den privaten Handelskonsortien begonnen. Andere haben die Idee eines Verbands bevorzugt, um ein französischsprachiges Ensemble zu schaffen – innerhalb eines hauptsächlich englischsprachigen Angebots.
- Die Wahl eines auf sieben Zeitschriften¹⁷ beschränkten Corpus, die den Anfangskern des Portals bilden und die schnelle Entwicklung eines Prototyps erlauben sollen.
- Die Begrenzung auf die retrospektive Digitalisierung dieser Zeitschriften, im Rahmen des derzeitigen Kompetenzfeldes der *Direction de l'enseignement supérieur*.

Ein wissenschaftliches Konzept:

- Eine enge und ständige Kooperation mit den Forschern/innen und den Zeitschriftenredaktionen, um das Profil und die Funktionen des Portals ihren spezifischen Bedürfnissen anzupassen.
- Die Festlegung einer Auswahl an Zeitschriften verschiedener Disziplinen der französischen wissenschaftlichen Forschung.

Ein offenes Konzept:

- Im Sinne des „öffentlichen Dienstes“: ein frei und kostenlos zugängliches Portal in einer nicht exklusiven Verbreitungslogik.
- Eine Wahl offener technologischer Optionen: Informatikentwicklungen in Open Source, Wahl von international anerkannten Standards und Normen, um die notwendige Kompatibilität mit ähnlichen oder ergänzenden Initiativen im Bereich der retrospektiven Digitalisierung oder des elektronischen Publizierens zu garantieren.
- Eine internationale Öffnung, die eine aktive Partnerschaft mit ERUDIT, einem Projekt aus Quebec¹⁸, bevorzugt, dessen politische Zielsetzungen, wissenschaftliche und technologische Optionen, sowie linguistischer und disziplinarischer Umkreis jenen ähnlich sind, die soeben dargelegt wurden.

17 Die *Annales*, die *Bibliothèque de l'école des chartes*, *L'Homme*, *Matériaux pour l'histoire de notre temps*, die *Revue de l'art*, die *Revue économique*, die *Revue française de science politique*.

18 ERUDIT, vgl. <<http://www.erudit.org>>.

Die Schaffung des Konsortiums PERSEE

Nachdem diese drei allgemeinen Grundsätze und die entsprechende strategische Wahl festgelegt worden waren, wurde das Projekt unter der Leitung eines Ausschusses, der vom Direktor der Hochschulbildung geleitet wurde, gestartet, wobei drei Phasen unterschieden werden können:

- eine Phase der Studie und der Analyse der Hauptmodelle des Verlagswesens (eine Studie, die dem *Institut des sciences du document numérique*, auf der Webseite des ENSSIB zugänglich, anvertraut wurde)¹⁹,
- eine Abstimmungsphase mit bevorzugten Partnern: Arbeitssitzungen mit den potentiell am Projekt Beteiligten (Redaktionen der Zeitschriften, Herausgeber); ausführliche Analyse von ERUDIT,
- eine Vorbereitungsphase des Aktionsplans, die mit einem Aufruf an die französischen Hochschulen zu einer Zusammenarbeit im Bereiche der Digitalisierung und der Verbreitung der Zeitschriftensammlungen nach deutlich aufgestellten Anforderungen abgeschlossen wurde.

Am Ende dieser Studien- und Vorbereitungsphase wurde am 7. März 2003 der Aufruf zur Zusammenarbeit veröffentlicht. Die Wahl der Einrichtung, die für das Endprojekt sorgen sollte, ist am 6. Juni 2003 erfolgt. Mit der Möglichkeit, allein oder im Konsortium zu antworten, haben mehrere öffentliche Einrichtungen beschlossen, gemeinsam einen Vorschlag auszuarbeiten. Insgesamt gingen vier Vorschläge ein. Nach Beratung des Auswahl Ausschusses wurde die Bewerbung der Universität von Lyon 2 in Partnerschaft mit der *Maison de l'Orient et de la Méditerranée* und der Universität von Nizza-Antipolis ausgewählt. Danach wurde ein Übereinkommen zwischen dem Staat, der Einrichtung, die für das Projekt sorgte, den Rechteinhabern der Zeitschriften und dem CINES²⁰ getroffen, um das Portal innerhalb eines Jahres nach der Unterzeichnung dieses Übereinkommens, die am 15. Oktober 2003 erfolgt ist, zu verwirklichen.

19 ENSSIB: Ecole nationale supérieure des sciences de l'information et des bibliothèques, <www.enssib.fr>. Institut des sciences du document numérique, vgl. <http://revues.enssib.fr>.

20 Das CINES, Centre informatique national de l'enseignement supérieur, ist Host der Webseite und für Langzeitarchivierung der Dateien verantwortlich. Das CINES liegt in Montpellier.

Das Portal PERSEE

Infolge der Unterzeichnung des Übereinkommens konnte das Konsortium unter Leitung der Universität von Lyon 2 mit der Phase der Realisierung des Projekts beginnen. Das Projekt-Team wurde damit beauftragt, eine Digitalisierungsplattform aufzubauen und die Software des Portals zu entwickeln sowie die Verwaltung dieses Portals zu übernehmen.

Für diesen Auftrag sollten folgende Elemente geschaffen werden:

- Die sieben Zeitschriften sollten als Bild sowie als Text digitalisiert werden, vom ersten bis zum letzten Faszikel des Jahres 2002 (circa 410.000 Seiten);
- die Dateien sollten automatisch verarbeitet werden, um genormte XML-Notizen zu schaffen;
- die digitalisierten Dokumente sollten auf einer dafür spezifisch entwickelten Webseite online verfügbar gemacht werden;
- die Langzeitarchivierung der Dateien sollte vom CINES entwickelt werden.

Dem Projekt PERSEE ist es gelungen, innerhalb eines Jahres auf XML-Basis die erforderlichen technischen Mittel als Open Source Software – eine strategische Entscheidung, die durch das Ministerium getroffen wurde – zu entwickeln. Die technischen und dokumentarischen Lösungen machen die Besonderheit dieses Projekts aus. Sie erlauben, die zwei Recherche-möglichkeiten des Portals zu optimieren: *Browsen* und *Suchen*.

- Mit dem *Browsen* können die Benutzer/innen in derselben Zeitschrift von einem Faszikel zum anderen übergehen, die interaktiven Inhaltsverzeichnisse durchgehen und so den gewünschten Artikel finden. Die Digitalisierung der Zeitschrift als Bild erlaubt, die visuelle Identität der Zeitschrift zu respektieren, da jede befragte Seite mit dem Original auf Papier identisch ist.
- Bei der *Suche* liegt der Artikel im Zentrum der Recherche. Jeder Artikel wird neben einer Digitalisierung als Bild auch dank einem OCR als Text verfügbar gemacht und in XML nach einem Modell strukturiert, das erlaubt, die die Recherche betreffenden Daten (Name der Zeitschrift, Titel des Artikels, Name des Autors/der Autorin, Veröffentlichungsdatum, Zusammenfassung...) zur Geltung zu bringen.

Die Benutzer/innen können so eine präzise Suche in einem (oder mehreren) Gebiet(en) in einem bestimmten Feld machen oder diese auf den Volltext über die Gesamtheit des verfügbaren Corpus erweitern. Sie können auch andere Plattformen, mit denen Partnerschaften aufgestellt worden sind, befragen.

Die Wahl der Strukturierung und der Digitalisierung wurde von der Gemeinschaft der Forscher/innen getroffen, die an der Ausarbeitung des Lastenheftes des Projekts beteiligt waren. Die Verbindung eines ausgedehnten Corpus und einer mächtigen Suchmaschine begünstigt so die Transdisziplinarität, die den Geistes- und Sozialwissenschaften inhärent ist. Die zahlreichen Online-Dienste, die das PERSEE-Portal anbietet, wurden für die entsprechenden Bedürfnisse der Wissenschaftler/innen entwickelt. Außer den Browser- und Suchfunktionen werden auch personalisierte Dienste angeboten, die gemäß dem Profil des Nutzers/der Nutzerin aufgestellt wurden: persönlicher Raum, der eine Archivierung ausgewählter Artikel ermöglicht; automatische Bibliografieerstellung; Verwaltung des geschichtlichen Überblicks der Forschungsergebnisse; sofortige Information, wenn ein neuer interessanter Artikel online gestellt wird, usw. Die aktive Teilnahme der Redaktionen der Zeitschriften an diesem Prozess, sowohl in der Entscheidungsphase durch ihre Präsenz im Leitungsausschuss als auch in der konkreten Zusammenarbeit, was die materielle und intellektuelle Beschreibung ihrer Zeitschrift betrifft (um die Digitalisierungsoperationen und die Angemessenheit der Indexierung zu optimieren) und in der systematischen Suche nach den Autoren/innen oder ihren Rechteinhabern, muss auch erwähnt werden. Das französische Gesetz verpflichtet PERSEE nämlich, die Autoren/innen um eine offizielle Genehmigung zu bitten, um deren Werke online verfügbar machen zu dürfen. Die systematische Suche nach den Autoren/innen oder ihren Rechteinhabern, die von den Redaktionen der Zeitschriften verlangt wird, bildet also einen wesentlichen Teil des Programms PERSEE. In diesem Sinne konnte PERSEE in seiner Verwirklichungsphase von einer ständigen Rechtsbegleitung profitieren, die durch das CECOJI (*Centre d'Etudes sur la Coopération Juridique Internationale*) vom CNRS gewährleistet wurde, ein Zentrum, das auf geistiges Eigentumsrecht sowie auf das Recht der Infor-

mation und der Kommunikation spezialisiert ist.

Die Entwicklungsperspektiven von PERSEE

Die offizielle Eröffnung von PERSEE hat am 21. Januar 2005 das Ende der ersten Phase des Projekts markiert, die der Verwirklichung der Prototypen und dem progressiven Onlinestellen der Sammlungen der Pionierzeitschriften gewidmet war. Sie hat eine zweite Phase (2005-2007) eröffnet, deren Ziel die Entwicklung neuer Funktionen, die Einführung neuer Partnerschaften und die Integration neuer Zeitschriften ist. Parallel dazu wurde Ende 2005 damit begonnen, an eine Verdoppelung der Produktionskapazitäten zu denken, die dank der Öffnung eines zweiten Digitalisierungszentrums die Vermehrung der Zeitschriften auf PERSEE ermöglichen soll, da die potentielle Zahl der französischen Zeitschriften in den Geistes- und Sozialwissenschaften, und dadurch die potentielle Zahl der Zeitschriften auf PERSEE, sich auf über 200 beläuft.

Seit der Öffnung des Portals im Januar 2005 hat PERSEE ständig Inhalte hinzugefügt. Momentan verbreitet PERSEE die digitalisierten Sammlungen von zwölf Online-Zeitschriften mit einer Gesamtzahl von circa 30.000 Volltextartikeln und -berichten, alle kostenlos zugänglich, und ist damit die größte französischsprachige Webseite für die Verbreitung von Volltextartikeln in den Geistes- und Sozialwissenschaften. Während des Jahres 2005 wurden die Kandidaturen von 16 weiteren Zeitschriften von dem Leitungsausschuss akzeptiert, von fünf weiteren im März 2006, was die Gesamtzahl der Partnerzeitschriften auf 28 erhöht. Die Vorbereitung dieser neuen Zeitschriften ist jetzt im Gange, und die Online-Bereitstellung wird in den nächsten Monaten stattfinden. Für das Jahr 2006 sind noch sieben bis zehn weitere Titel geplant, so dass im Jahre 2007 mehr als 40 Zeitschriften online verfügbar sind.

PERSEE hat bis jetzt großen Erfolg beim Publikum erlebt. In knapp über einem Jahr wurden mehr als 20 Millionen Seiten heruntergeladen, circa 5.000 individuelle Benutzerkonten sind entstanden, der Durchschnitt liegt bei 200.000 Sitzungen pro Monat, der Höhepunkt im November 2005

lag bei mehr als 400.000 Verbindungen.²¹ Die Tatsache, dass PERSEE für die Benutzer/innen vollständig kostenfrei ist und vielseitige Dienste anbietet, und dass PERSEE auch gut von den nationalen und internationalen Suchmaschinen und Portalen (SUDOC, *Google Scholar*, OAIster...) nachgewiesen wird, kann diesen Erfolg sehr wahrscheinlich erklären, genauso wie die einfache, benutzerfreundliche und effiziente Schnittstelle. Die Partnerschaften mit anderen Portalen von Zeitschriften in den Geistes- und Sozialwissenschaften werden die Attraktivität von PERSEE wahrscheinlich weiter erhöhen. Seit 2003 hat sich nämlich die Landschaft der französischsprachigen Portale von Zeitschriften in den Geistes- und Sozialwissenschaften stark entwickelt. Zu „den schon alten“ Projekten, wie ERUDIT in Quebec und Revues.org²² in Frankreich sind öffentliche sowie private Initiativen gekommen.

Zu den privaten Initiativen gehört CAIRN²³, ein Portal, das aus der Zusammenarbeit von vier belgischen und französischen Verlagshäusern stammt, die auf Geistes- und Sozialwissenschaften spezialisiert sind (Belin, de Boeck, La Découverte, Erès). CAIRN entstand im September 2005 und bietet zurzeit etwa 60 Zeitschriften als kostenpflichtiges Abonnement an. Eine weitere bemerkenswerte Initiative ist diejenige der Verlagshäuser Armand Colin und Sedes, die ein eigenes Portal für ihre laufende Produktion aufgebaut haben.²⁴ Was den öffentlichen Sektor angeht, muss man außer Revues.org und PERSEE das vom CNRS geführte CENS (*Centre d'édition numérique scientifique*) erwähnen. Das CENS, das im Dezember 2005 eröffnet wurde, ist noch in der Anfangsphase, was das Onlinestellen der Zeitschriften betrifft.²⁵

Das französischsprachige Angebot im Bereich der Zeitschriften in den Geistes- und Sozialwissenschaften erscheint daher in mehrere Akteure aufgeteilt, die jeweils verschiedene wirtschaftliche Modelle verfolgen.

21 Seit der Redaktion dieses Beitrags in Februar 2006 ist die Zahl der Sitzungen auf bis zu 600.000 monatlich gestiegen (September 2006).

22 Vgl. <<http://www.revues.org>>.

23 Vgl. <<http://www.cairn.info>>.

24 Vgl. <<http://www.armand-colin.com/revues.php>>.

25 Vgl. <<http://www.cens.cnrs.fr/>>.

Außerdem drängt keine von diesen Portalen auf eine exklusive Verbreitung, das heißt, dass die Redaktionen der Zeitschriften die Wahl haben, auf mehr als einem einzigen Portal anwesend und zugänglich zu sein, insbesondere für ihre laufende Produktion, weil PERSEE zur Zeit das einzige Portal ist, das sich auf die retrospektive Sammlungen spezialisiert hat. Trotz dieser gesplitterten Lage haben sich diese Akteure in einem informellen Netz zusammengetan, in dem sie das Ziel verfolgen, über die technologische Interoperabilität ihrer Mittel und über die Verknüpfung zwischen ihren verschiedenen Angeboten nachzudenken, um einerseits die Kontinuität der auf den verschiedenen Portalen verbreiteten Sammlungen zu gewährleisten, andererseits es den Besuchern/innen zu ermöglichen, durch eine einzige Suche die Gesamtheit der verfügbaren Sammlungen auf diesen Portalen zu durchsuchen. Somit sollen die technologischen und dokumentarischen Lösungen, die vom Ministerium für PERSEE gefunden wurden, eine Verknüpfung zwischen den retrospektiven Sammlungen auf PERSEE und der laufenden Produktion der gleichen Zeitschrift auf einem anderen Portal erlauben.

Dem Konzept des freien Zugangs zu den retrospektiven Sammlungen der Zeitschriften – als Gegenleistung für die aus öffentlichen Mitteln finanzierte Digitalisierung – folgend, hat das Programm PERSEE eine wirkliche Umstrukturierung des Angebotes von Zeitschriften in den Geistes- und Sozialwissenschaften bewirkt; so konnte sich ein wirtschaftliches Modell etablieren, das auf einem Gleichgewicht zwischen einem freien und kostenlosen Zugang zu den retrospektiven Sammlungen (mit den von jeder Zeitschrift beschlossenen *Moving Walls*) und einem filterten und kostenpflichtigen Zugang zu den laufenden Veröffentlichungen der Zeitschriften beruht.

Dieses Portal, das den retrospektiven Sammlungen von Zeitschriften in den Geistes- und Sozialwissenschaften gewidmet ist, erweist sich als eine konkrete Antwort auf die Erwartungen der wissenschaftlichen Gemeinschaft. Als Frucht einer engen Kooperation mit den Forschern/innen, den Herausbergremien und den Fachleuten der Informationswelt, berücksichtigt PERSEE die derzeitigen Nutzungsmodelle, aber auch neue Bedürfnisse, um es der französischen Forschungsgemeinschaft zu erlau-

ben, von einer Verbreitung ihres wissenschaftlichen Kulturgutes angemessen zu profitieren. Von der *Direction de l'enseignement supérieur* und der *Direction de la Technologie* unterstützt, erlaubt PERSEE die Verbreitung eines außergewöhnlichen wissenschaftlichen Corpus im Netz, in einer Logik der Aufbewahrung, die den Wünschen der internationalen wissenschaftlichen Gemeinschaft entspricht.

Schlussfolgerung

In mehreren Aspekten spiegelt sich die gesplitterte Landschaft der Digitalisierung in den Hochschulbibliotheken wider, einem Bereich, in dem Frankreich anderen Ländern wie Deutschland oder den USA gegenüber noch im Rückstand ist. Lokale Initiativen nehmen zu, manchmal mit einer Logik der reinen Aufbewahrung, manchmal in Verbindung mit den Themen der „Open Archives“ und der elektronischen Dissertationen. In diesem Zusammenhang ist es die Rolle des Ministeriums, die Projekte zu begleiten, die ihm unter finanziellem wie unter wissenschaftlichem Gesichtspunkt vorgelegt werden, und diese verschiedenen Initiativen zu koordinieren, um ihnen die bestmögliche Sichtbarkeit zu gewährleisten. Das SUDOC-Portal, das in der *Agence bibliographique de l'enseignement supérieur* (ABES) untergebracht ist und durch das Ministerium finanziert wird, trägt dazu bei, diese Sichtbarkeit zu erhöhen: dieses Portal soll also das Inventar der digitalisierten Sammlungen der Hochschulen aufnehmen.

In einer internationalen Lage, die von den Projekten großer privater Akteure (*Google, Yahoo, Microsoft*) oder öffentlicher Einrichtungen (der europäischen Kommission – Programm i2010) geprägt wird, entwickeln die französischen Hochschulbibliotheken ein digitales Angebot, das auf einen unvergleichbaren kulturellen und wissenschaftlichen Korpus zurückgreifen kann. Alle diese Initiativen werden dazu beitragen, langfristig die linguistische und kulturelle Vielfalt der wissenschaftlichen Dokumentation im Internet aufrechtzuerhalten und zu erweitern.

Frédéric Blin ist im Ministère de l'Éducation nationale, de l'Enseignement supérieur et de la Recherche in der Direction générale de l'enseignement supérieur, Sous-direction des bibliothèques et de l'information scientifique, beschäftigt. Er ist für die Internationale Kooperation zuständig. E-Mail: frederic.blin@education.gouv.fr

Valérie Néouze ist ebenfalls dort beschäftigt und für den Bereich Digitalisierung zuständig. E-Mail: valerie.neouze@education.gouv.fr

KOPAL – EIN DIGITALES ARCHIV ZUR DAUERHAFTEN ERHALTUNG UNSERER KULTURELLEN ÜBERLIEFERUNG

von Thomas Wollschläger

Der Wandel des Publikationsmarktes hin zur elektronischen Publikation in Virtuellen Netzen verändert das Aufgabenprofil insbesondere der Bibliotheken, die für die Bewahrung, Sicherung und Überlieferung des kulturellen Erbes einer Nation mitverantwortlich sind, erheblich. Ziel und Zweck des durch das Bundesministerium für Bildung und Forschung (BMBF) geförderten Projekts kopal¹ ist die praktische Erprobung und Implementierung eines kooperativ erstellten und betriebenen Langzeitarchivierungssystems für digitale Dokumente und Daten als eine nachnutzbare Lösung für sowohl die Langzeiterhaltung als auch die Sicherstellung der Langzeitverfügbarkeit digitaler Ressourcen. Der Artikel dokumentiert die Schwerpunkte und den Projektverlauf des Vorhabens kopal zwischen 2004 und 2007.

Herausforderungen der digitalen Langzeitarchivierung

Der internationale und auch der deutsche Publikationsmarkt ist geprägt von einem weiterhin rasch wachsenden elektronischen Publikationssektor. Sowohl das wissenschaftliche Publizieren als auch allgemein gesellschaftliche und kulturelle Aktivitäten finden mit einem zunehmenden Anteil im Netz statt. Heute angewandte und machbare Selektionskriterien bei der Sammlung elektronischen Materials entscheiden über die Tradierung unseres kulturellen Erbes. Dieser „Nachlass“ ist – auch gesetzlich fixiert – eine Aufgabe von National- und Landesbibliotheken und vergleichbaren

¹ „Kooperativer Aufbau eines Langzeitarchivs digitaler Informationen“; Projekt-Website <<http://kopal.langzeitarchivierung.de/>> (27.02.2006).

Einrichtungen. Auch die aktuelle Novellierung des Gesetzes über Die Deutsche Bibliothek steht in diesem Kontext: Sie erweitert den Sammelauftrag der künftigen Nationalbibliothek² um nicht-trägergebundene digitale Materialien. Teile dieser Materialien, wie Online-Hochschulschriften, Online-Newsletter oder elektronische Zeitschriften einiger Verlage, sammelt Die Deutsche Bibliothek seit Jahren auf freiwilliger Basis. Der neue Auftrag bezieht nunmehr sowohl sämtliche Formen elektronischer Publikationen ein als auch einen enorm gewachsenen Umfang dieser Materialien (zum Beispiel im Rahmen von Web-Archiving).

Um dem erweiterten Sammelauftrag gerecht zu werden, benötigen die Bibliotheken und Gedächtnisorganisationen entsprechend geeignete und dimensionierte Archive für elektronische Materialien. Die bestehenden Archive erfüllen jedoch die Anforderungen an „vertrauenswürdige Archive“³ nicht oder nur teilweise. Im Fokus der Bemühungen, ein geeignetes Archivsystem für das elektronische Material zu finden, standen unter anderem folgende Überlegungen:

- Als Grundlage müssen die binären Daten erhalten werden, denn kein existierender Datenträger ist ewig oder auch nur langfristig genug haltbar. Das Archiv müsste also regelmäßige Erhaltungsmaßnahmen (Datenträgermigrationen) durchführen können.
- Der rasante Technologiewechsel erschwert immer wieder aufs Neue den Zugriff auf ältere Datenformate. So gibt es bereits zahlreiche Formate, zu denen ständig neue hinzukommen, während bisherige Formate verschwinden oder obsolet werden. Somit entstehen komplexe Abhängigkeiten von aktueller Soft- und Hardware. Das Archiv müsste daher in der Lage sein, sowohl Format-Migrationen (regelmäßiges Konvertieren) als auch Emulationen (Nachstellen benötigter Systeme) zu ermöglichen.

Bei der Migration werden alte Dateiformate rechtzeitig, das heißt solange

² Die Gesetzesnovellierung ist mittlerweile zum 29.06.2006 in Kraft getreten. Die Deutsche Bibliothek heißt nunmehr Deutsche Nationalbibliothek (DNB).

³ „Trusted Digital Repositories“. Siehe hierzu vor allem die Materialien der nestor-AG „Vertrauenswürdige Archive“, <<http://nestor.cms.hu-berlin.de/tiki/tiki-index.php?page=AG%20Vertrauensw%C3%BCrdige%20Archive>> (27.02.2006).

man die Formate noch vollkommen lesen kann, in aktuellere Formate konvertiert. Dies erfolgt periodisch und bedeutet eine Überführung aus bestimmten Hardware- und Software-Konfigurationen in eine neuere Umgebung, die aktuell technisch und organisatorisch beherrscht werden kann. Zweck der Migration ist es also, die Integrität und die Verfügbarkeit digitaler Ressourcen trotz des stetigen Wandels der technischen Umgebung zu erhalten. Problematisch bei diesem Verfahren ist jedoch der möglicherweise nicht erkannte Verlust von (Teil-)Informationen bei automatisch ablaufenden Migrationsroutinen; je komplexer das vorliegende Datenformat ist, desto größer wird die Wahrscheinlichkeit, dass bei längeren Migrationsketten (über Jahre hinweg) bestimmte Features nicht mehr vorhanden bzw. lauffähig sind. Demgegenüber ist die Migration ein geeignetes Verfahren für große Datenmengen, und es dürfte vor allem für eher statische Formate die vernünftigste Strategie sein.

Bei der Emulation versucht man mit einem speziellen Programm (dem Emulator) oder gegebenenfalls mit mehreren Programmen auf einer aktuellen Systemumgebung eine alte Systemumgebung nachzubilden. Zweck der Emulation ist es also, auf einem (aktuellen) System Daten und Programme zu verarbeiten, die ursprünglich für ein anderes, historisches System bestimmt waren. Dies kann sehr aufwändig sein und es setzt vor allem eine sehr genaue Definition der hard- und softwareseitigen Systemanforderungen voraus. Ein Vorteil der Emulation ist es jedoch, dass sie für sehr komplexe Formate (zum Beispiel Multimedia-Anwendungen) eine geeignete Lösung darstellt, um die Features dieser Formate auf lange Zeit nutzbar zu halten.

Beide Herangehensweisen haben also Vor- und Nachteile, weshalb ein Archivsystem in der Praxis zur Langzeiterhaltung und Langzeit-Verfügbarmachung digitaler Publikationen eine Kombination beider Verfahren einsetzen können müsste.

Grundprinzipien des kopal-Projekts

Das Projektziel von kopal ist daher der Aufbau einer von Gedächtnisorganisationen nachnutzbaren technischen und organisatorischen Infrastruktur zur Sicherung der Langzeitverfügbarkeit elektronischer Publikationen. Es

ist ein Förderprojekt des Bundesministeriums für Bildung und Forschung (BMBF) über eine Laufzeit von drei Jahren (bis Mitte 2007, Fördervolumen: 4,2 Millionen Euro). Den Kern des kopal-Archivsystems bildet das von IBM in Zusammenarbeit mit der Nationalbibliothek der Niederlande (Koninklijke Bibliotheek, KB) entwickelte Digital Information Archiving System (DIAS)⁴. Innerhalb des Projekts werden digitale Materialien aller Art der DDB und der Niedersächsischen Staats- und Universitätsbibliothek (SUB Göttingen) im Massenverfahren in das Langzeitarchiv eingestellt.

Im Wesentlichen beruht kopal dabei auf den drei Prinzipien *Kooperation*, *Universalität* und *Nachnutzbarkeit*. Im Rahmen einer nationalen Kooperation arbeiten im Projekt zwei sehr unterschiedliche Bibliotheken zusammen. Dies betrifft nicht nur den jeweiligen Auftrag (Nationalbibliothek bzw. Sondersammelgebietsbibliothek der DFG), sondern auch Schwerpunkte bei den bisherigen Aktivitäten (Online-Dissertationen und Netzpublikationen bei der DDB, Digitalisate und im naturwissenschaftlich-technischen Bereich gängige Datenformate bei der SUB Göttingen). Auf diese Weise werden sehr unterschiedliche Aspekte in das Projekt eingebracht, die die Erfahrungspalette spürbar erweitern und die Nachnutzungsmöglichkeiten für weitere Nutzer/innen nochmals deutlich erhöhen. Außerdem arbeiten die Projektpartner eng mit dem „Kompetenznetzwerk Langzeitarchivierung“ (nestor)⁵ zusammen. Der technische Betrieb des Langzeitarchivs ist beim Rechenzentrumspartner „Gesellschaft für wissenschaftliche Datenverarbeitung Göttingen“ (GWDG) angesiedelt. Der Entwicklungspartner IBM Deutschland GmbH ermöglicht eine professionelle Anpassung der Softwarekomponenten und bietet eine langfristig stabile Unterstützung.

Im internationalen Rahmen arbeiten die Projektpartner zum einen eng mit der Königlichen Bibliothek der Niederlande zusammen, die das originale DIAS-Archivsystem als In-House-Lösung betreibt. Zusammen mit ihr werden Anforderungen an künftige Weiterentwicklungen von DIAS sowie

4 Siehe <<http://www-5.ibm.com/nl/dias/index.html>> (27.02.2006).

5 Ausführliche Informationen unter: <<http://www.langzeitarchivierung.de/>> (27.02.2006).

Strategien eines „Preservation Planning“ entwickelt. Außerdem beteiligt sich kopal an der internationalen Entwicklung und dem Austausch von Metadaten zur Langzeitarchivierung und kooperiert mit verschiedenen internationalen Projekten zur Langzeitarchivierung.

Das Projekt erfüllt den Anspruch an ein universell nutzbares Archivsystem, indem zum einen die Sicherstellung der langfristigen Verfügbarkeit durch Migration und Emulation unterstützt wird. Zum anderen gibt es dabei in kopal keinerlei Einschränkungen sowohl für die Art des Materials, welches in das Archiv eingespielt werden kann (Text, Bilder, Audio, Video), als auch für die Bandbreite der möglichen Dateiformate (PDF, HTML, TIFF usw.). Wiewohl das kopal-System für den Projektzeitraum eine begrenzte Gesamtkapazität hat, ist die Größe des einzelnen Archivobjekts nicht begrenzt. Jeder der Partner ist zunächst völlig frei in der Auswahl und Regelfestlegung beim Einspielen der von ihm gesammelten Objekte. Gleichzeitig ist es möglich, unter Einbeziehung des betreibenden Partners GWDG Regeln für die gezielte Migration einzelner Objektgruppen zu definieren und sie auch über alle Bestände im Archiv umzusetzen.

Ein wesentlicher Aspekt des Systems ist schließlich die dezidierte Nachnutzbarkeit. Um diese zu gewährleisten, werden etablierte Standards genutzt. Der Transfer der Objekte in ein digitales Archiv über standardisierte Formate, Transportwege und Systemschnittstellen ist dabei ein wichtiges Erfordernis. Ohne Standardisierung ist die Bewältigung großer Mengen von Objekten nicht möglich und ein arbeitsteiliges Zusammenwirken kooperierender Archivsysteme wäre deutlich erschwert.

Das Projekt kopal hat dafür mit dem „Universellen Objektformat“ ein Austausch- und Archivformat vorgelegt und implementiert, mit dem digitale Objekte zusammen mit Metadaten archiviert und zwischen Institutionen und Archivsystemen ausgetauscht werden können. Es basiert auf den Formaten METS⁶ in der Version 1.4 und den Langzeitarchivierungsmetadaten für elektronische Ressourcen (LMER)⁷ in der Version 1.2. Ein

6 Metadata Encoding & Transmission Standard; <<http://www.loc.gov/standards/mets/>> (27.02.2006).

7 Informationen unter: <<http://www.ddb.de/standards/lmer/lmer.htm>> (27.02.2006).

Beispiel für die Metadatenstruktur eines Archivobjekts in kopal gibt die folgende Abbildung wieder:

```
<?xml version="1.0" encoding="utf-8" standalone="no" ?>
< mets OBJID="1" TYPE="Kopal Submission Information Package" PROFILE="DDB" xmlns="http://www.loc.gov/METS/" xmlns:xlink="http://www.w3.org/1999/xlink"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns:lmer="http://www.ddb.de/LMER/" xmlns:lmerProcess="http://www.ddb.de/LMERprocess"
  xmlns:lmerProcessXsd="http://www.ddb.de/LMERprocess" xsi:schemaLocation="http://www.loc.gov/METS/ http://www.loc.gov/standards/mets/mets.xsd
  http://www.ddb.de/LMERfile http://www.ddb.de/standards/lmer/lmer-file.xsd http://www.ddb.de/LMERObject http://www.ddb.de/standards/lmer/lmer-
  object.xsd http://www.ddb.de/LMERprocess http://www.ddb.de/standards/lmer/lmer-process.xsd" />
< metaHdr CREATEDATE="2006-02-16T10:26:25" RECORDSTATUS="PRODUCTION" />
< agent ROLE="ARCHIVIST" TYPE="ORGANIZATION" />
< name>Deutsche Bibliothek</name>
< note>Automatisch generierte Metadaten. Für weitere Informationen: kopal.langzeitarchivierung.de</note>
</agent>
< metaHdr />
< amdSec ID="AmdSec-0001" />
< techMD ID="TechMD-LMER-Object" />
< mdWrap ID="TechMD-LMER-Object-MdWrap" MDTYPE="text/xml" MDTYPE="OTHER" OTHERMDTYPE="lmerObject" LABEL="lmerObject" />
< mdData />
  < lmerObject:name>Ruckenbiel, Jan ____ Soziale Kontrolle im NS-Regime</lmerObject:name>
  < lmerObject:persistentIdentifier:urn:nbn:de:160206-970693761</lmerObject:persistentIdentifier>
  < lmerObject:objectVersion>1</lmerObject:objectVersion>
  < lmerObject:masterCreationDate>2006-02-16T10:26:24</lmerObject:masterCreationDate>
  < lmerObject:metadataCreationDate>2006-02-16T10:26:24</lmerObject:metadataCreationDate>
  < lmerObject:metadataRecordCreator>KOPAL DIAS</lmerObject:metadataRecordCreator>
  < lmerObject:numberOfFiles>3</lmerObject:numberOfFiles>
</mdData>
</mdWrap>
< techMD />
< techMD ID="TechMD-File-1" />
< techMD ID="TechMD-File-2" />
< techMD ID="TechMD-File-3" />
< amdSec />
< fileGrp ID="ASSET" AMDID="TechMD-LMER-Object" />
  < file ID="FILE0001" MDTYPE="application/pdf" SIZE="8742" CREATED="2006-02-16T10:26:26" CHECKSUM="7da2bb756d5d40f680de072dc6c7e74f1ef2ec"
    CHECKSUMTYPE="SHA-1" AMDID="TechMD-File-1" />
    < local LOCTYPE="URL" xlink:type="simple" xlink:href="file:///./abstract.pdf" />
  </file>
  < file ID="FILE0002" MDTYPE="application/pdf" SIZE="1629224" CREATED="2006-02-16T10:26:26"
    CHECKSUM="acc02ca3e6c74ff05535c567510959dfdb034" CHECKSUMTYPE="SHA-1" AMDID="TechMD-File-2" />
    < local LOCTYPE="URL" xlink:type="simple" xlink:href="file:///./41/970693761.pdf" />
  </file>
  < file ID="FILE0003" MDTYPE="application/octet-stream" SIZE="3187" CREATED="2006-02-16T10:26:26"
    CHECKSUM="fe58675eb95a34b6418550be373335de35f7b1" CHECKSUMTYPE="SHA-1" AMDID="TechMD-File-3" />
    < local LOCTYPE="URL" xlink:type="simple" xlink:href="file:///./meta.txt" />
  </file>
</fileGrp>
< fileSec />
< structMap TYPE="ASSET" />
  < div ORDER="1" LABEL="File list" TYPE="ASSET" />
</structMap>
< structMap TYPE="LOGICAL" />
  < div />
  < div ORDER="1" LABEL="Text" />
  < div ORDER="2" LABEL="Abstract" />
  < div ORDER="3" LABEL="Metadaten" />
</div>
</structMap>
</mets>
```

Abbildung 1: Beispiel für eine mets.xml-Datei, die Metadaten zum Archivobjekt im Universellen Objektformat abbildet. Es handelt sich in dem Fall um eine Online-Dissertation, bestehend aus den Dateien Hauptdokument, Abstract und Checksumme.

Zum anderen wird die zuverlässige Kern-Software durch flexible Open Source-Module erweitert. Das kopal-Archivsystem besteht aus dem weiterentwickelten DIAS der IBM und den von den Partnern SUB Göttingen und der DDB entwickelten kopal-Tools. Dabei beruht DIAS auf Standardsoftware-Komponenten der IBM und richtet sich in seiner Architektur

streng nach dem OAIS-Modell.⁸ Hierfür erstellen die DDB und die SUB Göttingen auf DIAS abgestimmte Softwareprodukte: die kopal-Tools, die als „kopal Library for Retrieval and Ingest“ (koLibRI) unter einer Open Source-Lizenz veröffentlicht werden. Die Systementwicklung wird dabei so offen angelegt, dass eine Ausdehnung der kooperativen Nutzung um weitere Archivbibliotheken sowie Nachnutzer aus dem Kreis aller „Gedächtnisorganisationen“ (also auch Archive und Museen) und wissenschaftlichen Datenarchive möglich ist.

Die Erweiterbarkeit von DIAS für neue Nutzerinstitutionen und präzise definierte Import- und Exportschnittstellen ermöglichen es, eine Archivnutzung in verschiedenste Workflows unterschiedlichster Institutionen zu integrieren. Durch den Einsatz von vielfach bewährter IBM-Standardsoftware, wie der DB2-Datenbank, dem Content Manager und dem Tivoli Storage-Manager, sind langfristige Stabilität, Performanz und Skalierbarkeit gewährleistet.

Die Deutsche Bibliothek und die Staats- und Universitätsbibliothek Göttingen entwickeln zurzeit ein Software-Paket für die Benutzung der kopal-Solution: die „kopal Library for Retrieval and Ingest“ (koLibRI). Bei diesen *kopal-Tools* geht es hauptsächlich um den Bereich des Einspiels von Objekten in das DIAS sowie um den Zugriff auf die archivierten Objekte (siehe Abbildung 2).

⁸ OAIS = Open Archival Information System (ISO Standard 14721). DIAS verwaltet – dem OAIS-Standard für Archivsysteme entsprechend – die gespeicherten Daten und stellt eine Umgebung für unterschiedlichste Erhaltungsstrategien zur Verfügung. Dabei wird das Objekt als Submission Information Package (SIP) verarbeitet und als Dissemination Information Package (DIP) geliefert (Terminologie gemäß dem OAIS-Modell). Innerhalb des DIAS wird ein Datenteil (Archival Information Package oder auch AIP) abgetrennt und gelangt in einen Storage (Festplatte, Magnetband usw.). Bestimmte Metadaten werden in einer Datenbank (Data Management) abgelegt, auf die über eine Administrationschnittstelle zugegriffen werden kann. Zum OAIS-Standard siehe <<http://ssdoo.gsfc.nasa.gov/nost/wwwclassic/documents/pdf/CCSDS-650.0-B-1.pdf>> (27.02.2006).

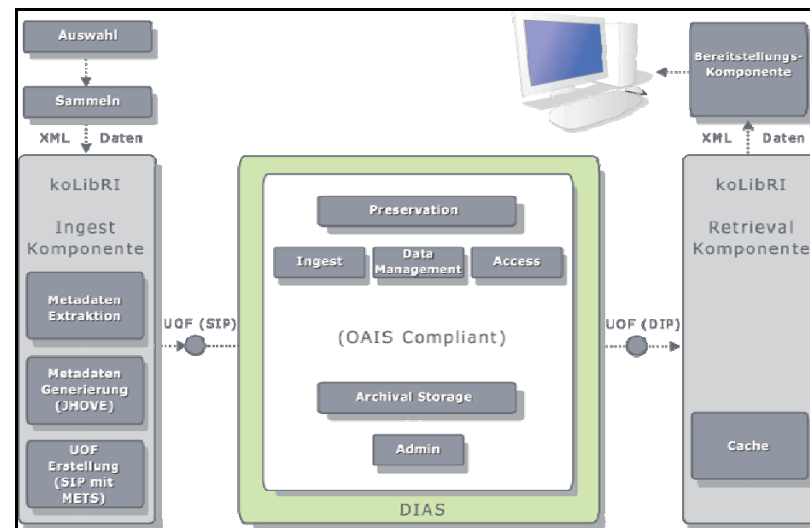


Abbildung 2: Der Aufbau des kopal-Archivsystems richtet sich nach dem OAIS-Modell. Die koLibRI-Software übernimmt die Erstellung der Archivobjekte und deren späteres Retrieval.

Die Schnittstelle für den Datenimport (Ingest) erfüllt folgende Anforderungen:

- Der Ingest kann automatisiert erfolgen.
- Das universelle Objektformat (kopal-UOF) wird unterstützt.
- Die flexible Schnittstelle ermöglicht eine Integration in verschiedenste Umgebungen und Informationssysteme.
- Es werden internationale Standards angewandt.
- Eine Nachnutzbarkeit durch Dritte ist gewährleistet.
- Eine grafische Oberfläche ist in Entwicklung.

Die Anforderungen an den Datenexport (Access) sind bei der DDB und der SUB Göttingen sehr unterschiedlich, so dass generische Module genutzt werden, die von den einzelnen Institutionen je nach Bedarf erweitert werden können. Mit Hilfe dieser Software soll künftig auch eine Administration des kopal-Systems realisiert werden.

Aufgrund der sehr unterschiedlichen und teils heterogenen Systemstrukturen der beiden Institutionen wird eine flexible Software benötigt. Das so genannte Workflow-Tool bietet eine gemeinsam nutzbare Infrastruktur für Module. Dieses Workflow-Tool kann als Asset-Builder zum Erstellen von Archivpaketen dienen. Weiterhin kann es als eine zentrale Vermittlungsinstanz zum DIAS-System, als Client-Loader, genutzt werden, indem es Archivpakete von mehreren Asset-Buildern sammelt und an DIAS übermittelt. Weitere Nutzungsmöglichkeiten lassen sich zudem problemlos integrieren.

Der bisherige Projektverlauf von kopal

Nach einer Pilotphase zur Systemevaluierung und Entwicklung des Universellen Objektformats befindet sich kopal derzeit inmitten der Entwicklungsphase, die auf verschiedene Arbeitspakete aufgeteilt worden ist. Die DIAS-Basisentwicklung und -Bereitstellung sowie die DIAS-Schnittstellen-Implementierung sind dabei bereits abgeschlossen worden. Die Kernkomponente DIAS wurde durch IBM an die Anforderungen der Partner angepasst. Verantwortlich für die Bereitstellung des erforderlichen Hardwareumfeldes in dieser Phase war die GWDG. Die Einrichtung der Mandantenfähigkeit hat zusätzlich einen flexiblen und skalierbaren Zugriff auf das System gesichert und der Fernzugriff lässt nun auch einen nicht ortsgebundenen Zugang eines Mandanten auf das System zu. Dadurch wurde auch eine zukünftige Aufnahme neuer Teilnehmer/innen ermöglicht.

Die Partner DDB und die SUB Göttingen haben auf der Basis des DIAS Software implementiert, die die Voraussetzung für ein Überführen bestehender und zukünftiger Bestände sowie für Zugriffe auf das Langzeitarchiv geschaffen hat. Diese äußeren Dienste sind insbesondere zur Nachnutzung durch Dritte gedacht und deshalb an internationalen Standards ausgerichtet.

Ende 2005 erfolgte eine Testphase zur Abnahme des bei der GWDG installierten Systems und für die Erprobung der von der DDB und der SUB Göttingen entwickelten Ingest- und Retrieval-Tools. Beide Bibliotheken haben Testobjekte aus ihren elektronischen Beständen in die Tests eingebracht. Die GWDG stellte die Testumgebung zur Verfügung und führte

Tests bezüglich betriebsüblicher Störungen und Systemausfälle durch. Die Testphase verlief erfolgreich und das System wurde schließlich durch das Projekt kopal abgenommen.

In der aktuellen Projektphase (Stand: Februar 2006) ist nun auch das Aufsetzen des Produktivsystems abgeschlossen worden. In den kommenden Monaten spielen DDB und die SUB Göttingen eine Auswahl ihrer vorhandenen Datenbestände ein. Ziel ist es, anhand einer Vielzahl von Objekten in verschiedenen Formaten die praxisnahe Nutzung der kopal-Lösung aufzuzeigen. Gleichzeitig werden bei Projektende die exemplarisch übernommenen Datenbestände den Grundstock der dauerhaften Archivierung der elektronischen Materialien der Bibliotheken bilden.

Einzuspielende Daten für kopal

Die beiden Bibliotheken in Frankfurt und Göttingen werden im Rahmen der exemplarischen Datenübernahme Bestände einspielen, die sowohl bisher auf freiwilliger Basis gesammelte Online-Publikationen umfassen als auch verschiedene digitalisierte Quellen, CD-Rom-Images und weitere Datenarten.

Für die DDB ist die Übernahme folgender Daten bzw. von Teilen folgender wissenschaftlich und kulturell bedeutsamer Datenbestände vorgesehen:

- Online-Dissertationen (Originär digitale Dokumente ab 1997); Anzahl: derzeit 41.000; Datenmenge: 300 Gigabyte
- Netzpublikationen (elektronische Zeitschriften, zum Beispiel SpringerLINK); Datenmenge: 300 Gigabyte
- CD-ROMs (Quellensammlungen, Multimedia-Anwendungen, Filme, Spiele, Datensammlungen (zum Teil wissenschaftliche Primärdaten), sonstige Publikationen); Anzahl: 25.000; Datenmenge: 14.300 Gigabyte
- Digitalisate der Herzog-August-Bibliothek Wolfenbüttel; Datenmenge: 1.500 Gigabyte
- Exilpresse Digital und Jüdische Zeitschriften in NS-Deutschland (Bestand der Deutschen Bibliothek, Frankfurt); Datenmenge: (vorerst) 150 Gigabyte

- Digitalisate des Deutschen Buch- und Schriftmuseums, Leipzig; Datenmenge: (vorerst) 10.000 Gigabyte
- Digitale Musikdaten (von Schellackplatten bis Notendigitalisaten, Bestand des Deutschen Musikarchivs, Berlin); Datenmenge: 544.000 Gigabyte

Bis auf die CD-Rom-Images und die Digitalen Musikdaten sollen diese Datenbestände im Projektzeitraum möglichst vollständig nach kopal übernommen werden. Dabei hat sich im bisherigen Projektverlauf gezeigt, dass die ursprünglich für kopal geplanten Kapazitäten trotz der damals sehr großzügigen Anlage des Systems selbst für die bis Ende 2007 vorhandenen Materialien nicht ausreichen werden. Einen Eindruck des Verhältnisses von Materialmengen, die eingespielt werden könnten, und den geplanten Kapazitäten gibt die Abbildung 3 wieder.

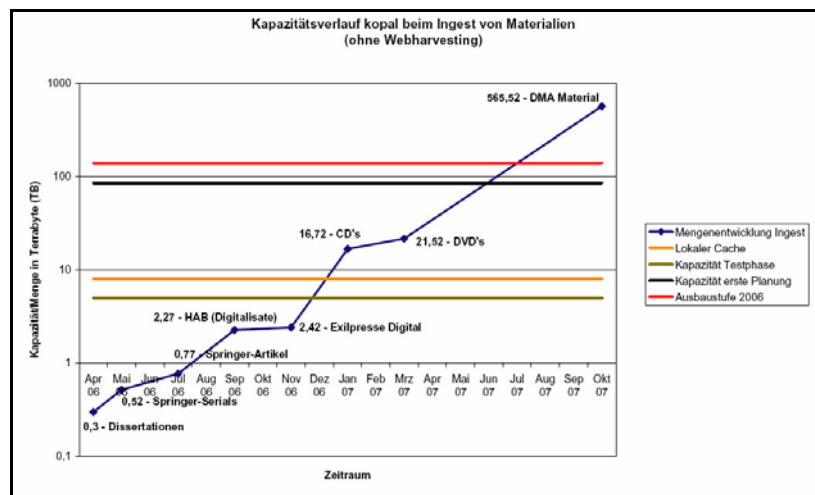


Abbildung 3: Geschätzter Mengenzuwachs bei einzuspielenden Materialien in kopal in Beziehung zu den bei Beginn des Projekts geplanten Kapazitätsgrenzen (logarithmische Darstellung)

Die Grenzen für den lokalen Cache spielen nur für die Optimierung der

Performanz des Systems beim Abruf von Daten eine Rolle und sind daher für das Einspielen der Daten kein Hindernis. Um die maximale Kapazität des bei der GWDG installierten Systems zu erhöhen und um so viele Materialien wie möglich übernehmen zu können, plant man derzeit, einen erheblichen Speicherzuwachs vorzunehmen. Dies ist umso wichtiger, da ein Bereich in der oben dargestellten Architektur noch nicht eingeschlossen ist, nämlich die Archivierung geharvesteter Websites.

Im Jahre 2005 hat die Australische Nationalbibliothek ein Domain-Komplettharvesting für die Domain .au (= Australien) durchgeführt. Dabei wurden rund 185 Millionen eindeutig unterscheidbare Dokument(seiten) von 811.000 Sites geharvestet, mit einem Gesamtvolumen von 6,69 Terabyte (in komprimierter Form zuzüglich Archivierungsmetadaten: 4,6 Terabyte).⁹ Dies zeigt, dass selbst für das einmalige Harvesten einer kleineren nationalen Domain enorme Datenmengen anfallen. Auch die DDB hat zur Erprobung von Harvesting-Strategien bereits Pilotmaßnahmen durchgeführt. Es werden derzeit Strategien entwickelt, welche den Umfang der zu harvestenden Domain(s) bzw. Sites für Deutschland, die Frequenz des Harvesting und die Frage von gezieltem oder generellem Harvesting adressieren.

Herausforderung: Preservation Planning und Service

Neben der exemplarischen Datenübernahme steht derzeit die Adressierung des Arbeitspakets Preservation Planning und Bitstream-Preservation. Die Partner IBM, DDB und SUB Göttingen entwickeln dabei unter Einbeziehung der KB (Niederlande) ein tragfähiges Konzept zum Preservation Planning. Dabei geht es um die Erkennung der digitalen Objekte, die bedingt durch den technologischen Wandel in der Zukunft nicht mehr zugänglich sein werden. Dazu gehören die Anbindung einer funktionierenden internationalen Format Registry, die performante Migration großer Datenmengen und die erfolgreiche Umsetzung von Emulationsmechanismen.

⁹ Bericht unter: <http://pandora.nla.gov.au/documents/domain_harvest_report_public.pdf> (27.02.2006).

Das entwickelte Konzept soll innerhalb der kopal-Solution umgesetzt werden. Eine Grundlage dafür ist die Bitstream-Preservation, das sichere und verlustfreie langfristige *Erhalten* von digitalen Objekten. Die dafür nötigen Speicherstrategien sowie die Migrations- und Emulationsstrategien, welche das eigentliche Preservation Planning im Sinne der langfristigen *Verfügbarmachung* ausmachen, werden von den Projektpartnern entwickelt.

Um eine sichere Speicherung und langfristige Verfügbarkeit digitaler Daten zu gewährleisten, muss die Archivierungslösung in bestehende Informationssysteme und Arbeitsabläufe integriert werden. Hierbei ist die Nachnutzung durch weitere Kulturerbeinstitutionen sowie durch sonstige Institutionen, die eine Langzeitarchivierung benötigen, ausdrücklich erwünscht.

Die kopal-Solution ist von vornherein auf unterschiedliche Bedürfnisse ausgerichtet. Kopal verfolgt dabei das Ziel, künftig verschiedene Nutzungsmodalitäten bereitzustellen: Es wird einerseits die Möglichkeit geboten werden als Mandant mit einem eigenen „Schließfach“ das bestehende System mit abgesichertem Speicherplatz und eigener Verwaltung der Daten zu nutzen. Diese Lösung ist besonders geeignet für Institutionen mit geringerem Archivgutaufkommen. Andererseits wird die Möglichkeit einer Nachnutzung der kopal-Solution durch eine eigene Installation des DIAS-Systems bestehen. Dies wird in der Regel zusammen mit einer Nutzung der koLibRI-Software erfolgen.

Perspektive

Bereits seit März 2006 steht für Testzwecke ein vorläufiger Release der neu erstellten Ingest- und Retrieval-Software (kopal-Tools bzw. koLibRI) in einer Beta-Version zur Verfügung. Auf der Website von kopal gibt es – neben weiteren Informationen zum Projekt, den Standards und Downloads von Dokumentationen – ebenfalls seit März 2006 den „kopal-Demonstrator“ – eine interaktive Animation, die die Grundlagen, Funktionalitäten und Module des Systems anschaulich darstellt.

Gemeinsam mit der Königlichen Bibliothek der Niederlande haben die kopal-Partner eine DIAS-User-Group ins Leben gerufen. Die beteiligten

Institutionen halten mehrmals im Jahr Arbeitstreffen ab, um Informationen und Erfahrungen auszutauschen, notwendige Anpassungen zu diskutieren und gemeinsame Strategien für künftige Nutzenanwendungen zu entwickeln. Außerdem steht dem Projekt ein Beirat mit Experten aus Wirtschaft, Wissenschaft und öffentlicher Verwaltung zur Seite, die den Gesamtfortschritt und Ablauf des Projektes beratend begleiten. Die Zusammensetzung des Beirats gewährleistet, dass die Ziele von kopal gut unterstützt und die strategische Zielsetzung erreicht werden kann.

Mit dem Projektabschluss im Juni 2007 werden ein nachnutzbares System und ein endgültiges Release der vollständig entwickelten koLibRI-Software bereitgestellt. Hierauf bauen die unterschiedlichen kopal-Dienste auf.

Dr. Thomas Wollschläger ist Wissenschaftlicher Mitarbeiter im Projekt kopal in der Abteilung Informationstechnik der Deutschen Bibliothek Frankfurt am Main. Innerhalb des Themenkomplexes Langzeitarchivierung elektronischer Daten und speziell des Projekts sind seine Aufgabenschwerpunkte die Vernetzung des Archivsystems mit den Workflows für elektronische Materialien innerhalb der Bibliothek, die Koordination der Fachabteilungen und externer Partner sowie die Unterstützung des Projektcontrollings. E-Mail: t.wollschlaeger@d-nb.de

DAS PROJEKT NESTOR – KOOPERATIVE STRUKTUREN
FÜR DIE ÜBERLIEFERUNGSBILDUNG – BEISPIELE AUS DEM
BESTAND MASCHINENLESBARER DATEN DES BUNDESARCHIVS

von Andrea Hänger und Karsten Huth

Der Aufsatz beschreibt den Beitrag des Bundesarchivs im Rahmen des Projekts nestor. Zunächst werden das Projekt und der Bestand maschinenlesbarer Daten des Bundesarchivs vorgestellt. Dann wird der aktuelle Stand der Forschung anhand von konkreten Beispielen dargestellt. Zur Sprache kommen Fragen zur Vertrauenswürdigkeit von Archiven, zur Authentizität von digitalen Objekten und zu Metadaten für die Langzeitarchivierung (PREMIS).

Einleitung

In seinem Artikel „Ensuring the Longevity of Digital Documents“ beginnt der Autor Jeff Rothenberg mit einer kurzen Geschichte, die in der Zukunft spielt. Seine Enkel finden im Jahr 2045 einen Brief und eine CD ihres verstorbenen Großvaters. Der Brief besagt, dass auf der CD der Weg zum Erbe des Großvaters beschrieben ist. Leider können seine Enkel die CD nicht verwenden, weil es schon seit dreißig Jahren keine CD-Laufwerke mehr gibt und die entsprechende Software, die das Testament lesbar machen könnte, schon längst nicht mehr verfügbar ist. Somit ist das Erbe für immer verloren.¹ Dieser Artikel von 1995 stammt aus einer Zeit, in der der technologische Fortschritt scheinbar grenzenlos schien. Dennoch markiert er auch den Anfangspunkt, an dem sich weite Kreise von Fachleuten der

¹ Rothenberg, Jeff, Ensuring the Longevity of Digital Documents, in: Scientific American 272 (1995), S. 24-29, hier S. 24.

Möglichkeit eines digitalen Desasters für die zukünftige Überlieferungsbildung bewusst wurden. Es folgten zehn Jahre an kooperativen Bemühungen, um Lösungen und Strategien für das Problem der digitalen Langzeitarchivierung zu finden. Dieser Artikel soll an Beispielen aus der Praxis zeigen, wie Kooperationen neue Entwicklungen auf dem Gebiet der historischen Überlieferungsbildung fördern können.

Die Langzeitarchivierung digitaler Objekte ist eine Herausforderung, der sich nicht nur die klassischen Gedächtnisorganisationen wie Archive, Bibliotheken und Museen stellen müssen, sondern auch die Produzenten und Dienstleister. Viele der wissenschaftlichen, kulturellen und administrativen Ressourcen werden heute ausschließlich digital produziert. Hinzu kommen die Ressourcen aus Digitalisierungsprojekten, die mit erheblichem Aufwand, zum Schutz der analogen Objekte, ins Leben gerufen worden sind. Das Problem der digitalen Langzeitarchivierung ist deshalb so komplex, weil man es mit einer Fülle von unterschiedlichen Objekten zu tun hat. Von reinem Text hin zu digitalisierten, hochauflösenden Bildern, hin zu Datenbanken und wissenschaftlichen Rohdaten, komplexen Webseiten, elektronischen Akten, E-Journals, Multimedia und Computerprogrammen. Hinzu kommen komplizierte rechtliche Bestimmungen des Urheberrechts. Aufgrund der Komplexität des Problems sind kooperative Strukturen bei der Lösungsfindung unabdingbar.

Die kooperative Struktur – nestor

Eine solche kooperative Struktur stellt das Network of Expertise in Long-Term Storage of Digital Resources (nestor), das „Kompetenznetzwerk Langzeitarchivierung und Langzeitverfügbarkeit Digitaler Ressourcen“, dar.² „Nestor wurde initiiert,

- um diejenigen zusammenzubringen, die sich hierzulande mit dem Problem der dauerhaften Erhaltung digitaler Objekte beschäftigen,
- um eine Informations- und Kommunikationsplattform als zentrale Anlaufstelle für alle Fragen der digitalen Langzeitarchivierung zu schaffen und

² Vgl. <<http://www.langzeitarchivierung.de>>.

- um einen nach außen sichtbaren Focus in Deutschland zu bilden, der Kooperationen vermittelt und als Einstiegspunkt für internationale Allianzen zur Verfügung steht.
- Langfristiges Ziel von nestor ist der nachhaltige Aufbau einer kooperativen Infrastruktur, in der vielfältige Fachkompetenzen zusammenwirken.³ Das Bundesarchiv ist seit 2005 ein Partner von nestor.

Maschinenlesbare Bestände des Bundesarchivs

Das Bundesarchiv beschäftigt sich seit vielen Jahren mit dem Thema der Archivierung elektronischer Unterlagen. Die ältesten archivierten Daten stammen aus dem Jahr 1970. Sie sind heute noch vollständig lesbar und interpretierbar. 35 Jahre sind für Archive/innen eine ausgesprochen kurze Zeitspanne. Für die Zeitrechnung der Informationstechnologie hingegen ist diese Zeit schon beachtlich. Rund 200 digitale Archivobjekte mit ungefähr 300.000 Dateien, die im Bundesarchiv verwahrt werden, sind in der Deutschen Demokratischen Republik in der Zeit von 1970 bis 1990 entstanden. Datenbanken aus den Bereichen Statistik, Wirtschaft, Landwirtschaft, Bildung, Strafvollzug und Arbeit konnten erhalten werden. Diese Daten werden regelmäßig von wissenschaftlichen Benutzern/innen, aber auch von Betroffenen genutzt, die mit Hilfe der Daten versuchen, Rechtsansprüche geltend zu machen.

Nach der Wiedervereinigung der beiden deutschen Staaten im Oktober 1990 wurden ostdeutsche Behörden, die nicht von Bundes- oder Länderbehörden übernommen wurden, entweder privatisiert oder aufgelöst. Viele staatliche Rechenzentren wurden geschlossen und in diesem Auflösungsprozess wurden zahlreiche Datenbestände systematisch zerstört oder in neue private Unternehmen verschoben. Das Bundesarchiv musste sofort eine Übernahmestrategie entwickeln und konnte so eine große Anzahl von historisch und rechtlich relevanten Unterlagen retten. Es war vor allem wichtig, nicht nur die Daten selbst, sondern auch die begleitenden Doku-

3 Dobratz, Susanne; Neuroth, Heike; Schoger, Astrid; Strathmann, Stefan, nestor – Entwicklungsstand des Kompetenznetzwerkes zur Langzeitarchivierung digitaler Ressourcen in Deutschland, in: Zeitschrift für Bibliothekswesen und Bibliographie Jahrgang 52 (2005), S. 151-162, hier S. 152.

mentationsunterlagen zu sichern. Nur mit diesen Unterlagen konnten die Daten interpretiert werden. Die Übernahme war und ist bis heute ein fortdauernder Prozess. Über die Jahre wurden hunderte von Magnetbändern auf Dachböden und in verlassenen Büros aufgefunden. Selbst heute werden noch Disketten in Papierakten im Zuge der Erschließung aufgefunden. Die jüngste Entdeckung sind die Daten eines Untersuchungsausschusses der letzten Volkskammer der DDR. Dieser Untersuchungsausschuss wurde eingesetzt, um Korruptionsvorwürfe gegen die Regierung aufzuklären. Die Daten waren auf 5.25 Zoll Disketten im Format Redabas gespeichert, der sozialistischen Version des Datenbankprogramms dBase.

Einer der wichtigsten Datenbestände ist der Zentrale Kaderdatenspeicher. Dieser enthält Daten, die in den 1980er Jahren im Auftrag des DDR-Ministerrates erfasst wurden. Es sind Informationen über mehr als 700.000 Personen einschließlich der funktionalen Elite und der Entscheidungsträger der DDR (mit Ausnahme der Vollzeit-Parteifunktionäre, der Militärs und der Bediensteten der Staatssicherheit). Diese einzigartige Quelle liefert vollständige Angaben zum sozialen und familiären Hintergrund der hochrangigen Parteimitglieder, zu ihrer Familiensituation in den achtziger Jahren, ihrem beruflichen Werdegang, ihrer Ausbildung, der Zugehörigkeit zu Partei- und Massenorganisationen, ihrer Stellung im Parteiapparat und weitere Informationen bezüglich ihrer Stellung als Kader (wie Fremdsprachenkenntnisse, Berechtigung für Reisen in das nicht-sozialistische Ausland) usw. Diese Daten werden – in anonymisierter Form – schon seit vielen Jahren von einem Sonderforschungsbereich an der Universität Jena ausgewertet, bei dem es um die Erstellung einer kollektiven Biografie der funktionalen Elite der DDR geht.

Andere Beispiele wichtiger Datenbestände sind die Volkszählungsdaten von 1971 und 1981 oder die Daten über die Eingaben der Bürger beim Staatsrat der DDR; welche die Petitionen von mehr als einer Million DDR-Bürgern aus der Zeit von 1979 bis 1989 enthalten.

Langzeitarchivierungsstrategien

Im Zuge der ersten Datenübernahmen musste auch eine Langzeitarchivierungsstrategie aufgestellt werden. Nach Margret Hedstrom bedeutet Lang-

zeiterhaltung, dass die Fähigkeit erhalten wird, digitale Informationen abzuspielen, wieder zu finden, zu verändern und zu benutzen angesichts sich beständig verändernder Technologien.⁴ Für archivische Zwecke geht es nicht nur um die Erhaltung für die kommenden Monate oder Jahre, sondern um die Langzeiterhaltung. Langzeit meint dabei, dass die Zeitspanne lang genug ist, um von den Auswirkungen sich ändernder Technologien, einschließlich neuer Medien oder Formate, oder von einer sich ändernden Nutzergruppe betroffen zu sein. Langzeit kann sich bis in die Unendlichkeit ausdehnen.

Im Allgemeinen gibt es zwei verschiedene Strategien für die Langzeiterhaltung originär digitaler Daten: Migration und Emulation. Migration kann im weitesten Sinne definiert werden als der Transfer digitaler Informationen von einer Hardware oder Software-Plattform zu einer anderen oder von einer Generation der digitalen Technologie zu der nächsten. Migration ist immer ein fortlaufender Prozess. Denn sie kann eigentlich niemals in einem permanent gültigen Datenformat resultieren. Stattdessen müssen Migrationen in regelmäßigen Abständen durchgeführt werden, um sicherzustellen, dass die Informationen interpretierbar bleiben. Emulation als zweite Möglichkeit versucht, die originäre Hard- oder Softwareumgebung zu erhalten. Emulatoren rekonstruieren den so genannten „look and feel“ des Originals, indem sie die Software und Hardware, mit der das Dokument entstanden ist, simulieren. Die Benutzer/innen können somit das Dokument in seiner ursprünglichen Form mit allen Funktionalitäten nutzen. Aus archivwissenschaftlicher Perspektive ist dies ohne Zweifel der beste Weg, weil er soviel vom Original beibehält wie möglich. Jedoch ist es auch der komplizierteste und teuerste Weg.

Die Wahl der Strategie hängt schließlich von dem Grad der Funktionalität, die erhalten werden soll, und von den Ressourcen, die ein Archiv in die digitale Langzeiterhaltung investieren kann, ab. Man kann heute feststellen, dass sich keine der beiden Strategien als die eindeutig bessere

⁴ Hedstrom, Margaret, Preserving Digital Information, in: Long Term Preservation of Electronic Materials. A JISC/British Library Workshop as part of the Electronic Libraries Programme (eLib), vgl. <<http://www.ukoln.ac.uk/services/papers/bl/rdr6238/paper.html#hedstrom>> (16.02.2006).

Lösung darstellt. Langzeiterhaltung ist und bleibt ein technisch komplexer und kostenintensiver Prozess. Das vorrangige Ziel sollte sein, die Daten für alle möglichen Fragen offen zu halten, die im Lauf der Zeit gestellt werden könnten. Langfristig können neue Benutzergruppen entstehen, deren Erwartungen sich deutlich von denen heutiger Nutzer/innen unterscheiden. Migration und Emulation bedeuten beide fortlaufende Verpflichtungen bei der Langzeiterhaltung – ein Archiv muss in der Lage sein, Ressourcen in regelmäßigen Abständen bereitzustellen, um die Benutzbarkeit der Daten zu sichern. Selbst wenn man die Fragen der logischen Interpretierbarkeit der Daten außer Acht lässt, sind die physischen Speichermedien dem Verfall ausgesetzt und müssen regelmäßig erneuert werden.

Migration im Bundesarchiv

Das Bundesarchiv hat sich bei den DDR-Daten für eine Migrationsstrategie entschieden. Als Langzeiterhaltungsformat wurde ASCII gewählt. ASCII ist der sicherste Weg, um Daten langfristig zu speichern, aber es wird auch lediglich der Bitstream ohne jede Information über die ursprünglichen Funktionalitäten bewahrt. Diese Information muss durch zusätzliches Dokumentationsmaterial gesichert werden. In vielen Fällen mussten bei den DDR-Daten die Datenstrukturen rekonstruiert werden. Viele Recherchen waren erforderlich, um Datensatzstrukturen und Kodierungen wieder herzustellen. Um Speicherplatz zu sparen, der in den 1970er und 1980er Jahren sehr teuer war, wurden die Informationen so weit es ging kodiert, gepackt oder komprimiert. Die Konvertierung in ein langzeitarchivierungsfähiges Format konnte daher nicht mit marktüblichen Konvertierungsprogrammen durchgeführt werden. Jeder Datenbestand erforderte die Programmierung spezieller Konvertierungsprogramme. In einigen Fällen mussten die Programmierer/innen ausfindig gemacht werden, welche die Komprimierungsalgorithmen geschrieben hatten, weil sie die einzigen waren, die die Komprimierungen auflösen konnten.

2003 hat das Bundesarchiv zusammen mit der Universität Koblenz ein Programm entwickelt, das es erlaubt, den Großteil der Datenbestände von EBCDIC nach ASCII oder auch XML zu konvertieren. Das Programm löst die gepackten und komprimierten Daten auf und konvertiert hexadezimale

in binäre Werte. Mit diesem Programm können die einzelnen Feldinhalte analysiert werden, so dass Datensatzstrukturen und Codierungen validiert werden können. Die wichtigste Funktion des Programms besteht darin, dass Daten, Datensatzstruktur und Kodierungen untrennbar miteinander verbunden werden können. Diese Funktion kann den Nachteil von ASCII, das lediglich die Speicherung des Bitstreams erlaubt, ausgleichen.

Vertrauenswürdigkeit, Authentizität – Metadaten

Die Bearbeitung der Daten wurde bisher in speziellen Findmitteln dokumentiert. Dazu gehören auch Informationen über Fehler oder mögliche Vorbehalte gegenüber der Zuverlässigkeit der Daten ebenso wie die Informationen über die Original Soft- und Hardware, die technische Bearbeitung usw. Im Moment erprobt das Bundesarchiv im Rahmen von nestor die Umsetzung dieser Informationen in internationale Metadatenstandards.

Die Dokumentation ist zwingend notwendig, um die Authentizität der Quellen langfristig zu wahren. Die Frage, wie die Authentizität elektronischer Unterlagen erhalten und nachgewiesen werden kann, ist eine der Kernfragen nicht nur für die Archivare/innen, sondern für die gesamte Informationsgesellschaft. Grundlegend ist somit der Begriff der Authentizität. Was macht ein Objekt zu einem authentischen Objekt? Cullen bemerkt hierzu: „An authentic object is one whose integrity is intact – one that is and can be proven or accepted to be what its owners say it is. It matters little whether the object is handwritten, printed, or in digital form.“⁵

Es gibt vielseitige Möglichkeiten, um die Authentizität eines analogen Objekts nachzuweisen. Am bekanntesten sind sicher die chemischen Analysen des Papiers oder der verwendeten Tinte. Man kann aus der chemischen Analyse das Herstellungsverfahren des Papiers oder der Tinte ableiten. Diese Herstellungsverfahren lassen sich einem Zeitraum zuordnen. Stimmt dieser nicht mit dem auf dem Dokument angegebenen Datum

5 Cullen, Charles T., Authentication of Digital Objects. Lessons from a Historian's Research, in: Authenticity in a Digital Environment, Washington, D.C. 2000, S. 1-7, hier S. 1, vgl. <<http://www.clir.org/pubs/abstract/pub92abst.html>> (16.2.2006).

überein, so sind Zweifel an der Authentizität des Dokuments angebracht. Dies ist nur eine Möglichkeit einer Authentizitätsprüfung eines analogen Objekts. Dass Information und Informationsträger, in den meisten Fällen Papier, eine untrennbare Verbindung eingegangen sind, ist hierbei ein großer Vorteil. Papier, Tinte und Schriftbild sowie der Allgemeinzustand sind Teil des intrinsischen Wertes eines Objekts und somit mit ihm verbunden.

Auch digitale Objekte müssen auf einem Datenträger, einem Medium, gespeichert werden. Im Gegensatz zum analogen Objekt ist die Wahl des Datenträgers jedoch unabhängig von der Art des Objekts. Während in der analogen Welt das Papier beinahe immer die preisgünstigste und einfachste Alternative darstellt, kann man in der digitalen Welt jede technisch mögliche Lösung auch scheinbar problemlos anwenden. Dieser Umstand und die zusätzliche Möglichkeit zur schnellen Datenübertragung macht die Bindung der Information an ihren Träger extrem flüchtig. Auch andere Kriterien zur Authentizität aus der analogen Welt bekommen das digitale Objekt nicht in den Griff. Deshalb muss man zunächst das „Wesen“ eines digitalen Objekts verstehen, bevor man sich auf die Suche nach seiner Authentizität macht. Die besondere Eigenschaft eines digitalen Objektes ist seine „Existenz“ auf drei Ebenen. Jedes digitale Objekt kann gleichzeitig als ein physisches, logisches und konzeptuelles Objekt betrachtet werden.⁶

- *physisches Objekt*: Zeichen, die auf einem Informationsträger eingeschrieben sind. Die Art und Weise der physischen Beschaffenheit dieser Zeichen kann aufgrund der unterschiedlichen Beschaffenheit des Trägers sehr unterschiedlich sein. Auf einer CD-ROM sind es die sogenannten „pits“ und „lands“ auf der Trägeroberfläche, bei magnetischen Datenträgern sind es Übergänge zwischen magnetisierten und nicht magnetisierten Teilchen (die sogenannten Flusswechsel). Auf der physischen Ebene haben die Bits keinerlei Bedeutung, das heißt, Bits die zu einem Text gehören unterscheiden sich in keinsten Weise von

6 Thibodeau, Kenneth, Overview of Technological Approaches to Digital Preservation and Challenges in Coming Years, in: The State of Digital Preservation. An International Perspective, Washington D.C. 2002, S. 4-31, hier S. 6-10, vgl. <<http://www.clir.org/pubs/abstract/pub107abst.html>> (16.02.2006).

Bits, die Teil eines Computerprogramms sind.

- *logisches Objekt*: Bits, die vom Informationsträger gelesen werden und nun von einer Software interpretiert werden. In dieser Ebene existieren die Objekte als Bitstrom, der als Dateiformat erkannt wird.
- *konzeptuelles Objekt*: Zeichen und Objekte, die vom Menschen wahrgenommen und sinnvoll interpretiert werden können. Das konzeptuelle Objekt ist die eigentliche, für die Betrachter/innen bedeutungsvolle Einheit (zum Beispiel ein Text, Musikstück, Film, Computerprogramm).

Ausgehend von dieser Grundeigenart, die für alle digitalen Objekte gilt, sind Verfahren, die die Authentizität eines digitalen Objekts anhand einer Autopsie überprüfen wollen, ungeeignet. Der Datenträger kann bereits tausendfach gewechselt haben. Selbst innerhalb eines Archivs wird das digitale Objekt den Träger noch mehrfach wechseln müssen. Seine Prüfung führt deshalb zu keinem Ergebnis. Auch das logische Objekt kann sich im Rahmen einer Datenmigration ohne eine augenscheinliche Veränderung des konzeptuellen Objekts geändert haben. Es kann aber auch der Fall sein, dass eben jene Datenmigration auch das konzeptuelle Objekt gravierend verändert hat, so dass beispielsweise bei einem E-Journal bestimmte Artikel nicht mehr lesbar sind. Wenn die Kriterien für die Authentizität eines Objekts nicht mehr am Objekt selbst abprüfbar sind, wo dann?

Da ein digitales Objekt ständig kopiert und verändert werden kann, müssen in einem Archiv die Veränderungen festgehalten und dokumentiert werden. Wesentliche Fragen zur Authentizität eines digitalen Objekts sind zum Beispiel

- Wer hat wann in welcher Art und Weise das Objekt verändert?
- In welchen Dateiformaten lag und liegt das Objekt vor?
- Sind die Eigenschaften des Objekts verändert worden?
- Aus welcher technischen Umgebung stammt das Objekt ursprünglich?
- Ist das Objekt manipuliert worden?

Diese Daten über digitale Objekte nennt man Metadaten. Sie enthalten Informationen über Informationsobjekte, ähnlich wie bibliografische Beschreibungen, die Informationen über Bücher enthalten. Geordnet werden Metadaten im Allgemeinen in einem Metadatenschema. Es gibt

heute viele Metadatenschemata, die für unterschiedliche Zwecke geschaffen wurden. Das Metadatenschema Dublin Core beschreibt größtenteils das konzeptuelle Objekt, besitzt aber auch Elemente, die das physische und das logische Objekt mit einbeziehen. Das Schema wurde zu einem besonders großen Anteil von Bibliotheken entwickelt, um Informationen zu finden, zu verwalten und unter Partnern zu teilen. Im Archivbereich ist das Encoded Archival Description (EAD) Metadatenschema zur Erstellung von Findmitteln sehr gebräuchlich. Bislang fehlte aber noch ein Metadatenschema, das speziell auf die Bedürfnisse der digitalen Langzeitarchivierung ausgerichtet ist. An diese Stelle tritt nun das Metadatenschema PREMIS.

PREMIS

Das Akronym steht für Preservation Metadata Implementation. PREMIS wurde 2003 mit Unterstützung der Research Libraries Group (RLG) und des Online Computer Library Center (OCLC) gegründet. Die Initiative ist ein internationaler Verbund von Archiven, Bibliotheken, Museen und Vertretern aus der privaten Wirtschaft. Das PREMIS Metadatenschema ist das Arbeitsergebnis einer kooperativen Struktur. Der Abschlussbericht der Arbeitsgruppe wurde im Mai 2005 veröffentlicht. Die Ziele der Gruppe waren:

- Die Definierung eines implementierbaren Kernsatzes von Metadaten zur digitalen Langzeitarchivierung;
- Die Erstellung eines Data Dictionaries zur Erklärung des Metadatenschemas;
- Die Evaluierung alternativer Strategien zur Formatwahl, Speicherung und Verwaltung von Langzeitarchivierungsmetadaten innerhalb eines digitalen Archivs und zum Austausch zwischen verschiedenen Archivsystemen;
- Die Organisation von Pilot Programmen zum Test der Empfehlungen unter unterschiedlichen Rahmenbedingungen;
- Die Erforschung von Möglichkeiten zum Erstellen und Teilen von Langzeitarchivierungsmetadaten.
- Die Arbeit am PREMIS Metadatenschema ist noch nicht abgeschlossen.

sen. Seit der Veröffentlichung durchläuft es eine geplante achtzehnmonatige Test- und Implementierungsphase.⁷

PREMIS ist ein geeignetes Werkzeug, um die Authentizität eines digitalen Objekts zu überprüfen und über einen langen Zeitraum auch zu erhalten. Die gespeicherten Metadaten geben Auskunft über das archivierte digitale Objekt. Natürlich können die Informationen, die in einem Metadatenschema gespeichert und geordnet werden, auch in anderer Form dargestellt werden. Bislang wurden sie im Bundesarchiv in den Dossiers in natürlicher Sprache auf Papier ausgedruckt, sie liegen aber auch als Text in elektronischer Form vor. Das folgende Beispiel beschreibt den Verlauf einer typischen Migration einer Datentabelle aus dem Bestand der DDR Volkszählungsdaten der Jahre 1971 und 1981. Der Text gibt Auskunft auf die Frage: „Wer hat wann in welcher Art und Weise das Objekt Datei 0009_A9881.B7101.N11 verändert?“

Die Datei mit dem Namen 0009_A9881.B7101.N11 wurde am 20.10.2005 um 15:00 Uhr mitteleuropäischer Zeit vom Programm „Konverter.exe“ konvertiert. Diese Aktion ist nach Wunsch verlaufen. Die Datei 0009_A9881.B7101.N11 liegt im EBCDIC Format vor. Nach der Migration sind zwei neue Dateien mit dem gleichen Informationsinhalt entstanden. Die beiden neuen Dateien heißen 0009_A9881.B7101.N11_1_0.fed und 0009_A9881.B7101.N11_2_0.fed und liegen im ASCII.csv Format vor.

Ein Vorteil des Dossiers ist seine leichte Lesbarkeit für das menschliche Auge. Der große Nachteil indes liegt in der schlechten maschinellen Verwertbarkeit. Einen reinen Text kann man nicht strukturiert in eine relationale Datenbank überführen. Zudem ist die Semantik der einzelnen Begriffe nicht maschinell erschlossen. Hier setzt nun ein Metadatenschema an. Es ist zwar für den Menschen auf den ersten Blick viel schwerer lesbar, in ihm wird die Information jedoch in immer gleicher Art und Weise strukturiert und semantisch erschlossen. Dadurch lassen sich die Informationen ma-

⁷ Data Dictionary for Preservation Metadata. Final Report of the PREMIS Working Group, Dublin, Ohio 2005 hier S. vii, vgl. <<http://www.oclc.org/research/projects/pmwg/premis-final.pdf>> (16.02.2006).

schinell verarbeiten. PREMIS nutzt die Auszeichnungssprache XML als Format zur Speicherung und Vermittlung der Metadaten. Die Strukturen und Datentypen von PREMIS wurden innerhalb von XML-Schema Dateien definiert, welche zur freien Verfügung bereitstehen. Diese Schema-Dateien bilden die Grundlage für die Tests am Bundesarchiv. Sie werden zur Zeit sukzessive an die Bedürfnisse eines Archivs angepasst.

Der obige Text verteilt sich auf die folgenden PREMIS-Elemente:

```
<eventType>Migration</eventType>
<eventDateTime>2005-10-20T15:00:00
+01:00</eventDateTime>
<eventDetail>EBCDIC Datei wird in zwei ASCII csv
Dateien migriert. Dateiendungen aus technischen
Gründen fed.</eventDetail>
<eventOutcomeInformation>
<eventOutcome>OK</eventOutcome>
</eventOutcomeInformation>
```

Die Elemente sind bereits im XML-Format dargestellt. Die Elementnamen sind in spitze Klammern gesetzt. Die Metadaten werden durch zwei so genannte Tags eingerahmt, die jeweils den entsprechenden Elementnamen beinhalten, wobei der zweite Tag (End-Tag) mit einem Schrägstrich beginnt. Dem Element mit dem Namen „eventType“ (Ereignis Typ) wurde hier der Begriff Migration zugewiesen. Das Element eventDateTime gibt den genauen Zeitpunkt des Ereignisses wieder. Unter „eventDetail“ wird der Vorgang näher beschrieben und das Element „eventOutcome“ sagt uns, dass die Migration technisch korrekt verlaufen ist. Somit wurde dokumentiert, welche Veränderung am digitalen Objekt vorgenommen wurde, wann die Veränderung durchgeführt wurde und ob die Prozedur nach Plan verlaufen ist.

Die folgenden Elemente verweisen auf das Programm, mit dem die Migration durchgeführt wurde, sowie auf seine Rolle innerhalb des Migrationvorgangs. Das Programm diente als Konvertierungsprogramm, das heißt, es war verantwortlich für die Übersetzung der Ursprungsdatei in ein anderes Dateiformat.

```
<linkingAgentIdentifier>
<linkingAgentIdentifierType>Barch_ID_B2_Agent-
```

```

    nameschlüssel </linkingAgentIdentifizierType>
<linkingAgentIdentifizierValue>ans_konverterEXE
  </linkingAgentIdentifizierValue>
<linkingAgentRole>Konvertierer</linkingAgentRole>
</linkingAgentIdentifizier>

```

Die folgenden Zeilen beschreiben das Konvertierungsprogramm näher. Sie enthalten die Information, dass es sich um ein Programm mit dem Namen Konverter.exe handelt.

```

<agentIdentifizier>
<agentIdentifizier-
  Type>Barch_ID_B2_Agentnameschlüssel</
  agentIdentifizierType>
<agentIdentifizierValue>ans_konverterEXE</agent
  IdentifizierValue>
</agentIdentifizier>
<agentName>Konverter.exe</agentName>
<agentType>Programm</agentType>

```

Alle bisher genannten Informationen werden über Schlüssel (linkingObjectIdentifizier) mit den Informationen über die von der Migration betroffenen Dateien verknüpft. In diesem Beispiel sind drei Dateien von der Migration betroffen. Aus der Quelldatei fns_0009_A9881.B7101.N11 entstehen die Dateien 0009_A9881.B7101.N11_1_0.fed und 0009_A9881.B7101.N11_2_0.fed im ASCII/CSV Format.

```

<linkingObjectIdentifizier>
<linkingObjectIdentifizierType>Barch_ID_B2_file-
  nameschlüssel</linkingObjectIdentifizierType>
<linkingObjectIdentifizierValue>fns_0009_A9881.
  B7101.N11</linkingObjectIdentifizierValue>
</linkingObjectIdentifizier>
<linkingObjectIdentifizier>
<linkingObjectIdentifizier-
  Type>Barch_ID_B2_filenameschlüssel</linking
  ObjectIdentifizierType>
<linkingObjectIdentifizierValue>0009_A9881.B7101.
  N11_1_0.fed</linkingObjectIdentifizierValue>
</linkingObjectIdentifizier>
<linkingObjectIdentifizier>
<linkingObjectIdentifizierType>Barch_ID_B2_file-

```

```

    nameschlüssel</linkingObjectIdentifizierType>
<linkingObjectIdentifizierValue>0009_A9881.B7101.
  N11_2_0.fed</linkingObjectIdentifizierValue>
</linkingObjectIdentifizier>

```

Bei den Metadaten zu den jeweiligen Dateien findet man Informationen über die Dateigröße in Kilo Byte und das entsprechende Dateiformat. Das Element „formatRegistry“ gibt einem die Möglichkeit, auf ein Formatverzeichnis außerhalb des Archivs zu verweisen. Innerhalb eines solchen Verzeichnisses werden die unterschiedlichen Dateiformate technisch ausführlich beschrieben und ihre Datenstruktur wird definiert. Zur Zeit sind einige Formatverzeichnisse in Planung. In Harvard wurde gerade ein Projekt begonnen, aus dem einmal ein globales Verzeichnis entstehen soll.⁸ Das Beispiel verweist auf das Verzeichnis PRONOM⁹ des britischen Nationalarchivs. Hinter dem Schlüssel „id 45“ verbirgt sich die technische Beschreibung des Dateiformats ASCII/CSV.

```

<size>72471552</size>
<format>
<formatDesignation>
<formatName>ascii/csv</formatName>
<formatVersion>1.0</formatVersion>
</formatDesignation>
<formatRegistry>
<formatRegistryName>PRONOM </formatRegistryName>
<formatRegistryKey>id 45</formatRegistryKey>
</formatRegistry>
</format>

```

Veränderungen und Manipulationen an Dateien können über Hash-Werte erkannt werden. Dabei wird eine spezielle Hash-Funktion über dem Bitstrom einer Datei ausgeführt. Dadurch wird ein typischer Hash-Wert erzeugt. Dieser wird üblicherweise durch eine Zeichenkette im Hexadezimalcode dargestellt, der als digitaler Fingerabdruck verwendet werden kann. Bereits geringfügigste Veränderungen an der Datei durch Manipula-

8 Global Digital Format Registry (GDFR), vgl. <<http://hul.harvard.edu/gdfr/about.html>> (16.2.2006).

9 The technical registry PRONOM, vgl. <<http://www.nationalarchives.gov.uk/pronom/>> (16.02.2006).

tion oder fehlerhafte Datenübertragung führen zu einem völlig veränderten Hash-Wert. Da es verschiedene Hash-Funktionen gibt, muss der Name der Funktion ebenso bei den Metadaten gespeichert werden, wie der entsprechende Hash-Wert. PREMIS hat die entsprechenden Elemente dafür. Die Hash-Funktion des Beispiels hat den Namen MD5. Der entstandene Wert hat im Hexadezimalcode den Wert „7c9b35da4f2ebd436f“. Die entsprechende Integritätsprüfung wurde vor der Abgabe der Datei an das Bundesarchiv im Statistischen Bundesamt durchgeführt.

```
<fixity>
<messageDigestAlgorithm>MD5</messageDigest-
  Algorithm>
<messageDigest>7c9b35da4f2ebd436f</messageDigest>
<messageDigestOriginator>StatistischesBundesamt
  </messageDigestOriginator>
</fixity>
```

Die vorgestellten Elemente sind nur ein kleiner Ausschnitt aus dem Metadatenschema PREMIS. Dennoch zeigen sie, wie die Authentizität von digitalen Objekten innerhalb eines Archivs überprüfbar bleibt, obwohl das Objekt gravierende Veränderungen erfährt. Festzuhalten ist:

- Wir können nachvollziehen, inwiefern ein Objekt mit der Zeit verändert wurde und welche Konsequenzen dies für seinen Informationswert hat.
- Wir können nachvollziehen, wann und warum diese Änderungen vorgenommen wurden und wer dafür verantwortlich ist.
- Wir können durch spezielle Funktionen überprüfen, ob die Objekte ungewollt verändert wurden.
- Wir können die digitalen Objekte technisch umfassend beschreiben.

Bisheriger Umgang mit Authentizität

Authentizität ist kein Selbstzweck. Wie bereits erwähnt, werden die Datenbestände des Bundesarchivs häufig zu Nachweiszwecken benutzt. Bürger/innen der ehemaligen DDR nutzen die Daten, um ihre Ansprüche auf Wiedergutmachung, Entschädigungen, Rentenzahlungen durchzusetzen. Einer der wichtigsten Datenbestände ist zum Beispiel der Datenbe-

stand „Datenspeicher gesellschaftliches Arbeitsvermögen“, der Informationen zu einem Großteil der Beschäftigten in der ehemaligen DDR enthält, einschließlich der Informationen über Ausbildung und Beschäftigungsverhältnisse von mehr als 7,25 Millionen Bürger/innen. Diese Quelle dient oft als Nachweis von Beschäftigungsnachweisen für Rentenansprüche. Die größte Anzahl von Anfragen bezieht sich auf den Datenbestand „Strafgefangenen- und Verhaftetendaten“. Diese Daten wurden im Auftrag des Innenministeriums der DDR erhoben. Sie enthalten Informationen über die familiäre Situation, die Haftzeit, die Straftat und über alle Ereignisse, die sich während der Haftzeit zugetragen haben, unter anderem über die Krankheiten.

Besonders die Nachweise über während der Haft erlittene Krankheiten oder Unfälle können ehemaligen Häftlingen heute dabei helfen, Entschädigungen zu bekommen. Diese Informationen gibt es nur in elektronischer Form und es ist die einzige Quelle, auf die sich Betroffene berufen können, wenn sie eine Entschädigung beantragen. In erster Instanz richten sie ihre Forderungen an die lokalen Behörden, die über den Antrag entscheiden. Wenn diese den Antrag ablehnen, besteht die Möglichkeit, den Anspruch vor Gericht einzuklagen, was häufig vorkommt. Wie gezeigt, entsprach die Übernahme der Daten in das Bundesarchiv in den 1990er Jahren keineswegs den Regeln eines ordnungsgemäßen, sicheren Datentransfers. Aber keine Behörde und kein Richter, keine Richterin hat bis heute die Authentizität der vorgelegten Daten angezweifelt. Scheinbar hat die sorgfältige Dokumentation der Übernahme und Bearbeitung der Daten Behörden und Richter/innen zufrieden gestellt.

Fazit und Ausblicke

Diese Erfahrung bestätigt die Politik des Bundesarchivs, anstelle von kurz- oder mittelfristigen Lösungen für die technische Sicherung der Authentizität, wie zum Beispiel durch elektronische Signaturen, auf organisatorische Lösungen für die Langzeitsicherung zu setzen. Das Bundesarchiv arbeitet aktiv bei der nestor Arbeitsgruppe „Vertrauenswürdige Archive – Zertifizierung“ mit. Hierbei handelt es sich um ein interdisziplinäres Projekt, das nach dem Vorbild der Research Library Group Kriterien für eine Zertifizie-

rung vertrauenswürdiger Archive entwickelt. Das Ziel des Projektes besteht darin, Anforderungen für die vertrauenswürdige und damit authentische Speicherung, Bearbeitung und Nutzbarmachung elektronischen Archivguts aufzustellen. Diese Kriterien kann keine Institution allein aufstellen, dazu bedarf es einer breiten Kooperation, wie sie mit nestor gegeben ist.

Zusammenfassend können wir sagen, dass die Wahrung der Authentizität von digitalen Objekten grundlegend für eine zukünftige Überlieferungsbildung ist. Vertrauenswürdige Archive, die in der Lage sind, digitale Objekte authentisch zu speichern, werden in Zukunft an Bedeutsamkeit zunehmen. Kooperative Strukturen wie nestor, die Kompetenzen und Material aus verschiedenen Bereichen zusammenbringen, sind wichtig zur Erforschung und Diskussion von neuen Kriterien und Techniken in digitalen Archiven. Sie tragen somit ihren Teil zu einer gesicherten geschichtlichen Überlieferung bei.

Dr. Andrea Hänger arbeitet seit 2000 am Bundesarchiv in Koblenz. Dort leitet sie das Referat für elektronische Archivierung. Sie arbeitet in zahlreichen nationalen und internationalen Gremien zur digitalen Langzeitarchivierung mit, unter anderem vertritt sie das Bundesarchiv als Partner im Projekt nestor, war Leiterin der Arbeitsgruppe, die das DOMEA-Aussonderungskonzept erstellt hat, und ist Vorstandsmitglied im europäischen DLM Forum. E-Mail: a.haenger@barch.bund.de

Karsten Huth arbeitet im Rahmen des Projekts nestor am Bundesarchiv in Koblenz. Zur Zeit entwickelt er ein Metadatenschema zur digitalen Langzeitarchivierung. E-Mail: k.huth@barch.bund.de

LANGFRISTIGE VERFÜGBARKEIT VON QUELLENEDITIONEN IM DIGITALEN ZEITALTER – TEXT ENCODING INITIATIVE, GESCHICHTSWISSENSCHAFT UND BIBLIOTHEKEN

von Stefan Cramme

Editionen gehören zum grundlegenden Handwerkszeug der historischen Wissenschaften. Aus verschiedenen Gründen ist für sie in den letzten Jahren die elektronische Publikation attraktiv geworden. Zur Gewährleistung der langfristigen Verfügbarkeit bietet sich die Verwendung der Richtlinien der Text Encoding Initiative (TEI) an, die auf offenen Standards basieren und unabhängig von möglichen Ausgabeformaten sind. In diesem Zusammenhang kann auch Bibliotheken eine große Rolle zukommen, die als Dienstleister bei der Erstellung und dauerhaften Archivierung digital aufbereiteter Quellenbestände mitwirken können. Vorgestellt werden die Erfahrungen der Bibliothek für Bildungsgeschichtliche Forschung mit der TEI-Auszeichnung für Werkausgaben von Adolf Reichwein und Briefe Friedrich Fröbels.

Historische Quellen im digitalen Zeitalter

Texteditionen stellen nicht nur in den philologischen Disziplinen, sondern auch in den historisch arbeitenden Wissenschaften nach wie vor eine unverzichtbare Arbeitsgrundlage dar. Sie gewährleisten insbesondere die dauerhafte Verfügbarkeit und Überprüfbarkeit schriftlicher Quellen¹,

¹ Die Begriffe Quelle und Edition, ihr Sinn oder möglicherweise auch Unsinn für die verschiedenen Bereiche der Geschichtswissenschaft können hier nicht weiter reflektiert werden. Siehe etwa Gall, Lothar; Schieffer, Rudolf (Hgg.), Quelleneditionen und kein Ende? (Historische Zeitschrift, Beihefte 28), München 1999;

unabhängig von der Möglichkeit, eventuell weit entfernte Archive oder Bibliotheken aufsuchen zu müssen, oder der Gefahr des Verlustes des Originals.² Die jahrhundertlang bewährte Veröffentlichung in Form gedruckter Bände stößt aber in letzter Zeit auf vor allem finanzielle Schwierigkeiten, weil Verlage angesichts der schlechten Absatzmöglichkeiten mit Recht sehr zurückhaltend sind, insbesondere bei umfangreichen und über einen langen Zeitraum geplanten Editionsprojekten, wie sie oftmals üblich waren.

In den letzten Jahren hat sich für Editionen die elektronische Erfassung und Bearbeitung der Texte durchgesetzt, zunächst als Vorstufe einer traditionellen Druckausgabe. Zunehmend tritt allerdings auch die direkte Bereitstellung elektronischer Fassungen für die Nutzenden ins Blickfeld.³ Solche elektronischen Editionen⁴ bieten gegenüber gedruckten Bänden klar

Quellen und Quellenedition im neuen Medienzeitalter (Fundus – Forum für Geschichte und ihre Quellen 2), unter <http://webdoc.sub.gwdg.de/edoc/p/fundus/html/heft_2.html> (alle Links wurden zuletzt überprüft am 26.09.2006); Fellner, Fritz, Die historische Quelle – Instrument der Geschichtsforschung und Baustein des Geschichtsbewußtseins oder Baustein der Geschichtsforschung und Instrument des Geschichtsbewußtseins?, in: Klingenstein, Grete; Fellner, Fritz; Hye, Hans Peter (Hgg.), Umgang mit Quellen heute. Zur Problematik neuerzeitlicher Quelleneditionen vom 16. Jahrhundert bis zur Gegenwart (Fontes rerum Austriacarum, 2. Abteilung Diplomataria et acta 92), Wien 2003, S. 19-33. Zur literaturwissenschaftlichen Editionsphilologie: Nutt-Kofoth, Rüdiger (Hg.), Text und Edition. Positionen und Perspektiven, Berlin 2000; Kraft, Herbert, Editionsphilologie, 2. neubearb. Aufl., Frankfurt am Main 2001. Alle in den Anmerkungen genannten URLs wurden zuletzt am 28. August 2006 überprüft.

2 Zur elektronischen Bereitstellung von nicht editorisch bearbeitetem Archivmaterial siehe beispielsweise Jucker-Kupper, Patrick; Koller, Christophe; Ritter, Gerold (Hgg.), Digitales Gedächtnis. Archivierung und die Arbeit der Historiker der Zukunft (Geschichte und Informatik 13/14), Zürich 2004.

3 Eine aktuelle Übersicht über geschichtswissenschaftliche Editionsprojekte ist nur schwer zu erhalten. Eine auf Österreich bezogene Umfrage 1999/2000 ergab, dass weniger als 10 Prozent eine elektronische Edition anstrebten (Klingenstein u.a. (Hgg.), Umgang (wie Anm. 1), S. 211-325; kurze Auswertung: Ehmer, Josef; Steidl, Annemarie, Überlegungen zur digitalen Edition historischer Quellen und Vorstellung des Editionsprojekts Wiener Innungsarchivalien auf CD-ROM, ebd. S. 191-192). Ankündigungen neu begonnener Projekte weisen jedoch in den letzten Jahren regelmäßig zumindest auf die Möglichkeit digitaler Publikationsformen hin.

4 Zur noch unscharfen Terminologie: Kamzelak, Roland, Hypermedia – brauchen wir eine neue Editionswissenschaft?, in: Ders. (Hg.), Computergestützte Text-

erkennbare Vorteile:

- Suchmöglichkeiten im Volltext, insbesondere, wenn sie von der Auszeichnung bestimmter Elemente (Namen, Schlagwörter, Sprachliches usw.) unterstützt werden;
- in digitaler Form können prinzipiell größere Textmengen verarbeitet werden als bei gedruckten Bänden;
- Herstellen von Bezügen und Verweisen innerhalb des Textes, aber auch auf andere elektronisch vorliegende Texte, auf Kommentare oder ergänzendes Material⁵; die Benutzeroberfläche kann dabei, muss aber nicht, vom bisher Gewohnten abweichen⁶;
- Weiterverarbeitung, insbesondere Umwandlung in verschiedene Ausgabeformen;
- leichtere Publikation von Zwischenergebnisse und korrigierten/überarbeiteten Fassungen. (Dies ist gerade bei Editionen, die eine zitierfähige, stabile Textfassung gewährleisten sollen, natürlich nicht unproblematisch.)

Diesen unverkennbaren Vorteilen stehen allerdings auch gewichtige Einwände gegenüber:

- das Rezeptionsverhalten zumal bei längeren Texten ist immer noch auf das Lesen einer gedruckten Fassung ausgerichtet⁷;
- die dauerhafte Verfügbarkeit elektronischer Dokumente ist durch die ständige Weiterentwicklung von Soft- und Hardware gefährdet;
- insbesondere in der Übergangs- und damit auch Experimentierphase sind keine zeitlichen und finanziellen Einsparungen zu erwarten.

Diesen Nachteilen kann mit einer Doppelstrategie begegnet werden:

Edition (Beihefte zu Editio 12), Tübingen 1999, S. 119-126, hier S. 120; zur Abgrenzung von nicht editierten elektronischen Texten im Internet: Pianos, Tamara, Das Canterbury Tales Project, in: Bibliothek und Wissenschaft 36 (2003), S. 95-150, hier S. 110.

5 Siehe Eibl, Karl; Jannidis, Fotis; Willems, Marianne, Der Junge Goethe in neuer Ausgabe. Einige Präliminarien und Marginalien, in: Kamzelak (Hg.), Text-Edition (wie Anm. 4), S. 69-78, hier S. 73-74.

6 Jenks, Stuart, Dekonstruktion und Rekonstruktion der Quellenedition, in: Mitteilungen aus dem Bundesarchiv 11 (2003), Heft 1, S. 5-13.

7 Siehe Eibl u.a., Junge Goethe (wie Anm. 5), S. 71-72.

- Hybridausgaben, die gedruckte und elektronische Fassung kombinieren, verbinden die Vorteile beider Medienarten;
- zur Auszeichnung und Speicherung der elektronischen Fassungen ist auf offene und standardisierte Formate zurückzugreifen.

Offene Formate und Auszeichnungssprachen

Die Verwendung proprietärer, also von einem Hersteller oder einem bestimmten Softwareprodukt verwendeter, nicht offen dokumentierter und anpassbarer Formate ist unter dem Gesichtspunkt der langfristigen Verfügbarkeit auf jeden Fall zu vermeiden und durch echte, das heißt offen gelegte und breit unterstützte Formate zu ersetzen, die unabhängig von bestimmter Anwendungssoftware oder Betriebssystemen genutzt werden können. Eine aktuell weite Verbreitung kann dabei kein Gegenargument sein, weil etwa das Dateiformat der Microsoft-Office-Programme mehrfach geändert worden ist und erst jetzt unter dem Druck von Mitbewerbern eine zögerliche Offenlegung der Formate erfolgt, aber nur für zukünftige Versionen. (Bei Multimediaformaten sind die möglichen Gefahren noch größer.⁸)

Etwas besser sieht es bei dem von Adobe eingeführten Portable Document Format (PDF) aus, das zwar kein offener Standard im eigentlichen Sinne ist, aber zumindest von einer Vielzahl von Anwendungen plattformübergreifend geschrieben und gelesen werden kann. Der Nachteil von PDF liegt aber darin, dass es vor allem das Erscheinungsbild eines Dokuments beschreibt, nicht die zugrunde liegenden logischen Strukturen. PDF ist ein Endformat, das sich nicht sinnvoll weiterverarbeiten lässt. Für Texte ist daher die Verwendung von Auszeichnungssprachen anzuraten, die eine Trennung der inhaltlichen Struktur von der Erscheinungsform ermöglichen. Schon seit den 1980er Jahren gibt es SGML (Standard General

8 Beispielhaft sei auf das Projekt der BBC für eine digitale Version des Domesday Books hingewiesen, das keine zwei Jahrzehnte nach seiner Erstellung nur noch unter Verwendung musealer Hardware zu verwenden war: Darlington, Jeffrey; Finney, Andy; Pearce, Adrian, Domesday redux. The rescue of the BBC Domesday Project videodiscs, in: Ariadne 36 (2003), unter <<http://www.ariadne.ac.uk/issue36/tna/>>.

Markup Language) als international standardisiertes Meta-Format zur Definition von Auszeichnungssprachen. Es hat weite Verbreitung vor allem beim technischen Publizieren gefunden. Als Weiterentwicklung, teilweise auch Vereinfachung, in spezieller, aber nicht ausschließlicher Hinsicht auf die Anwendung im World Wide Web (WWW) ist jetzt vor allem XML (Extensible Markup Language) zu nennen⁹, das SGML in den meisten Fällen ersetzt hat, insbesondere bei Anwendungen, die in den letzten Jahren begonnen worden sind. XML verwendet den kompletten Zeichenvorrat von Unicode und ist damit zumindest theoretisch für alle Sprachen und Schriftsysteme gerüstet. Wie bei SGML muss aber die konkrete Anwendung noch spezifiziert werden. Dies geschieht durch eine Document Type Definition (DTD) oder zunehmend durch ein XML-Schema. In beiden Fällen wird festgelegt, welche Elemente in welcher Verschachtelungsform und mit welchen Inhaltstypen vorkommen können. Inzwischen gibt es für XML und die damit zusammenhängenden Standards wie XSLT¹⁰ (eine Skriptsprache für die Umwandlung von XML-Dokumenten) eine große Anzahl von Werkzeugen zu Erstellung und Weiterverarbeitung, so dass keine Hersteller- oder Plattformabhängigkeit zu befürchten ist. XML kann nicht nur, wie im Folgenden weiter beschrieben, zur Auszeichnung kompletter Texte verwendet werden, sondern auch als Codierung für Metadaten-Formate wie RDF oder EAD.

Die wohl bekannteste Anwendung von SGML und XML, nämlich die im WWW verwendete Hypertext Markup Language (HTML), ist für Texteditionen kaum brauchbar, weil ausschließlich auf die Online-Anzeige ausgerichtet. Inzwischen wird XML auch von generellen Office-Programmen als Grundlage des Dokumentenformats verwendet. So beruhen die Formate von OpenOffice.org/StarOffice, iWork von Apple und in der nächsten Programmversion auch Microsoft Office auf XML. Allerdings gewährleistet dies nicht automatisch eine Trennung von Struktur und Form.

Besser berücksichtigt wird diese Forderung bei verschiedenen DTDs,

9 1. Version 1998, 4. Ausgabe 2006, vgl. <<http://www.w3.org/TR/xml/>>.

10 Vgl. <<http://www.w3.org/TR/xslt/>>.

die speziell zur Auszeichnung von Texten entwickelt wurden; beispielhaft sei DocBook genannt.¹¹ Gezielt mit der wissenschaftlichen Beschreibung und Edition von (bereits vorliegenden) Texten hat sich die Text Encoding Initiative (TEI) beschäftigt, die im Folgenden näher vorgestellt werden soll.

Text Encoding Initiative

Geschichte

Die TEI¹² bildete sich 1987, um auf der Grundlage der damals verfügbaren Meta-Auszeichnungssprache SGML ein Auszeichnungssystem für Texte jeder Art zu entwickeln. Ziel war dabei ein gemeinsames Austauschformat, das aber so flexibel sein sollte, dass jede/r Herausgeber/in die Möglichkeit hat, sich aus dem von der TEI bereitgestellten Grundlagen ein für seine/ihre Bedürfnisse passendes Format zusammenzustellen. Nach einigen vorläufigen Versionen lag die erste stabile Fassung der TEI-Richtlinien (Guidelines) mit der Kennung P3 im Jahr 1994 vor, noch auf SGML-Basis. Sie wurde seitdem kontinuierlich weiterentwickelt. Die 2002 veröffentlichte Version P4 beruht erstmals auch auf XML.¹³ Aktuell wird an Vorversionen von P5 gearbeitet, das nur noch auf XML ausgerichtet ist und neben einer DTD auch Schemas verwenden kann.¹⁴

11 Vgl. <<http://www.docbook.org/>>; siehe auch Bunke, Hendrik, *Schreibt strukturiert! XML und Docbook in Sozial- und Geisteswissenschaften* (2005), unter <<http://hbxt.org/edutech/docbook-in-geisteswissenschaften>>.

12 Vgl. <<http://www.tei-c.org/>>. Siehe auch Vanhoutte, Edward, *An introduction to the TEI and the TEI Consortium, 2004*, unter <<http://www.kantl.be/ctb/pub/2004/teillc.htm>>. Allgemeine Einführungen in deutscher Sprache: Schmidt, Frieder, *Neuland für die Buchgeschichte – Quellenaufbereitung im Zeitalter des WWW*, in: *Leipziger Jahrbuch für Buchgeschichte* 7 (1997), S. 343-365, besonders S. 351-361; Bader, Winfried, *Was ist die Text Encoding Initiative (TEI)?*, in: Kamzelak (Hg.), *Text-Edition* (wie Anm. 4), S. 9-20; Bruvik, Tone Merete, „Yesterday's information tomorrow“. *Die Text Encoding Initiative (TEI)*, in: *Sichtungen online*, 19.05.2002, unter <<http://purl.org/sichtungen/bruvik-tm-1a.html>>.

13 DeRose, Steven, *XML and the TEI*, in: *Computers and the humanities* 33 (1999), S. 11-30.

14 Zum Stand Oktober 2005: [Burnard, Lou,] *TEI P5 progress report*, unter <<http://www.tei-c.org/Talks/2005/Wuerzburg/p5report.pdf>>.

Die TEI wird durch verschiedene Fachgesellschaften und Forschungseinrichtungen unterstützt, die sich im TEI-Konsortium zusammengefunden haben. Es besteht derzeit (August 2006) aus mehr als 80 Mitgliedern, unter denen erst seit 2006 eines aus dem deutschsprachigen Raum vertreten ist¹⁵, obwohl auch hier die TEI-Richtlinien in zahlreichen Projekten verwendet werden. TEI beschränkt sich nicht mehr auf die Rolle des Austauschformats, sondern wirkt auch in andere Bereiche der Datenaufbereitung (die W3C-Standards XLink und XPointer gehen auf TEI-Entwicklungen zurück) und hat eine eigene *Research community* hervorgebracht.¹⁶ Die Textauszeichnung nach den TEI-Richtlinien ist insbesondere in den Literaturwissenschaften international weit verbreitet. Aber auch viele für Historiker/innen relevante Materialien sind entsprechend ausgezeichnet.¹⁷

Prinzipien der TEI-Richtlinien

Die TEI-Guidelines¹⁸ basieren auf den Erfahrungen bei der Erstellung konventioneller Texteditionen, zwingen also nicht dazu, Formate, die ursprünglich für einen ganz anderen Einsatzzweck entwickelt wurden, mühselig anzupassen. Sie sind modular aufgebaut und können deshalb für verschiedene Textarten adaptiert werden.¹⁹ Neben einem für alle TEI-Dokumente verbindlichen Grundbestand gibt es eigenständige Teile für Prosa, Drama, Lyrik, für kritische Apparate oder analytische Strukturen. Bei zahlreichen Projekten wird eine ursprünglich für didaktische Zwecke

15 Kooperationsprojekt TextGrid, siehe <<http://www.textgrid.de/>>, noch nicht aufgeführt auf der Mitgliederliste unter <<http://www.tei-c.org/membership.php>>.

16 Mylonas, Elli; Renear, Allen, *The Text Encoding Initiative at 10. Not just an interchange format anymore – but a new research community*, in: *Computers and the humanities* 33 (1999), S. 1-9.

17 Die Zahl der TEI-verwendenden Projekte (unter Einschluss derer, die keine vorhandenen Texte, sondern aktuelle Publikationen damit aufarbeiten), liegt weit über jenen 131, die aktuell (28.08.2006) auf der TEI-Website aufgeführt sind, vgl. <<http://www.tei-c.org/Applications/>>.

18 Einführung in die TEI-Richtlinien: Jannidis, Fotis, *TEI in der Praxis*, unter <<http://computerphilologie.uni-muenchen.de/praxis/teipraxis.html>>; Ders., *Wider das Altern elektronischer Texte. Philologische Textauszeichnung mit TEI*, in: *Editio 11* (1997), S. 152-177.

19 Auch online mit Hilfe des Pizza Chefs für P4, vgl. <<http://www.tei-c.org/pizza.html>>, bzw. von Roma für P5, vgl. <<http://tei.oucs.ox.ac.uk/Roma/>>.

konzipierte Auswahl verwendet, TEI lite, die die am häufigsten verwendeten Elemente zusammenfasst und bei Bedarf auf der Grundlage des vollen Umfangs modifiziert werden kann.²⁰

Die TEI-Guidelines beschäftigen sich ausschließlich mit der Struktur des auszuzeichnenden Textes, während im Normalfall keine Festlegungen über Aussehen oder Ausgabeform getroffen werden. Grundprinzip ist es, implizite Informationen explizit zu machen (wie bei konventionell erstellten Texteditionen liegt hier also ein Element der Interpretation vor). Die Richtlinien sehen folgende Bestandteile vor:

- Metadaten (TEI-Header);
- Grundstrukturen des Textes (Kapitel, Strophen, Briefe, Anmerkungen);
- speziell zu formatierende Passagen;
- editorische und semantische Codierung (Anmerkungen des Herausgebers/der Herausgeberin, kritischer Apparat, Namen, sprachliche Besonderheiten usw.);
- durchgehend Spezifizierung durch Attribute;
- möglich ist auch die Integration von Abbildungen und Tabellen sowie Verweise auf weitere Materialien (dazu ist ein eindeutiger Identifier für die Elemente erforderlich).

Voraussetzung für die sinnvolle Anwendung sind eine gewisse Textanalyse und die konsequente Durchführung der Auszeichnung. Wie bei allen Auszeichnungssprachen ist ein systematisches Vorgehen erforderlich. Gerade am Beginn eines Projekts kann dabei der Erfahrungsschatz, der in TEI eingeflossen ist, durchaus hilfreich sein. Eine TEI-Grundstruktur ist relativ schnell zu erzeugen; eine sinnvolle semantische Codierung erfolgt normalerweise mit einem XML-fähigen Editor. Die Auszeichnung ist je nach Bedürfnis und Kapazität verschieden tief möglich; von amerikanischen Bibliotheken entworfene Empfehlungen sehen fünf Ebenen vor, von einer automatischen Konversion bis zur textkritischen Edition.²¹

20 Burnard, Lou; Sperberg-McQueen, C. M., TEI U5. Encoding for interchange. An introduction to the TEI, 1995/2002, unter http://www.tei-c.org/Lite/teiu5_split_en.html.

21 TEI text encoding in libraries. Guidelines for best encoding practices, Version 2.1,

Die Umwandlung in eine dem gewohnten Lese- und Rezeptionsverhalten entsprechende Form geschieht in der Regel mit Hilfe von Style-Sheet-Sprachen, insbesondere XSLT und XSL-FO²²; unter Umständen kann auch das eigentlich für HTML entwickelte CSS²³ herangezogen werden. Die Stylesheets sind normalerweise nicht nur für ein Dokument, sondern für eine Vielzahl von Texten zu verwenden. Die Ausgabeformate können dynamisch auf Anfrage (etwa auf einem Webserver) oder im Vorhinein erzeugt werden. Auf diese Weise ist etwa die Herstellung von HTML- oder PDF-Fassungen möglich, aber auch die Generierung von Konkordanzen oder E-Book-Formaten. Die offenen Standards sollten dabei auch für zukünftige Formate offen sein.

Nicht weiter eingehen möchte ich hier auf die Möglichkeit des Zusammenspiels von TEI-codierten Daten und älteren, aber leistungsfähigen Werkzeugen wie TUSTEP²⁴, T_EX²⁵ oder Folio Views. Letzteres wird als Retrieval-Oberfläche in zahlreichen elektronischen Editionsprojekten verwendet, steht aber unter den genannten Vorbehalten bezüglich der langfristigen Verfügbarkeit.²⁶

27. März 2006, vgl. <http://www.diglib.org/standards/tei.htm>.

22 Vgl. <http://www.w3.org/TR/xsl/>.

23 Vgl. <http://www.w3.org/Style/CSS/>.

24 Vgl. <http://www.uni-tuebingen.de/zdv/tustep/index.html>.

25 Vgl. <http://www.tug.org/>. Ein darauf beruhendes Editionssystem ist etwa CET (Karasch, Bernt, Critical Edition Typesetter (CET). Ein Programmsystem zum Satz textkritischer Editionen auf PCs, in: Kamzelak (Hg.), Text-Edition (wie Anm. 4), S. 87-99), nicht zu verwechseln mit CTE, der ein eigenes Dokumentformat verwendet, aber auch HTML und TEI erzeugen kann (Hagel, Stefan, Zur druckfertigen Edition mit dem eigenen PC – der „Classical Text Editor“, in: Klingenstein u.a. [Hgg.], Umgang [wie Anm. 1], S. 198-201).

26 Die Zuversicht von Monika Seekircher (Der Wittgenstein-Gesamtbriefwechsel in maschinenlesbarer Form, in: Bauer, Werner M.; John, Johannes; Wiesmüller, Wolfgang (Hgg.), „Ich an Dich“. Edition, Rezeption und Kommentierung von Briefen (Innsbrucker Beiträge zur Kulturwissenschaft, Germanistische Reihe 62), Innsbruck 2001, S. 189-204, hier S. 202), ihre mit Folio Views erstellte Wittgenstein-Briefausgabe noch im Jahr 2020 erweitern zu können, erscheint angesichts der wechselvollen Geschichte dieses Programms, das aktuell zur Suchmaschinenfirma FAST gehört (vgl. <http://www.fastsearch.com/l3a.aspx?m=497&amid=3116>), eher zweifelhaft.

Anwendungen der TEI-Guidelines

Die TEI-Richtlinien wurden ursprünglich vor allem im Bereich der Literaturwissenschaft und Linguistik angewendet (insbesondere für große Textcorpora oder Wörterbücher). Ihre Flexibilität macht sie aber auch geeignet für Editionen. Dabei gehen sie über die ursprüngliche Funktion als Austausch- und Sicherungsformat hinaus.

Einige Beispiele aus unterschiedlichen thematischen Bereichen seien genannt, wobei Projekte den Vorzug erhielten, die eine Online-Version anbieten:

- CAMENA (Universität Mannheim/Universität Heidelberg)²⁷: neulateinische Texte; HTML auf Grundlage von TEI, Einbindung von Images;
- Jean Pauls Exzerptheft (Universität Würzburg)²⁸: diplomatische Transkription;
- Sanger Papers (New York University)²⁹: Transformation ›on-the-fly‹ mit PHP/Sablotron³⁰;
- Newton Project (Imperial College London)³¹: Benutzer/innen können zwischen diplomatischer und Lese-Transkription wählen;
- eScholarship (University of California)³²: keine Editionen, sondern aktuelle wissenschaftliche Publikationen auf TEI-Basis;

27 Vgl. <<http://www.uni-mannheim.de/mateo/camena/docs/camena.html>>. Siehe auch Schibel, Wolfgang, CAMENA – Neulateinische Dichtung im World Wide Web, in: Neulateinisches Jahrbuch 3 (2001), S. 211-219; Ders.; Kredel, Heinz, Vom Spezialbestand alter Drucke zum kommentierten Volltextcorpus im WWW. Das Projekt CAMENA (Neulateinische Dichtung Deutschlands), in: Burch, Thomas; Fournier, Johannes; Gärtner, Kurt; Rapp, Andrea (Hgg.), Standards und Methoden der Volltextdigitalisierung. Beiträge des Internationalen Kolloquiums an der Universität Trier, 8./9. Oktober 2001 (Akademie der Wissenschaften und der Literatur, Abhandlungen der Geistes- und Sozialwissenschaftlichen Klasse, Einzelveröffentlichung 9), Stuttgart 2003, S. 187-209.

28 Vgl. <http://www.uni-wuerzburg.de/germanistik/neu/jp-arbeitsstelle/jpa_exz01.htm>.

29 Vgl. <http://wilde.acs.its.nyu.edu/sanger_dev/documents/search.php>.

30 Zimmerman, Matthew, Publishing XML files on the web (2003), vgl. <http://www.nyu.edu/its/pubs/connect/fall03/zimmerman_xml.html>.

31 Vgl. <<http://www.newtonproject.ic.ac.uk/>>. Zur Technik: Dunning, Alastair, The Newton Project. Implementing and exploiting XML, 7. November 2005 <http://ahds.ac.uk/_print_/creating/case-studies/newton/index.htm>.

32 Vgl. <<http://www.cdlib.org/programs/escholarship.html>>.

- Sichtungen (Österreichisches Literaturarchiv der Österreichischen Nationalbibliothek)³³: elektronische Zeitschrift mit TEI-Master, HTML- und Druckfassung generiert durch XSL-Stylesheets.

Nur am Rande sei darauf hingewiesen, dass auch das umfangreiche, aber wissenschaftlichen Kriterien bisher nicht genügende Textarchiv des Project Gutenberg³⁴ seit kurzem TEI-codierte Texte aufnimmt.

Erfahrungsbericht: Computereditionen der Bibliothek für Bildungsgeschichtliche Forschung

Die Bibliothek für Bildungsgeschichtliche Forschung (BBF)³⁵ ist die größte pädagogische Spezialbibliothek Deutschlands und gehört zum Deutschen Institut für Internationale Pädagogische Forschung, einer Einrichtung der Leibniz-Gemeinschaft. Als Forschungsbibliothek hat sie außer den eigentlichen bibliothekarischen und dokumentarischen Aufgaben einen Schwerpunkt bei der Zusammenarbeit mit der Wissenschaft in Form von Forschungs- und Publikationsprojekten, Tagungen und Ausstellungen. Neben Digitalisierungs- und Fachinformationsprojekten ist die BBF seit 2001 auf dem Gebiet der Computereditionen aktiv.

Werkausgabe Adolf Reichwein

Die Werkausgabe des Pädagogen und Widerstandskämpfers Adolf Reichwein (1898–1944) war ursprünglich als konventionelle Edition geplant, die in fünf Bänden alle pädagogischen Schriften umfassen sollte.³⁶ Bei Beginn der von der Deutschen Forschungsgemeinschaft (DFG) und weiteren Sponsoren geförderten Arbeiten wurde jedoch eine Erfassung aller, auch der nicht-pädagogischen Schriften Reichweins ins Auge gefasst, die in Form einer CD-ROM den Bänden der Werkausgabe beigegeben werden sollen. Von vornherein war für die Erfassung aller Texte die Verwendung von TEI vorgesehen, das somit auch als Grundlage der weiterhin vorgesehenen Druckausgabe dienen soll. Das wissenschaftliche und technische Editions-

33 Vgl. <<http://purl.org/sichtungen/>>.

34 Vgl. <<http://www.gutenberg.org/>>.

35 Vgl. <<http://www.bbf.dipf.de/>>.

36 Wissenschaftliche Bearbeiter: Ullrich Amlung, Karl Christoph Lingelbach. Die Edition wird vom Adolf-Reichwein-Verein unterstützt.

konzept wurde im Februar 2003 auf einem Expertenworkshop in der BBF vorgestellt, der zu einigen Modifikationen führte. Insbesondere der Umfang der Kommentierung wurde reduziert, um den durch die DFG-Finanzierung gesetzten Zeitrahmen nicht zu überschreiten. Die ersten beiden Bände sollen Ende 2006 in druckfertiger Form vorliegen. Noch nicht abschließend geklärt ist, ob Teile oder das gesamte Textcorpus auch über das Internetangebot der BBF zur Verfügung gestellt werden können.

Die Werke Reichweins sind zum großen Teil nur noch in der publizierten Fassung erhalten, selten als Schreibmaschinentypskript; Handschriftliches gibt es praktisch überhaupt nicht (Reichwein wurde während des Krieges ausgebombt). Die Texterfassung erfolgte teilweise über OCR, bei schlechten oder frakturgedruckten Vorlagen über manuelle Erfassung. Die daraus resultierenden Dokumente im DOC-Format von Microsoft Word wurden, nachdem Möglichkeiten der automatischen Konvertierung getestet, aber verworfen wurden, als unformatierter Text in den XML-Editor XMetaL geladen und dort mit einer zunächst elementaren Auszeichnung nach TEI lite versehen. Neben den eigentlichen Textabschnitten wurden dabei auch Fußnoten und Tabellen berücksichtigt. Diese Phase, die auch mit Verwendung von studentischen Hilfskräften erfolgte, ist 2004 abgeschlossen worden, so dass seitdem alle Texte Reichweins zumindest in einer minimalen TEI-Struktur vorhanden sind. XMetaL hat sich, nicht zuletzt durch seine wahlweise an eine normale Textverarbeitung erinnernde Benutzeroberfläche, in diesem und den im Folgenden vorgestellten Projekten bewährt, auch wenn die Wahl eines XML-Editors immer auch eine Frage der persönlichen Vorliebe ist.³⁷

Eine inhaltliche Auszeichnung wurde bisher bei einem Teil der Texte vorgenommen. Dabei sind nach Entscheidung der wissenschaftlichen Bearbeiter signifikante Personen- und Ortsnamen, gelegentlich auch Institutionen, durch die entsprechenden TEI-Elemente gekennzeichnet worden. Vorgesehen ist die automatische Generierung eines entsprechenden Index

³⁷ Vgl. <http://www.xmetal.com/>. Wer das Arbeiten in einer code-orientierten Ansicht nicht scheut, findet etwa bei oXygen unter <http://www.oxygenxml.com/> eine gute Unterstützung der TEI und aktueller Technologien wie XSLT und XML Schema.

für die Druckausgabe, der verbunden ist mit einem Glossar der Personennamen, das Teile des ursprünglich geplanten Kommentars ersetzt.

Briefausgabe Friedrich Fröbel

Die bisher nur auszugsweise edierten Briefe des Pädagogen Friedrich Fröbel (1782–1852) wurden seit 1991 mit Unterstützung der DFG durch Helmut Heiland (Universität Duisburg) erfasst und transkribiert. Die mit einem knappen textkritischen Apparat, aber keinem Kommentar versehene Transkription erfolgte im DOC-Format, aufgrund der langen Bearbeitungszeit in verschiedenen Versionen.

Das Corpus der Fröbel-Briefe besteht aus 1.661 bekannten Briefen oder Brieffragmenten, die im Word-Format etwa 7.300 Seiten umfassen.

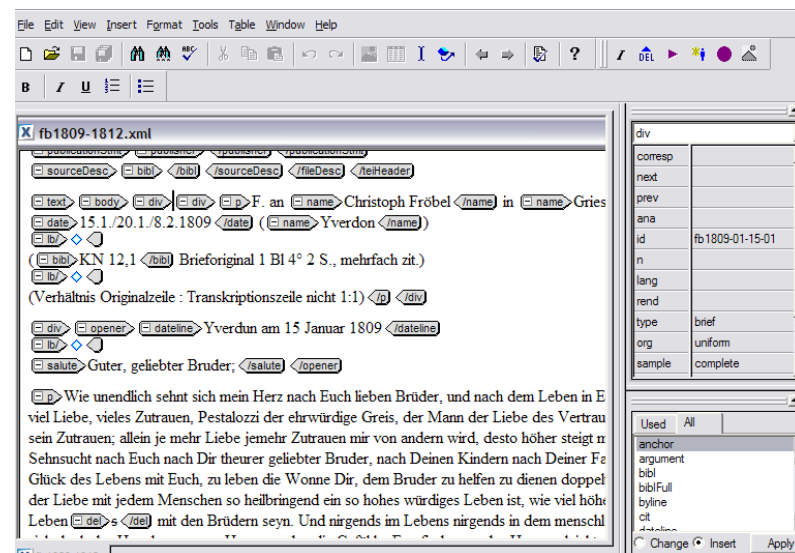


Abbildung 1: Text eines Fröbelbriefes im Editor XMetaL

Angesichts dieses Umfangs war eine komplette Veröffentlichung in Druckform nicht praktikabel, so dass eine Computeredition ins Auge gefasst wurde, bei der die BBF ihre bei der Reichwein-Ausgabe gewonnenen

technischen Erfahrungen einbringt. Seit September 2005 werden in einem zweijährigen, von der DFG geförderten Projekt die in etwa 50 Word-Dateien vorliegenden Texte in TEI lite konvertiert. Dabei wird als erster Schritt eine automatische Konversion mit Hilfe des TEIOO-Filters von Sebastian Rahtz (Universität Oxford) durchgeführt.³⁸ Da die Dateien im Word-Format aufgrund der langen Bearbeitungszeit aber zahlreiche Formatierungsunregelmäßigkeiten aufweisen und rein visuell orientiert waren, müssen umfangreiche Nacharbeiten erfolgen. Dabei werden die einzelnen Briefe voneinander abgetrennt, innerhalb der Briefe editorische Einleitung und Transkription, bei dieser wiederum Einleitungs- und Schlussformeln gekennzeichnet.

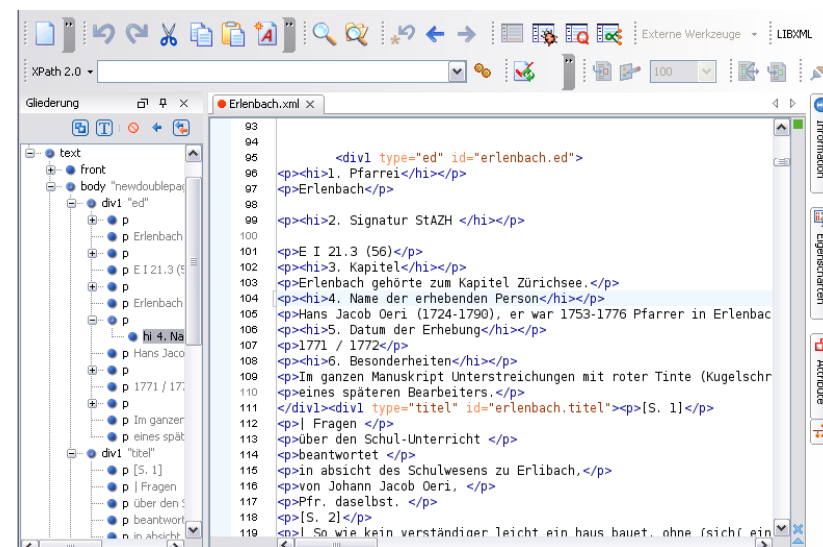
Die Transkription hatte ursprünglich versucht, das Erscheinungsbild der handschriftlichen Vorlagen zumindest näherungsweise wiederzugeben. Daher finden sich zahlreiche Einzüge, horizontale oder vertikale Leerräume und Schriftformatierungen wie hochgestellt, unterstrichen (einfach, doppelt und vereinzelt sogar dreifach) oder kursiv (als Kennzeichen des Wechsels zwischen deutscher und lateinischer Handschrift). Bei der Umsetzung in TEI-Stukturen ist dabei oftmals die Frage aufgetaucht, wie sich diese Formatierungseigenschaften in logische Strukturen umsetzen lassen. Eine 1:1-Identität des Endergebnisses mit der Word-Fassung (die wiederum ja schon eine Abstraktion der Handschrift darstellt) wird sich nicht erreichen lassen und ist auch nicht angestrebt. Der textkritische Apparat, der Streichungen, Einfügungen und unsichere Lesungen betrifft, wird weitgehend in die entsprechenden TEI-Elemente umgesetzt. Hier hat sich auch die Notwendigkeit gezeigt, möglichst früh während der Bearbeitung in Absprache mit den wissenschaftlichen Bearbeitern/innen klare Entscheidungen für Zweifelsfälle zu treffen. Eine gewisse Vertrautheit mit den Eigenheiten der handschriftlichen Vorlage ist dabei sinnvoll, auch wenn die eigentliche Transkription bereits erfolgt ist. Aktuell ist noch die Frage offen, ob weiterhin mit TEI lite gearbeitet wird oder die DTD auf der Grundlage des Gesamtumfangs der TEI modifiziert werden soll. Die Veröffentlichung ist für 2007 über das Internetangebot der BBF vorgese-

38 Vgl. <<http://www.tei-c.org/Software/teioo/>>.

hen. Als Ergänzung ist ein Einführungsband vorgesehen, der auch ausgewählte Briefe abdrucken wird, aber kein eigentlicher Bestandteil der Computeredition ist.

Weitere Projekte und Planungen

Bei einem Editionsprojekt des Instituts für Historische Bildungsforschung Pestalozzianum Zürich³⁹ gibt die BBF im Rahmen einer schon länger bestehenden Kooperation technische Unterstützung. Die Transkription einer 1771/1772 durchgeführten Fragebogenaktion bei den Pfarrern des Kantons Zürich über den Zustand ihrer Landschulen wird, ähnlich wie bei den Fröbelbriefen, aus einer vorliegenden Word-Fassung in TEI lite umgesetzt. Die Veröffentlichung des gesamten Textes erfolgt 2006 als CD-ROM, die einem Kommentarband beigelegt ist.



Die TEI-Auszeichnung und elektronische Veröffentlichung weiterer

39 Vgl. <<http://ihbf.phzh.ch/>>.

bildungshistorischer Quellenbestände aus dem 19. und 20. Jahrhundert ist geplant, nachdem mehrere Wissenschaftler/innen Interesse an einer Kooperation geäußert haben. In einem Pilotprojekt wird derzeit ferner ein Mitteilungsblatt nach den TEI-Richtlinien ausgezeichnet, um die Möglichkeit der langfristigen Archivierung auch aktueller Publikationen nach diesem Modell zu testen.

Erste Ergebnisse und Erfahrungen

Die Erfahrungen der BBF zeigen, dass es möglich ist, sich rasch in die Eigenheiten der Textauszeichnung nach TEI einzuarbeiten. Bei den bisherigen Projekten gab es immer nur eine/n Hauptbearbeiter/in, so dass kaum Probleme bei der Arbeitsorganisation und Versionskontrolle anfielen. Zukünftig wird erwartet, dass mehrere Personen gleichzeitig oder nacheinander an Texten arbeiten; damit ergibt sich die Notwendigkeit für ausführliche und eindeutige schriftliche Auszeichnungsanweisungen, quasi „Anwendungsbestimmungen“ für die TEI-Guidelines.

Die in der Community schon lange diskutierten Mängel des SGML/XML-basierten TEI-Markups werden auch bei den Editionsprojekten der BBF deutlich: Logische Strukturen lassen sich oft nicht von physischen trennen⁴⁰; auch die Anordnung der Schrift und weiterer grafischer Elemente auf dem Papier kann Bedeutung tragen. Die TEI-Richtlinien waren bisher weitgehend auf die inhaltlichen Strukturen ausgerichtet, enthalten im neuesten Entwurf P5 aber auch einen Abschnitt zur „Manuscript Description“.⁴¹ Überlappende Strukturen (etwa Sinnabschnitte und Manuskript- oder Druckseiten) lassen sich auf XML-Grundlage nicht gleichermaßen gut auszeichnen; es gibt aber einige Versuche, dieser

40 Siehe dazu auch: Vanhoutte, Edward, Display or argument? Markup and visualization for electronic scholarly editions, in: Burch u.a. (Hgg.), Standards (wie Anm. 27), S. 71-96.

41 Die andauernde Weiterentwicklung der TEI-Richtlinien und der zugrunde liegenden Standards können natürlich auch Zweifel an der Dauerhaftigkeit einer darauf aufbauenden Edition erzeugen: Rischer, Tobias, Eine säurefreie elektronische Edition des Ulysses. Bestandsaufnahme und einige Vorschläge, in: Henkes, Christiane u.a. (Hgg.), Schrift – Text – Edition. Hans Walter Gabler zum 65. Geburtstag (Beihefte zu Editio 19), Tübingen 2003, S. 339-348.

bekannten Beschränkung zu begegnen.⁴² Eine mögliche Abhilfe für die genannten Mängel wäre die Integration von Images (elektronischen Faksimiles), wie bei CAMENA und vielen anderen Projekten bereits erfolgt.⁴³

Ausblick: Zusammenarbeit von Wissenschaft und Bibliotheken in der elektronischen Quellenedition

Viele Editions- und Textauszeichnungsprojekte leiden unter mangelhafter technischer Unterstützung. Es ist aber auch kaum vorstellbar, jedes Projekt mit der entsprechenden Ressourcenausstattung, insbesondere der personellen, zu versehen. Hier ist die Unterstützung durch Kompetenzzentren denkbar, die sinnvollerweise in einem klar definierten disziplinären Kontext agieren. Sie können dadurch unter Umständen näher an den konkreten Bedürfnissen der Wissenschaftler/innen sein als allgemeine „Electronic Text Center“⁴⁴, auch wenn etwa das Kompetenzzentrum für elektronische Erschließungs- und Publikationsverfahren in den Geisteswissenschaften an der Universität Trier erfolgreich Projekte aus verschiedenen Fachgebieten betreut.⁴⁵

Eine wichtige Rolle kann hier den Forschungsbibliotheken zukommen, die sich durch die enge Verbindung mit der Wissenschaft definieren.⁴⁶ (Wenn im Folgenden die Bibliotheken im Vordergrund stehen, ist das dem persönlichen Hintergrund geschuldet. Natürlich können auch Computerzentren, Archive oder spezielle Einrichtungen ähnliche Dienstleistungen erbringen.) Durch diese Zusammenarbeit mit den Erzeugern elektronischer Angebote, hier also besonders von Editionen, befinden sie sich in einer anderen Situation als etwa regional definierte Pflichtexemp-

42 Ein Vorschlag, die materiellen Eigenschaften alter Drucke mit zu erfassen: Bauman, Syd; Catapano, Terry, TEI and the encoding of the physical structure of books, in: Computers and the humanities 33 (1999), S. 113-127.

43 Siehe auch Flanders, Julia, Trusting the electronic edition, in: Computers and the humanities 31 (1998), S. 301-310.

44 Seaman, David, The Electronic Text Center at the University of Virginia Library, in: Burch u.a. (Hgg.), Standards (wie Anm. 27), S. 97-107.

45 Vgl. <<http://germazope.uni-trier.de/Projects/KoZe2>>.

46 Grundlegend: Knoche, Michael, Die Forschungsbibliothek. Umriss eines in Deutschland neuen Bibliothekstyps, in: Bibliothek. Forschung und Praxis 17 (1993), S. 291-300.

larbibliotheken, die mit heterogenen digitalen Publikationen rechnen müssen, auf deren Gestaltung sie nur beschränkt, wenn überhaupt, Einfluss haben. Neben einer disziplinären Zuordnung ist natürlich auch eine Spezialisierung auf bestimmte Epochen oder Materialgattungen denkbar.⁴⁷

Bibliotheken können ihre Dienste insbesondere auch bei der Bereitstellung der Publikationen und der Einbindung in umfassendere Angebote erbringen; hier kommen ferner Themen wie Metadaten und eindeutige Identifier ins Spiel, die ebenfalls in den Kompetenzbereich von Bibliotheken und Archiven fallen, in diesem Abriss aber nicht behandelt werden konnten. Gleiches gilt schon bei der Textauszeichnung für die Regularisierung von Namensformen, wo die Verwendung von Normdaten sehr hilfreich sein kann.⁴⁸ Besonders wichtig ist natürlich der Bereich der Langzeitarchivierung: Bei selbst betreuten Veröffentlichungen ist eine datenträgerunabhängige Migration anders als bei vielen Fremdprodukten sinnvoll durchführbar, weil dieser Aspekt von vornherein berücksichtigt werden kann. Das Problem proprietärer CD-ROMs, das die Archivbibliotheken bereits erreicht hat (von den Disketten früherer Jahre zu schweigen) ist damit zwar nicht beseitigt, aber zumindest ein möglicher Lösungsweg aufgezeigt, der idealerweise Vorbildcharakter haben könnte.

Das aktuelle Schlagwort des Open Access, also des freien Zugangs zu wissenschaftlichen Publikationen, ist im Bereich der Quelleneditionen, die bestenfalls einen marginalen kommerziellen Wert haben, natürlich äußerst relevant. Neben der besseren Sichtbarkeit im Internet, insbesondere in allgemeinen Suchmaschinen, erlauben offene Formate und entsprechende Lizenzen auch die Spiegelung auf weiteren Servern, was unter Umständen ebenfalls ein Element einer langfristigen Archivierungsstrategie sein

47 Für Altbestandsbibliotheken siehe Stäcker, Thomas, XML für alte Drucke, in: Benkert, Hannelore; Rosenberger, Burkard; Dittrich, Wolfgang (Hgg.), 92. Deutscher Bibliothekartag in Augsburg 2002. Die Bibliothek zwischen Autor und Leser (Zeitschrift für Bibliothekswesen und Bibliographie, Sonderheft 84), Frankfurt am Main 2003, S. 259-275.

48 Diesen Punkt hebt etwa hervor Sukovic, Suzana, Beyond the scriptorium. The role of the library in text encoding, in: D-Lib magazine 8 (2002), Nr. 1, unter <<http://webdoc.sub.gwdg.de/edoc/aw/d-lib/dlib/january02/sukovic/01sukovic.html>>.

könnte. Unbedingt erforderlich ist es dafür, neben einer aufbereiteten (HTML- oder PDF-) Fassung auch den TEI-codierten XML-Text bereitzustellen. Die bisher noch oft übliche Veröffentlichung auf einer (separaten oder meist zusammen mit einem Buch angebotenen) CD-ROM gewährleistet zwar eine stabile Version, ist aber unter langfristiger Betrachtungsweise nicht weniger problematisch, als es eine potenziell verschwundene Website wäre.

In der Praxis wird man natürlich Rücksicht nehmen müssen auf die gewohnten Rezeptionsgewohnheiten, Vorbehalte wegen der Manipulierbarkeit digitaler Publikationen und traditionelle Verlagsbeziehungen, die man aber durchaus auch für Open Access und die Verwendung standardisierter Formate ausnutzen kann. Die Hybridedition in gedruckter und elektronischer Form hat sich in der wohl noch lange andauernden Übergangszeit, in der wir uns befinden, durchaus bewährt, und offenbar können sich viele Wissenschaftler/innen und Bibliothekar/innen allmählich mit dem Gedanken anfreunden, dass eine auf lange Dauer angelegte Veröffentlichung wie eine Edition nicht notwendigerweise ein gedrucktes Buch sein muss.

Dr. Stefan Cramme ist stellvertretender Bibliotheksleiter der Bibliothek für Bildungsgeschichtliche Forschung des Deutschen Instituts für Internationale Pädagogische Forschung; Arbeitsschwerpunkte sind derzeit: Computereditionen, Fachinformation für die historische Bildungsforschung. E-Mail: cramme@bbf.dipf.de

VERTEILTE DIGITALE INKUNABELBIBLIOTHEK – EIN BAUSTEIN ZUR GESAMTDIGITALISIERUNG ALLER INKUNABELAUSGABEN

von Timo Steyer

Auf der Internetseite des Forschungsprojektes „Verteilte Digitale Inkunabelbibliothek“ (vdIb) werden bis Mitte des Jahres 2006 über 1.200 digitalisierte und mit Metadaten erschlossene Inkunabeln der Forschung zur Verfügung gestellt sein.¹ Der Artikel berichtet über die Projektarbeit und die Möglichkeiten, welche die Projektseite für die Benutzung der Digitalisate offeriert. Anhand dieser Ausführungen sollen ferner die Grenzen, welche eine quantitative Digitalisierung bietet, diskutiert werden. Dabei wird auch der Frage nachgegangen, inwieweit das Projekt als Muster für eine zukünftige Gesamtdigitalisierung aller Inkunabelausgaben dienen kann.

Einleitung

Die retrospektive Digitalisierung historischer Quellen erfährt in jüngster Zeit eine Konjunktur, die nicht nur in der Förderung durch die Deutsche Forschungsgemeinschaft (DFG) zum Ausdruck kommt, sondern auch in der Vielzahl von internationalen Digitalisierungsprojekten sichtbar wird.² Aufgrund des mittlerweile erreichten technischen Stands können diese Vorhaben quantitative und qualitative Ergebnisse erzielen, die vor wenigen Jahren noch als wesentlich längerfristige Zukunftsperspektiven galten. Ein

¹ Projekt-Website: <<http://inkunabeln.ub.uni-koeln.de/>>.

² Vgl. dazu die Linksammlung der Zentral- und Landesbibliothek Berlin zu Digitalisierungsprojekten: <<http://linksammlungen.zlb.de/1.2.2.61.0.html>> (14.02.2006).

Ende dieser Entwicklung ist bei weitem noch nicht in Sicht. Doch diese Dynamik wirft auch Fragen auf: Wie korrespondiert die Möglichkeit der quantitativen Digitalisierung mit der Herstellung einer möglichst großen Benutzerfreundlichkeit und inhaltlicher Erschließung der digitalisierten Quellen? Und wie können bereits bestehende Internetressourcen in neue Projekte eingebunden werden, um Synergieeffekte zu erzielen? Die folgenden Ausführungen sollen anhand des Forschungsprojektes der „Verteilten Digitalen Inkunabelbibliothek“ (vdIb) diesen Fragen nachgehen und daher zum einen aufzeigen, wie das Projekt mit anderen Ressourcen verknüpft ist. Zum anderen soll durch die Darstellung der inhaltlichen und technischen Vorgehensweisen bei der Projektarbeit ein Eindruck vermittelt werden, wie eine Digitalisierung von Inkunabeln unter einem quantitativen Fokus vonstatten gehen kann. Dabei soll das Projekt als Diskussionsgrundlage, aber auch als Vorschlag für ähnliche Digitalisierungsprojekte dienen.

Gegenstand und Philosophie der vdIb

Die vdIb ist ein von der DFG gefördertes Kooperationsprojekt zwischen der Universität Köln, der Universitäts- und Stadtbibliothek Köln sowie der Herzog August Bibliothek (HAB) in Wolfenbüttel, welches eine möglichst repräsentative Digitalisierung der jeweiligen Inkunabelbestände unter Berücksichtigung konservatorischer Gutachten anstrebt.

In Wolfenbüttel wurden im Verlauf des Projektes 667 Inkunabeln mit insgesamt 123.241 Seiten digitalisiert. Dabei entspricht die Gesamtmenge der digitalisierten Inkunabeln etwa fünf bis sechs Prozent der weltweit noch existierenden Wiegendrucke. Diese Zahlen machen deutlich, dass der Schwerpunkt des Projektes auf der Erreichung quantitativer Zielvorgaben lag. Die Idee hierbei war, dass es der Forschung am ehesten nutzt, wenn eine möglichst hohe Anzahl von benutzbaren Quellen im Internet verfügbar ist, auch wenn dabei Beschränkungen in der inhaltlichen Erschließung in Kauf genommen werden müssen.

Zur Produktionsbeschleunigung der Digitalisierung wurde der Inkunabelbestand zwischen den beiden Projektpartnern entsprechend dem Druckjahr aufgeteilt. So wurden in Köln Inkunabeln ausgewählt, welche

vor dem Jahr 1485 gedruckt worden sind. In Wolfenbüttel dagegen wurden Inkunabeln aus der Zeit von 1485 bis 1500 digitalisiert. Die Auswahl des Jahres 1485 als Trennpunkt ergab sich aufgrund der zeitlichen Verteilung der Inkunabelbestände beider Sammlungen, die sich so zu fast gleich großen Stückzahlen auf die Projektteilnehmer verteilten. Der Schwerpunkt lag bei beiden Sammlungen auf Inkunabeln aus deutschen Druckorten. Freilich sind trotzdem zahlreiche Inkunabeln aus dem italienischen, niederländischen und französischen Raum vertreten. Thematisch vermitteln die Digitalisate in ihrer Gesamtheit aufgrund der zur Verfügung stehenden umfangreichen Sammlungen einen guten Einblick in die vielfältige Thematik der Inkunabelzeit. Neben der Digitalisierung bildete die Bereitstellung einer möglichst benutzerfreundlichen Handhabung der Inkunabeln einen zweiten Schwerpunkt im Rahmen des Projektes. Dieser Anspruch wurde nicht nur an die Projektseite im Internet gestellt, sondern auch auf die Digitalisate selber bezogen. Aus diesem Grund wurden Teile der Texte transkribiert und drucktypische sowie spezifische Eigenschaften der einzelnen Inkunabeln vermerkt (siehe dazu das Kapitel „Inhaltliche Bearbeitung der Inkunabeln“). Des Weiteren ist die Benutzung der bereitgestellten Inkunabeln maßgeblich von der Qualität der Digitalisierung abhängig, weshalb im Folgenden ausführlich auf die technischen Rahmenbedingungen der Digitalisierung eingegangen werden soll.

Technische Rahmenbedingungen

Die Digitalisierung der Inkunabeln inklusive der Einbände erfolgte nach den DFG- Praxisregeln mit mindestens 300 dpi.TIFF. Dabei sind die Digitalisate im JPEG-Format in den drei Auflösungsmodi 600, 1024 und 2000 Pixel verfügbar. Insgesamt werden allein für die Rohdatenmenge des Gesamtprojektes ungefähr 3,5 Terrabyte (TB) Speicherplatz benötigt, für die Sicherstellung der Daten noch einmal die gleiche Speichermenge. Die Digitalisate wurden in Form von 24 BitDateien langfristig gespeichert. Das langfristige Speicherkonzept dient nicht nur der Datensicherung, sondern durch die Speicherung der Rohdaten kann auf zukünftige technologische Fortschritte bei der Hardwareentwicklung Rücksicht genommen werden. Das folgende Beispiel aus einer französischen Ausgabe der Genealogiae

deorum von Boccaccio³ soll die Qualität der Digitalisate verdeutlichen:

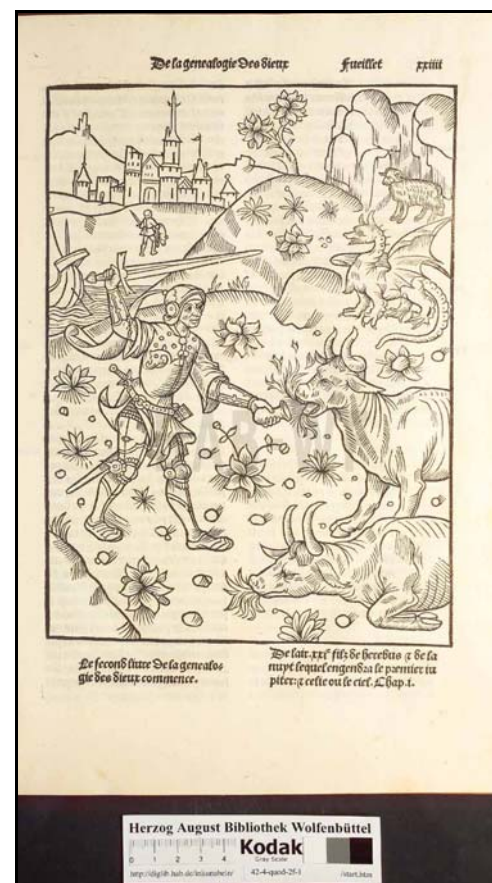


Abbildung 1: Digitalisat aus Giovanni Boccaccios *Boccace de la geneologie Des Dieux*, 1498

³ Giovanni Boccaccio, *Boccace de la geneologie Des Dieux*, Paris 1498, hier S. 49 der digitalen Ausgabe: <<http://diglib.hab.de/wdb.php?dir=inkunabeln/42-4-quod-2f-1>> (14.02.2006).

Voraussetzung für eine preiswerte und schnelle Digitalisierung des empfindlichen Schriftguts war eine spezielle Buchwippe, der so genannte Wolfenbüttler Buchspiegel. Diese Buchwippe stellt einen wesentlichen Beschleunigungsfaktor dar, weil sie zügige Aufnahmen mit einem 45-Grad-Öffnungswinkel erlaubt. Dieses Verfahren stellt eine schonende Bearbeitung des Materials sicher; bei herkömmlichen Methoden muss das Werk häufig in einem 180° Winkel geöffnet werden, was bei empfindlichen Inkunabelausgaben zu Beschädigungen des Einbandes geführt hätte. Das Material wurde farbig digitalisiert, da die Verwendung von farbigen Partien in den Inkunabeln bewusst zur Textstrukturierung eingesetzt wurde und diese Informationen so für die Forschung zugänglich sind. Auch für die buchhistorische Forschung sollte die Benutzung der Digitalisate einen Gewinn darstellen, da das Material in der angegebenen Qualität die normale Lektüre, auch bei kleinbuchstabigem Druck, ermöglicht. Gleichfalls sollten die Digitalisate in dieser Beschaffenheit für typenkundliche Untersuchungen ausreichen, solange keine speziellen visuellen Behandlungsschritte vorausgesetzt werden wie zum Beispiel Durchlichtaufnahmen zur Wasserzeichenbestimmung.

Inhaltliche Bearbeitung der Inkunabeln: Vom Rohdatensatz zum erschlossenen Digitalisat

Bereits eingangs wurde das Anliegen der vdlb erwähnt, die bereitgestellten Digitalisate möglichst anwendungsfreundlich im Netz zur Verfügung zu stellen. Aus diesem Grund wurden die fertigen Digitalisate zunächst unter Verwendung eines Image-Viewers mit dem Namen TOC (Table of Content)-Editor⁴ durchgesehen, um potenzielle Fehler bei der Digitalisierung festzustellen. Die Struktur der Dateien der Digitalisate nach dieser Bearbeitungsphase soll anhand der Seiten 15 bis 20 einer Ausgabe der Epigramma des Ausonius⁵ vorgestellt werden:

4 Eine Dokumentation zu den Einsatzmöglichkeiten und dem Aufbau des TOC-Editors befindet sich auf der Projektseite: <<http://inkunabeln.ub.uni-koeln.de/>> unter den Punkten Projekt vdlb/ Erschließung/TOC-Editor.

5 Ausonius, Avsonii Peonii Poetae Disertissimi Epigrammata, Venedig 1496.

```
[...]
    <div level="0" >
    <page>
    <image extref="00015.jpg" seqno="019"
nativeno="15" />
    </page>
    <page>
    <image extref="00016.jpg" seqno="020"
nativeno="16" />
    </page>
    <page>
    <image extref="00017.jpg" seqno="021"
nativeno="17" />
    </page>
    <page>
    <image extref="00018.jpg" seqno="022"
nativeno="18" />
    </page>
    <page>
    <image extref="00019.jpg" seqno="023"
nativeno="19" />
    </page>
    <page>
    <image extref="00020.jpg" seqno="024"
nativeno="20" />
    </page>
[...]
```

Wie der Name TOC-Editor schon nahe legt, wurden die Digitalisate mit diesem Programm nicht nur auf mögliche Digitalisierungsfehler untersucht, sondern auch weiter bearbeitet. Das Programm erlaubt ein gewisses Maß an inhaltlicher Erschließung der digitalisierten Inkunabeln, das allerdings vom Erschließungsausmaß nicht mit Handschriftenbeschreibungen verglichen werden sollte. Eine solch detaillierte Bearbeitung der Digitalisate hätte in keinem Verhältnis zu den Projektzielen und -vorgaben gestanden. Vielmehr ging es um die Bereitstellung von Navigationshilfen, die es den Benutzern/innen ermöglichen, sich schnell in die Textstruktur einzufinden und so den gewünschten Textabschnitt in kurzer Zeit zu finden.

In der Regel wurde daher mindestens alle zehn Seiten ein Textab-

schnitt des Digitalisates transkribiert. Diese Regel wurde jedoch nicht stringent beachtet: Wo texteigene Übersichtsstrukturen vorhanden waren, wurde auf diese zurückgegriffen, um so eine möglichst große Nähe zur Quellenstruktur und deren Inhalt zu wahren. Doch gab es vor allem bei den frühen Inkunabeln häufig Ausgaben, die weder Inhaltsübersichten noch markante Textabschnitte in verwendbaren Seitenabständen aufwiesen. In diesen Fällen erfolgte die Transkription der betreffenden Seitenanfänge. Maximal wurden drei Texteinträge pro Seite aufgeführt, also zum Beispiel die Überschrift des anfangenden Buchteils, dann die Überschrift des Kapitels und dessen erster Satz. Zusätzlich erlaubte es das Programm, bestimmte Kommentare und Schlagworte zu den einzelnen Seiten zu verfassen. Auf diese Weise konnten fehlende oder vertauschte Lagen, Druckfehler, das Vorhandensein eines Inhaltsverzeichnisses usw. festgehalten werden. In der gleichen Weise wurde bei denjenigen Inkunabeln, für die eine moderne Edition greifbar war, die Kommentierfunktion verwendet, um die transkribierten Passagen in der jeweiligen Edition nachweisen zu können. Bei einigen Literaturarten wurden auch themenspezifische Merkmale der dazugehörigen Inkunabeln mittels der Kommentier- bzw. Schlagwortfunktion aufgeführt. So wurden zum Beispiel bei der Sermonesliteratur in entsprechenden Abständen die Feier- und Heiligentage in der gebräuchlichen deutschen Bezeichnung wiedergegeben und zusätzlich die erste vorgeschlagene Bibelstelle für die Predigt des jeweiligen Tages entsprechend den Loccumer Richtlinien aufgeführt.

Außer den eigenen Eintragungen erlaubte es das Programm, vorhandene Eigenheiten der jeweiligen Inkunabel – wie Provenienzen, Kolophone, Musiknoten, Illustrationen, Schmuckinitialien, Druckerzeichen, Annotationen und das Titelblatt – festzuhalten. Diese Eigenschaften wurden allerdings – aufgrund der quantitativen Ausrichtung der vdIb – lediglich vermerkt und nicht ausgewertet. So erfahren die Nutzer/innen zwar, auf welchen Seiten sie zum Beispiel Annotationen finden, aber deren Inhalt und Gegenstand müssen sie selbst ergründen. Die aufgezählten Erschließungsfunktionen helfen ihnen bei der schnellen Orientierung innerhalb des Textes. Die eingegebenen Transkriptionen und die vermerkten Texteeigenschaften können entweder mit den entsprechenden Seiten

zusammen oder aber ohne diese in einer Übersicht aufgerufen werden. Die Digitalisate bieten somit den Benutzern/innen teilweise eine bessere Möglichkeit mit dem Inhalt des Textes zu arbeiten als die Originale. Als Beleg dafür sei der bereits vorgestellte Abschnitt der Ausonius-Ausgabe in seiner fertigen Bearbeitung angeführt, vorangestellt ist die 15. Seite der digitalen Ausonius-Ausgabe⁶ mit Bearbeitungsfenster des TOC-Editors:

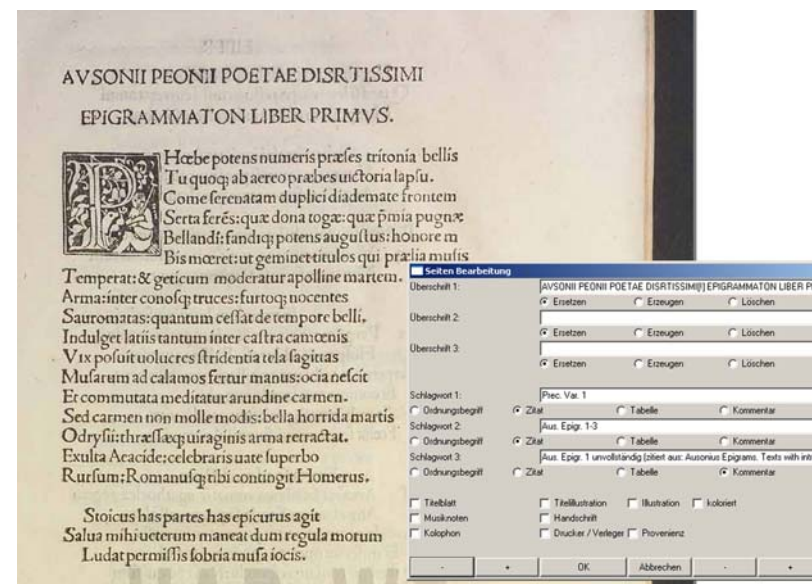


Abbildung 2: Seite der digitalen Ausonius-Ausgabe mit Bearbeitungsfenster des TOC-Editors

```
[...]
<div level="0" >
<head>AVSONII PEONII POETAE DISRTISSIMI[!]
  EPIGRAMMATON LIBER PRIMVS</head>
<div level="1" >
  <page>
```

6 <<http://diglib.hab.de/wdb.php?dir=inkunabeln/12-5-poet-2f-1>> (14.02.2006).

```

    <image extref="00015.jpg" seqno="019"
nativeno="15" />
<text type="reference" >Prec. Var. 1</text>
<text type="reference" >Aus. Epigr. 1-3</text>
<text type="comment" >Aus. Epigr. 1 unvollständig
(zitiert aus: Ausonius Epigrams. Texts with
introduction and commentary by N. M. Kay. -
London 2001)</text>
</page>
  <page>
    <image extref="00016.jpg" seqno="020" na-
tiveno="16" />
<text type="reference" >Aus. Epigr. 5-6, 8-9,
45</text>
<text type="comment" >Eingeschoben: Ignota
aeternae ne sint tibi tempora Romae/ Regibus et
patrum ducta sub imperiis</text>
</page>
  <page>
    <image extref="00017.jpg" seqno="021"
nativeno="17" />
<text type="reference" >Aus. Epigr. 4, 10-
11</text>
<text type="comment" >Eingeschoben: Hactenus
ascripsi fastos: si fors uolet: ultra / Adiiciam:
si non: qui legis: adiicies</text>
<text type="comment" >Eingeschoben: Urbis ab
aeternae deductam rege Quirino / Annorum seriem
cum procul accipies.</text>
</page>
  <page>
    <image extref="00018.jpg" seqno="022"
nativeno="18" />
<text type="reference" >Aus. Epigr. 12-16</text>
</page>
  <page>
    <image extref="00019.jpg" seqno="023"
nativeno="19" />
<text type="reference" >Aus. Epigr. 17-21</text>
</page>
  <page>
    <image extref="00020.jpg" seqno="024"

```

```

nativeno="20" />
<text type="reference" >Aus. Epigr. 22-26</text>
</page>
[...]
```

Das Beispiel verdeutlicht, wie die einzelnen Komponenten zusammenwirken und das Dokument erschließen. Im Folgenden sollen die daraus resultierenden Recherchemöglichkeiten vorgestellt werden.

Recherchemöglichkeiten innerhalb der vdlb

Aus den im vorangegangenen Kapitel dargestellten Bestandteilen der inhaltlichen Erschließung der Inkunabeln resultiert eine große Daten- und Informationsmenge, die zu jedem einzelnen Digitalisat verfügbar ist. Damit diese in ihrer ganzen Bandbreite genutzt werden können, sollen durch den Einbau von verschiedenen Recherchemöglichkeiten auf der Internetseite des Projektes alle eingetragenen Daten einzeln gesucht werden können.⁷ Es soll also möglich sein, sowohl nach den üblichen Formulareintragen als auch nach den Eintragungen der inhaltlichen Erschließung zu suchen. Grundsätzlich sind dafür drei Suchmöglichkeiten auf der Projektseite verfügbar: eine Indexsuche, eine Maskensuche und eine Einfache Suche. Das Suchinstrument der Einfachen Suche folgt „modifizierten Google-regeln“. Die Index- und Maskensuche erlauben die Recherche nach ausgewählten Kategorien, die sich an dem oben formulierten Anspruch orientieren. Ergänzend dazu unterstützen die Suchmaschinen die bereits bestehenden Ressourcen zur Inkunabelkunde.

Einbindung der vdlb in bestehende Ressourcen

Die vdlb ist eng verbunden mit bereits existierenden Internetressourcen zum Thema Inkunabelkunde. Dementsprechend kam den ISTC-Nummern (Incunabula Short Title Catalogue) eine wichtige Funktion bei der Digitalisierungskampagne zu, denn die ISTC-Titelaufnahmen wurden für alle nachgewiesenen Inkunabeln der Projektpartner in Form einer suchbaren Komponente des Gesamtserver implementiert. So wird nach dem Eintrag

⁷ Diese Arbeiten sind zurzeit (Stand: Februar 2006) noch nicht im vollen Umfang abgeschlossen, jedoch wird dies zeitnah geschehen.

der ISTC-Nummern in die XML-Struktur die komplette ISTC-Beschreibung inklusive der umfangreichen bibliografischen Angaben zugewiesen. Sie kann somit jederzeit auf der Projektseite abgerufen werden. Im Gegenzug wurden die digitalisierten Inkunabeln dem ISTC für die weitere Informationsverarbeitung gemeldet. Die vorgestellten Suchmaschinen unterstützen folglich auch die Recherche nach ISTC-Nummern. Des Weiteren kann auch nach den GW-Nummern⁸ (Gesamtkatalog der Wiegendrucke) und den BSB-Ink-Signaturen⁹ (Bayerische Staatsbibliothek, Inkunabelkatalog) der jeweiligen Inkunabeln gesucht und geordnet werden. Umgekehrt können die Produkte der vdIb über ISTC-online¹⁰ und INKA¹¹ (Inkunabelkatalog deutscher Bibliotheken) gefunden werden.

Im Laufe des Projektes wurde in Eigenleistung auch eine Neukatalogisierung der Inkunabeln im PICA-OPAC erbracht. Damit verbunden ist der unmittelbare Zugriff auf alle Digitalisate über den Katalog der HAB sowie den GBV möglich. Zusätzlich gibt es einen Linkapparat, in dem bereits bestehende digitalisierte Inkunabeln anderer Internetressourcen und auch digitalisierte Nachschlagewerke angegeben sind.

Fazit

Wie aus den vorangegangenen Ausführungen ersichtlich wird, versteht sich die vdIb nicht als ein von seiner Umwelt und von weiteren Entwicklungen losgelöstes Projekt, sondern vielmehr wird die Einbindung bzw. Verknüpfung in zukünftige Digitalisierungsvorhaben ausdrücklich erwünscht. Die Anzahl der existierenden Inkunabeln ist überschaubar, außerdem sind sie bibliografisch hervorragend erschlossen, so dass ein bundesweites „Inkunabelportal“ in der mittelbaren Zukunft möglich ist. Die vdIb begreift sich als ein Baustein auf dem Weg dazu. Das Angebot bietet Orientierungsmöglichkeiten, aber sicher auch die Chance der kritischen Auseinandersetzung. Der Grundsatz der vdIb, auf eine inhaltliche Erschließung zu verzichten und stattdessen eine größere Quantität an Digitalisaten zu erreichen, birgt

8 <<http://www.gesamtkatalogderwiegendrucke.de/>> (14.02.2006).

9 <<http://mdz1.bib-bvb.de/cocoon/bsbink/start.html>> (14.02.2006).

10 <<http://www.bl.uk/catalogues/istc/index.html>> (14.02.2006).

11 <<http://www.inka.uni-tuebingen.de/>> (14.02.2006).

sicherlich Gefahren in sich, macht aber auf der anderen Seite auch deutlich, dass es möglich ist, innerhalb von verhältnismäßig kurzer Zeit signifikante Bestände zu digitalisieren und dabei die Benutzungsfreundlichkeit nicht außer Acht zu lassen. Es wäre wünschenswert, wenn andere Projekte das Angebot der vdIb aufgreifen und es so in der näheren Zukunft möglich wäre, weltweit auf die bestehenden Inkunabelbestände zuzugreifen. Auch wenn dies zum gegenwärtigen Zeitpunkt noch eher eine Vision denn nahe Realität ist, so haben die Arbeit an dem Projekt und das Ergebnis gezeigt, dass es durchaus die Chance gibt, diese Vision zu verwirklichen.

Anhang

Zur weiteren Verdeutlichung seien an dieser Stelle zwei Screenshots von digitalisierten Inkunabelseiten mit Illustrationen angeführt, die der digitalen Bibliothek der HAB entstammen.

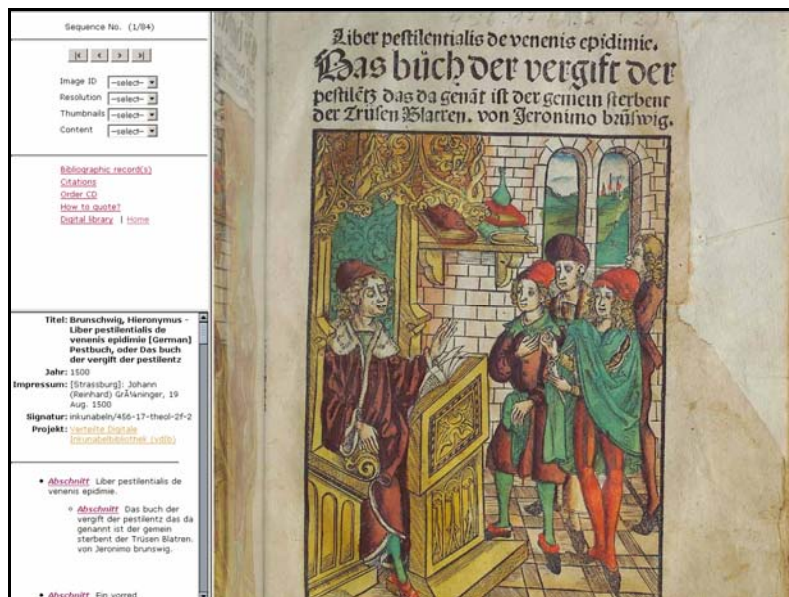


Abbildung I: Titelblatt des Pestbuches von Hieronymus Brunswig¹²

¹² Hieronymus Brunswig, Liber pestilentialis de venenis epidimie. Das buch der vergift der pestile[n]tz das da gena[n]t ist der gemein sterbent der Trüsen Blatren. Straßburg 1500. Hier S. 1 der digitalen Ausgabe: <<http://diglib.hab.de/wdb.php?dir=inkunabeln/456-17-theol-2f-2>> (14.02.2006).



Abbildung II: Der Anfang des Kapitels über unnütze Bücher im Narrenschiff des Sebastian Brant¹³

Timo Steyer ist Wissenschaftlicher Mitarbeiter im Forschungsprojekt der „Verteilten Digitalen Inkunabelbibliothek“ an der Herzog August Bibliothek Wolfenbüttel. Ein weiterer Interessenschwerpunkt des Autors liegt in der Erforschung von Ausdrucksformen der spätmittelalterlichen und frühneuzeitlichen Volksfrömmigkeit. E-Mail: Timo.Steyer@gmx.de

¹³ Sebastian Brant, Stultifera Nauis Narragonice p[ro]fect[i]onis nunq[ua]m satis laudata Nauis, Basel 1498. Hier S. 21 der digitalen Ausgabe: <<http://diglib.hab.de/wdb.php?dir=inkunabeln/30-4-poet-2>> (14.02.2006).

ZEITSCHRIFTEN DER AUFKLÄRUNG IM NETZ – RETROSPEKTIVE
DIGITALISIERUNG WISSENSCHAFTLICHER REZENSIONSORGANE
UND LITERATURZEITSCHRIFTEN

von Sabine Rahmsdorf

Mit dem an der Universitätsbibliothek Bielefeld durchgeführten und von der DFG geförderten Projekt „Retrospektive Digitalisierung wissenschaftlicher Rezensionsorgane und Literaturzeitschriften des 18. und 19. Jahrhunderts aus dem deutschen Sprachraum“ ist eine im Internet frei zugängliche Sammlung der vollständigen Textkorpora wichtiger Rezensionsorgane und Literaturzeitschriften der deutschen Aufklärung entstanden. Verfügbar sind bisher 45 Zeitschriften, die auf der Grundlage von Mikroverfilmungen des Georg Olms Verlags digitalisiert wurden. Die inhaltliche Erschließung wird weitgehend durch den an der Akademie der Wissenschaften zu Göttingen erstellten „Index deutschsprachiger Zeitschriften 1750-1815“ geleistet, der in das Projekt eingeflossen ist. Die Verbindung von digitaler Reproduktion der Zeitschriften mit hochwertigen inhaltserschließenden Daten eröffnet der Forschung neue Rezeptionsmöglichkeiten. Der Beitrag stellt die entstandene digitale Sammlung mit ihren Recherchemöglichkeiten vor und berichtet über die im Projektverlauf gewonnenen Erfahrungen mit der Bereitstellung und Erschließung historischer Quellen in digitaler Form.

Die Zeitschrift ist ein zentrales Kommunikationsmedium der Epoche der Aufklärung. Dies wird dokumentiert in einem vielfältigen Spektrum literarischer Zeitschriften und Rezensionsorgane, gelehrter oder historisch-politischer Journale, das sich im 18. und frühen 19. Jahrhundert im deut-

schen Sprachraum (und natürlich nicht nur hier) entfaltet.¹ Neben langjährigen und überregional bedeutsamen Zeitschriften stehen viele kurzlebige Unternehmungen oder solche mit nur einem lokal eng begrenzten Rezipientenkreis. Die für die Epoche der Aufklärung so kennzeichnende Blüte der Zeitschriften ist dabei Katalysator und Ergebnis einer sich konstituierenden gelehrten bürgerlichen Öffentlichkeit, die für ihr Bedürfnis nach öffentlicher Diskussion und Information in der Zeitschrift mit ihrer inhaltlich flexiblen Form und ihrer durch die periodische Erscheinungsweise angestrebten Aktualität das ideale Medium fand. Gerade die so dezentrale „Gelehrtenrepublik“ in den deutschen Ländern konnte in den literarischen und gelehrten Zeitschriften die wesentlichen Diskurse der Epoche führen.² Die Bedeutung der Zeitschriften für die Erforschung des Jahrhunderts der Aufklärung sowie der angrenzenden Epoche liegt daher seit langem auf der Hand. Neben die Analyse einzelner Zeitschriftenbeiträge tritt dabei die Betrachtung des Mediums selbst, denn die Untersuchung der Beziehungen zwischen Herausgebern/innen, Beiträgern/innen und Rezipienten/innen verschafft Aufschluss über persönliche Verbindungen und bestehende Netzwerke zwischen den Trägern der Aufklärungsdiskussion und der daran Anteil nehmenden Öffentlichkeit. Vergleichende Untersuchungen der oft genug aufeinander Bezug nehmenden Beiträge in den verschiedensten Zeitschriften können wertvolle Erkenntnisse über den Verlauf einzelner Debatten und Diskursthemen und ihrer Rezeption erbringen.³

-
- 1 Einen fundierten Überblick bietet Kirchner, Joachim, Die Zeitschriften des deutschen Sprachgebietes von den Anfängen bis 1830 (= Ders. (Hg.), Bibliographie der Zeitschriften des deutschen Sprachgebietes bis 1900, Bd. 1, Stuttgart 1969).
 - 2 Einführend zur Thematik siehe Wilke, Jürgen, Literarische Zeitschriften des 18. Jahrhunderts (1688-1789), Stuttgart 1978; Hocks, Paul; Schmidt, Peter, Literarische und politische Zeitschriften 1789-1805. Von der politischen Revolution zur Literaturrevolution, Stuttgart 1975.
 - 3 Auf den Stand der Forschung kann an dieser Stelle nicht eingegangen werden, beispielhaft seien einige neuere Arbeiten zur Thematik angeführt: Schneider, Ute, Friedrich Nicolais Allgemeine Deutsche Bibliothek als Integrationsmedium der Gelehrtenrepublik, (Mainzer Studien zur Buchwissenschaft 1), Wiesbaden 1995; Heinz, Andrea (Hg.), „Der Teutsche Merkur“ – die erste deutsche Kulturzeitschrift?, (Ereignis Weimar-Jena. Kultur um 1800 2), Heidelberg 2003; Matuschek, Stefan (Hg.), Organisation der Kritik. Die Allgemeine Literatur-

Der Bedeutung der Zeitschriften für die Aufklärungsforschung steht die Zugänglichkeit dieses Quellenmaterials entgegen. Das in seinem Erhaltungszustand fragile Material findet sich als historisch gewachsener und verstreuter Besitz in Bibliotheken und ist für die Forschung im Original nur erschwert zugänglich. Insbesondere bei vergleichenden Fragestellungen stellt die Arbeit mit Originalen an verschiedenen Standorten eine Erschwernis dar. Daher hat es in der Vergangenheit Bemühungen gegeben, dieses Quellenmaterial besser zu erschließen. Das Unternehmen „Index deutschsprachiger Zeitschriften 1750-1815“⁴ der Akademie der Wissenschaften zu Göttingen, von 1976 bis 1987 gefördert durch die Stiftung Volkswagenwerk und die Deutsche Forschungsgemeinschaft (DFG), hatte die „Erschließung der in den Zeitschriften der deutschen Aufklärung sachlich diskutierten Themen und dichterisch behandelten Motive in repräsentativer und dennoch möglichst breiter Auswahl“⁵ zum Ziel. Nicht erfasst wurden die moralischen Wochenschriften und Rezensionsorgane. Im Ergebnis konnten 195 Zeitschriften mit fast 100.000 Beiträgen erschlossen werden, dies entspricht „etwa 20% der in dieser Zeit publizierten Inhalte“⁶. Alle in den ausgewerteten Zeitschriften enthaltenen Beiträge einschließlich Rezensionen wurden bibliografisch erfasst, Autoren/innen und beteiligte Personen anonymer Beiträge wo möglich ermittelt und inhaltserschließende Schlagwörter vergeben. Ein weiteres Großprojekt ist die von der Kulturstiftung der Länder herausgegebene und vom Georg Olms Verlag verlegte Sammlung „Deutsche Zeitschriften des 18. und 19. Jahrhunderts“, die über 500 Zeitschriften als Mikrofiche-Volltextverfilmung umfasst.⁷

In den vergangenen rund zehn Jahren ist mit der Etablierung des In-

Zeitung in Jena 1785-1803, (Ereignis Weimar-Jena. Kultur um 1800 5), Heidelberg 2004.

4 Akademie der Wissenschaften zu Göttingen (Hg.), Index deutschsprachiger Zeitschriften 1750-1815. Erstellt durch eine Arbeitsgruppe unter Leitung von Klaus Schmidt, Hildesheim 1990.

5 Ebd., Begleitheft S.VII.

6 Ebd., Begleitheft S.VIII.

7 Kulturstiftung der Länder (Hg.), Deutsche Zeitschriften des 18. und 19. Jahrhunderts, Hildesheim 1994ff.

ternet als Informations- und Kommunikationsmittel auch der Wissenschaft die digitale Reproduktion historischer Literatur und anderen Quellenmaterials und dessen Veröffentlichung als Onlinepublikation zu einer wichtigen Möglichkeit geworden, auch schwer zugängliches Quellenmaterial prinzipiell weltweit zugänglich zu machen. Bibliotheken – auf kommerzielle Produkte auf diesem Sektor soll an dieser Stelle nicht eingegangen werden – stellen in verschiedenen Digitalisierungsprojekten wichtige Teile ihrer historischen Bestände als digitale Reproduktionen einer interessierten Öffentlichkeit frei im Internet zur Verfügung.⁸ In Deutschland hat die DFG durch ihr Förderprogramm „Retrospektive Digitalisierung von Bibliotheksbeständen“, das 2004 in dem neuen Förderprogramm „Kulturelle Überlieferung“ aufgegangen ist, maßgeblich zur Initiierung und Finanzierung dieser Aktivitäten beigetragen. Mit dem Aufbau von zwei nationalen Digitalisierungszentren, dem Göttinger Digitalisierungszentrum (GDZ) der Niedersächsischen Staats- und Universitätsbibliothek Göttingen und dem Münchner Digitalisierungszentrum an der Bayerischen Staatsbibliothek München, wurden zwei Kompetenz- und Servicezentren für retrospektive Digitalisierung von Bibliotheksbeständen geschaffen. Auf den Webseiten der beiden Digitalisierungszentren finden sich jeweils Übersichten über abgeschlossene und laufende Digitalisierungsprojekte.⁹ Um die Vielzahl der verschiedenen Digitalisierungsprojekte zusammenzuführen und einen zentralen Zugriff auf die in digitalisierter Form vorliegenden Drucke und anderes Quellenmaterial zu ermöglichen, befindet sich zurzeit, gefördert wiederum von der DFG, ein Zentrales Verzeichnis digitalisierter Drucke im Aufbau. Beteiligt sind hieran die AG Sammlung Deutscher Drucke (SDD), die Verbundzentrale des Gemeinsamen Bibliotheksverbundes (VZG) und das Hochschulbibliothekszentrum (HBZ) des Landes Nord-

8 Eines der international bekanntesten und bedeutendsten Digitalisierungsprojekte ist die von der französischen Nationalbibliothek erstellte Sammlung gallica, <<http://gallica.bnf.fr/>> (16.02.2006).

9 Göttinger Digitalisierungszentrum (GDZ), Niedersächsische Staats- und Universitätsbibliothek Göttingen, <<http://gdz.sub.uni-goettingen.de/de/index.html>> (21.02.2006); Münchner Digitalisierungszentrum, Bayerische Staatsbibliothek München, <<http://www.bsb-muenchen.de/mdz/>> (21.02.2006).

rhein-Westfalen. Eine erste Testversion ist im Internet verfügbar.¹⁰

Die Vorteile einer Bereitstellung digitalisierter Quellen oder historischer Buchbestände im World Wide Web liegen auf der Hand. Neben der verbesserten Zugänglichkeit kann durch die Erstellung virtueller Kompendien von Quellen nicht nur im Original verstreutes, aber thematisch zusammengehöriges Material zusammengeführt werden, durch übergreifende Suchmöglichkeiten lassen sich Einzelwerke und Sammlungen zudem in neuartiger Weise inhaltlich erschließen und können so auch die Bearbeitung neuer Fragestellungen ermöglichen. Einschränkend muss jedoch gesagt werden, dass digitalisierte historische Dokumente überwiegend noch als digitales Image, also als Bilddatei der einzelnen Seiten eines Dokuments angeboten werden. Die inhaltliche Erschließung digitaler Quellen konzentriert sich damit aber noch ganz überwiegend auf die Möglichkeit der Stichwortsuche in zusätzlich maschinenlesbar erfassten Inhaltsverzeichnissen und Registern oder in der Suche nach ergänzend angebotenen Schlagwörtern und Thesauri. Die Konvertierung der als Image vorliegenden Seiten in einen maschinenlesbaren Volltext mit Hilfe von Optical Character Recognition-Software (OCR) stößt bei historischen Quellen und Dokumenten aufgrund der verwendeten nicht normierten Schriften, Schreibweisen oder Abkürzungen nach wie vor auf große Schwierigkeiten. Genau dies ist aber die Voraussetzung, um eine unkomplizierte Möglichkeit der Volltextsuche zu ermöglichen, da sich zumal bei öffentlich geförderten Projekten die grundsätzlich denkbare aber kostenintensive Texterfassung des gesamten Inhalts eines digitalisierten Dokuments zumeist verbietet. In diesen Kontext bibliotheksseitiger Digitalisierungsprojekte ist auch das hier vorzustellende Projekt der Universitätsbibliothek Bielefeld einzuordnen.

Das Projekt

Das an der Universitätsbibliothek Bielefeld von 2000 bis 2003 durchgeführte Projekt „Retrospektive Digitalisierung wissenschaftlicher Rezensi-

10 Zentrales Verzeichnis digitalisierter Drucke, <<http://www.digitalisiertedrucke.de/>> (16.02.2006).

onsorgane und Literaturzeitschriften des 18. und 19. Jahrhunderts aus dem deutschen Sprachraum“ (Phase 1)¹¹ hatte die digitale Rekonstruktion der vollständigen Korpora zentraler wissenschaftlicher Rezensionsorgane und Literaturzeitschriften der deutschen Aufklärung und folgender Epochen zum Ziel. Berücksichtigt wurden Zeitschriften aus der zweiten Hälfte des 18. und aus dem frühen 19. Jahrhundert. Die digitalisierten Quellen sollten frei zugänglich über das Internet zur Verfügung gestellt werden und durch Suchmöglichkeiten sowie durch ein Browsing erschlossen werden. Das Projekt wurde von der DFG im Rahmen des Schwerpunkts „Retrospektive Digitalisierung von Bibliotheksbeständen“ gefördert. Im Ergebnis stehen 45 Zeitschriften zur Verfügung mit der Möglichkeit zur Recherche in insgesamt 82.000 Artikeln bzw. 460.000 Seiten. Digitalisiert wurden zum Beispiel die von Friedrich Nicolai herausgegebene Allgemeine deutsche Bibliothek, die Bibliothek der schönen Wissenschaften und der freyen Künste, das Deutsche Museum / Neues Deutsches Museum oder der Teutsche Merkur / Neue Teutsche Merkur.¹²

Das Projekt wurde durchgeführt in Kooperation mit dem Georg Olms Verlag und mit der Akademie der Wissenschaften zu Göttingen. Die Digitalisierung der Zeitschriften erfolgte nicht anhand der Originale, sondern auf der Grundlage von Mikrofiches aus der von der Kulturstiftung der Länder herausgegebenen Mikrofiche-Edition „Deutsche Zeitschriften des 18. und 19. Jahrhunderts“, die vom Olms Verlag zur Verfügung gestellt wurden. Die recherchierbaren Daten wiederum sind Ergebnis der Erschließungsarbeiten des Unternehmens „Index deutschsprachiger Zeitschriften 1750-1815“ der Akademie der Wissenschaften zu Göttingen. Damit ist es gelungen, die wertvollen Erschließungsdaten des „Index“ nachzunutzen und in neuer Form zu präsentieren und durch die Verbindung mit den Volltexten der Zeitschriften einen Mehrwert in dem digitalen Rechercheangebot zu erzielen.

11 Retrospektive Digitalisierung wissenschaftlicher Rezensionsorgane und Literaturzeitschriften des 18. und 19. Jahrhunderts aus dem deutschen Sprachraum, <<http://www.ub.uni-bielefeld.de/diglib/aufklaerung/>> (16.02.2006).

12 Eine Auflistung der zur Verfügung stehenden Zeitschriften ist unter der oben angegebenen URL des Projekts einsehbar.

Das Rechercheangebot

Zur Navigation und Recherche in den digitalisierten Zeitschriften werden sowohl Möglichkeiten zum Browsing als auch eine Suchoberfläche angeboten. Beim Browsing kann ausgehend von der Auflistung aller Zeitschriften eine Übersicht über die Bände einer Zeitschrift und auf der nächsten Ebene über die Inhalte des Einzelbandes angesteuert werden. Ausgehend vom Eintrag eines Beitrags im Inhaltsverzeichnis kann der zugehörige Volltext aufgerufen werden. Innerhalb der digitalisierten Dokumente kann Seite für Seite und in Schritten von fünf bzw. zehn Seiten vorwärts und rückwärts geblättert werden, auf diese Weise kann ein kompletter Zeitschriftenband auch Seite für Seite betrachtet werden.

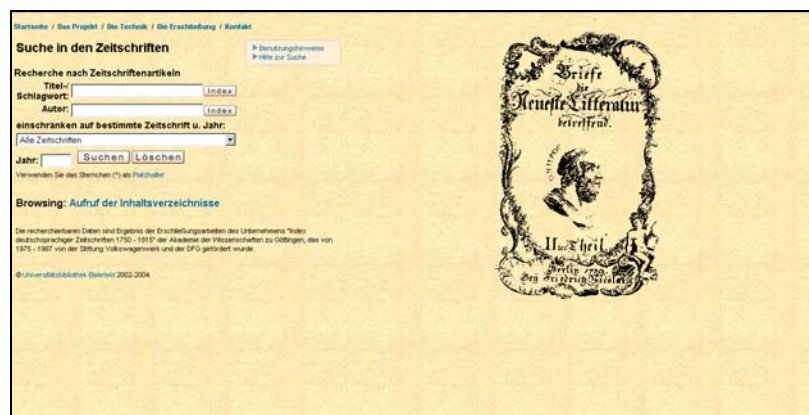


Abbildung 1: Die Suchoberfläche mit Möglichkeit zum Browsing über die Zeitschriften

Die Suchoberfläche bietet die Möglichkeit der personenbezogenen oder thematischen Suche nach einzelnen Zeitschriftenbeiträgen. Möglich ist die Suche nach dem Autor bzw. der Autorin eines Beitrags oder einer Rezension sowie nach sonstigen an der Erstellung eines Beitrags beteiligten Personen, wie beispielsweise einem Übersetzer oder einer Übersetzerin. Das Suchfeld „Titel-/Schlagwort“ ermöglicht die Suche nach Stichwörtern

aus dem Titel eines Beitrags und nach inhaltserschließenden Schlagwörtern, wie sie im „Index“ vergeben wurden. Die Schlagwörter bezeichnen zum Beispiel geografische Bezüge oder sind Personen- oder Sachschlagwörter. Jede Suchanfrage kann entweder über den gesamten Zeitschriftenbestand erfolgen oder auf eine bestimmte Zeitschrift und / oder bestimmte Erscheinungsjahre eingegrenzt werden. Als Ergebnis einer Recherche wird eine Kurztitelliste der Treffer angezeigt, aus der heraus die Vollanzeige jedes Treffers mit den vollständigen bibliografischen Angaben des Beitrags einschließlich Schlagwörter und dem Link zum zugehörigen Volltext aufgerufen werden kann. Die Recherche in den digitalisierten Zeitschriften erfolgt auf der Basis der erschließenden Metadaten, also anhand der bibliografischen Angaben einschließlich inhaltserschließender Schlagwörter zu jedem einzelnen Zeitschriftenbeitrag. Eine Suche in den Volltexten der Zeitschriften wird bisher nicht angeboten.

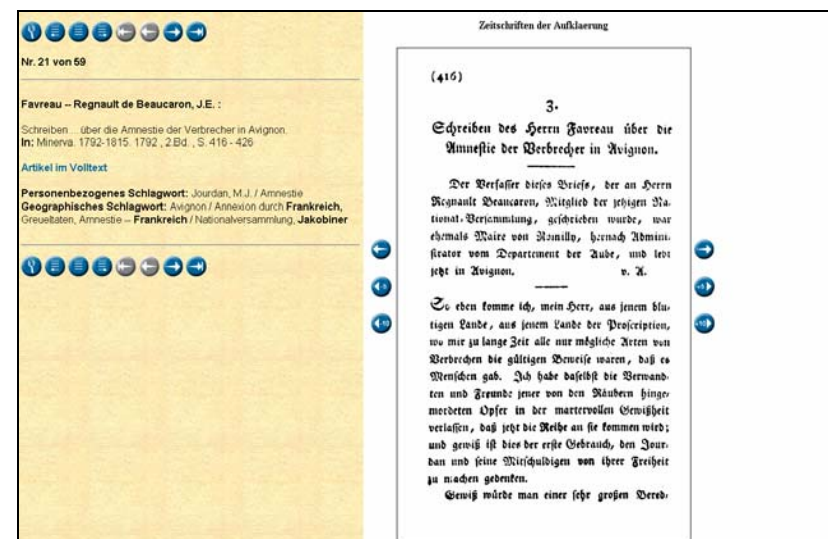


Abbildung 2: Trefferanzeige mit zugehörigem Volltext und Funktion zum Blättern im Dokument

Die technische Realisierung

Zur technischen Realisierung des Projekts waren verschiedene Schritte erforderlich, die im Folgenden skizziert werden sollen.

Scannen vom Mikrofiche und Speicherung von Einzelseiten als Bitmap-Dateien:

- Die technischen Parameter für das Scannen der vom Olms Verlag bereitgestellten Masterfiches durch einen externen Dienstleister folgten den üblichen und von der DFG empfohlenen Vorgaben¹³: Die Scans der Einzelseiten wurden bitonal (schwarz/weiß), mit 600 dpi Auflösung als Bitmap-Dateien im TIFF-Level-4-Format erstellt. Das Tagged Image File Format (TIFF) ist für die langfristige Speicherung von Bilddaten in der Regel das zu empfehlende Format, das zudem eine hohe Bildqualität gewährleistet und als Ausgangsformat für die Erstellung komprimierter Bildformate genutzt werden kann, die für die schnelle Datenübertragung über das Internet eher geeignet sind. Die Benennung der Bitmap-Dateien erfolgte rein numerisch, entsprechend der Reihenfolge der Originalseiten auf den vom Olms Verlag bereitgestellten Mikrofiches.
- Aufbereitung der inhaltserschließenden Daten:
Um die von der Akademie der Wissenschaften zu Göttingen als Textdateien bereitgestellten inhaltserschließenden Daten aus dem „Index deutschsprachiger Zeitschriften 1750-1815“ in einen durchsuchbaren Datenpool zu überführen und für die Einspielung in eine Datenbank vorzubereiten, wurden diese in das als internationaler Standard vorliegende Auszeichnungsformat SGML (Standard Generalized Markup Language) konvertiert. Für fünf Zeitschriften lagen keine oder nur unvollständige Erschließungsdaten aus dem „Index“ vor, darunter die Allgemeine deutsche Bibliothek. Hier mussten ersatzweise Inhaltsverzeichnisse bzw. Register auf der Grundlage geeigneter Erfassungsre-

13 Die aktuelle Version der „Praxisregeln im Förderprogramm ‚Kulturelle Überlieferung‘“ der DFG, die in erster Linie technische Hinweise zu den Bereichen Verfilmung und Digitalisierung geben, ist verfügbar unter: <http://www.dfg.de/forschungsfoerderung/formulare/download/12_151.pdf> (16.02.2006).

geln im Double-Key-Verfahren¹⁴ manuell erfasst und ebenfalls nach SGML konvertiert werden. Anschließend wurden alle in SGML vorliegenden Erschließungsdaten mit Hilfe von Perl-Skripten zur Einspielung in das eingesetzte Datenbanksystem BRS/Search (Bibliographic Retrieval System) konvertiert, mit dem zahlreiche Datenbanken der Universitätsbibliothek Bielefeld aufbereitet werden. Zu jedem Datensatz wurde die erste Seite des zugehörigen Aufsatzes entsprechend der Originalseitenzählung gespeichert. Die manuelle Erfassung von Inhaltsverzeichnissen und Registern und die Konvertierung nach SGML sowohl der Daten des „Index deutschsprachiger Zeitschriften 1750-1815“ wie auch der Daten der manuellen Erfassung war Bestandteil der Arbeiten eines externen Dienstleisters.

- Erzeugung der Paginierungskonkordanz:
Die Paginierungskonkordanz dient der Sicherstellung des seitenrichtigen Zugriffs von den inhaltserschließenden SGML-Dateien (Metadaten) via Hyperlink auf die jeweils zugehörigen TIFF-Dateien (Dokumentenvolltexte). Sie beinhaltet die exakte Zuordnung zwischen der rein numerischen Benennung der TIFF-Dateien einerseits und der Seitenzählung der Originale andererseits. Die Paginierungskonkordanz wurde für jeden Einzelband mittels intellektueller Prüfung erstellt.
- Generierung der Benutzeroberfläche:
Bei der Extraktion der in SGML vorliegenden Artikel-Metadaten mittels Perl-Skript für den Aufbau der BRS/Search-Datenbank wurden neben dem Datenbankeingabeformat statische, miteinander verknüpfte HTML-Seiten erzeugt, welche die Titel-, Band- und Heftstruktur widerspiegeln.
Die Web-basierte Benutzeroberfläche besteht aus statischen HTML-Eingangsseiten, die inhaltliche Informationen bereithalten und mit den

14 Beim Double-Key-Verfahren wird eine Vorlage von zwei Datentypisten unabhängig voneinander elektronisch erfasst (abgeschrieben) und der erfasste Text anschließend elektronisch gegeneinander abgeglichen. Durch diesen Abgleich können die bei jeder Datenerfassung auftretenden individuellen Fehler durch falsches Abschreiben minimiert und eine hohe Erfassungsgenauigkeit erzielt werden. Double-Keying ist ein anerkanntes Verfahren.

Funktionen für Datenbanksuche und Browsing in der bibliografischen Zeitschriftenstruktur verknüpft sind. Auf diesen Seiten und in der Datenbankergebnisanzeige werden Links auf ein CGI-basiertes Perlskript¹⁵ verankert, das als Parameter Pfad und Benennung der Startdatei erhält. Dieses Skript übernimmt die On-The-Fly-Umwandlung der als Speicherformat, nicht aber für die Darstellung der digitalisierten Dokumente im Internet geeigneten TIFF-Dateien in das WWW-gerechte Grafik-Austausch-Format GIF (Graphics Interchange Format) und steuert gleichzeitig die Navigation über die einzelnen Seiten der Dokumente. In der Benutzeroberfläche wird in einer Framestruktur im linken Fenster die Suche und Anzeige in der BRS/Search-Datenbank oder die Browsing-Navigation in der Struktur der Zeitschriften angeboten, im rechten Fenster erscheint die jeweils angewählte zugehörige Imagedatei mit der Darstellung des ausgewählten Zeitschriftenbeitrags.

Für jede Zeitschrift wurde eine eigene Frontdoor-Seite erstellt, die neben den bibliografischen Daten der Zeitschrift den Einstieg zu den HTML-Seiten mit der Übersicht über die Bandstruktur und den Inhaltsverzeichnissen der Einzelbände sowie zu den zugehörigen Volltexten der enthaltenen Artikel bietet. Die URL der Frontdoor-Seite wird auch in der Zeitschriftendatenbank (ZDB) und der Elektronischen Zeitschriftenbibliothek (EZB) nachgewiesen, so dass die digitalisierten Zeitschriftentitel auch in diesen überregionalen Katalogen recherchierbar sind. Über Internetsuchmaschinen sind die digitalisierten Zeitschriftentitel ebenfalls suchbar.

Blick zurück: Erfahrungen aus dem Projektverlauf

Nach Abschluss des Projekts ist es angebracht, aus den gewonnenen Erfahrungen mit der digitalen Bereitstellung eines historischen Quellenkorpus, wie es die zur Verfügung gestellten Zeitschriften der Aufklärung darstellen, ein Resümee zu ziehen. Insgesamt haben sich die gewählten technischen Vorgehensweisen und Arbeitsabläufe mit Blick auf Praktikabi-

¹⁵ Der Standard CGI (Common Gateway Interface) dient, allgemein gesprochen, der Ausführung externer Programme auf einem WWW-Server.

lität und erzielt es Ergebnis bewährt, einige Punkte jedoch illustrieren, welche besonderen Anforderungen und Schwierigkeiten die digitale Reproduktion gerade historischer Dokumente immer noch in sich birgt:

Das gewählte Verfahren der Digitalisierung vom Mikrofiche und nicht vom Original stellte im Kontext dieses Projekts ein sinnvolles Vorgehen dar, da auf diese Weise davon abgesehen werden konnte, die an verschiedenen Bibliotheksstandorten verfügbaren Originale für eine Digitalisierung mit hohem Aufwand erneut zusammenzustellen. Das Scannen von einer Sekundärform der Originalquelle erfordert jedoch ein besonders hohes Maß an Qualitätskontrolle für alle produzierten Imagedateien, um eine einwandfreie Präsentation zu gewährleisten. Entsprechend hoch war der Personal- und Zeitaufwand hierfür. Die Digitalisierung vom Mikrofiche ist daher in anderen Kontexten sorgfältig gegen andere Faktoren abzuwägen und kann nicht etwa generell als Alternative zum Scannen des Originals bezeichnet werden.

Der Workflow von der Digitalisierung der Mikrofiches bis zur Präsentation im Internet beinhaltete einen relativ hohen Anteil an manueller und intellektueller Bearbeitung: Für die manuelle Erfassung von Inhaltsverzeichnissen bzw. Registern der nicht im „Index“ erfassten Zeitschriften mussten genaue Erfassungsregeln definiert werden, die bereits die Logik der anschließenden SGML-Konvertierung berücksichtigen mussten. Dies war in besonderem Maße erforderlich, weil die manuelle Erfassung quasi „mechanisch“ durch Schreibkräfte ohne spezifische Sach- oder Sprachkenntnisse vorgenommen wurde und nicht etwa durch speziell geschultes Personal. Aufgrund der hohen Variabilität bei der Gestaltung der Inhaltsverzeichnisse und Register, beispielsweise hinsichtlich der verwendeten Abkürzungen und Verweisungszeichen, mussten diese Erfassungsregeln mit entsprechendem Personalaufwand jedoch jeweils für jede Zeitschrift und teilweise auch für jeden einzelnen Band individuell erstellt werden.

Die Erstellung der Paginierungskonkordanz als notwendigem Bindeglied zwischen inhaltserschließenden Metadaten und zugehörigem Volltext erfolgte durch den intellektuellen Abgleich zwischen den in den Metadaten enthaltenen Seitenzahlen, der Originalseitenzählung innerhalb der TIFF-Dateien und der numerischen Benennung der TIFF-Dateien, für

den studentische Hilfskräfte eingesetzt wurden. Trotz des verhältnismäßig hohen Personalaufwands für manuelle Erfassung, Erstellung von Erfassungsregeln und Erstellung der Paginierungskonkordanz zeigte sich im Projektverlauf, dass eine automatisierte Bearbeitung keine gangbare Alternative darstellte, wenn nicht deutliche Qualitätsabstriche in Kauf genommen werden sollten. Die denkbare Erfassung von Inhaltsverzeichnissen und Registern der nicht im „Index“ enthaltenen Zeitschriften durch Digitalisierung der entsprechenden Seiten und anschließende Konvertierung in maschinenlesbaren Volltext mithilfe von OCR war auf Grund zu hoher Fehlerraten bei der Schrifterkennung nicht möglich.

OCR-Software kam im Projekt bisher nicht zum Einsatz, da die Fehlerrate bei der Schrifterkennung als nicht tolerabel erschien und einen zu hohen Aufwand für Fehlerkontrolle und –korrektur bedeutet hätte. Die in den Zeitschriften der Aufklärung verwendeten Frakturschriften waren bisher generell kaum erfolgreich mit OCR zu bearbeiten. Erschwerend war zudem die stark variierende Druckqualität der Originale (durchscheinender Satzspiegel, unscharfes Schriftbild). Bei eigenen Tests war noch gegen Ende des Projekts die Erkennungsrate der verwendeten OCR nicht akzeptabel. Es bleibt abzuwarten, ob die neuesten Versionen geeigneter OCR-Software auch Frakturschriften künftig mit überzeugenden Erkennungsraten bearbeiten können. Die Übertragung der gescannten Images in maschinenlesbaren Volltext mit Hilfe von OCR bleibt ein interessantes Ziel, da sie Voraussetzung für die Realisierung einer Volltextrecherche ist. Dies würde ohne Frage eine Verbesserung der Recherchemöglichkeiten bedeuten, etwa bei der Suche nach in den Zeitschriftenbeiträgen behandelten Themen, erwähnten Personen oder Orten, die sich aus dem Titel nicht erschließen und den Rahmen der Erfassung durch notwendigerweise knappe Schlagwörter übersteigen, oder bei der gezielten Suche nach Begriffen für begriffsgeschichtliche Fragestellungen.

Die entstandene digitale Sammlung von 45 Zeitschriften der Aufklärung hat ihr wesentliches Ziel erreicht, denn das Angebot ermöglicht die komfortable und gezielte Recherche nach einzelnen Zeitschriftenbeiträgen. Insbesondere für eine vergleichende Arbeit mit diesen Quellen stellt das zur Verfügung stehende Korpus eine wichtige Erleichterung dar. Die

Sammlung erfährt eine ausgesprochen positive Benutzerresonanz, wie sich immer wieder an Kommentaren, die uns per E-Mail erreichen, zeigt. Für die positive Aufnahme spricht auch die rege Nutzung des Angebots, das in der Nutzungsstatistik des gesamten Webangebots der Universitätsbibliothek Bielefeld stets vordere Plätze belegt und circa 60.000 bis 80.000 Zugriffe (=Seitenbewegungen) pro Monat verzeichnet. Dabei überwiegt der Anteil externer Zugriffe, dies ist insbesondere auf die Recherchierbarkeit der Zeitschriften in Internetsuchmaschinen zurückzuführen.

Blick nach vorn: Fortsetzung des Projekts

Die Fortsetzung des Projekts wurde inzwischen durch die DFG bewilligt. Zielsetzung dieser zweiten Projektphase, in der die Kooperation mit dem Georg Olms Verlag und der Akademie der Wissenschaften zu Göttingen fortgesetzt wird, ist die digitale Bereitstellung aller übrigen im „Index deutschsprachiger Zeitschriften 1750-1815“ erschlossenen 150 Zeitschriften. Dabei werden die bisher eingesetzten technischen Verfahren und Workflows im Wesentlichen beibehalten. Ein weiterer Kooperationspartner wird die Klassikstiftung Weimar, Herzogin Anna Amalia Bibliothek sein, die ein eigenes DFG-Projekt zur Digitalisierung und Erschließung von Zeitschriften des Weimar-Jenaer-Literaturkreises um 1800 durchführen und die erstellten digitalen Reproduktionen in das Bielefelder Angebot der „Zeitschriften der Aufklärung“ einbringen wird.

Dr. Sabine Rahmsdorf hat mit einer Arbeit über Stadt und Architektur in der literarischen Utopie der Frühen Neuzeit im Fach deutsche Literaturwissenschaft an der Universität Hannover promoviert. Sie ist seit 2001 an der Universitätsbibliothek Bielefeld beschäftigt, wo sie nach der Mitarbeit in verschiedenen Projekten die Fachreferate Geschichte und Kunstgeschichte betreut. E-Mail: sabine.rahmsdorf@uni-bielefeld.de

DIE DIGITALISIERTEN „JAHRESBERICHTE FÜR DEUTSCHE
GESCHICHTE“ AUS DER ZWISCHENKRIEGSZEIT – EIN SPIEGEL DER
GESCHICHTSWISSENSCHAFT ZWISCHEN 1925 UND 1938

von Daniel Schlögl

Die Jahresberichte für deutsche Geschichte der Zwischenkriegszeit (Berichtsjahre 1925-1938) sind für die heutige Forschung in mehrfacher Hinsicht bedeutsam: Mit ihren bibliografischen Nachweisen und Forschungsberichten dokumentieren sie einerseits die zeitgenössische Fachliteratur. Zudem wird aus den Vorworten, dem Wandel der inhaltlichen Gliederung und insbesondere aus den Forschungsberichten die zunehmende Beeinflussung durch politische und ideologische Motivationen deutlich erkennbar. In einzigartiger Weise spiegeln die älteren Jahresberichte somit die Entwicklung der Historiografie in der Zeit der Weimarer Republik und des Nationalsozialismus wider. Die Bände wurden in einem DFG-Projekt digitalisiert und liegen, nach einer Überarbeitung, in einer XML-Datenbank vor.

Die *Jahresberichte für deutsche Geschichte* können auf eine über 125jährige Entwicklung zurückblicken und sind damit eine der traditionsreichsten Einrichtungen der geisteswissenschaftlichen Informationsinfrastruktur in Deutschland.¹ Neben die konventionellen gedruckten Jahresbibliografien² ist inzwischen die akkumulierende Online-Datenbank als Hauptpublikati-

onsform getreten; sie stellt mit derzeit etwa 290.000 Titeleinträgen ab dem Berichtsjahr 1985 die umfassendste Fachbibliografie zu allen Epochen und Sachgebieten der deutschen Geschichte dar.³

Die Umstellung auf eine EDV-gestützte Arbeitsweise erfolgte zu Beginn der 1990er Jahre zeitgleich mit der Integration der *Jahresberichte* als Langzeitvorhaben in die neu konstituierte Berlin-Brandenburgische Akademie der Wissenschaften. Von Beginn an wurde es als ein wichtiges Ziel angesehen, nicht nur die aktuelle Literaturproduktion zu dokumentieren, sondern auch die älteren Druckbände retrospektiv in die digitale Form zu überführen, um somit sukzessive einen möglichst weit zurückreichenden, vollständigen Literaturnachweis zur deutschen Geschichte anbieten zu können. Für die nach dem Zweiten Weltkrieg erarbeiteten Jahressbände der Bibliografie wird dies durch die manuelle Übertragung der Titeleinträge in die Datenbank praktiziert. Die zwischen 1927 und 1940 erschienenen Bände der *Jahresberichte für deutsche Geschichte* wurden hingegen in einem eigenständigen DFG-Projekt digitalisiert. Hierfür war entscheidend, dass die älteren *Jahresberichte* mit ihrer Parallelität von Bibliografien und Forschungsberichten eine grundlegend andere Struktur als die Bände der nach dem Zweiten Weltkrieg begonnenen „Neuen Folge“ aufweisen. Hinzu kamen weitere Beweggründe, für deren Herleitung ein kurzer Blick auf die Geschichte geworfen werden soll.

Die *Jahresberichte* starteten 1880 als *Jahresberichte der Geschichtswissenschaft*.⁴ Es handelte sich um ein Referatsorgan mit universalgeschichtlicher Ausrichtung, wobei die jeweils besprochene Literatur in den Fußnoten nachgewiesen wurde. Im Verlauf des Ersten Weltkriegs mussten die Arbeiten an dem Informationsmittel, das inzwischen auch internationale Reputation gewonnen hatte, eingestellt werden. Nach Kriegsende startete ein erster Versuch der Wiederaufnahme, wobei nicht zuletzt aufgrund der finanziellen und personellen Engpässe und der abgerissenen internationalen Kontakte eine inhaltliche Einengung erfolgte, die sich im neuen Titel

1 Vgl. Winfried Enderle, *Geschichtswissenschaft, Fachinformation und das Internet*, in: eForum zeitGeschichte 3/4, 2001 <http://www.eforum-zeitgeschichte.at/3_01a7.pdf> (25.10.2006), S. 3.

2 Zuletzt erschienen: Berlin-Brandenburgische Akademie der Wissenschaften (Hg.), *Jahresberichte für deutsche Geschichte, Neue Folge*, 56 (2004), Berlin 2005.

3 Homepage: <<http://www.jdg-online.de/>> (25.10.2006).

4 Vgl. Historische Gesellschaft zu Berlin (Hg.) *Jahresberichte der Geschichtswissenschaft*, 36 Bände, Berichtsjahre 1878-1913, Berlin 1880-1916.

Jahresberichte der deutschen Geschichte widerspiegelt.⁵ Das Unternehmen war allerdings organisatorisch und finanziell zu schwach ausgestattet, um auf Dauer erfolgreich arbeiten zu können, so dass mit Unterstützung mehrerer Ministerien Preußens und des Reichs Mitte der 1920er Jahre eine grundlegende Neuorganisation stattfand: Es begann die Geschichte der *Jahresberichte für deutsche Geschichte* im engeren Sinne⁶, die als eine Art nationales Pendant zur zeitgleichen Initiative für die Gründung der *Internationalen Bibliographie der Geschichtswissenschaften* konzipiert waren. Als Herausgeber fungierten Albert Brackmann und Fritz Hartung; in die Trägerorganisation waren die Direktoren der Preußischen Staatsarchive und der Preußischen Staatsbibliothek sowie die Deutsche Bücherei Leipzig eingebunden.⁷ Bereits für den ersten Band konnten sich Brackmann und Hartung auf eine breite Basis von 74 Berichterstellern stützen.⁸ Rasch wurden die *Jahresberichte für deutsche Geschichte* zu einem unverzichtbaren Bestandteil der Fachinformation – die ab 1928 erarbeitete Neuauflage des *Dahlmann-Waitz* stützte sich nicht nur in inhaltlicher Hinsicht maßgeblich auf die *Jahresberichte*, sondern wurde auch im Auftrag der Gesellschaft *Jahresberichte für deutsche Geschichte* herausgegeben⁹, die sich damit gleichsam anschickte, zu einer zentralen Trägerstruktur für die geschichtswissenschaftliche Fachinformation in Deutschland zu werden.

Vom konzeptionellen Aufbau der Bände her wiesen die *Jahresberichte* der Zwischenkriegszeit gegenüber ihren Vorläufern eine wesentliche Änderung auf: Waren die bibliografischen Nachweise zuvor lediglich in den Fußnoten zu den Forschungsberichten enthalten, so wurden sie nun

5 Vgl. Loewe, Victor; Stimming, Manfred (Hgg.), *Jahresberichte der deutschen Geschichte*, 7 Bände, Berichtsjahre 1918-1924, Breslau 1920-1926.

6 Vgl. Brackmann, Albert; Hartung, Fritz (Hgg.), *Jahresberichte für deutsche Geschichte*, 16 Bände, Berichtsjahre 1925-1940, Leipzig 1927-1942.

7 Vgl. Brackmann u.a. (Hgg.), *Jahresberichte*, 3 (1927), Leipzig 1929 (wie Anm. 6), S. Vf. Vgl. auch Dietrich, Dieter, *Der Beitrag der Deutschen Bücherei zur Erarbeitung historischer Fachbibliographien seit 1929*, in: *Jahrbuch der Deutschen Bücherei* 25 (1989), S. 50-68.

8 Vgl. Brackmann u.a. (Hgg.), *Jahresberichte*, 1 (1925), Leipzig 1927 (wie Anm. 6), S. XIIIff.

9 Vgl. Haering, Hermann (Hg.), *Dahlmann; Waitz* (Hgg.), *Quellenkunde der Deutschen Geschichte*, 9. Auflage, Leipzig 1931, S. IIIff. u. VI.

von diesen separiert und traten, zusammengefasst in bibliografische Abschnitte, gleichberechtigt neben diese. Mit den entsprechenden Querverweisungen entstand somit ein Informationsmittel, dessen komplexe Struktur auch für die Digitalisierung besondere Herausforderungen und Möglichkeiten mit sich brachte.

In inhaltlicher Hinsicht besitzen die *Jahresberichte* der Zwischenkriegszeit als Nachweis der damaligen Fachliteratur in den bibliografischen Teilen, samt zeitgenössischer Einordnung in den Forschungskontext in den Berichtsteilen, nach wie vor einen ungebrochenen Wert: Sie bieten in ganz besonderer Weise einen Zugang zur historischen Forschung in der Weimarer Republik und der NS-Zeit, der für zahllose Fragestellungen gewinnbringend genutzt werden kann. Über die aus den Literaturzitierten und Forschungsberichten implizit und explizit sichtbar werdenden Tendenzen und Wertungen hinaus sind die *Jahresberichte* aber auch mit ihrer inhaltlichen Gliederung und den Vorworten ein Spiegel dafür, wie politische und ideologische Motivationen die Historiografie beeinflussten: Die Erfahrung der Niederlage im Ersten Weltkrieg und der Gebietsabtretungen infolge des Versailler Vertrags fand nicht zuletzt darin ihren Widerhall, dass die *Jahresberichte*, „wie es sich von selbst versteht, von dem historischen Deutschland, nicht vom heutigen Umfang des Deutschen Reiches“ ausgingen. Sie „berücksichtigen daher auch die Gebiete, die einst zum Deutschen Reiche gehört haben, aber im Laufe der Geschichte von ihm getrennt worden sind. Wir behandeln ferner diejenigen Glieder des deutschen Volkes, die außerhalb des politischen Reichsverbandes stehen, bringen also einen Abschnitt über das Deutschtum im Auslande, der in den alten Jahresberichten fehlte, und ziehen auch die ausländische Literatur zur deutschen Geschichte in weitestem Umfang heran ...“¹⁰ Die auch in der Geschichtswissenschaft zunehmenden völkischen Tendenzen ließen die *Jahresberichte* nicht unberührt: Die Herausgeber sahen sich 1931 mit der Aufforderung konfrontiert, Forschungen zur „Geschichte der deutschen Bevölkerung, die nach sozialen, statistischen, rasse- und siedlungsgeschichtlichen Gesichtspunkten untersucht werden solle“, in einem eigenen

10 Brackmann u.a. (Hgg.), *Jahresberichte*, 1 (1925) Leipzig 1927 (wie Anm. 6), S. V.

Bericht zu behandeln, was jedoch „aus praktischen wie auch aus sachlichen Gründen“ zunächst noch abgelehnt wurde.¹¹ Demgegenüber tritt im ersten nach der nationalsozialistischen Machtübernahme erschienenen Band deutlich das Bemühen hervor, die bisherige Arbeitsweise als von vornherein konform mit den neuen politischen Gegebenheiten darzustellen. Der „epochemachenden Bedeutung des Jahres 1933“ wurde durch eine Erweiterung des bislang mit dem Ersten Weltkrieg endenden Berichtgegenstandes um die „nunmehr abgeschlossene Periode von 1919 bis 1933“ Rechnung getragen.¹² In den Folgejahren wurde die inhaltliche Gliederung sukzessive um Abschnitte wie „Rasse und Bevölkerung“, „Die Juden in Deutschland“ oder „Staatsanschauungen und völkischer Gedanke“ erweitert; der Bereich „Bevölkerungsgeschichte“ erfuhr ab 1936 eine weitere Auffächerung.¹³

Die *Jahresberichte* der Zwischenkriegszeit können somit als Quelle von weitreichender Relevanz gelten. Um sie für die Forschung besser nutzbar zu machen, wurden die Bände in einem DFG-geförderten Kooperationsprojekt zusammen mit dem Trierer Kompetenzzentrum für elektronische Erschließungs- und Publikationsverfahren in den Geisteswissenschaften bis 2003 vollständig digitalisiert. Ziel war es dabei einerseits, die Bände unter weitgehender Beibehaltung ihrer originalen Struktur in das elektronische Medium zu überführen und zugleich dessen technische Möglichkeiten für verschiedene Recherchezugriffe nutzbar zu machen. Das Projektergebnis konnte 2003 in einer ersten Version zur Verfügung gestellt werden. Nach Projektende wurde das Informationsangebot in einem weiteren Arbeitsschritt an der Berlin-Brandenburgischen Akademie der Wissenschaften überarbeitet und im März 2005 als „TELOTA-Projekt des Monats“ vorgestellt.¹⁴ Kern der Maßnahme war eine Migrierung der ursprünglich SGML-codierten Daten in eine native XML-Datenbank.

11 Siehe Brackmann u.a. (Hgg.), *Jahresberichte*, 5 (1929), Leipzig 1931 (wie Anm. 6), S. Vf.

12 Siehe Brackmann u.a. (Hgg.), *Jahresberichte*, 7 (1931), Leipzig 1934 (wie Anm. 6), S. Vf.

13 Siehe Brackmann u.a. (Hgg.), *Jahresberichte*, 8 (1931), Leipzig 1934; 9/10 (1933/1934), Leipzig 1936; 12 (1936), Leipzig 1937 (alle wie Anm. 6).

14 Siehe <http://www.bbaw.de/pom/langtext_200503.html> (25.10.2006).

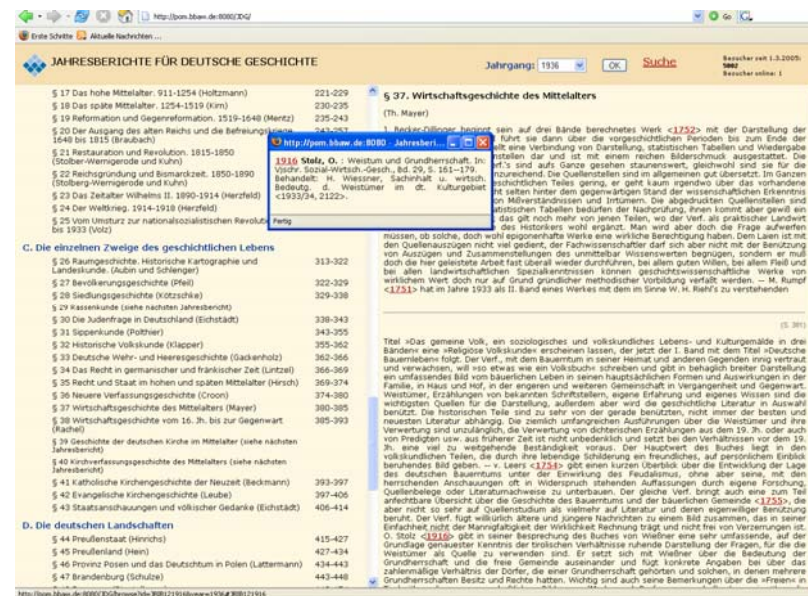


Abbildung 1: Blättern Zugriff auf Forschungsberichte, Verknüpfungen zu bibliografischen Angaben

Die digitalen *Jahresberichte für deutsche Geschichte aus der Zwischenkriegszeit* (Bände 1-14, Berichtsjahre 1925-1938)¹⁵ bieten zunächst einen blättern Zugriff über die einzelnen Bände, wobei Inhaltsverzeichnisse und Register weitestgehend in der Originalform beibehalten wurden (siehe Abbildung 1). Einen erheblichen Mehrwert gegenüber der Druckausgabe stellen die bandinternen Verknüpfungen dar, die nicht nur eine schnelle Navigation ermöglichen, sondern darüber hinaus in den bibliografischen Teilen sichtbar machen, ob ein Titel in den Forschungsberichten Berücksichtigung fand (siehe Abbildung 2). Durch die Implementierung bandübergreifender Suchmöglichkeiten ist es möglich, sich unter anderem einen Überblick über die Entwicklung bestimmter Forschungsgegenstände zu verschaffen (siehe Abbildung 3).

15 Siehe <<http://pom.bbaw.de:8080/JDG/>> (25.10.2006).

The screenshot shows the website interface for 'Jahresberichte für Deutsche Geschichte'. The main content is a table of contents for the year 1936. The table is organized into several sections, including 'C. Einzelne Zweige des geschichtlichen Lebens' and '7. Rassenkunde'. Each entry includes a title and a page range. For example, under '7. Rassenkunde', there are entries for '1690 Ekkehart, W.', '1699 Gross, W.', '1700 Reche, O.', '1701 Reiche, K.', '1702 Hülle, W.', '1703 Grahmann, R.', '1704 Reche, O.', '1705 Müller, Gertrude', '1706 Grau, R.', '1707 Kammer, H.', '1708 Eickstedt, E. Frh. v.', '1709 Richter, B.', '1710 Jajima-Jejmiz, S.', '1711 Müller, Karl Otto', and '1712 Lebzelter, V.'.

Abbildung 2: Bibliografischer Abschnitt mit Verknüpfungen zu Forschungsberichten

Die Arbeiten an dem Projekt sind noch nicht abgeschlossen: Redaktionelle Schwierigkeiten, die aufgrund der Komplexität des Informationsangebots, der großen Mitarbeiterzahl etc. schon bei der Erstellung der Druckbände vorhanden waren, machen sich noch heute bei der Überführung in die digitale Form bemerkbar. Die automatische Erkennung und Codierung uneinheitlich gestalteter bzw. teilweise auch fehlerhafter Bezüge wirft vielfache Probleme auf, so dass ein hoher Aufwand für die Nachbearbeitung erforderlich ist. Als weitere Arbeitsschritte sind neben der Überarbeitung die Einrichtung einer XQuery-Schnittstelle für individuelle Rechercheanfragen sowie die Konvertierung der bibliografischen Daten geplant, um diese, zusätzlich zu dem in sich geschlossenen Angebot, auch innerhalb der laufenden Datenbank der *Jahresberichte für deutsche Geschichte* bereitstellen zu können.

The screenshot shows the search interface for the year 1925. The top navigation bar includes the year '1925' and search buttons. Below this, there are search filters for 'Autor', 'Titel', and 'Volltext'. The search results section is titled 'Suche über alle Jahrgänge' and shows a list of search results for the year 1925. The results include a list of authors and titles, such as 'c) Saller.' and 'd) Stafer.'.

Abbildung 3: Bandübergreifende Recherchemöglichkeiten

Dr. Daniel Schlögl ist Arbeitsstellenleiter des Akademienvorhabens Jahresberichte für deutsche Geschichte an der Berlin-Brandenburgischen Akademie der Wissenschaften. E-Mail: schloegl@bbaw.de

VERFAHREN DER QUELLENSCHLIEßUNG
UND SUCHSTRATEGIEN

EIN TOR ZU VIELEN QUELLEN – DAS BAM-PORTAL

von Frank von Hagel

Das von der Deutschen Forschungsgemeinschaft (DFG) geförderte BAM-Projekt schafft einen einfachen und direkten Zugang zu den Quellen aus Archiven, Bibliotheken und Museen für die Forschung und die interessierte Öffentlichkeit. Hierzu werden digitale Kataloge, Findmittel und Inventare aus Bibliotheken, Archiven und Museen zusammengeführt und so eine einheitliche und gleichzeitige Recherche nach Literatur, Archivalien und Exponaten ermöglicht. In diesem Werkstattbericht sollen die gegenwärtigen Möglichkeiten des BAM-Portals, die sich vor allem durch die neuen Projektpartner ergeben, beleuchtet werden.

Im von der Deutschen Forschungsgemeinschaft (DFG) geförderten Projekt „Gemeinsames Portal für Bibliotheken, Archive und Museen. Ein Online-Informationportal (BAM)“ wird seit Mai 2001 eine Metasuchmaschine entwickelt¹, die digital verfügbare Erschließungsleistungen aus Bibliotheken, Archiven und Museen vorhält. Die im BAM-Portal zusammengeführten digitalen Kataloge, Findmittel und Inventare werden in einem gemeinsamen Index gebündelt, sodass die circa 32 Millionen Datensätze (Stand Mai 2006) schnell einrichtungs- und spartenübergreifend recherchiert werden können. Dem heterogenen Nutzerkreis aus Privatpersonen, Wissenschaftlern/innen usw. wird somit ein gemeinsamer Erstzugang zu diesen Informationsquellen ermöglicht und der Zugriff auf vorhandene Fachinformationssysteme der einzelnen Einrichtungen erleichtert; denn aus der gemeinsamen Trefferliste heraus wird direkt in diese Systeme verwie-

1 Vgl. <<http://www.bam-portal.de/>>.

sen.

BAM Portal für
Bibliotheken Archive Museen

Das BAM-Portal liefert Ihnen Informationen zu Literatur, Archivalien und Objekten aus deutschen Bibliotheken, Archiven und Museen in einem gemeinsamen Suchergebnis

Erweiterte Suche
Einstellungen

Suche Neue Suche

Ausgewählte Datenbestände im BAM-Portal

[Fußball](#)
Alles zum Thema Fußball im BAM-Portal

[Kalliope-Portal](#)
Die Datenbank für Autographen und Nachlässe

[Antike Kulturen](#)
Digitaler Katalog des Badischen Landesmuseums Karlsruhe der teilnehmenden Museen

[Glassammlungen](#)

[Home](#) [Projekt](#) [Teilnehmer](#) [Kontakt](#) [Impressum](#) [FAQ](#)

Gefördert durch die Deutsche Forschungsgemeinschaft **DFG**

powered by **BSZ**

Abbildung 1: Startseite BAM-Portal

Die bisher einmalige Form der Zusammenarbeit der Kulturgutträger Bibliotheken, Archive und Museen ist vor allem auf die immer deutlicher werdende Notwendigkeit für diese Einrichtungen zurückzuführen, den Benutzern/innen umfassende, spartenübergreifende und einfache Zugänge zu ihren Katalogen, Beständen und Sammlungen, dem digitalen Kulturerbe, anzubieten.

Auf internationaler Ebene haben Projekte zur Bereitstellung von digitalen Informationen zum nationalen oder regionalen Kulturgut begonnen. Außer *ArBiMus.dk*² oder *ABM-utvikling*³ sehen aber derzeit keine weiteren Projekte einen bibliothekarisch, archivarisch und museologisch übergrei-

2 Vgl. <<http://www.arbimus.dk/>>.

3 Vgl. <<http://www.abm-utvikling.no/>>, dort ist eine entsprechende Datenbank in Vorbereitung.

fenden Datenzugriff vor. Vielmehr sind die Angebote in der Regel spartenbezogen begrenzt. Insofern hat das BAM-Portal nach wie vor eine Vorreiterrolle inne.

Neben nationalen Projekten werden von der Europäischen Union verschiedene Projekte und Förderprogramme zum Aufbau eines europaweiten Zuganges zum digitalen Kulturerbe initiiert. Stichworte sind in diesem Kontext MINERVAPlus und Michael sowie MichaelPlus. Über diese Themen wurde im Verlaufe dieser Tagung bereits durch den Vortrag von Frau Hagedorn-Saupe, Institut für Museumskunde, Staatliche Museen zu Berlin – Stiftung Preußischer Kulturbesitz, informiert.⁴ Ziel dieser Programme ist die Förderung der Digitalisierung in Kultureinrichtungen und die Schaffung eines europaweiten Zuganges zu digitalen Beständen. Das BAM Projekt geht mit seinen Zielen einen Schritt weiter und führt das digitale Kulturerbe auf Objekt-, Titel- oder Verzeichnisebene zusammen.

Mit dem Beginn der dritten Förderungsphase durch die DFG im Oktober 2005 wurden, aufbauend auf den bereits gesammelten Erfahrungen,

- die Optimierung der Suchgeschwindigkeit,
- der Ausbau der Datenbasis,
- die Überarbeitung des Designs und
- der Usertools für die Benutzer/innen sowie
- die Etablierung eines tragfähigen Benutzermodells

als zentrale Aufgaben für die dauerhafte und erfolgreiche Nutzbarkeit des BAM-Portals definiert.

Wichtige Etappen dieser Aufgabenstellung konnten bereits bewältigt werden, weitere Schritte stehen an. In diesem Vortrag soll der aktuelle Stand des BAM-Projektes aufgezeigt und ein Ausblick auf die zukünftigen Entwicklungen gegeben werden. Die bereits vorliegenden Arbeitsergebnisse beweisen, dass es möglich ist, mittels einer Metasuchmaschine die digitalen Erschließungsleistungen aus Archiven, Bibliotheken und Museen miteinander zu vernetzen. Ein grundsätzliches Problem bei der Vernetzung der genannten Ressourcen resultiert allerdings aus der Verschiedenheit der Arbeitstraditionen, die sich aus der Arbeit mit unterschiedlichen „Quellen“,

4 Siehe dazu den Beitrag von Monika Hagedorn-Saupe im selben Band.

den Archivalien, Büchern und Objekten entwickelt haben. Diese Arbeitstraditionen müssen aufgrund der Besonderheiten des „Arbeitsgegenstandes“ unterschiedliche Aspekte im fachlichen Umgang mit den Materialien in den Vordergrund stellen und führen zum Beispiel dazu:

Dass es für einen Bibliothekskatalog in der Regel nicht als notwendig erachtet wird, die Materialeigenschaften des Buches zu erfassen, es sei denn, es handelt sich um Sondermaterialien.

Dass die Archive sich von den Bibliotheken und Museen durch ihr Zustandekommen unterscheiden. Nicht das Sammeln, sondern organisches Zuwachsen ist ihr Kennzeichen. Modernere Definitionen stellen jedoch nicht mehr nur den hoheitlichen Aspekt der Archivtätigkeit in den Vordergrund⁵, sondern sehen eine klare Abgrenzung zum Sammlungsgut darin, dass Sammlungsgut nicht dem Provenienzprinzip unterliegt, sondern gerade durch andere Entstehungszusammenhänge gekennzeichnet ist.

Dass Museen sich als Institutionen verstehen, die sich vornehmlich der Sammlung und Erforschung von Unikaten widmen. In diesem Zusammenhang gehören Materialeigenschaften, Abmessungen, ikonografische Angaben und vieles andere mehr zum Grundbestand jeglicher Objektbeschreibung.

Diese nur scheinbar eindeutige Abgrenzung der Museen zu Archiven und Bibliotheken wird im Alltag aber bereits dann problematisch, wenn es sich beim Sammlungsgegenstand um Objekte aus „Massenproduktionen“ wie Münzen, aus der industriellen Fertigung oder um Druckgrafiken bzw. Einblattdrucke und -serien (zum Beispiel Kartenwerke) handelt. In diesen Fällen ist eine so differenzierte Betrachtungsweise nicht immer sinnvoll und möglich, es sei denn, die Nutzungsgeschichte, also die Provenienz des Sammlungsobjektes, begründet dessen Einzigartigkeit. Dann befinden sich Museen in einer ähnlichen Situation wie Archive, die ebenfalls vornehmlich Unikate in ihrem Bestand haben und jedes Stück als Teil eines Geschäftsganges, der die Einzigartigkeit des Stückes betont, ansehen.

5 Vgl. <<http://staff-www.uni-marburg.de/~mennehar/datii/germanterms.htm>>.

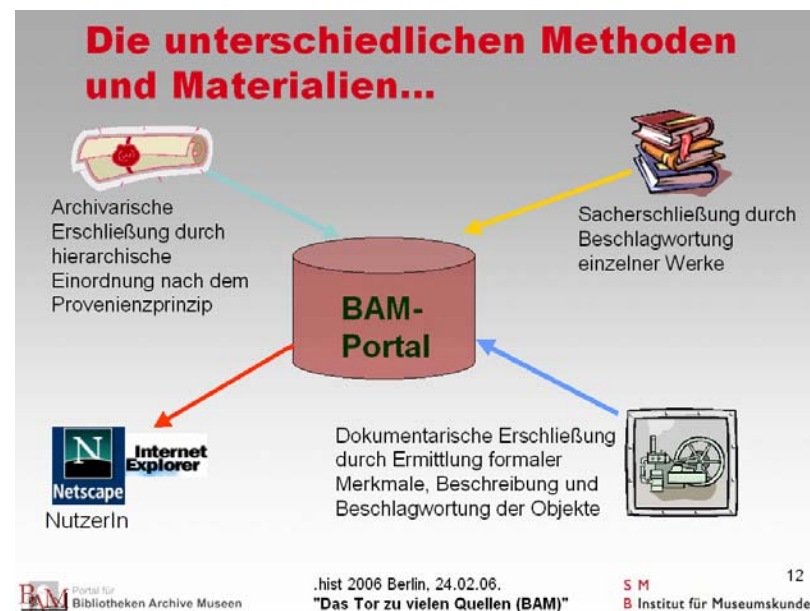


Abbildung 2: Umgang mit den Quellen

Auch die Beschränkung der Zuständigkeit für bestimmte Materialien ist kein eindeutiges Identifikationsmerkmal. Jeder Mitarbeiter und jede Mitarbeiterin einer Bibliothek, eines Archivs oder eines Museums ist in der Lage, Ausnahmen von dieser auf den ersten Blick eindeutigen Trennung der Arbeitsfelder zu benennen. Die Grenzen werden immer wieder überschritten: So finden sich in Museen häufig Archivalien, die im Kontext mit den Objekten überliefert wurden (zum Beispiel Korrespondenz zur Objektentstehung aus der Hand des Künstlers/der Künstlerin, Konstruktionszeichnungen, etc.) oder die im Kontext mit dem Objekt entstanden sind (zum Beispiel Informationen zur Erwerbung, dem Fund). Daneben sind in Archiven und Bibliotheken Objekte, unter anderem als „Anlagen“, zu finden.

Eine eindeutige Abgrenzung der drei Institutionen ist folglich nicht

möglich.⁶ Vielmehr scheint nur eine funktionale Trennung der Bereiche sinnvoll.⁷ Neben den dargestellten Unterschieden gibt es – wie angedeutet – eindeutig zu identifizierende Gemeinsamkeiten von Bibliotheken, Archiven und Museen, die sich in immer stärkerem Maße als *memory institutions*, insbesondere in den Erörterungen über kulturelles Gedächtnis und Erinnerungskünste im Zeichen elektronischer Speicher, verstehen.⁸ Daraus ergeben sich neue Aufgabenstellungen bei der Bereitstellung und Vermittlung ihrer Informationen für diese „Gedächtniseinrichtungen“.

Aus Sicht der Geschichtsforschung ist es von besonderer Bedeutung, die Zeugnisse der Geschichte, unabhängig davon, ob sie als Archivalie, Buch oder Objekt überliefert sind, möglichst umfassend in ihrem historischen Kontext rekonstruieren zu können.

Die Online-Stellung der Erschließungsinformationen ermöglichen den Nutzern/innen vorbereitende ortsunabhängige Recherchen. Die im BAM-Portal eingesetzten Technologien befähigen Wissenschaft und interessierte Öffentlichkeit darüber hinaus zu einer ersten, spartenübergreifenden Recherche und zur Orientierung über das vielfältige kulturelle Erbe.

Noch mehr als für die wissenschaftliche Nutzung ist für die Benutzer/innen mit Interesse am digitalen Kulturerbe ein leichter Zugang zu den Inhalten von entscheidender Bedeutung. Hierzu ist es notwendig, dass sich die unterschiedlichen Informationen in den Erschließungsdaten auf die für die Bedürfnisse relevanten Suchaspekte nach Zeit-, Personen- bzw. Körperschafts-, Orts- und Schlagwortsuche zusammenführen lassen.

6 Rogalla von Bieberstein, Johannes, Archiv, Bibliothek und Museum als Dokumentationsbereiche. Einheit und gegenseitige Abgrenzung (Bibliothekspraxis 16), München 1975, S. 19f.

7 Leonhardt, Holm A., Was ist Bibliotheks-, was Archiv- und Museumsgut? Ein Beitrag zur Kategorisierung von Dokumentationsgut und -institutionen, in: Der Archivar 42 (1989), S. 214-224.

8 Haverkamp, Anselm; Lachmann, Renate, Text als Memotechnik – Panorama einer Diskussion, in: Dies., Gedächtniskunst. Raum – Bild – Schrift, Frankfurt am Main 1991, S. 7-22; Dempsey, Lorcan, Scientific, Industrial, and Cultural Heritage. A Shared Approach: A Research Framework for Digital Libraries, Museums and Archives, in: Ariadne, Issue 22, 1999, vgl. <<http://www.ariadne.ac.uk/issue22/dempsey/>>.

führen zu unterschiedlichen Inhalten in den Erschließungsdaten...

- **Suchaspekt Zeit:**
 - Bibliothek: Erscheinungsdatum, **behandelter Zeitraum**;
 - Archiv: **Laufzeit der Archivalie**; Museum: **Herstellungszeit, Verwendungszeit**, Zeitpunkt der Übereignung etc.
- **Suchaspekt Person / Körperschaft:**
 - Bibliothek: Autor, **Personen- und Körperschaftsschlagworte**;
 - Archiv: Provenienzangaben, **personen- und körperschaftsbezogene Inhalte**; Museum: **Hersteller, Lieferant**
- **Suchaspekt Schlagwort:**
 - Bibliothek: Schlagwörter (nach SWD); Archiv: **Stichwörter, Textvorkommen**; Museum: **Sachschlagworte, Kontexterweiterung**
- **Suchaspekt Ort:**
 - Bibliothek: **Schlagwörter**; Archiv: **Stichwörter, Textvorkommen**; Museum: **Herstellungsort, Wohnort des Lieferanten**

BAM Portal für Bibliotheken Archive Museen .hist 2006 Berlin, 24.02.06. "Das Tor zu vielen Quellen (BAM)" S M 13 Institut für Museumskunde

Abbildung 3: Suchaspekte des BAM-Portals

Diese Suchaspekte müssen mit Hilfe einer einfachen Suchmöglichkeit recherchiert werden können. Die primäre Aufgabe der Ergebnisanzeige ist es, der Benutzerin und dem Benutzer eine Orientierungs- und Entscheidungshilfe über die Materialien zu bieten, mit denen sie/er sich weiter beschäftigen möchte: Jedes einzelne Suchergebnis muss daher direkt mit den Detailinformationen des Herkunftskatalogs verlinkt sein.

Das BAM-Portal bietet allen Benutzern/innen genau diese Möglichkeiten, erste Informationen über die überlieferten Quellen online zu erhalten, ohne die einzelnen Lagerungsorte der Quellen in den Bibliotheken, Archiven und Museen der Bundesrepublik mühevoll bereisen zu müssen. Es bietet einen ersten Überblick, es kann natürlich nicht die kritische Quellenforschung und Sichtung der Originale vor Ort ersetzen. Aber ebenso wenig wie heute der Verzicht auf die Nutzung eines Bibliotheks-OPACs für die wissenschaftliche Recherche vorstellbar ist, wird dies in

einigen Jahren für die Online-Findbücher der Archive und die Objektdatenbanken der Museen der Fall sein.

Da jedoch nicht alle Einrichtungen in der Lage sind, eigene Datenbanken über ihre Materialien im Internet bereit zu stellen, bietet das BAM-Portal diesen teilnehmenden Einrichtungen eine Möglichkeit, ihre Daten online zugänglich zu machen. Die notwendigen Detailinformationen werden auf dem BAM-Server gehostet. Weiterhin ist es möglich, so genannte „individualisierte Zugänge“ zu definieren, die eine Suche auf Datenausschnitte begrenzen. Hierdurch kann der Datenbestand einer oder mehrerer Einrichtungen über ein Suchformular in externe Angebote integriert werden. Dies bietet einzelnen Teilnehmern oder Gruppen von BAM-Teilnehmern aus dem Bereich der Bibliotheken, Archive oder Museen die Möglichkeit, auf einfache Art und Weise ihre Datenbanken miteinander zu vernetzen.

Darüber hinaus bietet die Bereitstellung von Informationen zum Bibliotheksgut, den Archivalien und Museumsobjekten im Internet grundsätzlich den Einrichtungen die Möglichkeit, der interessierten Öffentlichkeit und der wissenschaftlichen Forschung bislang unbekannte Bestände aus den Magazinen, die Tätigkeitsfelder der Einrichtung sowie die Highlights einer Sammlung oder eines Bestandes näher zu bringen. Von der Information über die Öffnungszeiten bis hin zur Unterstützung von Arbeitsabläufen wie zum Beispiel der Online-Bestellung von Archivalien vor Ort zur Einsicht sind vielfältige weitere Informationsangebote denkbar. Ein Bereich des BAM-Portals stellt daher die Teilnehmer am BAM-Projekt vor und dient dazu, auf die im Vorfeld eines Besuches hilfreichen Informationen wie Link zur Homepage, kurze Beschreibung der Einrichtung usw. hinzuweisen.

Für Museen ist ein weiterer Aspekt von Bedeutung. Da nur ein kleiner Teil der Objekte mit Hilfe der Dauerausstellung oder verschiedener Sonderausstellungen den Besuchern/innen und der Forschung zugänglich gemacht werden kann, ist es außerordentlich sinnvoll, zusätzliche „Schätze der Depots“ über das Internet verfügbar zu machen. Aus Sicht der Museen eröffnet sich die Möglichkeit, mit einer Objektdatenbank ein Werkzeug anzubieten, mit der ein Museumsbesuch vor- oder nachbereitet werden

kann, neue „Fragestellungen an die Objekte“ formuliert oder Schulen und Universitäten bei der Planung und der Umsetzung von Lehrveranstaltungen unterstützt werden können. Das aktiv mit Museumsobjekten arbeitende Fachpersonal erhält seinerseits ein Werkzeug, das die Planung von Ausstellungen und Forschungsarbeiten erleichtert.

Voraussetzung für all diese Nutzungsmöglichkeiten ist die umfassende EDV-basierte Erschließung der Bestände, eine Voraussetzung, die in Bibliotheken selbstverständlich erscheinen mag, in den meisten Museen und Archiven aber noch in weiter Ferne liegt. Der Anteil der bereits digital erschlossenen Bestände steigt jedoch stetig an. Auch wenn erst ein kleiner Teil der circa 6.500 Museen über ein entsprechendes Online-Angebot verfügt und auch im Archivbereich nur ein geringer Teil der Einrichtungen in der Lage ist, den für einen Online-Katalog notwendigen technischen Aufwand zu betreiben, so ist doch bereits jetzt das Angebot sehr groß. Versuche, mit Hilfe von traditionellen Suchmaschinen wie *Google*, *YAHOO* oder *METAGER* eine Übersicht über das digitale Kulturerbe zu erhalten, scheitern. Auch die gezielte und spartenübergreifende Suche nach Büchern, Archivalien und Museumsobjekten mit Hilfe dieser Werkzeuge führten zu keinem befriedigenden Ergebnis.

Um diesen Mangel zu beheben, ist es von großer Bedeutung, mit einem spezialisierten Werkzeug wie dem BAM-Portal auf bundesweiter Ebene einen zentralen Zugang zu den im Internet angebotenen Digitalisaten und Erschließungsinformationen zu schaffen. Daher ist es das Ziel, die im BAM-Portal geführten Nachweise auf die gesamte Bundesrepublik auszudehnen und das BAM-Portal auch nach Ende der Förderung durch die DFG dauerhaft zu betreiben. Um dieses Ziel zu erreichen, haben die alten und neuen Projektpartner einen Konsortialvertrag vorgelegt, in dem sie die für den dauerhaften Betrieb erforderlichen Aufgaben gemeinschaftlich tragen. Der Kreis der Projektpartner besteht aus:

- Bibliotheksservice-Zentrum Baden-Württemberg,
- Bundesarchiv,
- Landesarchiv Baden-Württemberg,
- Stiftung Landesmuseum für Technik und Arbeit Mannheim,
- Stiftung Preußischer Kulturbesitz: Staatliche Museen und Staatsbiblio-

thek zu Berlin sowie Institut für Museumsforschung. Zudem konnten durch die Beteiligung neuer Teilnehmer am BAM-Projekt neue Datenbestände integriert werden. Als wesentliche Ergänzungen zu den bisher primär baden-württembergischen Projektpartnern und den oben genannten neuen Partnern sind die Beteiligung des Gemeinsamen Bibliotheksverbands (GBV) und zahlreicher Museen zu nennen. Naturgemäß überwiegt hierbei die Anzahl der Titelnachweise von Büchern gegenüber den Nachweisen von Archivalien und Museumsobjekten deutlich. Dies ist aus dem unterschiedlichen Ausmaß digital verfügbarer Beschreibungen von Museumsobjekten und Archivalien gegenüber dem Bibliotheksgut erklärbar und im bibliothekarischen Bereich finden sich dabei nicht nur „flache“ Titelansetzungen. Im Bereich der Sondersammelgebiete sind die erfassten Informationen zum Teil sehr umfangreich, dies gilt zum Beispiel für Nachlässe und Autografen. Eine wichtige Ressource aus diesem Bereich ist das Kalliope-Portal⁹, die Nachweisdatenbank für Nachlässe und Autografen, sie ist mit ihren circa 800.000 Datensätzen aus rund 150 Einrichtungen in das BAM-Portal eingebunden.

Wie oben ausgeführt, reflektieren die recherchierbaren Dateninhalte auch unterschiedliche Arbeitstraditionen der beteiligten Sparten, dies wird in strukturellen und syntaktischen Unterschieden in den Erschließungsdaten deutlich, die in das BAM-Portal einfließen. Die Berücksichtigung und die fachgerechte Darstellung dieser unterschiedlichen Erschließungsleistungen ist ein wichtiger Aspekt des BAM-Projekts. Nicht zuletzt aus diesem Grund werden die Benutzer/innen für die detaillierte Information über die sie interessierenden Treffer auf die Fachsysteme der Bibliotheken, Archive und Museen weitergeleitet.

Über das BAM-Portal sind somit drei Informationsebenen für die Benutzer/innen erreichbar, die wie folgt aufeinander aufbauen: Nach der spartenübergreifenden Suche des BAM-Portals wird ein gemeinsames Ergebnis angeboten, von dort wird auf die vertiefenden Informationen in den einzelnen Fachinformationssystemen (OPACs, Online-Findbüchern oder Objektdatenbanken) verlinkt. Diese formalen Erschließungsdaten

9 Vgl. <<http://www.kalliope-portal.de>>.

werden oftmals durch Voll-Digitalisate, in der Regel Digitalfotos der Archivalie oder des Objektes, seltener Ton- und Videodateien oder Volltexte von Publikationen, ergänzt.

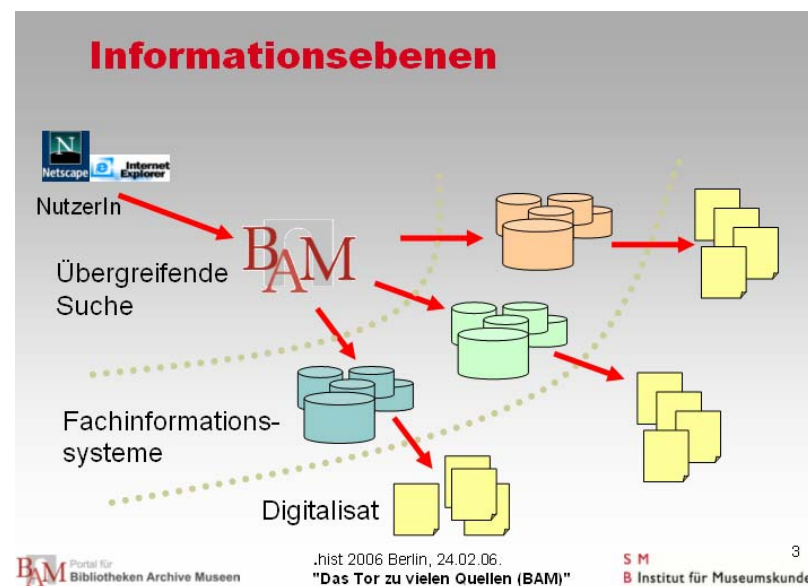


Abbildung 4: Informationsebenen des BAM-Portals

Um einen möglichst umfassenden Zugang zu den Daten zu gewährleisten, stehen für die Suche zwei Möglichkeiten zur Verfügung. Mit der „Einfachen Suche“, einer Volltextsuche nach dem Vorbild der Internetsuchmaschine *Google*, kommt das BAM-Projekt den Suchgewohnheiten eines großen Teils der Benutzer/innen entgegen. Durch die Verwendung Boolescher Operatoren, mit denen Suchbegriffe durch „und“, „oder“ und „nicht“ verknüpft werden können, kann die Suche präzisiert und somit der Umfang der Trefferanzeige verfeinert werden. Dennoch sind Ungenauigkeiten und Unschärfen in den Suchergebnissen bei einer derart unpräzisen Suchstrategie kaum zu vermeiden. Daher steht eine weitere Suchmöglichkeit zur Verfügung.

BAM Portal für
Bibliotheken Archive Museen

Start [Erweiterte Suche](#) [Projekt](#) [Teilnehmer](#) [Kontakt](#) [Hilfe](#)

Erweiterte Suche

Bibliothek Archiv Museum

SWD Synonyme Ergebnisse pro Seite: 10

Suchkriterium	Suchtext	Operator
Titel		mit allen Wörtern
Person		mit allen Wörtern
Schlagwort		mit allen Wörtern

Suche Zurücksetzen

Abbildung 5: „Erweiterte Suche“ des BAM-Portals

Die „Erweiterte Suchoption“ ermöglicht die Realisierung kombinierter und differenzierter Suchvorgänge über Teilbereiche der Daten des BAM-Portals (Titel, Personen, Jahreszahlen, Institutionen oder Sparten). Natürlich steht auch in der „Erweiterten Suchoption“ eine Volltextsuche zur Verfügung. Durch die Kombination der Suchfelder ist es möglich, sehr zielgerichtet den umfangreichen Datenbestand des BAM-Portals zu durchsuchen und die Treffergenauigkeit wesentlich zu erhöhen.

Als zusätzliche Funktion zur Unterstützung der Recherche im BAM-Portal besteht die Möglichkeit, diese um die Option einer „Semantischen Suche“ zu erweitern. Zu diesem Zweck können auf Wunsch in der „Erweiterten Suche“ Synonyme zum Suchbegriff berücksichtigt werden. Hierbei handelt es sich um einen entscheidenden Vorteil gegenüber der reinen Volltextsuche. Während die Volltextsuche nur eine definierte Abfolge von Zeichen sucht, wird mit dieser semantischen Suche ein Wort gemäß seinem Begriff gesucht. Bei der Sucheingabe „Äther“ wird somit beim Einsatz der Option „Synonyme“ auch die Schreibweise „Ether“ berücksichtigt. Ebenso können bei der Suche nach dem Wort „Auto“ auch die Synonym-

und Unterbegriffe „Personenkraftwagen“, „PKW“ und eine Reihe weiterer Worte, die alle die Bedeutung von „Auto“ haben, gefunden werden. Das BAM-Portal verwendet für die Ermittlung von Synonym- und Unterbegriffen die Schlagwortnormdatei der Deutschen Bibliothek (SWD).

Wie die Erfahrungen von Internet-Suchmaschinenbetreibern und die statistische Auswertungen über die Nutzung der Suchformulare zeigen, besteht nach wie vor eine große Zurückhaltung der Nutzer/innen bei der Verwendung dieser Suchoption. Um über deren Verhalten und Suchstrategien genauere Auskünfte zu erhalten, werden selbstverständlich auch im BAM-Projekt die Logfiles für den gesamten Webservice sowie die Logfiles der Suchmaschine ausgewertet.

Von zentraler Bedeutung für die Akzeptanz einer Suchmaschine sind Eindeutigkeit der Suchformulare, die hohe Funktionalität der Recherche und vor allem die Lieferung umfassender und rascher Suchresultate. Grundlage der Recherche im BAM-Portal ist der Einsatz einer hoch entwickelten und schnellen Suchtechnik. Im Gegensatz zu den ersten Projektphasen, die sich ausschließlich auf die Realisierung verteilter Suchen gestützt haben, wurde das BAM-Portal auf die zeitgemäße Textretrievalsoftware Apache-Lucene umgestellt. Gründe für diese Umstellung waren neben der zu geringen Performanz des Portals die mangelnde Stabilität des Angebotes in den ersten Projektphasen. Grundproblem der damaligen Vorgehensweise war die Tatsache, dass alle Daten von den angeschlossenen Datenbanken während der Abfrage-, sprich der Wartezeit der Benutzer/innen, übertragen werden mussten. Zusätzlich musste während dieser Zeit teilweise auch noch die Relationierung der Datenbestände vorgenommen werden. Diese ist aufgrund der unterschiedlichen Herkunft und Erschließungsweise der Daten erforderlich, um eine leicht nachvollziehbare Präsentation der Suchergebnisse im BAM-Portal zu ermöglichen. Selbstverständlich ist das BAM-Portal weiterhin in der Lage, auch diesen Weg zu beschreiten und Datenquellen mittels einer verteilten Suche einzubinden, doch ist dies derzeit nicht mehr die bevorzugte Vorgehensweise. Es hat sich als vorteilhafter erwiesen, einen zentralen Index aufzubauen.

Das BAM-Portal nutzt die Vorteile der zentralen Datenhaltung. Die Daten werden von den Datenquellen in den BAM-Index geladen und dort

vor allem für die Datenanalyse und das Retrieval gespeichert. Dies hat eine erhebliche Beschleunigung der Suche zur Folge. Der vermeintliche Nachteil der redundanten Datenhaltung im BAM-Portal kann durch die regelmäßige Aktualisierung der Daten gemildert werden und ist für die Partner ohne eigenes Fachinformationssystem zur Darstellung der detaillierten Informationen ohnehin zwingend erforderlich. Auf die Detailansichten der Partner mit eigener Datenbank wird aus den Ergebnissen des BAM-Portals verlinkt.

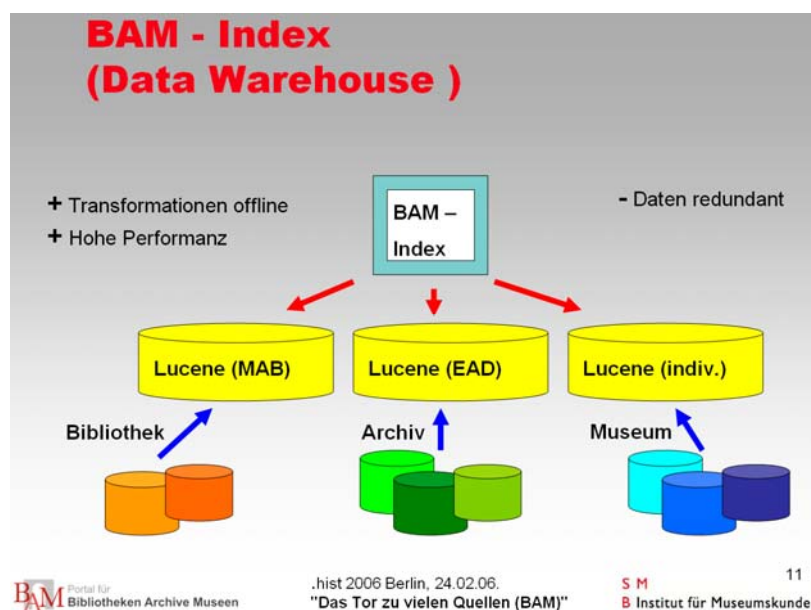


Abbildung 6: Datenhaltung des BAM-Portals

Diese Lösung ermöglicht des Weiteren die Verwendung von XML-Technologien zur Relationierung der unterschiedlichen Datenformate für Bibliotheken, Archive und Museen. Zum Einsatz kommen nach Möglichkeit Standardformate. Es handelt sich hierbei für die Bibliotheken um das „Maschinenlesbare Austauschformat für Bibliotheken (MAB2)“ und für

die Archive um „Encoded Archival Description (EAD)“. Während das MAB-Format bereits seit längerem in den deutschen Bibliotheken eingesetzt wird, ist der Einsatz des EAD-Formates als Austauschformat und Schnittstelle in deutschen Archiven relativ neu. Da sich ein Standardaustauschformat für Museen noch in der Entwicklung befindet, werden gegenwärtig individuelle Mappings aus den lokalen Museumsdokumentationssystemen in das BAM-Portal vorgenommen.

Als Open-Source-Software erlaubt Apache-Lucene spezielle Anpassungen an die Anforderungen des BAM-Portals, insbesondere im Hinblick auf die Heterogenität der Daten:

- Den Suchergebnissen wird eine Rangbewertung zugeordnet, die nach Datenherkunft konfigurierbar ist.
- Es werden computerlinguistische Verfahren zur Wortstambildung genutzt.
- Merkliste und Suchhistorie sind integriert.
- Eine XML-Open-Search-Schnittstelle, über die das BAM-Portal seinerseits in Metasuchmaschinen eingebunden werden kann, lässt sich implementieren.

Insbesondere die Verwendung linguistischer Methoden wie der oben beschriebenen Berücksichtigung von Synonymen und Unterbegriffen bietet interessante Möglichkeiten für die Ergänzung der Suchen. Die hier verwendeten Verfahren erhalten im gesamten Kulturbereich, aber auch in der Industrie, wachsende Beachtung.

Der Vorteil des Einsatzes linguistischer Methoden und von Normdaten für die semantische Recherche besteht in der Gewährleistung einer gemeinsamen Recherchebasis. Dies wird durch die Berücksichtigung von Synonymen sowie sprachlicher Variationen, die im Falle des BAM-Portals Bestandteile der zugrunde gelegten SWD sind, erreicht. Ein weiterer Vorteil bei der Nutzung der SWD ist die kontinuierliche Weiterentwicklung in Hinblick auf internationale Austauschbarkeit oder internationale Relationierbarkeit. Bekannte Beispiele dieser Bestrebungen sind die Projekte der Deutschen Bibliothek (DDB) mit den Namen MACS (Multilin-

gual Access to Subject Headings)¹⁰ und seit 2006 CrissCross.¹¹ Von diesen Entwicklungen profitieren alle Nutzer/innen der Normdaten.


Automatische Indexierung

```

189 Patent <8>
190 des <1>
191 Ernst -> ernst <10>
191 Ernst <18>
191 Ernst <6>
192 Mayer <17>
193 auf <1>
*194 eine -> einen <5>
*194 eine -> ein <14>
195 Glätt-Apparatur <6> :500: Apparatur <6> ## (1) Glätt-Apparatur
195 Glätt-Apparatur <6> :503: Glätten <6> ## (1) Glätt-Apparatur
196 und <1>
197 Pßebnaschinen -> Pressmaschinen -> Pressmaschine <6>

```

- Erkennung von Mehrwortbegriffen
- Wortzerlegung und Wortableitung bei Komposita
- Abgleich der ermittelten Begriffe mit Normdaten

 Portal für
Bibliotheken Archive Museen

.hist 2006 Berlin, 24.02.06.
 "Das Tor zu vielen Quellen (BAM)"

S M
 B Institut für Museumskunde

23

Abbildung 7: Automatische Indexierung

Als möglicher Nachteil an der gegenwärtigen Gestalt der SWD wird in den Archiven und Museen oftmals der Mangel an Sachbegriffen, historischen Ortsnamen sowie archiv- oder museumsrelevantem Vokabular empfunden. Zudem ist die Nutzung von Normvokabular bzw. der Einsatz von Schlagwörtern bei der Erschließung in Archiven und Museen unüblich und für bereits erschlossene Bestände nachträglich kaum realisierbar. Um dieser Situation zu begegnen und dennoch die Vorteile der Verschlagwortung für die Recherche im BAM-Portal verwenden zu können, wurden Verfahren zur automatischen Indexierung getestet. Hiermit ist es möglich, große

¹⁰ Vgl. <<http://www.ddb.de/wir/projekte/macs.htm>>.

¹¹ Vgl. <<http://www.ddb.de/wir/projekte/crisscross.htm>>.

Datenbestände zu bearbeiten, Mehrwortbenennungen zu erkennen und die Wortzerlegungen bei Komposita zu testen, die ermittelten Begriffe mit den Normdaten abzugleichen und somit Schlagwörter nachträglich zu erzeugen. Gleichzeitig müssen aber auch Möglichkeiten geschaffen werden, notwendiges Vokabular aus den Archiven und Museen in die Normdateien einzubringen. Hierzu wurden unter Beteiligung des BAM-Projekts bereits erste Tests durchgeführt und Erfahrungen gesammelt, die nun verfeinert werden.

Neben den vielfältigen technischen Möglichkeiten, die das BAM-Portal einsetzt, um erfolgreiche Recherchen sicherzustellen, ist es auch notwendig, Werkzeuge (neudeutsch: Usertools) bereitzustellen, die den Benutzern/innen die Arbeit mit dem BAM-Portal und den dort gewonnenen Ergebnissen erleichtern. Durch sie werden zum Beispiel die Speicherung der Suchergebnisse und die Zusammenstellung von Recherchelisten usw. ermöglicht. Hierzu dienen Merklisten und Suchhistorien. Zudem werden weitere Usertools für registrierte Benutzer/innen wie zum Beispiel die dauerhafte Speicherung von Suchergebnissen, der Versand von Ergebnislisten per Mail, die Möglichkeit der Ergänzung und Kommentierung der angebotenen Informationen, die Anreicherung der Suchergebnisse mit Verweisen auf verwandte Internetressourcen getestet.

Alle Angebote des BAM-Portals, die verschiedenen Suchfunktionen ebenso wie die Usertools, können auf ihre Nutzung hin evaluiert werden. Parallel findet mit Hilfe eines Fragebogens auch eine kurze Nutzerbefragung statt. Diese Bemühungen erlauben es den Betreibern, Nutzerbedürfnisse zu erkennen. Zudem wird dazu aufgefordert, sich auch aktiv an der weiteren Entwicklung zu beteiligen. Über diese Wege ist es möglich, das Angebot den Wünschen und Erfordernissen entsprechend weiter zu entwickeln.

Ein nächster Schritt, um diese neuen Angebote auch optimal bereitstellen zu können, ist die Überarbeitung des bestehenden Layouts. Mit der steigenden Teilnehmerzahl am BAM-Portal ist es zwingend erforderlich, den Benutzern/innen auch einen visuellen Eindruck über die Inhalte zu vermitteln und die Möglichkeit zu schaffen, sich navigierend durch die große Treffermenge zu bewegen. Auch wenn die Abbildung 8 der Ergeb-

nisansicht nicht das endgültige Design zeigt, wird deutlich, wie die Treffer nach den relevanten Einrichtungen geordnet und aufbereitet werden.

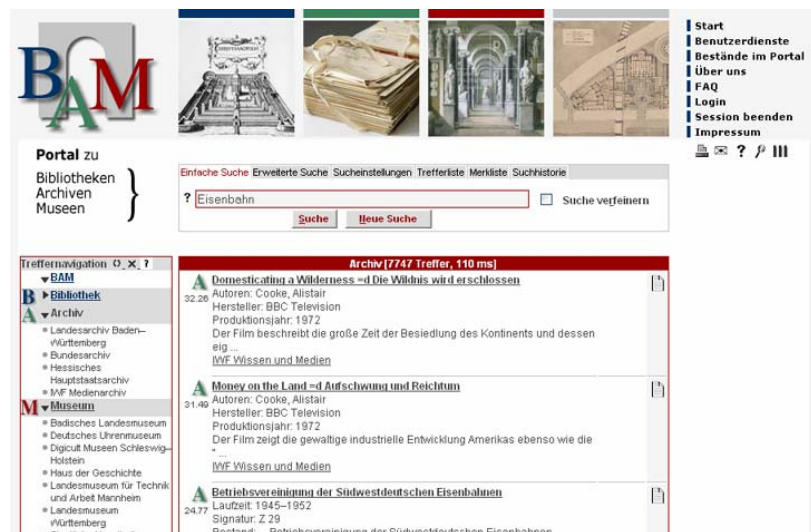


Abbildung 8: Ergebnisansicht

Selbstverständlich kann und will das BAM-Portal keinesfalls einen Museums-, Archiv- oder Bibliotheksbesuch ersetzen. Das besondere Erlebnis beim Betrachten eines Exponates, beim Lesen einer Archivalie oder eines Buches kann durch eine Online-Präsentation nicht ersetzt werden. Sehr wohl aber kann es dazu dienen, Interesse zu wecken, über die Arbeit der teilnehmenden Einrichtung zu informieren, einen Museums-, Archiv- und Bibliotheksbesuch vor- oder nachzubereiten und Wissen zu vertiefen. In den Museen kann eine online zugängliche Datenbank die Aufgaben übernehmen, die ehemals ein Bestandskatalog wahrnahm. Hierbei sind geringere Kosten, mehr Informationsgehalt und mehr Aktualität als positive Effekte zu vermerken.

Dem BAM-Projekt ist es gelungen, in einem Internetportal Quellen unterschiedlichster Herkunft zusammenzuführen und neue Zugänge zu

Objekten, Archivalien und Literatur und somit zu einem großen Teil des digitalen Kulturerbes zu schaffen. Die Erreichung dieses Zieles wird durch die Partizipation vieler neuer Teilnehmer immer weiter gefördert; die Grundlagen sind gelegt, und wir freuen uns über jeden neuen Interessenten.

Frank von Hagel ist Historiker und seit 1997 im Bereich der Museumsdokumentation tätig. Neben der internen Anwendung der Museumsdokumentation ist die Bereitstellung digitalen Kulturgutes ein Schwerpunkt seines Interesses. 2003 und 2004 arbeitete er in der Staatsbibliothek zu Berlin für das Kalliope-Portal, der Nachweisdatenbank für Nachlässe und Autographe, und wechselte von dort an das Institut für Museumsforschung. E-Mail: f.v.hagel@smb.spk-berlin.de

QUELENER SCHLIEßUNG FÜR DIE MONTANGESCHICHTE –
EIN WERKSTATTBERICHT AUS DEM MONTAN.DOK

von Stefan Przigoda

Seit Anfang 2001 vereint das Montanhistorische Dokumentationszentrum (montan.dok) beim Deutschen Bergbau-Museum (DBM) mit dem Bergbau-Archiv, der Bibliothek/Fotothek und den Musealen Sammlungen die klassischen Funktionsbereiche im Dokumentationswesen unter einem Dach.¹ Zentral für das institutionelle Selbstverständnis ist der Ansatz, alle schriftlichen und audio-visuellen Primär- und Sekundärquellen sowie alle dinglichen Objekte begriffserweitert als montanhistorisch relevante Dokumente anzusehen. Voraussetzung für einen integrierten Zugriff auf die Quellen in den verschiedenen Dokumentationsbereichen ist eine vernetzte Erschließung unter Wahrung bereichsspezifischer Anforderungen. Dies impliziert eine Integration und Harmonisierung der unterschiedlichen Erschließungsstrategien. Der Beitrag skizziert das Konzept, den Stand der Arbeiten und die Gewährleistung der Interoperabilität mithilfe entsprechender Standards bzw. Regelwerke. Anschließend wendet er sich der mittelfristig geplanten Dokumentation von Kontextinformationen zu Personen, Körperschaften usw. zu und thematisiert die Frage geeigneter Regelwerke bzw. Standards für eine übergreifende Vernetzung solcher Metadaten.

Einleitung

Die in den letzten Jahren nachhaltig forcierten Digitalisierungsanstrengun-

1 Das Folgende war Grundlage für den gleichnamigen Vortrag des Verfassers im Rahmen der Tagung „hist 2006. Geschichte im Netz. Praxis, Chancen, Visionen“, Berlin, 22.-24.02.2006.

gen von Bibliotheken, Archiven und Museen haben die Zugänglichkeit zu geschichtswissenschaftlich relevanten Fachinformationen grundlegend verbessert. Die Zahl der Digitalisierungsprojekte und Online-Angebote ist mittlerweile nur noch schwer zu überschauen.² So begrüßenswert diese Vielfalt ist, so schwer macht sie es den Benutzern/innen, die für sie jeweils relevanten Informationen zu finden. Damit ist die 2002 vom Bundesministerium für Bildung und Forschung unter dem programmatischen Titel „Informationen vernetzen – Wissen aktivieren“ formulierte Forderung nach einer strategischen Neuausrichtung wissenschaftlicher Informationsversorgung heute vielleicht aktueller denn je.³ Bibliotheken, Archive und Museen haben die Forderungen nach einer engeren Kooperation in jüngerer Zeit aufgegriffen und sich um eine bereichsübergreifende Vernetzung und um die Überwindung tradierter Abgrenzungen zwischen den Dokumentationsbereichen bemüht.⁴ Ein Problem sind dabei nicht zuletzt die unterschiedlichen Erschließungsstrategien und -muster. Dabei stand die ehemals strenge Trennung zwischen den Bereichen seit jeher in einem gewissen Kontrast zur alltäglichen Praxis. Dies bezieht sich zum einen auf die zum Teil gleichartigen Dokumente in den verschiedenen Dokumentationsbereichen. Es sei nur an die Nachlässe in Archiven, Bibliotheken oder Museen gedacht. Zum anderen sind gerade auf kommunaler Ebene und in der Wirt-

-
- 2 Die EUBAM-Liste der deutschen Digitalisierungsprojekte, Stand 28.07.2005, <http://www.dl-forum.de/dateien/EUBAM-Liste_Digitalisierungsprojekte_28-07-05.pdf> (10.02.2006) führt allein 230 laufende und abgeschlossene Projekte auf.
- 3 Bundesministerium für Bildung und Forschung (Hg.), Strategisches Positionspapier. Informationen vernetzen – Wissen aktivieren, Bonn 2002, unter: <http://www.bmbf.de/pub/information_vernetzen-wissen_aktivieren.pdf> (10.02.2006).
- 4 Exemplarisch seien nur die Projekte „Gemeinsames Portal für Bibliotheken, Archive und Museen“ (BAM) – vgl. den Beitrag von Frank von Hagel in diesem Band; <<http://www.bam-portal.de>> (14.02.2006) –, die „Online-Zentralkartei der Autographen“ (Kalliope) – vgl. <<http://kalliope.staatsbibliothek-Berlin.de>> (14.02.2006) – sowie die Kooperation zwischen Bundesarchiv und Clio-online – vgl. <<http://www.bundesarchiv.de/aktuelles/projekte/00100/index.html>> (14.02.2006) – genannt. Vgl. ferner Kluttig, Thekla; Kretzschmar, Robert; Lupprian, Karl-Ernst, Die deutschen Archive in der Informationsgesellschaft. Standortbestimmung und Perspektiven, in: Der Archivar 57 (2004), S. 28-36, hier S. 35.

schaft zahlreiche, meist kleinere oder mittelgroße Misch-Einrichtungen entstanden, die alle drei Dokumentationsbereiche in sich vereinen.

Das Montanhistorische Dokumentationszentrum (montan.dok) beim Deutschen Bergbau-Museum Bochum (DBM) ist solch eine Misch-Einrichtung. Zunächst sollen der organisatorische Zuschnitt, die Aufgaben und die Bestände des montan.dok skizziert und damit der institutionelle Rahmen für das Nachfolgende umrissen werden. Sodann stehen die im montan.dok verfolgten Lösungsansätze und Konzepte für eine bereichsübergreifende Vernetzung bei Erschließung und Verfügbarmachung montanhistorischer Quellen im Fokus. Abschließend stellt sich die Frage nach dem Umgang mit Kontextinformationen.

montan.dok: Archiv, Bibliothek und Museale Sammlungen unter einem Dach

Das DBM ist eines von sieben Forschungsmuseen in der Wissenschaftsgemeinschaft Gottfried Wilhelm Leibniz (WGL). Damit nimmt es nicht nur traditionelle Museumsaufgaben wahr, sondern fungiert zugleich als außeruniversitäres Forschungsinstitut für Montangeschichte. Diese Doppelfunktion kennzeichnet gleichfalls die Arbeit und den Zuschnitt des montan.dok. Es wurde im Zuge einer grundlegenden Neustrukturierung der wissenschaftlichen Informationsversorgung und Forschung im DBM Anfang 2001 als zentrale Servicestelle und als Forschungseinrichtung für die Montangeschichte seit der Industrialisierung gegründet und fasste die bis dahin eigenständigen Bereiche Bergbau-Archiv, Bibliothek/Fotothek und Museale Sammlungen unter einem gemeinsamen Dach zusammen.⁵ Ein konstitutiver Grundgedanke war das Verständnis, alle schriftlichen und audio-visuellen Primär- und Sekundärquellen sowie alle dinglichen Objek-

5 Siehe Farrenkopf, Michael, Bergbau-Archiv und montan.dok, in: Slotta, Rainer (Hg.), 75 Jahre Deutsches Bergbau-Museum Bochum (1930 bis 2005). Vom Wachsen und Werden eines Museums (Veröffentlichungen aus dem Deutschen Bergbau-Museum Bochum, Nr. 134), Bochum 2005, Bd. 1, S. 173-240; Ders., Das Bergbau-Archiv beim Deutschen Bergbau-Museum Bochum. Quellen für eine Technikgeschichte des Bergbaus, in: Rasch, Manfred; Bleidick, Dietmar (Hgg.), Technikgeschichte im Ruhrgebiet – Technikgeschichte für das Ruhrgebiet, Essen 2004, S. 39-54.

te im DBM begriffserweitert als montanhistorisch relevante Dokumente anzusehen, wobei „jedes Dokument für alle drei Dokumentationsfunktionen mehr oder weniger ausgeprägte Nutzungsmöglichkeiten aufweist“.⁶ Wenngleich ein solch umfassendes Quellenverständnis in der Geschichtswissenschaft keineswegs neu ist, so kann es doch nicht zuletzt für die Montangeschichte seit der Industrialisierung neue Perspektiven und Erkenntnispotenziale eröffnen.

Ein kurzer Blick auf die Bestände der einzelnen Teilbereiche des montan.dok mag die Bandbreite und Vielgestaltigkeit der montanhistorisch relevanten Dokumente im DBM illustrieren. Als ältester Teilbereich sind zuerst die Musealen Sammlungen zu nennen. Sie gehen in ihrem Kern auf die Lehr- und Schausammlungen der 1864 gegründeten Westfälischen Berggewerkschaftskasse zurück. Heute umfassen sie gut 250.000 unterschiedliche bergbauliche Relikte aus aller Welt. Das Spektrum reicht von der Briefmarke bis hin zum Fördergerüst über dem DBM als dem größten und wohl markantesten Objekt des Museums.

Wenige Jahre nach Gründung des DBM 1930 wurde eine hauseigene Arbeitsbibliothek für die Museumsmitarbeiter/innen eingerichtet, deren Funktionsprofil mit der gezielten Erwerbung wichtiger montanhistorischer Werke schon bald eine Erweiterung erfuhr. Diese Doppelfunktion besitzt heute noch Gültigkeit: Als wissenschaftliche Spezialbibliothek sammelt sie einschlägige Publikationen zum Montanwesen, bewahrt diese dauerhaft auf und macht sie internen wie externen Benutzern/innen zugänglich. Darüber hinaus hält sie die von den Mitarbeitern/innen und den Wissenschaftlern/innen des DBM für ihre Arbeiten benötigte Literatur vor bzw. beschafft diese. Derzeit umfassen die Bestände circa 30.000 Monografien sowie noch einmal so viele Zeitschriftenbände. Hinzu kommt die angegliederte Fotothek mit weit über 20.000 inventarisierten Fotos sowie einigen

6 Leonhardt, Holm A., Was ist Bibliotheks-, was Archiv- und Museumsgut? Ein Beitrag zur Kategorisierung von Dokumentationsgut und -institutionen, in: Der Archivar 42 (1989), Sp. 213-224, Zit. Sp. 220 f. In diesem Sinn wird der Begriff des Dokumentes nachfolgend als Sammelbegriff für die unterschiedlichen Gattungen und Formausprägungen historischer Überlieferungen, seien es nun Akten, Fotos, Filme oder aber dingliche Objekte, verwendet.

noch nicht erschlossenen Fotosammlungen.

Der jüngste Teilbereich ist das Bergbau-Archiv. Es wurde auf einem Höhepunkt der Bergbaukrise am 1. Juli 1969 gegründet, als die zahlreichen Zechenstilllegungen das Problem einer langfristigen Aufbewahrung der historisch relevanten Unterlagen des Bergbaus akut werden ließen. Als zentrales Wirtschafts- und Branchenarchiv ist das Bergbau-Archiv überregional tätig und sichert und erschließt Unterlagen aus dem gesamten deutschen Bergbau. Gegenwärtig beherbergt es über 250 Bestände und knapp 30 Sondersammlungen mit insgesamt über 4.500 laufenden Metern Archivgut. Infolge der überregionalen Zuständigkeit betreffen die Bestände zunächst verschiedene Bergbausparten, nämlich Erz-, Steinkohlen-, Braunkohlen- und Kalibergbau. Ferner kann man die Bestände in vier Tektonikgruppen untergliedern, nämlich in die Überlieferungen von Konzernen, von Einzelzechen, von Bergbauverbänden und von Privatpersonen.⁷

Mit der Zusammenfassung von Bergbau-Archiv, Bibliothek/Fotothek und Musealen Sammlungen traten die oben angesprochenen Probleme der Abgrenzung und Überschneidungen zwischen den drei Dokumentationsbereichen auch innerhalb des montan.dok zutage. Dies betraf zum einen die gleichartigen Dokumente in den verschiedenen Bereichen, wie zum Beispiel Fotografien oder Karten, Pläne und Risse. Je nachdem, ob sie in einem Provenienzzusammenhang überliefert worden sind oder nicht, wurden sie einem Dokumentationsbereich zugeordnet, was naturgemäß eine Erschließung nach je bereichsspezifischem Muster bedingte. Mit den verschiedenen Erschließungstraditionen bzw. EDV-gestützten Erschließungssystemen und den daraus resultierenden Schwierigkeiten einer Verknüpfung dieser Informationen ist ein zweiter Problembereich benannt. In der Bibliothek/Fotothek und den Musealen Sammlungen wurde eine Eigenentwicklung des DBM auf Basis von MS ACCESS eingesetzt. Sie

⁷ Vgl. Kroker, Evelyn, Das Bergbau-Archiv und seine Bestände (Veröffentlichungen aus dem Deutschen Bergbau-Museum Bochum, Nr. 94; Schriften des Bergbau-Archivs, Nr. 10), Bochum 2001. Eine aktualisierte Beständeübersicht ist über das Portal Archive in Nordrhein-Westfalen <<http://www.archive.nrw.de>> (10.02.2006) abrufbar.

war bei Gründung des montan.dok jedoch veraltet und entsprach den fachlichen Anforderungen neuerer Regelwerke nicht mehr. Im Bergbau-Archiv hatte man sich hingegen „zu einem EDV-bezogenen archivischen Sonderweg“⁸ entschlossen und eine auf archivische Erfordernisse abgestimmte Lösung auf Basis der Software FAUST entwickelt. Angesichts dieser unbefriedigenden und den Zielen des montan.dok in keiner Weise genügenden Situation fiel die Entscheidung zu einer grundlegenden Neustrukturierung und Reorganisation der EDV-gestützten Erschließung und Verwaltung, über die im Folgenden berichtet werden soll.

Bereichsübergreifende Erschließung montanhistorischer Dokumente im montan.dok

Zwei grundsätzliche Vorbemerkungen seien gestattet: Erstens ist das Folgende als Werkstattbericht zu verstehen. Das heißt: Vieles ist bereits umgesetzt, anderes noch in einer Test- oder Evaluierungsphase und manches muss erst noch im Detail erarbeitet werden. Zweitens handelt es sich um einen von vielen denkbaren Ansätzen. Demnach sind die vorgestellten Lösungen auf die spezifischen Erfordernisse des montan.dok als thematisch auf das Montanwesen konzentrierte Dokumentationsstelle mittlerer Größenordnung abgestimmt und lassen sich nicht ohne Weiteres auf andere Bereiche mit anderen Dimensionen oder aber gar auf archivübergreifende Ansätze übertragen.

Zentraler Ausgangspunkt für die EDV-bezogenen Arbeiten und Planungen im montan.dok ist der interdisziplinäre oder ganzheitliche Ansatz mit seinem umfassenden Grundverständnis von montanhistorisch relevanten Dokumenten. Es ist zu betonen, dass damit keine Nivellierung zwischen den klassischen Funktionsbereichen im Dokumentationswesen, etwa durch die Einführung einheitlicher Erschließungsgrundsätze und -muster, angestrebt wird. Vielmehr sind die unterschiedlichen Funktionen der Dokumentationsbereiche und die Verschiedenartigkeit der Dokumententypen integrale Bestandteile des Ansatzes. Es geht um eine Integration und Harmonisierung im Sinn einer gegenseitigen Annäherung und unter Ak-

⁸ Farrenkopf, Michael, Bergbau-Archiv und montan.dok (wie Anm. 5), S. 190.

zeptanz einer Methodenpluralität. Folgerichtig werden die Dokumente anhand einschlägiger Kriterien einem der drei Teilbereiche des montan.dok physisch zugeordnet. Bedingte das bisher jedoch zwangsläufig eine bereichsspezifische Erschließung, so soll diese Zwangsläufigkeit tendenziell aufgehoben werden.⁹

Ziel ist also eine bereichsübergreifende Vernetzung von Erschließungsinformationen unter Wahrung bereichsspezifischer Besonderheiten. Zwei Ebenen mit unterschiedlicher Reichweite sind hierbei zu differenzieren. In den meisten Fällen bleiben die Zuordnung zu einem Dokumentationsbereich und die daraus resultierende Erschließung sachlich-fachlich geboten. Überspitzt formuliert: Keiner wird einen Aktenbestand etwa nach bibliothekarischen Regeln katalogisieren. Mithin gilt es auf einer ersten Ebene, Schnittstellen für Verknüpfungen zwischen den Dokumentationsbereichen zu finden.

Das schließt zunächst eine Harmonisierung der inhaltlichen Erschließung mithilfe gemeinsamer Klassifikations- und Schlagwortkataloge ein, wobei die notwendige semantische Angleichung durch die thematische Konzentration der Dokumentations- und Forschungsarbeiten im montan.dok begünstigt wird. Ein gewisses Problem ist hierbei die Heterogenität und unterschiedliche Qualität der Altdaten. Im Bergbau-Archiv erfolgte eine kategorial differenzierte und kontrollierte Verschlagwortung anhand vorgegebener Begriffslisten. Demgegenüber wurde in den Bereichen Bibliothek/Fotothek und Museale Sammlungen lange Zeit eine vergleichsweise undifferenzierte und freie Verschlagwortung vorgenommen. Allein um Daten- bzw. Qualitätsverluste bei den Altdaten zu vermeiden, musste das neue Erschließungssystem diese Heterogenität in gewissen

9 Unter Erschließung wird hier vergleichsweise undifferenziert und pragmatisch die inhaltliche und formale Beschreibung eines Bestandes oder eines Dokumentes verstanden. Zum unterschiedlichen Gebrauch des Begriffs im Archiv- und im Bibliothekswesen und zu den entsprechenden Implikationen vgl. Nimz, Brigitta, Die Erschließung im Archiv- und Bibliothekswesen unter besonderer Berücksichtigung elektronischer Informationsträger. Ein Vergleich im Interesse der Professionalisierung und Harmonisierung (Texte und Untersuchungen zur Archivpflege, Bd. 14), Münster 2001, hier insbesondere S. 97-104, 205-208 und 277 f.

Grenzen abbilden und sollte zugleich Möglichkeiten für eine, soweit als möglich, automatisierte Harmonisierung bieten.¹⁰

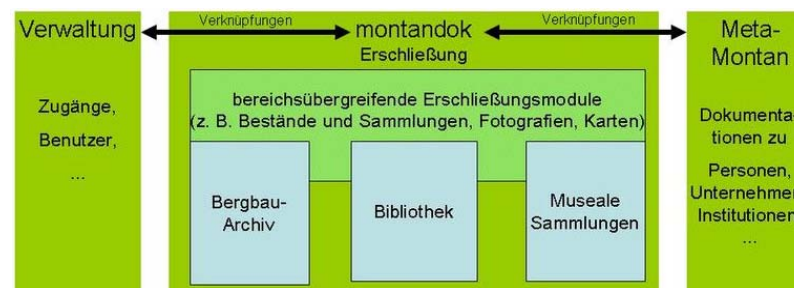


Abbildung 1: Datenbankstruktur im montan.dok

Auf einer zweiten Ebene scheint für bestimmte Dokumenttypen, wie zum Beispiel Fotografien oder Karten, eine über die bloße Verknüpfung von Datensätzen hinausgehende Vernetzung und Harmonisierung praktikabel und sinnvoll. Besonders auf der Ebene einzelner Verzeichnungseinheiten legen die durchaus bestehenden Gemeinsamkeiten zwischen den jeweiligen Erschließungspraktiken eine bereichsübergreifende Nutzung dokumentenspezifischer Erfassungsmasken nahe, mit denen inhaltlich-strukturell und materiell-technisch gleichartige Dokumente in den verschiedenen Teilbereichen aufgrund eines gemeinsamen Kategorienschemas erschlossen werden.

Dies birgt zweifelsohne die Gefahr einer unerwünschten Verwischung von Grenzen zwischen den Dokumentationsbereichen mit ihren spezifischen Funktionen und Anforderungen. Deshalb ist die Verortung jedes Dokumentes bzw. Erschließungsdatensatzes in der Tektonik des montan.dok mit seinen drei Teilbereichen unerlässlich. Darüber hinaus müssen

10 Die Struktur der Altdaten insbesondere aus den Bereichen Bibliothek/Fotothek und Museale Sammlungen lässt eine retrospektive Bearbeitung und zum Teil sogar Neuerschließung unumgänglich erscheinen. Entscheidend war, dass die Altdaten zunächst im System verfügbar und damit recherchierbar sind und hier nun sukzessive angeglichen werden können. Dies wird allerdings noch geraume Zeit in Anspruch nehmen.

bereichsspezifische Erschließungsstrukturen und -informationen, wie zum Beispiel zu Provenienz oder Klassifikation im Archiv, abgebildet werden und bei der Präsentation für die Benutzer/innen erkennbar bleiben. Somit kann für eine bereichsübergreifende Erfassung nur ein Kernset an Kategorien obligatorisch sein, während zusätzlich eine mehr oder minder große Menge bereichsspezifischer Informationen optional abbildbar sein muss.

Um einen integrierten, von der Zuordnung zu einem Dokumentationsbereich entkoppelten Zugriff auf gleichartige Dokumente zu ermöglichen, wird eine virtuelle, quer zu den Bereichen liegende Differenzierung nach Dokumenttypen vorgenommen. Zusammen mit den dokumentspezifischen Erfassungsmasken soll so eine adäquate Zuordnung und Erschließung in einem Dokumentationsbereich sowie zugleich eine übergreifende Recherche ermöglicht werden.



Abbildung 2: Beispiel für eine einfache Stichwortsuche im Intranet-Angebot des montan.dok

Für die Recherche ergeben sich somit mehrere Kombinationsmöglichkeiten: Es kann sowohl bereichsübergreifend nach einem bestimmten Dokumenttyp bzw. nach dessen spezifischen Merkmalsausprägungen gesucht

werden. Ebenso kann natürlich eine typenübergreifende Recherche auf einen Dokumentationsbereich begrenzt oder aber nur ein bestimmter Dokumenttyp in einem bestimmten Dokumentationsbereich gesucht werden.

Ein rechercherelevantes Problem stellt die angesprochene Heterogenität der Altdaten dar. Mit Blick auf die vorhandenen Personalressourcen haben wir uns hier für eine pragmatische und arbeitsökonomisch vertretbare Lösung entschieden. Aus allen rechercherelevanten Feldern, wie zum Beispiel Titelseinträgen, inhaltliche Beschreibungen oder Indexierungen, wird eine Wortliste generiert, die über eine einfache Stichwortsuche abgefragt werden kann. Darüber hinaus sind natürlich differenziertere Recherchen sowie eine gezielte Suche zum Beispiel in der Klassifikation eines Bestandes im Archiv oder aber in den Indizes zu Sachbegriffen, Personen, Institutionen usw. möglich.

Jedoch kann es nicht nur darum gehen, unterschiedliche Erschließungsmuster nebeneinander abzubilden. Im Sinn einer weiter gehenden Integration ist zu prüfen, ob, inwieweit und auf welchen Ebenen in bestimmten Fällen eine Adaption von Methoden, Mustern und Erfahrungen aus anderen Dokumentationsbereichen sinnvoll und möglich ist. Zu denken ist insbesondere an die im Archiv- und – mit Abstrichen – auch im Bibliotheksbereich verbreitete Erschließung auf verschiedenen Hierarchieebenen und an die Erfassung von Kontextinformationen.¹¹ Das bezieht sich natürlich zunächst auf die Beschreibung von Beständen oder Sammlungen, ist aber ebenso auf anderen Erschließungsebenen denkbar.

¹¹ Aus archivischer Sicht hat Reininghaus, Wilfried, Das Erschließungsprojekt „Sachthematisches Inventar zur preußischen Berg-, Hütten- und Salinenverwaltung 1763-1865“. Projektmanagement, Erschließungsstandards und Gesamtindex – Erfahrungsbericht und Desiderata, in: Black-Veldtrup, Mechthild; Farrenkopf, Michael; Reininghaus, Wilfried (Hgg.), Die Überlieferung der preußischen Bergverwaltung. Erfahrungen und Perspektiven zur Bearbeitung des sachthematischen Inventars der preußischen Berg-, Hütten- und Salinenverwaltung, 1763-1865 (Veröffentlichungen aus dem Deutschen Bergbaumuseum Bochum, Nr. 131; Schriften des Bergbau-Archivs, Nr. 17; Veröffentlichungen des Landesarchivs Nordrhein-Westfalen, Bd. 1), Bochum 2005, S. 12-18, hier S. 16, auf die wachsende Bedeutung solcher Kontext- oder Metainformationen hingewiesen.

Ein Beispiel mag dies illustrieren: Im montan.dok finden sich sowohl im Archiv als auch in der Fotothek zahlreiche Dia-Serien. Nicht nur aus arbeitsökonomischen Gründen scheint hier die Erschließung der Serien als solche ausreichend, zumal allen Dias einer Serie bestimmte formale Merkmale und sachliche Bezüge gemeinsam zu Eigen sind. Eine Erschließung einzelner Aufnahmen ist hier eher selten angebracht, muss aber gleichwohl möglich sein, wobei die eindeutige Zuordnung zum übergeordneten Dokument gewahrt bleiben muss.¹² Als ähnlich gelagerte Beispiele aus dem Bereich der Musealen Sammlungen seien mehrteilige Objekte, wie eine Uniform oder ein Kaffeeservice, oder ganze Objektsammlungen genannt. Hier ist es ebenfalls sinnvoll, Angaben zu Zeit und Zweck der Zusammenstellung oder zu Herkunft und Gebrauch auf einer übergeordneten Ebene zu dokumentieren.

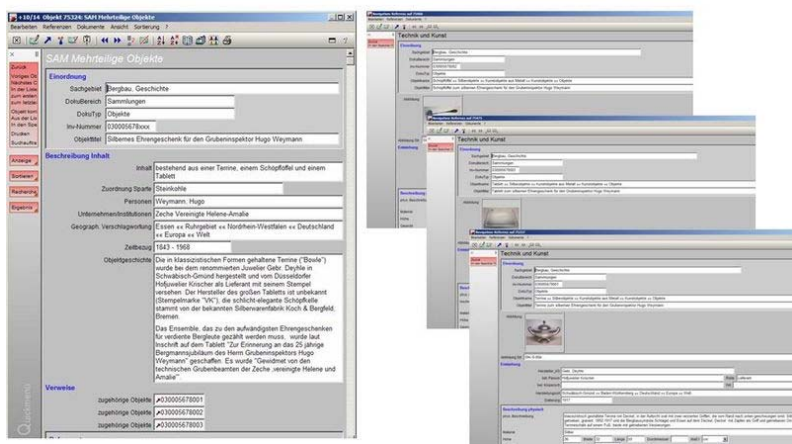


Abbildung 3: Erschließung mehrteiliger Objekte

Eine weitere zentrale Anforderung bei der Neukonzeption der Erschließungsdatenbank war die Kompatibilität der Daten zu gängigen Standards

12 Ähnliches gilt zum Beispiel auch für Fotoalben, die zudem Eigenschaften eines Musealen Objekts besitzen können. Vgl. Jäger, Jens, Photographie: Bilder der Neuzeit. Einführung in die Historische Bildforschung (Historische Einführungen, Bd. 7), Tübingen 2000, S. 155.

oder Best-Practice-Regelwerken. Dadurch sollen eine möglichst plattformunabhängige Datenhaltung und die Interoperabilität der Daten als entscheidende Voraussetzungen für eine langfristige Sicherung und für den Datenaustausch gewährleistet werden. Wie so viele marktgängige Datenbanken ist auch das im montan.dok eingesetzte Programm FAUST eine proprietäre Lösung. Zudem bietet es großen Spielraum für individuelle Anpassungen. Das macht das Programm einerseits attraktiv. Andererseits kann aber gerade diese Flexibilität für die Interoperabilität der Daten große Probleme aufwerfen. Um die daraus resultierenden Gefahren zu minimieren, wurde bei der Konzeption des Systems auf einschlägige Regelwerke und Standards zurückgegriffen. Sie sind die Basis für den Aufbau der Erfassungsmasken und für Schnittstellen, die ein Auslesen der Daten in entsprechend strukturierte XML-Dateien ermöglichen sollen. Auf diese Weise glauben wir die angestrebte Plattformunabhängigkeit und Interoperabilität zu erreichen.

Für eine bereichsübergreifend einsetzbare Schnittstelle ist zunächst das Dublin Core Metadata Element Set zu nennen.¹³ Die geringe Zahl an zentralen Datenkategorien macht es überschaubar und ermöglicht dabei doch die Abbildung wesentlicher Erschließungsinformationen zu einem Dokument. Kontextinformationen, die gerade, aber nicht nur im archivischen Bereich rechercherelevant sind, lassen sich aber nur unzureichend darstellen.¹⁴ Ein Weiteres kommt hinzu: Für eine Online-Präsentation von Erschließungsdaten mag man mit einer relativ geringen Zahl zentraler Datenkategorien auskommen. Für eine umfassende Interoperabilität im Sinn einer plattformunabhängigen Datenhaltung benötigt man hingegen

13 Vgl. <<http://dublincore.org/>> (14.02.2006), hier insbesondere Dublin Core Metadata Element Set, Version 1.1: Reference Description <<http://dublincore.org/documents/dces/>> (14.02.2006).

14 Zur Nutzung des Elementesets von Dublin Core in seiner Ausprägung als DLmeta-XML im Projekt „Gemeinsames Portal für Bibliotheken, Archive und Museen“ (BAM) vgl. 3. Arbeitsbericht an die Deutsche Forschungsgemeinschaft, September 2001 <<http://www.bam-portal.de/rep/BAM-Bericht03.pdf>> (14.02.2006), S. 4 f. und 4. Arbeitsbericht an die Deutsche Forschungsgemeinschaft, September 2003 <<http://www.bam-portal.de/rep/BAM-Bericht04.pdf>> (14.02.2006), S. 2f.

umfassendere Austauschformate.

Deshalb sind weitere, bereichsspezifische Formate herangezogen worden. Die Objekte in den Musealen Sammlungen werden gemäß den Guidelines for Museum Object Information des International Committee for Documentation of the International Council of Museums (ICOM), nach der französischen Bezeichnung Comité International pour la Documentation kurz CIDOC genannt, erschlossen.¹⁵ Für den Bibliotheksbereich sind Dublin Core und MAB II zu nennen. Allerdings können die Kriterien des MAB II mit Blick auf die Qualität der Altdaten und auf die personellen Kapazitäten für deren Überarbeitung auch langfristig nur teilweise erfüllt werden. Insofern sind hier die Elemente des Dublin Core maßgebend.¹⁶ Für die Musealen Sammlungen und die Bibliothek konnte also auf einschlägige Regelwerke zurückgegriffen werden. Problematischer war der archivische Bereich. Hier gibt es zwar mit der Encoded Archival Description (EAD) ein Regelwerk zur Beschreibung von Findmitteln.¹⁷ Allerdings handelt es sich dabei primär um ein Präsentationsformat, das vornehmlich im anglo-amerikanischen Raum genutzt wird und nur begrenzt auf deutsche Verzeichnungsstraditionen anwendbar ist.¹⁸ Aus diesem Grund greifen wir auf

15 Vgl. International Guidelines for Museum Object Information: The CIDOC Information Categories, 1995, Online-Ausgabe unter <<http://www.willpowerinfo.myby.co.uk/cidoc/guide/>> (14.02.2006).

16 Inwieweit eine entsprechende Schnittstelle auf dem Datenmodell DLmeta aufsetzen kann, wird geprüft. Vgl. zu der DLmeta-Initiative und dem Datenmodell <<http://www.dlmeta.de>> (14.02.2006).

17 Vgl. zur Encoded Archival Description (EAD) die offizielle Webseite <<http://www.loc.gov/ead/>> (14.02.2006) mit einer Vielzahl von Materialien, darunter die aktuelle Tag Library. Vgl. zu den Änderungen gegenüber Version 1.0 von 1998 Encoded Archival Description Tag Library. Version 2002, prepared and maintained by the Encoded Archival Description Working Group of the Society of American Archivists and the Network Development and MARC Standards Office of the Library of Congress, Chicago 2002, S. V f.

18 Das Bundesarchiv setzt für die Einbindung von Erschließungsdaten in internationale Portale und Suchmaschinen zwar auf eine Konversion von Erschließungsdaten nach EAD, benutzt aber für die eigene Online-Präsentation von Findmitteln ein eigenes Austauschformat. Vgl. DFG-Projekt „Präsentation von Online-Findbüchern unter Berücksichtigung des EAD-Systems“ <<http://www.bundesarchiv.de/aktuelles/projekte/00005/>> (14.02.2006). Vgl. ferner Black-Veldtrup, Mechthild, EAD und die deutsche Verzeichnungsstradition: Probleme und Chancen auf dem Weg zu einem Austauschformat, in: Dies.;

das im von der Deutschen Forschungsgemeinschaft (DFG) geförderten Projekt „Entwicklung von Werkzeugen zur Retrokonversion archivischer Findmittel“ erarbeitete Austauschformat zurück, das eine differenzierte Abbildung von Erschließungsinformationen als XML-codierte ASCII-Datei erlaubt.¹⁹

Wie weit sind wir mit den Arbeiten gediehen und welche Schritte sind auf längere Sicht geplant? Derzeit sind die Module für die Erschließung von Publikationen in der Bibliothek, von Objekten in den Musealen Sammlungen, von Akten im Bergbau-Archiv sowie für die bereichsübergreifende Erfassung von Beständen und Sammlungen und von Fotografien implementiert, befinden sich zum Teil aber noch in der Erprobung. Ein Modul für die Erschließung von Karten, Plänen und Rissen soll demnächst folgen. Online-Recherchen in den Beständen des montan.dok sind im

Dascher, Otfried; Koppetsch, Axel (Hgg.): Archive vor der Globalisierung? Beiträge zum Symposium des Nordrhein-Westfälischen Hauptstaatsarchivs in Verbindung mit den Allgemeinen Reichsarchiven in Brüssel (Belgien) und Den Haag (Niederlande) vom 11. bis 13. September 2000 in Düsseldorf (Veröffentlichungen der staatlichen Archive des Landes Nordrhein-Westfalen: Reihe E, Beiträge zur Archivpraxis, Bd. 7), Düsseldorf 2001, S. 129-138; Grau, Bernhard, Das deutsch-amerikanische Projekt „Gemeinsames Fachkonzept Online-Erschließung“. Zur Übertragbarkeit der Encoded Archival Description (EAD) auf die archivische Praxis in Deutschland, in: Menne-Haritz, Angelika (Hg.), Online-Findbücher, Suchmaschinen und Portale. Beiträge des 6. Archivwissenschaftlichen Kolloquiums der Archivschule Marburg (Veröffentlichungen der Archivschule Marburg, Institut für Archivwissenschaft, Bd. 35), Marburg 2002, S. 49-65. Die beiden letztgenannten Beiträge beruhen auf der Version 1.0 von EAD, dürften aber trotz der Änderungen in der Version 2002 noch im Großen und Ganzen gültig sein.

19 Das Austauschformat als XML-DTD ist unter <<http://www.archive.nrw.de/findbuch-digital/austauschformat/Austauschformat.dtd>> (14.02.2006) verfügbar; eine Erläuterung der Elemente findet sich unter <http://www.archive.nrw.de/dok/DFG_Abschluss/TagLibrary.doc> (14.02.2006). Vgl. ferner Przigoda, Stefan, Das Ziel- und Austauschformat als universelle Findbuch-DTD, in: Entwicklung von Werkzeugen zur Retrokonversion archivischer Findmittel. Tagung des Nordrhein-Westfälischen Hauptstaatsarchivs zum DFG-Projekt „Entwicklung von Werkzeugen zur Retrokonversion archivischer Findmittel“, Düsseldorf, 22.-23. 09. 2003 <<http://www.archive.nrw.de/dok/tagung-retro/>> (14.02.2006); zum Projekt und den Projektergebnissen die Beiträge von Black-Veldtrup, Mechthild, Motive und Ziele des DFG-Projekts; Richter, Olaf, Präsentation der Software docWORKS-Findbuch; Meusch, Matthias, Das Retrokonversionstool im Echtbetrieb: Voraussetzungen, Bedingungen und Kosten der Nachnutzung, ebd.

Moment nur begrenzt möglich. Lediglich die Beständeübersicht des Bergbau-Archivs ist über das Portal „Archive in NRW“ im Internet verfügbar. Ein Zugriff auf die neue Erschließungsdatenbank kann nur im Intranet des DBM erfolgen, so dass externe Benutzer/innen noch auf herkömmliche Art anfragen bzw. vor Ort recherchieren müssen. Die Freischaltung im Internet soll aber bis Ende 2006 erfolgen. Des Weiteren ist die Integration der Erschließungsdaten des Bergbau-Archivs in das neue nordrhein-westfälische Archivportal vorgesehen, das ebenfalls bis Ende 2006 auf die Findbuchebene erweitert werden soll.²⁰ Basis für eine entsprechende Import-Schnittstelle sind das erwähnte Austauschformat aus dem Retrokonversionsprojekt und eine darauf aufsetzende, im Anschluss an das Projekt entwickelte Archiv.dtd, auf die gleich zurückzukommen sein wird.

Ausblick: Erfassung von Kontextinformationen

Mit Blick auf die angesprochene Relevanz von Kontextinformationen für die Erschließung und Verfügbarmachung montanhistorischer Dokumente soll langfristig eine bereichsübergreifende Dokumentation von Informationen zum Beispiel zu Personen, Körperschaften oder aber Ereignissen erfolgen. Gerade für das montan.dok als interdisziplinäre Service- und Forschungseinrichtung scheint dies ein lohnenswertes Ziel. Dabei stellt sich zwangsläufig die Frage nach geeigneten Regelwerken bzw. Standards. Einen Ansatzpunkt bieten natürlich die in der bibliothekarischen Erschließung gebräuchlichen Normdaten zu Personen und Körperschaften.²¹ Allerdings lassen diese Formate nur eine flache Abbildung einer begrenzten Zahl an Informationskategorien zu. Im Mittelpunkt unserer Überlegung

20 Vgl. Black-Veldtrup, Mechthild, ... Und jetzt die Findbücher. Der Ausbau des Portals www.archive.net, in: Burckhardt, Daniel; Hohls, Rüdiger; Ziegeldorf, Vera (Hgg.), .hist 2003. Geschichte und Neue Medien in Forschung, Archiven, Bibliotheken und Museen. Tagungsband .hist 2003, Bd. 2, S. 499-511 (Historisches Forum, Bd. 7, Teilbd. II) <http://edoc.hu-berlin.de/e_histfor/7_II/PDF/HistFor_7-2005-II.pdf> (24.05.2006).

21 Vgl. zur Personennamendatei (PND) und der Gemeinsamen Körperschaftsdatei (GKD) <<http://www.ddb.de/standardisierung/index.htm>> (16.02.2006); ferner Hengel, Christine, Normdaten und Metadaten. Die Idee eines Internationalen Authority File, in: Zeitschrift für Bibliothekswesen und Bibliographie 50 (2003), H. 4, S. 210-214.

gen stehen deshalb vor allem archivische Formate zur Strukturierung vergleichsweise komplexer Informationen zu Provenienzstellen und Archivbeständen. Zunächst sind hier natürlich die Bemühungen einer internationalen Arbeitsgruppe zur Entwicklung von EAC (Encoded Archival Context) als logische Ergänzung zu EAD zu nennen.²² Inwieweit sich dieses Format für die Zwecke des montan.dok eignet, bleibt abzuwarten und ist noch eingehend zu prüfen.

Eine den Verzeichnungstraditionen in der deutschen Archivlandschaft eher entsprechende Alternative könnte ein im Anschluss an das DFG-Retrokonversionsprojekt entwickeltes Strukturformat sein. Anlass für die Definition dieser so genannten Archiv.dtd war die den Zielen des Retrokonversionsprojektes geschuldete Begrenzung des Austauschformates auf mehr oder minder bestandsbezogene Findmittel. Die Abbildung von Beständeübersichten mit ihren Kontextinformationen ist hingegen nicht oder doch nur sehr unzureichend möglich. Als sinnvoll erscheinende Ergänzung wurde das Austauschformat deshalb in die Archiv.dtd integriert, die ihrerseits eine strukturierte Darstellung solcher Kontextinformationen erlaubt.²³ So finden sich Containerelemente für eine differenzierte Wiedergabe von Informationen zu dem verwahrenden Archiv, zu Be-

22 Vgl. zu EAC die Materialien unter <<http://www.iath.virginia.edu/eac/>> (16.02.2006). Hier ist unter anderem der Arbeitsentwurf vom August 2004 (Betaversion) als XML-DTD oder XML-Schema abrufbar. Eine deutsche Übersetzung der Tag Library von Sebastian Barteleit u.a. findet sich unter <<http://www.staff.uni-marburg.de/~mennehar/mellonprojekt/uebersetzungen/eac.htm>> (16.02.2006). Vgl. ferner Pitti, Daniel V., Creator Description. Encoded Archival Context <http://www.sba.unifi.it/ac/relazioni/pitti_eng.pdf> (16.02.2006); Löbnitz, Anke, Digitalisiertes Archivgut im Internet. Internationaler Workshop, in: Mitteilungen aus dem Bundesarchiv, 13 (2005), H. 1, S. 78 f., auch unter <<http://www.bundesarchiv.de/imperia/md/content/abteilungen/abt/mitteilungen1-05/8.pdf>> (16.02.2006).

23 Vgl. Przigoda, Stefan, Das Ziel- und Austauschformat als universelle Findbuch-DTD (wie Anm. 19). Die Archiv.dtd ist in der Rubrik Austauschformat auf den Seiten der DFG-Vorstudie Retrokonversion archivischer Findmittel <<http://www.archive.nrw.de/findbuch-digital/index.html>> (14.02.2006) abrufbar. Das Austauschformat des DFG-Retrokonversionsprojektes (s. Anm. 19) ist als DTD integriert, so dass das Vorhandensein beider DTDs Voraussetzung für die Funktionsfähigkeit der Archiv.dtd ist. Letztere wurde vom Verfasser, der bis Ende 2002 Mitarbeiter des DFG-Retrokonversionsprojektes war, entwickelt.

standsinhalt oder Bestandsgeschichte. Ebenso können einzelne Tektonik- und Klassifikationsstufen rekursiv ineinander verschachtelt und somit hierarchisch abgebildet und beschrieben werden. Gleiches gilt für Bestände und Teilbestände.

In unserem Zusammenhang sind nicht zuletzt die Möglichkeiten zur strukturierten Abbildung von Metadaten zu Provenienzstellen (Personen und Institutionen) von Interesse. Das Containerelement „Prov_Info“ lässt zum Beispiel differenzierte Angaben zu natürlichen und juristischen Personen bzw. Institutionen sowie zu deren Kompetenzen, zu Gründung und Auflösung oder zu Vorgängern und Nachfolgern zu. Als einfaches Beispiel sei die Beschreibung einer natürlichen Person angeführt.

```
<?xml version="1.0" encoding="UTF-8" ?>
<!DOCTYPE Archiv (View Source for full doctype...)>
-< Archiv >
-< Datei_Info >
-< Dateiname >Beispiel_Person.xml</Dateiname >
-< Erstellung >
-< Datum >2006-02-16</Datum >
-< Bearbeiter >Przigoda</Bearbeiter >
-< Erstellung >
-</Datei_Info >
-< Archivinfo />
-< Bestand >
-< Bestandsname />
-< Provenienz >Walther Birnbaum</Provenienz >
-< Prov_Info >
-< Person >
-< Pers_Name >
-< Vorname >Walther</Vorname >
-< Nachname >Birnbaum</Nachname >
-</Pers_Name >
-< Pers_Name Art="alternative Schreibweise" >
-< Vorname >Walter</Vorname >
-< Nachname >Birnbaum</Nachname >
-</Pers_Name >
-< Datum Dat_Fkt="Geburt">1905-07-19</Datum >
-< Ort Ort_Fkt="Geburtsort">Laubach</Ort >
-< Datum Dat_Fkt="Tod">1985-09-27</Datum >
-< Ort Ort_Fkt="Sterbeort">Hannover</Ort >
-< Beruf_Funktion >1936-1939 Hilfsarbeiter</Beruf_Funktion >
-< Institution Inst_Art="Unternehmen">Mansfeld AG für Bergbau und Hüttenbetrieb</Institution >
-< Beruf_Funktion >1942-1944 Betriebsführer</Beruf_Funktion >
-< Institution Inst_Art="Unternehmen">Karpaten-Öl AG</Institution >
-< Geschlecht >m</Geschlecht >
-< Konfession >kath.</Konfession >
-< Familienstand >ledig</Familienstand >
-< Hilfsfeld Fkt="Biographie" >
-<p>Geboren am 19. Juli 1905 in Laubach (Hessen) - gestorben am 27. September 1985 in Hannover</p>
-<p>Dr. Walther Birnbaum trat nach dem Jurastudium in Gießen 1931 als juristischer Hilfsarbeiter in die Mansfeld AG für Bergbau und Hüttenbetrieb ein und war dort von 1936 bis zu seiner Einberufung zur Reichswehr 1939 als Justiziar tätig. [...]</p>
-</Hilfsfeld >
-< bibref bibre_Art="Literatur">Evelyn Kroker: Das Bergbau-Archiv und seine Bestände, Bochum 2001, S. 446 f.</bibref >
-</Person >
-</Prov_Info >
-</Bestand >
-</Archiv >
```

Abbildung 4: Abbildung von Metadaten zu einer natürlichen Person mittels Archiv.dtd

Bekanntlich steckt der Teufel im Detail, und das gilt in besonderem Maß

für die Entwicklung vergleichsweise komplexer Strukturformate. Insofern versteht sich vor allem die Archiv.dtd als ein Angebot, das gegebenenfalls anzupassen, zu erweitern und zu überarbeiten ist. Aus Sicht des montan.dok ist sie, ebenso wie die anderen Regelwerke, zunächst ein Bezugspunkt für die Gewährleistung der strukturellen Interoperabilität der Erschließungsdaten und damit für deren langfristige Sicherung. Der Weg zu der angestrebten Harmonisierung zwischen Bibliothek, Archiv und Musealen Sammlungen als Voraussetzung für einen umfassenden und integrierten Zugriff auf die gesamte Palette montanhistorisch relevanter Dokumente ist – das wird deutlich geworden sein – noch lang. Dabei ist der Rückgriff auf gemeinsame Regelwerke als Basis für eine Integration der Erschließungsdaten in übergreifende Suchmaschinen und Portale von zentraler Bedeutung. Das gilt zumal für kleinere und mittelgroße Einrichtungen mit vergleichsweise begrenzten Ressourcen wie das montan.dok.

Dr. Stefan Przigoda ist der Bereichsleiter der Bibliothek/Fotothek im Montanhistorischen Dokumentationszentrum beim Deutschen Bergbaumuseum Bochum. Interessenschwerpunkte: Fragen der Erschließung in Archiven, Bibliotheken und Museen; Geschichte des Bergbaus, insbesondere Bergbauverbände, Biografien und Geschichte des Industrie- und Bergbaufilms. E-Mail: stefan.przigoda@bergbaumuseum.de

VOM NUTZEN DER STRUKTUREN – ARCHIVISCHE RECHERCHE-STRATEGIEN IM INTERNET

von Anke Löbnitz und Jessica von Seggern

Die Erschließung von Archivgut als Voraussetzung für dessen Zugänglichkeit gehört zu den Kernaufgaben der Archive. Während in der Regel die Ergebnisse der Erschließung in Form von gedruckten Findbüchern zur Verfügung gestellt wurden, wird seit Mitte der 1990er Jahre auch das Internet für die Bereitstellung von Findmitteln genutzt. Um eine möglichst große Kompatibilität und damit Verwendbarkeit der Daten zu erreichen, müssen bestimmte archivfachliche und technische Standards entwickelt und angewendet werden. Diese ermöglichen einen komfortablen und flexiblen Zugang zu den Daten unter Anwendung verschiedener Recherchestrategien sowie die Einbindung in bestands- und institutionenübergreifende Suchmaschinen auch auf internationaler Ebene.

In den meisten Fällen ist der erste Schritt zur Vorbereitung eines Archivbesuchs die Recherche im Internet. Benutzer/innen können sich über die Internetseiten der Archive informieren, welche Archive für die von ihnen verfolgte Fragestellung relevant und wie beispielsweise die Öffnungszeiten und die Nutzungsmodalitäten geregelt sind. Darüber hinaus ermöglichen viele Archive ihren Nutzern/innen auch schon einen Überblick über die vorhandenen Bestände, indem Beständeübersichten online zur Verfügung gestellt werden.¹ Weitere Ebenen eines archivischen Internetangebots sind

¹ Zu den Möglichkeiten der Präsentation von Archiven im Internet siehe Flamme, Paul; Herkert, Udo; Viergutz, Volker, Hinweise zur Darstellung staatlicher Archive und Archivverwaltungen im WorldWideWeb des Internet, in: Der Archivar 51, Heft 2 (1998), Sp. 217-228.

Online-Findmittel und die Bereitstellung digitalisierten Archivguts.² Insbesondere über das Internet verfügbare Findmittel bieten die Möglichkeit, den Archivbesuch unabhängig vom jeweiligen Aufenthaltsort und den Öffnungszeiten effizient vorzubereiten und sich einen Überblick über die einschlägigen Akten zu verschaffen. Gegenüber den gedruckten Findmitteln erweitert das Internet die Zugänglichkeit des Archivguts durch einen mehrdimensionalen Zugriff mit komfortablen Navigationsmöglichkeiten. Dieser Mehrwert des Online-Zugangs soll beispielhaft an den Gewerkschaftsbeständen der Stiftung Archiv der Parteien und Massenorganisationen der DDR im Bundesarchiv (SAPMO) dargestellt werden. Diese umfassen den Bundesvorstand des FDGB (Freier Deutscher Gewerkschaftsbund) inklusive der Gewerkschaftshochschule und des Gewerkschaftsverlags sowie 20 Einzelgewerkschaften. Insgesamt handelt es sich dabei um 78.960 Akteneinheiten. Seit Ende Februar 2006 sind diese Bestände vollständig, wenn auch mit unterschiedlicher Erschließungstiefe, im Internet recherchierbar. Zu einem Teil wurden die Akten direkt mit dem Verzeichnungsprogramm MidosaXML³ erschlossen, der andere Teil wurde durch die Retrokonversion analoger Findhilfsmittel in Form von Karteikar-

² Der Präsident des Bundesarchivs hat als anzustrebende Zahl für die Zugänglichkeit allgemeiner Informationen über Archive und ihre Nutzungsmöglichkeiten von 100 Prozent gesprochen. Für die Bereitstellung von Findmitteln hält er dieses Ziel ebenfalls für wünschenswert, geht aber davon aus, dass realistisch etwa zehn Prozent des Archivguts über Online-Findmittel recherchierbar sein können. Angesichts des hohen Aufwands, der mit der Digitalisierung von Archivgut verbunden ist, geht er hier von einer Verfügbarkeit von einem Prozent aus. Vgl. Weber, Hartmut, Digitale Repertorien, virtueller Lesesaal und Praktikum im WWW – neue Dienstleistungsangebote der Archive an die Forschung, in: Fundus – Forum für Geschichte und ihre Quellen, <<http://webdoc.sub.gwdg.de/edoc/p/fundus/4/weber.pdf>> (27.03.2006).

³ Vgl. Menne-Haritz, Angelika, MidosaXML – der Findbucheditor – XML-basierendes Werkzeug zur Erfassung und redaktionellen Bearbeitung von Findbüchern entwickelt für PARSIFAL, in: Archivschule Marburg, <<http://www.archivschule.de/content/108.html>> (29.03.2006). Neben MidosaXML dient im Bundesarchiv außerdem BASYS-Fox zur Generierung von Online-Findmitteln, vgl. Sander, Oliver, Elektronisches Erschließen. Online-Findmittel des Bundesarchivs mit BASYS-Fox, in: Bundesarchiv, <<http://www.bundesarchiv.de/imperia/md/content/abteilungen/abt/47.pdf>> (27.03.2006).

ten oder Ablieferungsverzeichnissen in das Programm überführt.⁴

Archivische Findmittel im Internet müssen ebenso wie analoge Findbücher die Aufgabe erfüllen, den Benutzern/innen Antworten auf ihre Fragestellungen zu geben. Der übliche Zugang zu (analogem) Archivgut erfolgte bisher über die Klassifikation der Bestände oder einen Index. Allerdings erhalten die Benutzer/innen üblicherweise nicht wie beim Nachschlagen in einem Lexikon eine eindeutige und abschließende Antwort auf ihre Fragestellung, sondern sie müssen die ihnen angebotenen Informationen zunächst interpretieren. Dabei spielen die Strukturen und der Kontext, in den jede Archivalieneinheit eingebunden ist und der durch die Einordnung in die Klassifikation festgehalten wird, eine entscheidende Rolle.⁵

Während diese Zugangsmöglichkeiten auch bei einer Online-Recherche erhalten bleiben sollen, sind die Nutzer/innen des Internets daran gewöhnt, einen Suchbegriff in eine der zahlreichen Suchmaschinen einzugeben und dann aus einer mehr oder weniger großen Anzahl von Treffern die benötigten Informationen auszuwählen. Bei dieser Vorgehensweise scheinen Strukturen keine große Rolle zu spielen, da es voll-

4 Information zum Retrokonversionsprojekt des Bundesarchivs, vgl. Drei-Jahres-Projekt zur Retrokonversion von Findmitteln des Bundesarchivs, in: Bundesarchiv, <<http://www.bundesarchiv.de/aktuelles/projekte/00126/index.html>> (27.03.2006) sowie Rauschenbach, Petra, Retrokonversion von Findkarteien zu online - Findbüchern. Ergebnisse und Mehrwert einer digitalen Dienstleistung, in: Bundesarchiv, <<http://www.bundesarchiv.de/imperia/md/content/abteilungen/abtb/51.pdf>> (27.03.2006).

5 Menne-Haritz, Angelika, Internet und Archive – Die Wiederentdeckung der Strukturen, in: Dies. (Hg.), Online-Findbücher, Suchmaschinen und Portale. Beiträge des 6. Archivwissenschaftlichen Kolloquiums der Archivschule Marburg (Veröffentlichungen der Archivschule Marburg 35), Marburg 2002, S. 9-17; Zur Bedeutung der Verwendung strukturierter Informationen vgl. auch: DeRose, Steven J., Navigation, access, and control using structured information, in: *American Archivist* 60 (1997), S. 298-309. Ebenso betont Haworth die Notwendigkeit der Strukturen im Rahmen der archivischen Verzeichnung: „Descriptions of archival records must represent both the content and the context of their creation in order to convey to users two essential qualities; their impartiality and authenticity as evidence of actions and transactions by individuals and organizations.“, Haworth, Kent M., *Archival Description: Content and Context in Search of Structure*, in: Pitti, Daniel V.; Duff, Wendy M. (Hgg.), *Encoded Archival Description on the Internet*, New York 2001, S. 7-26, bes. S. 7.

kommen ausreicht, einen Begriff einzugeben und aus einer Trefferliste die gewünschten Informationen auszuwählen. Gerade für diese Auswahl benötigen Benutzer/innen Kontextinformationen, die sich aus der Struktur der Bestände ergeben, und die es ihnen ermöglichen, redundante Informationen auszuschließen und auch die Informationen auszuwählen, die nur zwischen den Zeilen stehen und die für ihre Fragestellung von Interesse sind. Mittels Navigation und Suchtechnologie sollen Benutzer/innen in die Lage versetzt werden, gezielt Daten abzufragen sowie Querverweisen nachzugehen und dabei jeweils die für ihre Fragestellung vorteilhafteste Strategie für die Suche zu wählen. Die Recherchemöglichkeiten, die den Benutzern/innen zur Verfügung gestellt werden sollen, umfassen die Volltextrecherche, die Navigation entlang der Klassifikation des Bestandes, das Blättern innerhalb der Findbücher sowie einen Zugang über Indices.⁶

Online-Beständeübersicht

Kurzinformationen zu Archivgut erhalten Benutzer/innen im Internet über Beständeübersichten. Diese umfassen idealerweise alle Bestände eines Archivs und enthalten knappe Informationen zur Geschichte des Bestandes und der Provenienzstelle sowie Angaben zu Umfang und Laufzeit. In der Beständeübersicht des Bundesarchivs haben Benutzer/innen die Möglichkeit, eine Volltextrecherche oder eine erweiterte Suche durchzuführen oder durch die Struktur frei zu navigieren.⁷ Der Zugang über die Beständeübersicht vermittelt Benutzern/innen zudem die Einordnung des einzelnen Bestandes in die Gesamtüberlieferung.

6 Zur Bedeutung der Fragestellungen der Benutzer/innen vgl. Coats, Lisa R., *Users of EAD Finding Aids: Who are they and are they satisfied?*, in: *Journal of Archival Organization* 2 (2004), S. 25-39.

7 Vgl. Beständeübersicht des Bundesarchivs, in: Bundesarchiv, <http://www.bundesarchiv.de/bestaende_findmittel/bestaendeuebersicht/index_frameset.html> (27.03.2006).

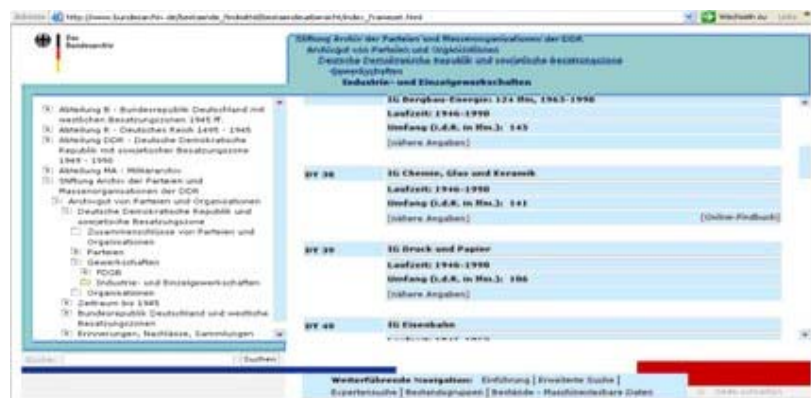


Abbildung 1: Beständeübersicht des Bundesarchivs

Online-Findbücher

Weiter gibt es zu jedem Gewerkschaftsbestand der ehemaligen DDR⁸ ein digitales Findbuch, das im Bundesarchiv unter anderem über die Beständeübersicht oder über eine Liste der verfügbaren Online-Findbücher aufgerufen werden kann.⁹

Durch die Beständeübersicht und die Online-Findbücher sind die Gewerkschaftsbestände der SAPMO vollständig im Internet recherchierbar. Beide Zugänge ermöglichen es den Benutzern/innen, sich Informationen über möglicherweise weitere für ihre Fragestellung interessante Bestände

- 8 Der Bestand des Bundesvorstandes des FDGB ist in Teilbestandsfindbücher unterteilt worden, so liegen beispielsweise für die Abteilung Sozialpolitik, für die Büros der Sekretäre für Sozialpolitik und für die Abteilung Arbeit und Löhne/Rahmenkollektivverträge separate Findbücher vor. Auch hier erweisen sich die Möglichkeiten einer Online-Präsentation als vorteilhaft gegenüber einer Präsentation in gedruckter Form, da ohne Probleme beispielsweise aus großen Beständen zunächst besonders stark nachgefragte Teile verzeichnet und zeitnah zur Verfügung gestellt werden können. Zu einem späteren Zeitpunkt können diese Datenbestände ohne Mehraufwand wieder zusammengefügt werden.
- 9 Auf der Startseite des Bundesarchivs (<<http://www.bundesarchiv.de>>) können die Optionen Beständeübersicht Online und Findmittel Online (<http://www.bundesarchiv.de/bestaende_findmittel/findmittel_online/index.html> (27.03.2006)) ausgewählt werden. Letztere Möglichkeit führt zu den Aufstellungen der Abteilungen über die verfügbaren Findmittel.

zu verschaffen und sie können die für ihre Fragestellungen und ihren Kenntnisstand vorteilhafteste Suchstrategie auswählen. Wenn bereits bekannt ist, in welchem Bestand sich die gesuchten Informationen befinden, so ist auch eine Beschränkung der Suche auf ein einzelnes Findbuch möglich.



Abbildung 2: Zugang zu den Gewerkschaftsbeständen über eine Liste der Online-Findmittel

Ein weiterer Vorteil, den die Verwendung der Internettechnologie mit sich bringt, ist die Möglichkeit einer Volltextsuche. Wenn man ein Online-Findbuch aufruft, kann man in ein Suchformular einen Begriff eingeben und erhält eine Trefferliste, die zunächst auf die Gliederungspunkte verweist, innerhalb derer sich relevante Verzeichnungseinheiten finden. Durch Anklicken der Gliederungspunkte wird man zu den Verzeichnungseinheiten geführt und gelangt schließlich über diese direkt an die Stelle im Findbuch, an der sich die Verzeichnungseinheit befindet. Damit erhalten die Benutzer/innen nicht nur die Information, wo sie Archivalien finden können, die mit ihrem Suchbegriff korrelieren, sondern sie erfahren gleichzeitig, in welchem Kontext diese Archivalien entstanden sind.

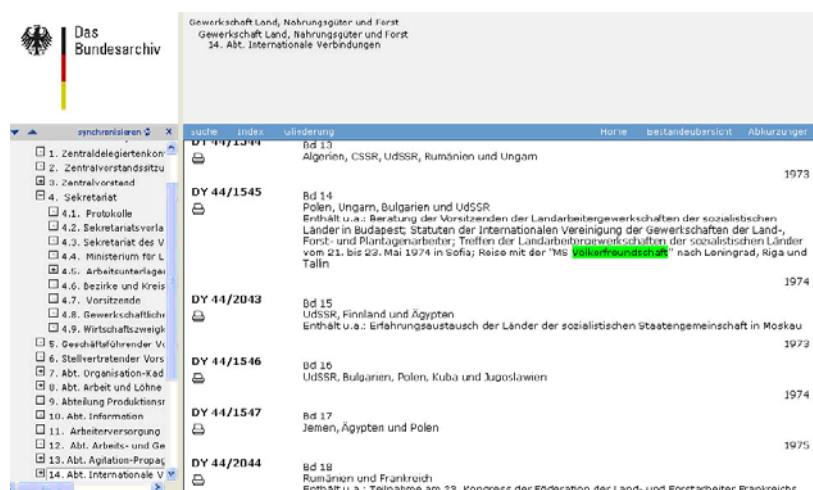


Abbildung 3: Ergebnis der Volltextsuche innerhalb eines Findbuchs

Im Kopfframe der Anzeige stehen genaue Informationen über die Stellung der Verzeichnungseinheit innerhalb der Klassifikation des Bestandes. Zudem besteht die Möglichkeit, die Suche nach weiteren relevanten Informationen durch das Navigieren innerhalb der im linken Frame angezeigten Klassifikation fortzusetzen.

Wie bei einem analogen Findbuch auch, können die Benutzer/innen den Zugang über einen Index wählen. In der Online-Version wird immer ein Mischindex zur Verfügung gestellt, der je nach vorheriger Eingabe beispielsweise einen Sach-, Personen- und Ortsindex umfassen kann. Der Zugang über den Index kann sinnvoll sein, wenn man sich einen Überblick über die in einem Bestand enthaltenen Stichwörter verschaffen möchte, oder wenn die Volltextsuche nicht zu einem Erfolg geführt hat, weil die Schreibweise eines Namens oder Begriffs unklar war.

Im Index werden die Bestellnummern der Verzeichnungseinheiten angegeben, die den gesuchten Begriff enthalten. Durch Anklicken der Bestellnummer gelangt man zu der jeweiligen Verzeichnungseinheit, in der der Begriff markiert ist und so leicht aufgefunden werden kann. Wie in

analogen Findbüchern auch erhalten Benutzer/innen zusätzliche Informationen über die Struktur des Bestandes und die Einordnung der für ihre Fragestellung wichtigen Verzeichnungseinheiten. Auch hier ist wiederum eine Fortsetzung der Suche über die Klassifikation im linken Frame der Anzeige möglich.

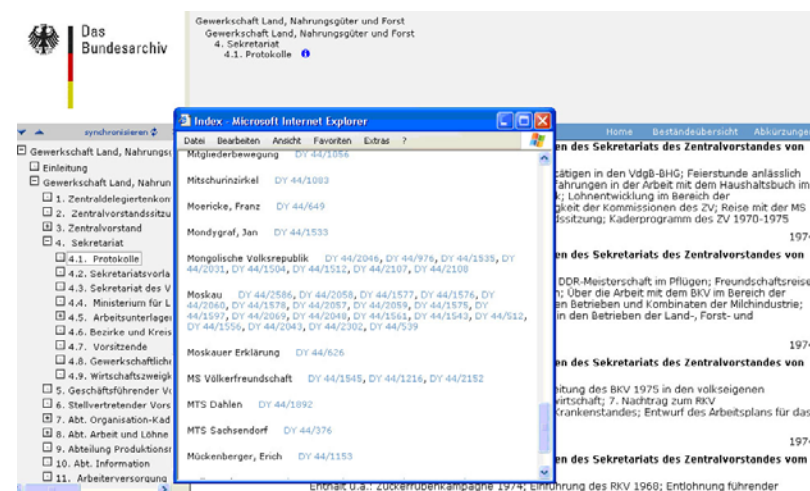


Abbildung 4: Suche innerhalb eines Findbuchs über den Index

Die Möglichkeit, innerhalb eines Findbuchs Links anzubringen, eröffnet zusätzlich die Gelegenheit, einzelne Verzeichnungseinheiten durch Abbildungen zu ergänzen. Diese Abbildungen können ein oder mehrere Blätter aus einer Akte enthalten, die besonders wichtig, wie zum Beispiel ein Organisationsplan, oder optisch besonders reizvoll sind.

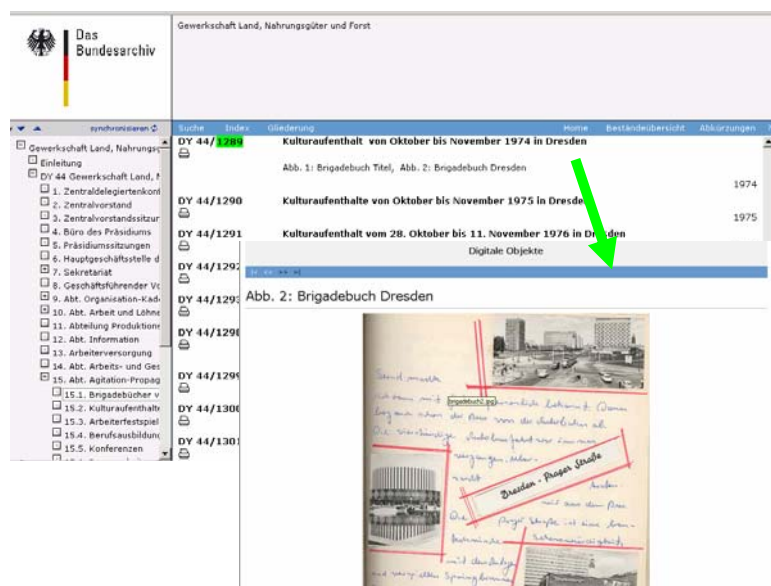


Abbildung 5: Einbindung digitaler Objekte

Auf diese Weise können ausgewählte digitalisierte Objekte im Kontext der Erschließungsinformationen dargestellt werden und den Benutzern/innen bereits im Internet erste Einblicke darüber bieten, was sie in den Beständen erwartet. Neben den Recherchemöglichkeiten seien einige Anmerkungen zu weiteren Funktionen erlaubt, die einen Mehrwert der Online-Findmittel gegenüber analogen Findbüchern begründen. Viele wichtige Informationen beispielsweise zur Geschichte eines Bestandes, zu Bewertungsentscheidungen oder zur Überlieferungsgeschichte werden in analogen Findmitteln in der Einleitung zusammengefasst. Diese lesen Benutzer/innen häufig dann, wenn sie ein Findbuch zum ersten Mal einsehen, aber sie erhalten die Informationen nicht in Zusammenhang mit den Verzeichnungseinheiten, bei denen sie diese tatsächlich benötigen. MidosaXML eröffnet die Möglichkeit, Zusatzinformationen auf allen Ebenen des Bestandes, also sowohl direkt bei den Verzeichnungseinheiten als auch in Zusammenhang mit den Gliederungspunkten, anzufügen. Auf diese Weise können Bewertungsent-

scheidungen transparent gemacht oder Lücken in der Überlieferung gekennzeichnet werden, und da die Benutzer/innen die Informationen parallel zu ihrer Recherche erhalten, geraten diese nicht so leicht in Vergessenheit.¹⁰

MidosaSEARCH

Innerhalb der Gewerkschaftsbestände bilden die Einzelgewerkschaften, der Bundesvorstand, ja sogar einzelne Abteilungen des Bundesvorstandes von einander getrennte Bestände. Für bestimmte Fragestellungen oder wenn Benutzer/innen bei Beginn ihrer Recherche noch nicht wissen, innerhalb welcher Bestände sie Archivalien für ihr Thema finden, ist deshalb eine Suchmöglichkeit über mehrere Bestände wichtig. Diese Suchfunktion wird durch eine Einbindung der Findbücher in MidosaSEARCH ermöglicht. Mit dieser Suchmaschine ist eine übergreifende Suche über alle mit MidosaXML erschlossenen Bestände möglich.¹¹ Gleichzeitig können Benutzer/innen aber auch die Suche auf ausgewählte Findbücher einschränken, also beispielsweise alle Gewerkschaftsbestände oder eine für ihr Thema relevante Gewerkschaft und zusätzlich den Bestand Bundesvorstand des FDGB durchsuchen.

10 Black-Veldtrup, Mechthild, Recherche via Internet: Neue Wege zum Archivgut, in: Fundus – Forum für Geschichte und ihre Quellen, <<http://webdoc.sub.gwdg.de/edoc/p/fundus/4/black.pdf>> (27.03.2006).

11 Siehe die Suchmaschine unter: <<http://212.88.135.193/MidosaSEARCH/web/>> (27.03.2006).

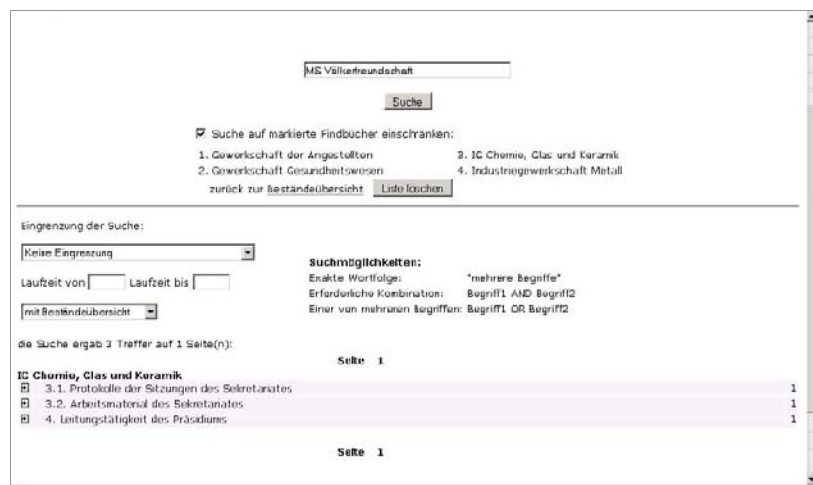


Abbildung 6: MidosoSEARCH: Suche über mehrere Bestände

Wie bei anderen Suchmaschinen auch kann hier mit Hilfe der Booleschen Operatoren nach einer bestimmten Wortkombination, zwei Begriffen, die gleichzeitig vorkommen müssen oder mehreren Begriffen, von denen mindestens einer vorkommen muss, gesucht werden. Außerdem lässt sich die Suche durch Abfrage der Laufzeit der für die Fragestellung relevanten Archivalien weiter eingrenzen, zudem können die Benutzer/innen entscheiden, ob sie die in der Beständeübersicht enthaltenen Angaben ebenfalls nach den eingegebenen Begriffen durchsuchen möchten. Als Ergebnis der Suche wird eine Liste der Findmittel mit den jeweiligen Gliederungsgruppen, in denen der Suchbegriff gefunden wurde, angezeigt. Die Treffer werden also bereits im Rahmen der wesentlichen Kontextinformationen angezeigt, so dass Benutzer/innen bereits an dieser Stelle eine Entscheidung darüber treffen können, ob der Treffer für ihr Forschungsvorhaben relevant ist. Durch Anklicken des Treffers gelangt man in das jeweilige Findbuch. Auch hier finden sich Benutzer/innen also unmittelbar in der Struktur des Bestandes wieder und haben die Möglichkeit, zusätzliche Informationen aufzunehmen und die Suche über die Klassifikation des

jeweiligen Bestandes fortzusetzen. Wenn man sich einmal in der Findbuchansicht befindet, kann man von Treffer zu Treffer navigieren und dabei auch von einem Findbuch in das nächste wechseln, wobei die Kontextinformationen jedes Mal aktualisiert werden. Gleichzeitig können Benutzer/innen – falls sie sich für ein Findbuch besonders interessieren – einen Großteil der erwähnten Suchoptionen nutzen, die innerhalb eines einzelnen Findbuchs zur Verfügung stehen.

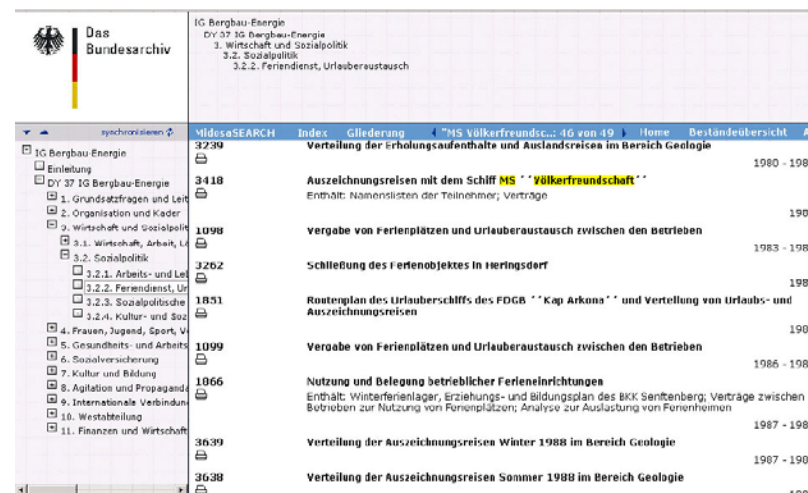


Abbildung 7: Hervorgehobene Treffer im Online-Findbuch

Die archivische Suchmaschine MidosoSEARCH wird derzeit in zwei Varianten eingesetzt. In der SAPMO ermöglicht sie den Benutzern/innen die Recherche über zwei Ebenen: die Ebene der Beständeübersicht und die Ebene der Online-Findbücher. Das heißt, alle Erschließungsinformationen können integriert durchsucht werden. Beim Netzwerk SED-Archivgut, das einen zentralen Sucheinstieg zu den SED-Beständen der neuen Bundesländer, des Landes Berlin und des Bundes anbietet, ermöglicht die Suchmaschine ebenfalls eine Suche über zwei Ebenen hinweg. Hier können Benutzer/innen übergreifend in den Kontaktdaten aller bereitstellenden

Archive und ihren Beständedaten recherchieren.¹² Damit ist nicht nur die Suche über verschiedene Bestände, sondern auch Institutionen hinweg möglich. Eine technisch realisierbare Erweiterung von MidosaSEARCH bei SED-Archivgut wäre es, eine dritte Ebene, nämlich die der Online-Findbücher, zu integrieren. Benutzern/innen stünde auf diese Weise ein zentraler Sucheinstieg zu allen Erschließungsdaten zum SED-Archivgut aus verschiedenen Archiven zur Verfügung. Nach dem gelungenen Einsatz von MidosaSEARCH für das Archiv- und Bibliotheksgut der SAPMO und im Rahmen von SED-Archivgut plant das Bundesarchiv, die Suchmaschine für alle Schriftgutbestände einzusetzen und so Benutzern/innen einen einheitlichen Einstieg anzubieten.

Der Strukturstandard EAD

Um den Benutzern/innen die Möglichkeiten einer findmittelübergreifenden Recherche sowie eines navigierenden Zugriffs bieten zu können, bedarf es einer Technologie, die dieses Vorgehen unterstützt. Gleichzeitig sollte die zugrundeliegende Technologie so weit standardisiert sein, dass ein Austausch der Daten auf nationaler oder internationaler Ebene, also auch die Einbindung in übergreifende Suchmaschinen ermöglicht wird. So gestattet die Standardisierung zugleich, dass das Archivgut einem größeren Benutzerkreis zugänglich gemacht wird.¹³

Ein internationaler Standard, der diese Anforderungen ebenso wie archivische Anforderungen an Verzeichnungsstandards erfüllt, ist Encoded Archival Description (EAD).¹⁴ Bei EAD handelt es sich um einen ameri-

12 Vgl. das Portal SED-Archivgut, in: Bundesarchiv, <<http://www.bundesarchiv.de/sed-archivgut/index.html>> (27.03.2006).

13 Haworth, Archival Description (wie Anm. 5), S. 10.

14 Grau, Bernhard, Die „Encoded Archival Description“ als standardisierter Weg zur Erstellung von Online-Findmitteln, in: Die staatlichen Archive in Bayern, vgl. <<http://www.gda.bayern.de/eadgrau.htm>> (27.03.2006); Black-Veldtrup, Mechthild, EAD und die deutsche Verzeichnungstradition: Probleme und Chancen auf dem Weg zu einem Austauschformat, in: Dies.; Dascher, Ottfried; Koppetsch, Axel (Hgg.), Archive vor der Globalisierung. Beiträge zum Symposium des Nordrhein-Westfälischen Hauptstaatsarchivs mit den Allgemeinen Reichsarchiven in Brüssel (Belgien) und Den Haag (Niederlande) vom 11. bis 13. September 2000 in Düsseldorf (Veröffentlichungen der

kanischen Standard, der ursprünglich im Bibliotheksbereich entstanden ist und von einer Working Group der Society of American Archivists (SAA) entwickelt wurde und von dieser auch gepflegt wird.¹⁵ Mittlerweile haben zahlreiche Tagungen unter internationaler Beteiligung sowie eine Adaption des Standards in Archiven verschiedener Länder stattgefunden, so dass inzwischen die Erfahrungen und Anforderungen internationaler archivischer Belange in die Entwicklung einbezogen werden konnten.¹⁶

EAD ist ein Strukturstandard, der ein Findmittel als einen Text ansieht und diesen mit Hilfe von Elementen und Attributen kodiert. Eine Liste mit Beschreibungen dieser Elemente und Attribute findet sich in der EAD-Tag-Library, eine Anleitung zur Erstellung von EAD-Findbüchern existiert in Form des auf den Webseiten der Library of Congress zugänglichen EAD-Anwenderleitfadens sowie des von Michael J. Fox erstellten EAD-Cookbook.¹⁷ Die Technologie, die EAD zugrunde liegt, ist die Standard Generalized Markup Language (SGML) bzw. die Extensible Markup Language (XML).¹⁸ Mit Hilfe von XML können Informationen über die

staatlichen Archive des Landes Nordrhein-Westfalen, E 7), Düsseldorf 2001, S. 129-138.; Dies., Recherche via Internet (wie Anm. 10).

15 Vgl. <<http://www.loc.gov/ead/eadwg.html>> (27.03.2006). Ausführliche Informationen, Materialien zum Download und Dokumentationen zur Entwicklung von EAD finden sich auf der Internetseite der Library of Congress, <<http://www.loc.gov/ead>> (27.03.2006) sowie auf den Internetseiten des Projektes daofind des Bundesarchivs und der Andrew W. Mellon Foundation, <<http://www.daofind.de>> (27.03.2006). Letzteres bietet zu einer Reihe von Materialien auch deutsche Übersetzungen an. Zur Akzeptanz von EAD, insbesondere des zur Unterstützung der Implementierung von EAD entwickelten „Kochbuchs“, vgl. Prom, Christopher, The EAD Cookbook: A Survey and Usability Study, in: The American Archivist 65 (2002), S. 257-275.

16 Pitti benennt die „archival community“ als verantwortlich für die Weiterentwicklung des Formats, vgl. Pitti, Daniel V., Encoded Archival Description. An Introduction and Overview, <<http://www.dlib.org/dlib/november99/11pitti.html>> (27.03.2006).

17 Die Tag-Library findet sich in: Library of Congress, vgl. <<http://www.loc.gov/ead/tglib/>> (27.03.2006), eine auszugsweise Übersetzung wird über das Projekt daofind in: Daofind, <<http://www.staff.uni-marburg.de/~mennehar/mellonprojekt/uebersetzungen/uebersetzung.htm>> (27.03.2006) zur Verfügung gestellt. Die Übersetzung des EAD-Anwenderleitfadens ist in Arbeit.

18 Pitti, Daniel V., Encoded Archival Description. Die Entwicklung eines Kodierungsstandards für Archivische Suchhilfen, in: Fundus – Forum für

Struktur eines Textes in diesen eingefügt und weltweit ausgetauscht werden.¹⁹ Auf diese Weise lassen sich archivische Informationen über das Internet zugänglich machen, wie auch gedruckte Findbücher erstellen, wobei die einmal eingegebenen Daten für beide Versionen verwendet werden können. Diese Wiederverwendbarkeit der Daten ist ebenso wie die Plattformunabhängigkeit und die damit verbundene Nachnutzbarkeit sowie die Flexibilität ein großer Vorteil des Formats, mit dessen Hilfe die Struktur eines Findbuchs beschrieben werden kann. Neben der Darstellung der Gliederungsgruppen und Verzeichnungseinheiten ermöglicht EAD zudem die Integration von Titelblatt und Einleitung sowie die Generierung eines Index.

EAD erlaubt die Darstellung miteinander verknüpfter Erschließungsinformationen sowie die Wiedergabe hierarchischer Zusammenhänge durch mehrfach gestufte Verzeichnungsebenen und bietet sowohl den Zugang über eine systematische Recherche durch eine Navigation entlang der Klassifikation der Bestände wie eine Volltextrecherche nach Stich- und Schlagwörtern.²⁰ EAD enthält in Form einer Document Type Definition (DTD) die Strukturanweisungen für maschinenlesbare Findmittel und gewährleistet die einheitliche Strukturierung verschiedener Dokumente.²¹ Während diese Einheitlichkeit einen hohen Wiedererkennungswert fördert und es den Benutzern/innen erleichtert, sich in den Findmitteln zurechtzu-

Geschichte und ihre Quellen, <<http://webdoc.sub.gwdg.de/edoc/p/fundus/4/pitti.pdf>> (27.03.2006).

- 19 Fox, Michael J., Encoded Archival Description: An EAD Primer, in: ICA, <http://www.wien2004.ica.org/imagesUpload/pres_109_FOX_CDS05.pdf> (27.03.2006).
- 20 Grau, Die „Encoded Archival Description“ (wie Anm. 14). Grau erwähnt als fünfte Anforderung an einen Standard für archivische Findmittel noch die Vermeidung redundanter Informationen. Einen Überblick über die an einen Standard wie EAD gestellten Kriterien gibt Anne Sexton, vgl. Dies., An introduction to Encoded Archival Description, <<http://www.ucl.ac.uk/leaders-project/Papers/EAD%20talk.ppt>> (27.03.2006). Hier findet sich auch eine Übersicht über die EAD-Struktur.
- 21 Ruth, Janice A., The Development and Structure of the Encoded Archival Description (EAD) Document Type Definition, in: Pitti, Daniel V.; Duff, Wendy M. (Hgg.), Encoded Archival Description on the Internet, New York 2001, S. 27-59.

finden, ermöglicht die Präsentation durch Style Sheets eine individuelle Darstellung der Bestände abhängig von den jeweiligen Institutionen. Die Erschließungssoftware MidosaXML enthält eine eigene DTD, die den deutschen Anforderungen an eine archivische Verzeichnung entspricht, gleichzeitig in ihrer Struktur jedoch stark an die Struktur von EAD angelehnt ist. Zudem erlaubt das Programm den Export einmal erfasster Daten in andere Datenformate, so dass jederzeit eine EAD-konforme Fassung erstellt werden kann.

Da EAD ein internationaler Standard ist, gestattet die Exportfunktion nach EAD in MidosaXML auch eine Einbindung der Online-Findmittel in die Datenbank der Research Libraries Group (RLG).²² Anders als der Name zunächst erwarten lässt, handelt es sich hierbei nicht nur um einen Zusammenschluss von Bibliotheken, sondern um Institutionen aus dem gesamten Wissenschaftsbereich, die mit Hilfe des Portals „Archival Resources“ Online-Findmittel zentral zugänglich machen. Mit Hilfe dieser Suchmaschine können weltweit mehr als 60.000 Online-Findbücher mit mehr als 800.000 Einträgen von über 2.500 Institutionen in eine Recherche einbezogen und mittels einer Volltextsuche oder einer erweiterten Suche durchsucht werden.

Als Ergebnis erhalten Benutzer/innen eine Liste der Findmittel, die den gesuchten Begriff enthalten, sowie erste Angaben über die Institution, in der das Archivgut aufbewahrt wird, so dass sie bereits hier eine erste Auswahl über die für ihr Forschungsthema relevanten Bestände treffen können. In dieser Liste kann geblättert werden, und durch Anklicken gelangt man direkt in die Findmittel, die mittels eines speziellen Style Sheets präsentiert werden. Hier werden über eine Einleitung und die Klassifikation des Bestandes die nötigen Kontextinformationen bereitgestellt, um die gelieferten Suchergebnisse einordnen und im Hinblick auf die jeweilige Fragestellung auswerten zu können. Die Präsentationsform der Online-Findmittel der RLG ist dabei denen, die mit MidosaXML in der

22 Vgl. <<http://www.rlg.org>> (27.03.2006). Der Zugang zu dieser Suchmaschine ist nicht frei, sondern nur über eine Lizenz möglich. Im Lesesaal des Bundesarchiv steht die Recherchefunktion zur Verfügung.

SAPMO erstellt wurden, in Struktur und Funktionalitäten sehr ähnlich.



Abbildung 8: Ergebnisliste einer Suche über RLG



Abbildung 9: EAD-konformes Online-Findbuch der RLG

Die RLG führt die Daten verschiedener Einrichtungen zentral mit Hilfe eines Harvesting-Verfahrens zusammen. Dieses läuft so ab, dass die Archive ihre EAD-Dateien im Internet an einer Stelle zur Verfügung stellen, die sie dort ständig aktualisieren oder um neue Dateien ergänzen können. Der RLG-Harvester sammelt einmal im Monat diese Daten ein und indiziert sie. Nach der Indexierung sind die Daten aller Archive über das Archival Resource-Portal durchsuchbar.

Neben dem Harvesting, das von der RLG angewandt wird, gibt es weitere praktizierte technische Verfahren, Erschließungsinformationen online zu recherchieren. Dies sind die Suchanfrage an eine zentrale Datenbank oder die Suche mittels einer Metasuche, die eine Suche in vorhandenen Suchmaschinen anstößt. Harvesting bietet den Vorteil, Daten verschiedener Einrichtungen unkompliziert integrieren und gleichzeitig Suchergebnisse schnell präsentieren zu können. Außerdem ermöglicht es, die Ergebnisse in der Präsentationsform des jeweiligen Archivs darzustellen.²³

Anzumerken ist, dass die Integration von EAD-Dateien in das Archival Resource-Portal von der RLG kostenlos vorgenommen wird.²⁴ Archive können so ihre Bestands- und Findbuchdaten im Kontext mit denen anderer Länder und Institutionen für Recherchen zugänglich machen. Das Bundesarchiv stellt mittlerweile über 500 Online-Findbücher für die RLG im EAD-Format zur Verfügung. Benutzer/innen, die Zugang zum Archival Resource-Portal haben, können so die Bestände der SAPMO gemeinsam mit verwandten Materialien durchsuchen, wie zum Beispiel die in amerikanischen Archiven aufbewahrten Nachlässe von emigrierten deutschen Kommunisten oder die Sammlungen zur Arbeitergeschichte, die das Institut für Sozialgeschichte in Amsterdam bereit stellt.

23 Dieses Verfahren wird momentan von der RLG eingesetzt. Unter der Bezeichnung ArchiveGrid besteht ein neuer Zugang zu den archivischen Erschließungsinformationen. Der bislang kostenpflichtige Zugang ist testweise frei bis zum 30. Juni 2006 zugänglich, in: RLG, <<http://archivegrid.org/web/jsp/index.jsp>> (24.03.2006). Es gibt Überlegungen, den Zugang dauerhaft kostenlos zur Verfügung zu stellen.

24 Anmeldung unter <<http://archivegrid.org/web/jsp/contribute.jsp>> (26.03.2006).

- 
1. [SED, ZK, Büro Lamberz](#)
 2. [Red Suit Case](#)
 3. [SED, ZK Büro Norden](#)
 4. [SED, ZK Büro Mittag](#)
 5. [Demokratischer Block](#)
 6. [SED, ZK Büro Honecker](#)
 7. [SED, ZK Abteilung Sicherheit](#)
 8. [Gewerkschaft der Angestellten](#)
 9. [IG Örtliche Wirtschaft](#)
 10. [Kongreß-Verlag](#)
 11. [Nachlass Buchmann](#)
 12. [DY 30 SED Agitation](#)
 13. [DY 30 SED Bauwesen](#)
 14. [DY 30 SED Fernmeldewesen](#)
 15. [DY 30 SED Akademie für Gesellschaftswissenschaften](#)
 16. [DY 30 SED Forschung und technische Entwicklung](#)
 17. [DY 30 SED Frauen](#)
 18. [DY 30 SED Gesundheitspolitik](#)
 19. [DY 30 SED Gewerkschaften und Sozialpolitik](#)
 20. [DY 30 SED Grundstoffindustrie](#)
 21. [DY 30 SED Landwirtschaft](#)
 22. [DY 30 SED Maschinenbau und Metallurgie](#)
 23. [DY 30 SED Büro des Politbüros](#)
 24. [DY 30 SED Planung und Finanzen](#)
 25. [DY 30 SED Propaganda](#)
 26. [DY 30 SED Transport und Nachrichtenwesen](#)
 27. [DY 30 SED Wirtschaftspolitik](#)
 28. [DY 30 SED Sozialistische Wirtschaftsführung](#)
 29. [DY 30 SED Wissenschaften](#)
 30. [DY 30 SED Zentrale Parteikontrollkommission](#)

Abbildung 10: Internetseite mit Liste von EAD-kodierten Findbüchern des Bundesarchivs für den RLG-Harvester

Projekt „daofind“ – weitere Standards und zusätzliche Online-Funktionalitäten

Neben dem EAD-Standard zur Abbildung der Struktur von Findbüchern können zwei weitere Standards von Nutzen sein. Zum einen ist hier der Encoded Archival Context (EAC) zu nennen, mit dessen Hilfe Informationen über Bestandsbildner, Personen oder Organisationen, also wichtige Kontextinformationen zu Beständeübersichten und Findbüchern kodiert und international ausgetauscht werden können.²⁵ Zum anderen ist über eine Einbeziehung des Metadata Encoding and Transmission Standards (METS)²⁶ nachzudenken, der die Verwaltung digitalisierter Objekte unter-

25 Pitti, Daniel V., Creator Description. Encoded Archival Context, <http://eprints.rclis.org/archive/00000316/01/pitti_eng.pdf> (27.03.2006).

26 Informationen finden sich ebenfalls auf der Internetseite der Library of Congress,

stützt und so neben den Erschließungsinformationen auch digitalisiertes Archivgut zugänglich machen kann. METS bietet eine Struktur, die digitale Bilder einer ganzen Akte zusammenhalten kann. Einzelne Seiten werden dadurch nicht aus dem Kontext gerissen und bleiben interpretierbar. METS erlaubt es zudem, ein Archivobjekt in unterschiedlichen Formaten zu erfassen. Beispielsweise können die Seiten einer digitalisierten Akte, die als TIF-, JPEG-, PNG-Dateien und mittels automatischer Texterkennung erstellte Text-Dateien vorliegen, mit den gleichen Beschreibungsdaten versehen werden. Die Metadaten zu einer einzelnen Seite, wie zum Beispiel die Follierung oder eine Beschriftung der Images können gleichzeitig mit allen vier Formaten verknüpft werden, so dass zwischen allen Bildern eine Verbindung hergestellt ist.

Um die drei archivischen Standards zu testen und miteinander zu kombinieren, hat das Bundesarchiv im November 2005 das Pilotprojekt „daofind“ gestartet.²⁷ Das Akronym steht für Digitales Archivgut in Online-Findbüchern. Ziel ist es, EAD, EAC und METS kombiniert einzusetzen, wodurch eine neuartige Webpräsentation von Erschließungsinformationen geboten wird, über die digitalisierte Akten komplett eingesehen werden können.

Die als Testversion vorliegende Präsentation bietet zunächst einen Einstieg über die Beständeübersicht, die in der gleichen Form präsentiert wird wie die mit MidosaXML erstellten Findbücher und die mit denselben Recherchefunktionalitäten ausgestattet ist. Trifft man in der Beständeübersicht auf einen für die eigene Forschungsfrage relevanten Bestand, kann man sich eine neue Übersicht mit allen Daten zu der Herkunftsstelle anzei-

in: Library of Congress, <<http://www.loc.gov/standards/mets/>>. Eine deutsche Übersetzung der Übersicht und Anleitung zur Verwendung von METS in: Daofind, <http://www.staff.uni-marburg.de/~mennehar/mellonprojekt/uebersetzungen/mets_erl.pdf> (27.03.2006).

27 Vgl. Pressemitteilung: Andrew-W.-Mellon-Foundation fördert Internetprojekt des Bundesarchivs, in: Bundesarchiv, <<http://www.bundesarchiv.de/aktuelles/pressemitteilungen/00100/index.html>> (24.03.2006), Webseite des Projekts <<http://www.daofind.de/>> (27.03.2006) sowie zum Projekt insgesamt Löbnitz, Anke, Erschließungsinformationen und Digitalisate im Internet. Internationaler Workshop des Bundesarchivs über Archivische Standards im Rahmen des Projekts <daofind>, in: Der Archivar 59 (2006), 1, S. 96-97.

gen lassen, aus deren Tätigkeit die Unterlagen entstanden sind oder direkt zum Online-Findbuch weiterklicken. Im Online-Findbuch kann man anschließend in den Verzeichnungseinheiten suchen und falls eine Akte besonders interessant erscheint, deren Inhalt komplett einsehen. Auf diese Weise werden nicht mehr nur Einzelseiten – wie in den bisherigen Online-Findbüchern – sondern Teile bzw. komplette Akten online bereit gestellt. Um bei einer umfangreichen Akte nicht den Überblick zu verlieren, kann die Eigenschaft von METS genutzt werden, für die Aktenstruktur bedeutsame digitalisierte Schriftstücke zu kennzeichnen. Die beispielsweise als Inhaltsverzeichnis, als Bild oder als Anfangsseite eines mehrseitigen Schriftstücks definierten Digitalisate werden dazu auf einer Orientierungsoberfläche jeweils mit dem oberen Drittel der Seite dargestellt und liefern auf diese Weise einen ersten Überblick über die Akte. Von dieser Orientierungsseite kann man dann je nach Interesse an unterschiedlichen Stellen in die Akte einsteigen und Teile oder die Akte insgesamt durchblättern. Bildlich kann man sich die neue Webpräsentation ungefähr so vorstellen: Die Beständeübersicht ist ein Straßennetz, dessen einzelne Wege – die der Gliederung entsprechen – auf Plätze zulaufen. Ein einzelner Platz gehört dabei zu einer Herkunftsstelle und ist gleichzeitig ein Online-Findbuch. Die Häuser am Platz stehen für die Verzeichnungseinheiten, in denen auf den einzelnen Etagen viele Zimmer (die Digitalisate) untergebracht sind. Ein Fahrstuhl, der der Orientierungsoberfläche entspricht, erlaubt es, die Zimmer – also die Digitalisate – anzusteuern. Hat man eine Etage – also ein besonders hervorgehobenes Schriftstück – des Hauses ausgewählt, fährt man hinauf. Wenn man möchte, kann man vom Ankunftsort aus alle Zimmer der Etage begehen – also vor und nach dem hervorgehobenen Bild befindliche Bilder ansehen – und die einzelnen Gegenstände in den Räumen – also den textlichen oder bildlichen Inhalt der Aktenseite – ganz genau betrachten.

Nach diesem kleinen Ausflug in die Phantasie stellt sich die Frage nach der Technik, durch die die neue Darstellungsform ermöglicht werden soll. Im Rahmen des daofind-Projekts wurde der bisher als Prototyp vorliegende XML-Editor MEX (MidosaEditor für XML-Standards) für die benutzergeführte Erstellung und Pflege von archivischen XML-Dateien

entwickelt, der es ermöglicht ohne detaillierte Kenntnisse über DTD oder Schemata EAD-, EAC- und METS-Dateien zu erstellen. Der Editor ist ein speziell für die Open Source-Entwicklungsumgebung Eclipse entwickeltes Plug-In und muss in diese nur eingeladen werden.²⁸ Je nachdem ob man Provenienzstellen beschreiben, Beständeübersichten bzw. Findbücher erstellen oder Digitalisate erfassen möchte, können Archivare/innen zwischen verschiedenen Ansichten wählen.²⁹ Eine Besonderheit ist die Ansicht bei der Erstellung von METS-Dateien, die es erlaubt Digitalisate anzusehen. Wenn beispielsweise zu einer digitalisierten Akte ein METS-Dokument angelegt werden soll, legt man zunächst ein leeres METS-Dokument an und lädt die Bilder zu einer Akte ein. Dies kann man Bild für Bild oder automatisch mit einem ganzen Ordner von Bildern durchführen. Die Bilder können anschließend eins nach dem anderen durchblättert werden und mit Beschriftungen versehen oder bestimmten Kategorien zugewiesen werden, die sich auf die Art und Weise der Online-Darstellung auswirken. Die Bedienungsfläche des Editors ist an diejenige von MidosaXML angelehnt. Im rechten Bereich lassen sich per Doppelklick neue Elemente in das Dokument einfügen, und die werden anschließend mit Metadaten befüllt. Wie in MidosaXML kann man eine HTML-Ansicht der erstellten Datei generieren.

28 Vgl. zum Aufbau und zur Funktionsweise von Eclipse und zur Plug-In Programmierung Shavor, Sherry; Fairbrother, Scott; D'Anjou, Jim u.a., Eclipse. Anwendungen und Plug-Ins mit Java entwickeln, München 2004 und <<http://www.eclipse.org/>> (24.03.2006).

29 Screenshots des Editors unter: <<http://www.daofind.de>> (31.03.2006).

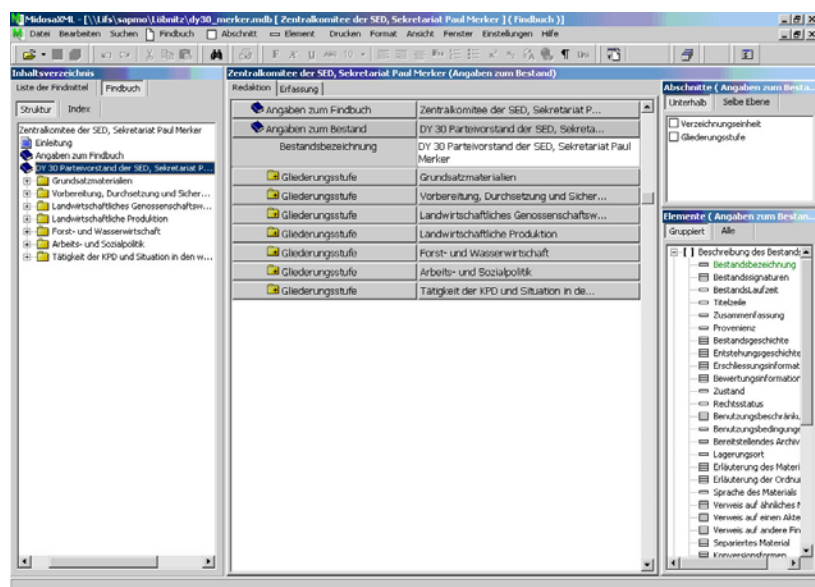


Abbildung 11: Redaktionsansicht von MidosaXML

XML-erfahrenen Anwendern/innen erlaubt MEX außerdem einen Blick in die Struktur des Dokuments. Man kann sich entweder das XML-Dokument insgesamt oder den Quelltext einzelner Elemente anzeigen lassen. Im Editor ist die Art und Weise der Verwendung der internationalen Fachstandards über so genannte Profile vordefiniert.

Anwender/innen haben die Möglichkeit, diese Profile zu ändern und zum Beispiel die Bezeichnung bestimmter Elemente oder auch die Struktur einer zu erstellenden EAD-, EAC- oder METS-Datei auf die Arbeitsweise des jeweiligen Archivs anzupassen. Die Profile sind in MEX in einer Exceltabelle definiert; dort können die Anpassungen vorgenommen werden. Dabei ist allerdings zu beachten, dass das erstellte Profil mit der DTD bzw. dem Schema des jeweiligen Standards übereinstimmt. Hierzu kann man die im Editor integrierte Validierungsfunktion nutzen, die Dateien auf Übereinstimmung einerseits mit dem definierten Profil, andererseits mit dem Schema bzw. der DTD prüft und Fehlerursachen ausgibt.

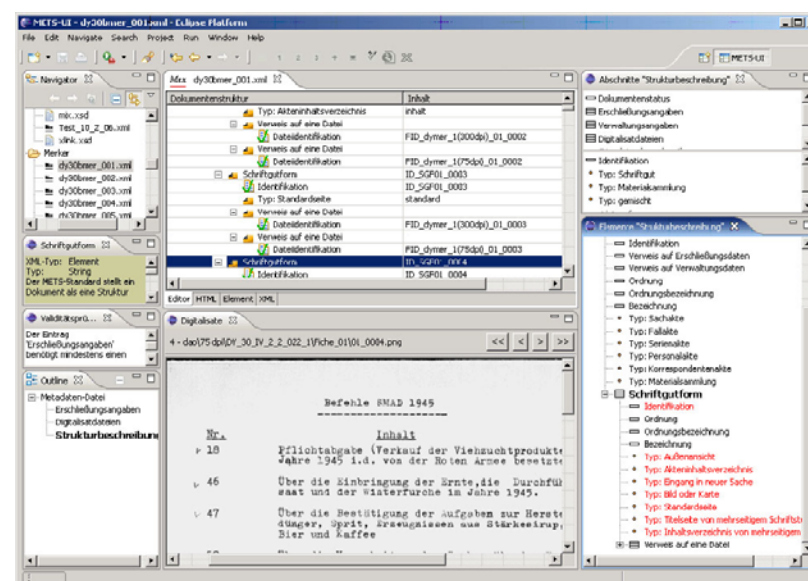


Abbildung 12: Erfassungsansicht von MEX

Ausblick

Für Forscher/innen bringt es große Vorteile, wenn sie Archivbesuche unabhängig von Orten sogar international am eigenen Schreibtisch vorbereiten und sich Informationen über Bestände und zudem digitalisiertes Archivgut im Internet beschaffen können. Sie können dazu verschiedene Recherchemöglichkeiten von Online-Findbüchern nutzen und mit Hilfe von archivischen Suchmaschinen bestände- und institutionenübergreifend suchen. Wichtig ist hier die Entwicklung von Standards, die einen Zugang zu Informationen auch über den schnellen Wechsel der Technologien hinweg gewährleisten.³⁰ Offene und plattformunabhängige Standards, wie EAD, EAC und METS, sind Voraussetzung für den integrierten Zugriff auf Erschließungsdaten, der über Portale erreicht wird. Sie bieten auch die Möglichkeit, zusätzliche Informationen, wie zum Beispiel digitalisiertes

30 Pitti, Encoded Archival Description (wie Anm. 18).

Archivgut, im Web in größerem Umfang als bisher zu präsentieren. Die Möglichkeiten des Zugangs zu Archivgut über das Internet müssen weiter ausgebaut und die Entwicklung und Anwendung internationaler Standards fortgeführt werden. Einfach zu bedienende und flexible Werkzeuge bieten dafür Archivaren/innen ein hilfreiches Mittel.

Dr. Jessica von Seggern war von 2000 bis 2002 Referendarin am Landesarchiv Schleswig-Holstein. Von 2002 bis 2003 war sie Wissenschaftliche Mitarbeiterin am Landesarchiv Schleswig-Holstein und von 2003 bis 2006 Projektleiterin in der Stiftung Archiv der Parteien und Massenorganisationen der DDR im Bundesarchiv. Seit 2006 ist sie Wissenschaftliche Mitarbeiterin bei der Edition „Akten zur Auswärtigen Politik der Bundesrepublik Deutschland“ des Instituts für Zeitgeschichte. E-Mail: jvonselggern@gmx.de

Anke Löbnitz war von Mai 2003 bis April 2005 Referendarin beim Bundesarchiv. Seit 2005 ist sie als Referentin in der Stabsstelle des Bundesarchivs in Berlin zuständig für Pressearbeit und arbeitet am Pilotprojekt „Digitalisiertes Archivgut in Online-Findbüchern“. E-Mail: a.loebnitz@barch.bund.de

STRUKTUR MEETS SCHLAGWORT – DER CLIO-ONLINE FINDMITTELKATALOG

von Sebastian Barteleit und Robert Zepf

Die Suche nach Quellenmaterial für historische Studien führt Wissenschaftler/innen in die unterschiedlichsten Archive. Die dort verwahrten Archivbestände wurden traditionell in analogen Findbüchern erschlossen, die zu einem großen Teil auch in Bibliotheken einzusehen sind. In den letzten Jahren treten zu den gedruckten Findbüchern auch verstärkt digitale Erschließungsinformationen hinzu, die direkt im Internet eingesehen werden können.

Das Bundesarchiv und die Staatsbibliothek zu Berlin werden mit einem Pilotprojekt im Rahmen des Fachportals Clio-online erstmals einen gezielten Überblick über die hybride Welt der publizierten Findmittel ermöglichen. Im Clio-online-Findmittelkatalog werden elektronische und gedruckte Bestandsübersichten, Findbücher und Inventare sowie von Verlagen lizenzierte archivische Findmittel und digitalisierte Archivbestände in einer einheitlichen Struktur entsprechend den Recherchebedürfnissen von Historikern/innen präsentiert und die Erschließungstraditionen von Archiven und Bibliotheken in innovativer Weise verbunden.

Einleitung

Einen großen Teil ihres professionellen Lebens verbringen Historiker/innen auf der Suche nach den Quellen, die ihre Thesen stützen, weiterentwickeln oder auch widerlegen können. Quellen, die sie in Bibliotheken, Museen, auf der Straße, in Ausgrabungen und natürlich auch in Archiven finden können. Archive stellen seit je her Inventare, Findbücher und

Repertorien bereit, die den Nutzern/innen den Weg zu den einzelnen Archivalien, den Urkunden und Akten, den Fotos und Plakaten und den vielen anderen Formen archivischer Überlieferung ebnet. Dies waren zunächst meist handschriftlich geführte Bücher oder Karteien, die in jüngerer Zeit mit der Schreibmaschine und dann mit dem Computer erstellt wurden. Viele dieser Inventare und Findbücher wurden und werden auch weiterhin als gedruckte Bände veröffentlicht, damit die Nutzer/innen auch außerhalb der Mauern des Archivs sich bereits über die Materialien in den Archiven informieren können. Diese publizierten Findmittel werden von den wissenschaftlichen Bibliotheken erworben, wo sie – oft in den Lesesälen – der Öffentlichkeit zur Benutzung zur Verfügung stehen, oder wohin sie schlimmstenfalls auch per Fernleihe bestellt werden können.

Mit dem Einzug der digitalen Informationstechnologie begannen die Archive seit Mitte der 1990er Jahre auch verstärkt, ihre Findmittel direkt im Internet zugänglich zu machen. Waren dies am Anfang vor allem mehr oder minder ausführliche Übersichten über die Bestände, so findet man heute auch vielerorts detaillierte Erschließungen von einzelnen Archivalien im Internet. Einige Archive stellen ganze Datenbanken im Internet zur Verfügung (zum Beispiel das Niedersächsische Landesarchiv oder die Landesarchivverwaltung Rheinland-Pfalz¹), andere wie das Bundesarchiv oder das Landesarchiv Baden-Württemberg einzelne Findbücher.² Darüber hinaus gibt es einzelne Archivbestände, deren Dokumente bereits umfassend digitalisiert worden sind und von Verlagen in Verbindung mit elektronischen Erschließungssystemen bereitgestellt werden wie zum Beispiel die Comintern Electronic Archives³ oder das Declassified Documents Reference System.⁴ Lizenzen zum Zugriff auf diese elektronischen Archive

-
- 1 Vgl. <<http://aidaonline.niedersachsen.de/>>, <<http://archivdatenbank.lha-rlp.de/>>.
 - 2 Vgl. <http://www.bundesarchiv.de/bestaende_findmittel/findmittel_online/index.html> sowie <http://www.landearchiv-bw.de/sixcms/detail.php?template=hp_artikel&id=6716&sprache=de>.
 - 3 Die Datenbank enthält ein vollständiges und frei zugängliches Verzeichnis der Bestände der Kominternarchive (55 Millionen Seiten) sowie ein laufend erweitertes Volltextarchiv der am häufigsten benutzten Dokumente als Grafikdateien; vgl. <<http://www.comintern-online.com>>.
 - 4 Das Declassified Documents Reference System (DDRS) bietet Zugriff auf

werden von einzelnen Bibliotheken erworben bzw. mit Unterstützung der Deutschen Forschungsgemeinschaft (DFG) als Nationallizenz erworben.⁵

In den günstigsten Fällen können Historiker/innen sich somit bereits vom Schreibtisch ihres Büros oder von zu Hause aus einen ersten Überblick über das in einzelnen Archiven zu findende Archivgut verschaffen oder direkt auf digitalisierte Archivalien zugreifen. Dabei sind sicherlich die zusätzlichen Suchmöglichkeiten, die Online-Findmittel und digitalisierte Archivbestände bieten, von großem Nutzen.⁶ Dennoch bleibt festzuhalten, dass auf diesem Weg nur ein Teil der möglichen Findmittel einzusehen ist. Solange nicht in großem Umfang alte Findmittel retrodigitalisiert werden und somit auch als Online-Findmittel angeboten werden können, bieten die in den Bibliotheken vorhandenen „analoge“ Findmittel einen zweiten, komplementären Sucheinstieg. Zudem bieten gedruckte Findmittel einen höheren Lesekomfort als das Lesen am Bildschirm – auch bei den Findmitteln werden voraussichtlich elektronische und gedruckte Form auf absehbare Zeit nebeneinander fortbestehen. Grund genug also, nach wie vor bei der Vorbereitung eines Archivaufenthalts den Besuch in einer Bibliothek einzuplanen.

Allerdings ist es bislang noch schwierig, sich einen Überblick über vorhandene, öffentlich zugängliche Findmittel zu verschaffen. Für die Recherche nach online verfügbaren Findmitteln muss man zumeist auf die Internetseiten der betreffenden Archive gehen, denn obwohl es sich um elektronische Publikationen handelt, die von Bibliotheken durchaus erschlossen werden könnten, sind die elektronischen Findmittel im Regelfall nicht in elektronischen Bibliothekskatalogen verzeichnet.

Aber auch die Suche nach gedruckten Findbüchern in den lokalen OPACs und übergreifenden Verbundkatalogen ist nicht ohne Tücken. Aufgrund der Erschließungskonventionen der Bibliotheken sind nämlich

-
- Dokumente, die von den US-Regierungsstellen freigegeben wurden. Die Datenbank enthält über 85.000 Dokumente mit einem Gesamtumfang von mehr als 525.000 Seiten. Vgl. <<http://infotrac.galegroup.com/itweb?=DDRS>>.
- 5 Vgl. <<http://www.nationallizenzen.de/>>.
 - 6 Siehe dazu auch den Beitrag von Anke Löbnitz und Jessica von Seggern in diesem Band.

für Historiker/innen entscheidende Fragen wie „Welche gedruckten Findmittel stellt ein Archiv zur Verfügung, um ortsunabhängig über seine Bestände zu informieren?“ und „Welches gedruckte Findbuch beschreibt Bestände, die für meine Fragestellung relevant sind?“ nicht immer ohne weiteres zu beantworten.

Angesichts des Aufwandes, den die Bibliotheken und ihre Mitarbeiter/innen in die Erschließung der von ihnen erworbenen Literatur investieren, ist dies eine eher deprimierende Feststellung. Als Gründe sind zum einen die wiederholten Änderungen von Regelwerken und Katalogisierungsgrundsätzen sowohl bei der formalen wie bei der inhaltlichen Erschließung zu nennen, zum anderen die heterogene Datenbasis der meisten Verbundkataloge, die – bei aller Bemühung um eine einheitliche Datenstruktur – eine Mischung von neuen Titelaufnahmen, retrokonvertierten Altdaten und Fremddaten unterschiedlicher Provenienz und Qualität sind. Im Falle der gedruckten Findmittel der Archive kommen jedoch häufig noch weitere, für das Material spezifische Problemfaktoren hinzu: Zum einen die Tatsache, dass viele Findmittel außerhalb des Verlagsbuchhandels erschienen sind und damit aufgrund der Erschließungspraxis Der Deutschen Bibliothek bislang von einer bundesweit einheitlichen Sacherschließung ausgenommen waren, zum anderen der häufig recht komplexe inhaltliche und bibliografische Aufbau von gedruckten Findbüchern und Findbuchreihen – Reihengliederungen mit mehreren Unterreihen, springende Bandzählungen, mehrbändige Werke innerhalb von Reihen sowie die Praxis, dass viele Findbücher im bibliografischen Sinne nicht vom besitzenden Archiv selbst, sondern von einer übergeordneten Institution (zum Beispiel einer Landesarchivdirektion) herausgegeben werden, sind keine Seltenheit und führen in Bibliothekskatalogen zu sehr komplexen Titelaufnahmen, deren Darstellung zumindest in der gegenwärtigen Generation von OPACs mitunter recht unübersichtlich gerät.

Die Ursache dafür, dass es Bibliothekskatalogen nur schwer gelingt, ihre Benutzer/innen zum richtigen archivalischen Findmittel zu führen, liegt jedoch tiefer: Bibliothekskataloge und bibliothekarische Regelwerke wurden ursprünglich entwickelt, um physische Informationsträger mit Metadaten formal und inhaltlich zu beschreiben und zu verwalten – Find-

mittel sind hingegen selbst Sammlungen von Metadaten; die Historikerin bzw. der Historiker interessiert sich in den seltensten Fällen für das Findbuch als Medium, sondern für den Bestand, die Ansammlung von Quellen, die es beschreibt.

Eine weitere, oft viel entscheidendere Hürde auf dem Weg zum richtigen Findmittel ist jedoch die vielfältige Archivlandschaft, die die Nutzer/innen zunächst durchschauen müssen, um zu dem richtigen Archiv und letztlich auch zu den richtigen Beständen zu finden. Denn vor der Arbeit in den Beständen heißt es für die Forscher/innen zunächst, das richtige Archiv zu finden; erst dann können Beständeübersichten und Findbücher gewälzt und anschließend an und mit den Quellen gearbeitet werden.

Ein Instrument, das die Orientierung in dieser Informationslandschaft ermöglichen würde, gibt es bisher nicht – weder für gedruckte noch für Online-Findmittel. Ansätze für solch eine Metasicht gibt es allenfalls auf „regionaler“ Ebene, wie etwa im Internetangebot des Geheimen Staatsarchivs Preußischer Kulturbesitz, das für die preußischen Behörden den Nachweis führt, welche Bestände im eigenen Haus zu finden sind und in welchen anderen Archiven weitere Unterlagen zu erwarten sind.⁷ Ähnliches bieten viele regionale Archivportale, zum Beispiel Nordrhein-Westfalen oder Hessen, die neben einer Übersicht der im jeweiligen Bundesland ansässigen Archive auch Beständeübersichten und Findbücher zur Konsultation bereitstellen.⁸ Für Historiker/innen, die Fragestellungen quer zu aktuell existierenden Ländergrenzen bearbeiten, bieten diese Portale dementsprechend allerdings nur einen eingeschränkten Nutzen.

Der Findmittelkatalog

Die Auffindbarkeit von Quellen zu erleichtern, ist eines der Hauptanliegen des Portals Clio-online – eine Umfrage ergab, dass dies eine der vier wichtigsten Funktionen ist, die das Portal aus der Sicht seiner Nutzer/innen

7 Vgl. <<http://www.gsta.spk-berlin.de/framesets/frameset.php>>.

8 Vgl. <<http://www.archive.nrw.de/>>, <<http://www.archive.hessen.de/>>. Fast alle Bundesländer bieten inzwischen Archivportale an, die Art und Tiefe der gebotenen Informationen sind aber sehr unterschiedlich.

erfüllen sollte.⁹ Das Bundesarchiv und die Staatsbibliothek zu Berlin – Preußischer Kulturbesitz haben sich daher zusammengetan, um als wichtige Clio-Partnerinstitutionen aus den Bereichen Archiv und Bibliothek gemeinsam eine Verbesserung der oben geschilderten Benutzungssituation herbeizuführen. In dem Antrag bei der DFG für die Weiterfinanzierung des Projektes Clio-online wurden daher Mittel für den Aufbau eines elektronischen Findmittelkatalogs beantragt.

Ziel dieses Findmittelkatalogs ist es, eine Plattform innerhalb des Clio-online-Portals zu schaffen, auf der Informationen über gedruckte und Online-Findmittel zusammengeführt und in einer für Benutzer/innen leicht zugänglichen Form vermittelt werden. Das Bundesarchiv vertritt dabei im Rahmen von Clio-online den archivischen Teil des Projektes, die Staatsbibliothek zu Berlin bringt ihre bibliothekarische Kompetenz und Erfahrung bei der Erschließung und Vermittlung gedruckter Findmittel sowie kommerziell lizenzierter elektronischer Archivressourcen ein. Als eine der nationalen Archivbibliotheken für die Sammlung amtlicher Publikationen und als größte wissenschaftliche Universalbibliothek Deutschlands verfügt sie – neben der Bayerischen Staatsbibliothek in München – über die größte Sammlung von gedruckten Findmitteln im deutschsprachigen Raum und wird ihre bibliografischen Daten in das Projekt einbringen.

Eingabe und Pflege bzw. Einspielung der Daten erfolgt in einer vom Bundesarchiv gewarteten Datenbank, die später für Zwecke der Recherche und Präsentation zum Clio-online-Portal gespiegelt wird. In einer weiteren Ausbaustufe ist geplant, die Pflege der Daten über Online-Findmittel und die Mitteilung neu erschienener gedruckter Findbücher allen Archiven selbst in die Hand zu geben. Dazu soll nach dem Vorbild der „Zentralen Datenbank Nachlässe“¹⁰ das Pflegemodul so erweitert werden, dass über gesonderte Zugangskennungen auch von außerhalb des Bundesarchivs und der Staatsbibliothek in die Datenbank eingepflegt werden kann.

9 Vgl. zu der Clio-online Nutzerstudie den Beitrag von Karsten Borgmann in diesem Band.

10 Vgl. <<http://www.nachlassdatenbank.de/>>.

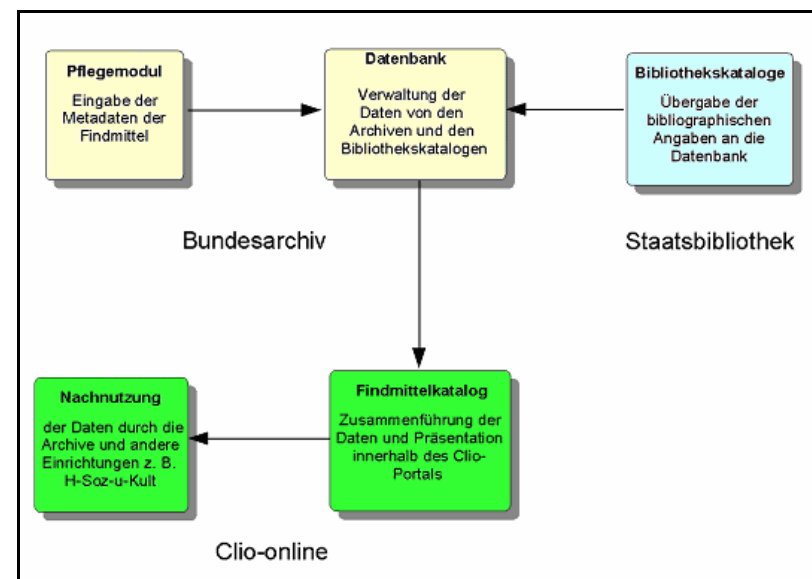


Abbildung 1: Skizze des geplanten Workflows

Metadatenschema

Entsprechend der Zielsetzung des Clio-Portals insgesamt geht es beim Findmittelkatalog nicht um eine umfassende und detaillierte Dokumentation aller Findmittel im Sinne einer Gesamtkatalogs oder einer Bibliografie, sondern vielmehr um die Schaffung eines auf die Recherchebedürfnisse von Historikern/innen abgestimmten Leitsystems, das an der Schnittstelle von Archiven und Bibliotheken den Weg zu den relevanten Ressourcen in elektronischer wie gedruckter Form weist. Grundlage der Datenbank ist ein von beiden Einrichtungen gemeinsam erarbeitetes Metadatenschema, das im Kern die für Historiker/innen relevanten Informationen beinhaltet. Es ist mit archivischen und bibliothekarischen Erschließungssystemen kompatibel, im Vergleich zu bibliothekarischen Katalogisierungsregelwerken jedoch relativ schlank. Im Interesse des Recherchekomforts, aber auch um den Erstellungs- und Pflegeaufwand zu reduzieren, wird zum Beispiel auf eine differenzierte Darstellung bibliografischer Hierarchien konsequent

verzichtet. Mit Ausnahme des Seitenumfangs enthält der Findmittelkatalog auch keine detaillierten Angaben zur physischen Form der Findbücher. Er weist auch nicht nach, wo Exemplare der gedruckten Findbücher verfügbar sind – über die ISBN und andere eindeutige Identifikatoren wird stattdessen eine Verlinkung in Bibliotheks-OPACs und Verbundkataloge angeboten, in denen die vollständigen bibliografischen Angaben verfügbar sind.

Im Vergleich zu Bibliothekskatalogen ergibt sich der Mehrwert des Findmittelkatalogs daher nicht aus der Vollständigkeit der bibliografischen Beschreibung, sondern aus der Hervorhebung und gezielten Suchbarkeit bestimmter Merkmale, die in den Katalogen nicht ohne Mühe ermittelt werden können – wie etwa das Archiv, in dem sich der Bestand befindet, sowie die Bestandssignatur – und die Einordnung in eine klassifikatorische Struktur, die den Forschungsbedürfnissen von Historikern/innen und der Organisation des deutschen Archivwesens besser entspricht als die gängigen bibliothekarischen Erschließungssysteme.

Bei der Erstellung des Metadatensets wurde Wert darauf gelegt, dass alle Datenfelder kompatibel zu dem internationalen Strukturstandard für Findbücher EAD (Encoded Archival Description) sind. Dadurch ergibt sich mittelfristig die Möglichkeit, die hier gesammelten Daten leicht in andere Kontexte einzubinden. EAD wird inzwischen vom Bundesarchiv als Austauschformat favorisiert und in verschiedenen nationalen (zum Beispiel BAM) und internationalen Projekten (zum Beispiel RLG) eingesetzt.¹¹

Historische Fragestellungen

Neben der reinen Bündelung von Informationen über Findmittel soll der Findmittelkatalog aber auch zu einer Orientierung der Historiker/innen in der vielfältigen deutschen Archivlandschaft dienen. Historische Fragestellungen bewegen sich stets in einem Rahmen, der durch die behandelte Zeitspanne, die Länder, Regionen oder Orte und durch ein konkretes inhaltliches Thema begrenzt wird. Für die Suche nach archivalischen

Quellen kommt deshalb diesen drei Kategorien besondere Bedeutung zu.

Archivische Erschließung

Die archivische Erschließung trägt diesen Kategorien zwar durchaus Rechnung, allerdings werden die Kriterien nur sehr bedingt als Sucheinstiege verwendet.

Das archivische „Sprengelprinzip“, das besagt, dass ein Archiv für eine bestimmte regional und verwaltungsrechtlich abgrenzbare Entität (Staat, Stadt, Landeskirche, Universität usw.) zuständig ist, bildet das Kriterium der Region zwar ab, macht aber im Einzelfall nicht deutlich, welche anderen Archive gegebenenfalls für Fragen, die die gleiche Region betreffen, noch relevant sein könnten. Dies umso mehr, als der regionale Sprengelzuschnitt der verschiedenen Archivtypen sehr unterschiedlich sein kann. So ist der Sprengel eines Staatsarchives zumeist ein Bundesland (oder ein Teil davon) samt der Vorgängerstaaten und somit nicht deckungsgleich zu einem Kirchenarchiv, das als Sprengel eine Diözese oder Landeskirche hat.

Die innerarchivische Ordnung der Bestände, die Tektonik, bildet hingegen zumeist die zeitliche Differenzierung sehr gut ab. Meist werden große historische Zäsuren als Maß genommen, um die Bestände zu gruppieren. Allein der Inhalt der Bestände spielt bei der archivischen Erschließung auf der Ebene der Bestände eine nachgeordnete Rolle. Bestände werden nicht nach inhaltlichen Gesichtspunkten gebildet, sondern nach der Art ihrer Herkunft (Provenienz). Den Nutzern/innen wird also ein Sucheinstieg über die Art der Institution gegeben, die Urheber der archivischen Überlieferung ist. Inhalt spiegelt sich hier also vor allem darin wieder, dass jede Institution einen bestimmten Zweck erfüllt, bestimmte Kompetenzen und Aufgaben hat. In der zum Teil sehr ausführlichen Beschreibung des Bestandes kommt dann hingegen auch der inhaltliche Gesichtspunkt zum Tragen.

Jedes Archiv versucht seinen Nutzern/innen also durchaus mit Hilfe dieser zentralen Kriterien die Suche nach relevanten Quellen zu ermöglichen. Als Problem bleibt bestehen, dass dieser Suchweg in weiten Teilen nur über das jeweilige Archiv ermöglicht wird. Die verschiedenen Archivportale der Länder geben zwar zum Teil bereits einen recht guten Über-

¹¹ Zu BAM siehe: <<http://www.bam-portal.de/>> und den Beitrag von Frank von Hagel in diesem Band, zur RLG siehe: <<http://www.rlg.org/>>.

blick über Archive und ihre Bestände, eine Perspektive, die ganz Deutschland einschließlich der historischen Grenzregionen in den Blick nimmt, fehlt bislang noch. Der Findmittelkatalog setzt an dieser Stelle an und versucht über differenzierte Sucheinstiege auch archivunabhängige Suchen zu ermöglichen.

Bibliothekarische Erschließung

Auch die bibliothekarische Erschließung von Findmitteln trägt seit jeher den Facetten Zeit, geografischer Bezug und sachlicher Bezug der Dokumente Rechnung – in einer Form allerdings, die nicht auf die Ordnung von Archivbeständen zugeschnitten ist, sondern auf die Erschließung inhaltlich heterogener Ressourcen aller Fächer, Länder und Zeiten.

Für Historiker/innen, die an Büchern und Quellen aus allen Zeiten interessiert sind, ist insbesondere die Diskontinuität der bibliothekarischen Sacherschließung von Nachteil. Fast alle deutsche Bibliotheken haben im 20. Jahrhundert einmal oder mehrmals ihr Sacherschließungssystem gewechselt – mit der Folge, dass ältere Bestände inhaltlich nur schwer recherchierbar sind, zumal die historischen Sachkataloge nur vereinzelt in den elektronischen Katalogen suchbar gemacht worden sind. Erst seit Mitte der 1980er Jahre setzt sich die Erschließung mit Schlagworten nach RSWK und der Schlagwortnormdatei durch – sie ermöglicht einen punktuellen Zugriff mit normierten Suchbegriffen, durch Einsatz der Schlagwortnormdatei auch mit Synonyma. Maßgeblich ist jedoch der Inhalt des gesamten Dokuments – im Falle eines Findbuchs wird daher versucht, den gesamten Inhalt in einer Schlagwortkette zu erfassen, die in der Regel geografische, sachliche, formale und zeitliche Komponenten enthält. Der Differenzierungsgrad einer ausführlichen Beschreibung des Bestands, wie sie die Archive vornehmen, kann dabei nicht erreicht werden.

Eine einheitliche klassifikatorische Komponente, die eine Orientierung über größere Zusammenhänge ermöglichen würde, fehlt jedoch bislang. Im deutschsprachigen Raum am weitesten verbreitet sind die Regensburger Verbundklassifikation¹² mit Schwerpunkt in Süd- und

12 Vgl. <<http://www.bibliothek.uni-regensburg.de/Systematik/systemat.html>>.

Ostdeutschland sowie die sehr grobe und für historische Fragestellungen zu wenig differenzierte Basisklassifikation des Gemeinsamen Bibliotheksverbundes¹³ mit Schwerpunkt in Norddeutschland. Eine neue Entwicklung ist die aktive Anwendung der aus den USA stammenden, aber seit dem Zweiten Weltkrieg zunehmend international verbreiteten Dewey Dezimal Klassifikation durch Die Deutsche Bibliothek¹⁴ seit Beginn dieses Jahres. Das geografische Strukturprinzip der DDC ist für mitteleuropäische Historiker/innen ungewohnt, denn es ist konsequent gegenwartsorientiert – zugespitzt kann man sagen, dass alle Themen und Ereignisse seit der Völkerwanderungszeit auf die aktuellen administrativen Strukturen abgebildet werden.

Suche im Findmittelkatalog

Kerngedanke der Suche im Findmittelkatalog ist es, den Nutzern/innen möglichst verschiedene Einstiege in die Recherche zu bieten, und alle verschiedenen Suchmöglichkeiten untereinander zu verbinden. Konkret bedeutet dies, dass nach jeder Suche die Möglichkeit besteht, über eines der anderen Kriterien die Suche weiter einzuschränken. Die Suchkriterien im Einzelnen sind:

Historische Geografie

Historische Fragestellungen implizieren stets eine regionale Komponente.

-
- 13 Zur Konzeption der Kooperativen Sacherschließung im GBV vgl. die Sacherschließungsrichtlinie unter: <<http://www.gbv.de/du/sacher/inhalt.shtml>>.
- 14 Die Deutsche Bibliothek ist federführendes Mitglied im 2000 gegründeten Konsortium DDC Deutsch, in dessen Auftrag die 22. Auflage der DDC ins Deutsche übersetzt wurde. Die Übersetzung wurde in den Jahren 2002 bis 2005 als gemeinsames Projekt Der Deutschen Bibliothek und der Fachhochschule Köln erstellt und erscheint Ende 2005 als Druckausgabe bei K. G. Saur. Gleichzeitig startet Die Deutsche Bibliothek den Online-Dienst Melvil, der die DDC Deutsch zum Klassifizieren zur Verfügung stellt und als Retrievaltool die Recherche nach DDC-erschlossenen Dokumenten ermöglicht. Seit 2004 sind alle Reihen der Deutschen Nationalbibliografie nach Sachgruppen, die auf der DDC beruhen, gegliedert. Ab Bibliografiejahrgang 2006 werden die Reihen B und H der Deutschen Nationalbibliografie mit DDC-Notationen erschlossen, ab 2007 ist dies auch als Ergänzung zur verbalen Sacherschließung mit RSWK/SWD für die Reihe A geplant. Vgl. <<http://www.ddc-deutsch.de/>>.

Im Findmittelkatalog wird deshalb ein Sucheinstieg über historische Territorien ermöglicht. Denn Geografie ist in historischer Hinsicht kein fest gefügtes System, sondern wandelt sich im Laufe der Zeit. Anstelle von Längen- oder Breitengraden und meteregenauen GPS-Systemen werden sich Historiker/innen eher an alten Länder- und Territorialgrenzen orientieren. Als Sucheinstieg ist deshalb die Geografie vor allem in der zeitlich gebundenen Ausprägung von Gemeinschaften zu verstehen.

Der Findmittelkatalog orientiert sich deshalb zunächst zeitlich an den Zäsuren der Geschichte, die sich in größeren Umstrukturierungen der Staatenverteilung in Deutschland manifestierten.

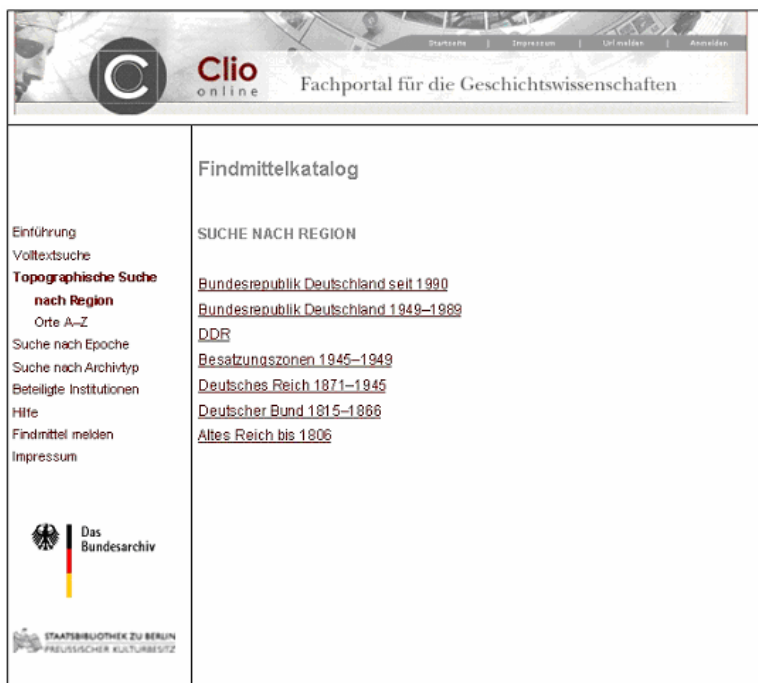


Abbildung 2: Entwurf für die Rechercheoberfläche des Findmittelkatalogs – Topografie

Auf der darunter folgenden Ebene wird es anschließend die Möglichkeit geben, die historischen Territorien dieser Zeit auszuwählen und eine Auswahl von Findbüchern zu erhalten, die für diese Region relevant sind. In einem ersten Schritt wird dies über Auswahllisten funktionieren, in einer weiteren Ausbaustufe soll zusätzlich mittels einer interaktiven Landkarte das entsprechende Gebiet ausgewählt werden können. Zusätzlich wird es die Möglichkeit geben, nach einzelnen Städten und Ortschaften zu suchen.



Abbildung 3: Entwurf für die Rechercheoberfläche des Findmittelkatalogs – Orte

Epoche

Neben der regionalen Einordnung einer historischen Frage steht natürlich

die zeitliche Einordnung. Wie gesehen wird die Suche nach der geografischen Verortung eines Themas auch zentral über zeitliche Zäsuren geleitet, da diese sich an den größeren Umbrüchen im Zuschnitt der Grenzen von Territorien und Staaten niederschlagen, sind sie jedoch möglichst weit gefasst. In diesem Sucheinstieg werden gängige Zeiteinteilungen Deutschlands bzw. Mitteleuropas verwendet (siehe Abbildung 4), um eine etwas kleinteiligere Suche zu ermöglichen.

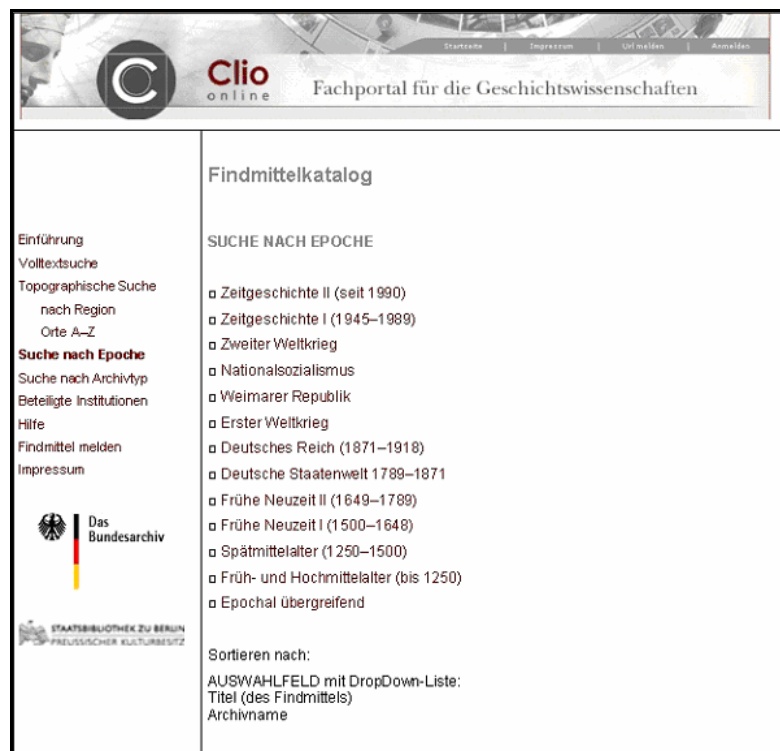


Abbildung 4: Entwurf für die Rechercheoberfläche des Findmittelkatalogs-Epochen

Archivtyp

Eine Aufschlüsselung nach Archivtypen ermöglicht einen weiteren Sucheinstieg. Dabei wird von der Tatsache ausgegangen, dass in unterschiedlichen Archiven unterschiedliche Archivalien zu finden sind. Die Archivtypen im Einzelnen sind:

- Staats- und Landesarchive
- Kreisarchive
- Stadt- und Gemeindearchive
- Kirchliche Archive
- Wirtschaftsarchive
- Unternehmensarchive
- Adelsarchive
- Parlamentsarchive
- Archive politischer Parteien und Verbände
- Medienarchive
- Hochschularchive
- Archive wissenschaftlicher Einrichtungen
- Literaturarchive
- Sonstige Archive

Diese Form der Suche ermöglicht es den Nutzern/innen, nach allen Findmitteln zum Beispiel aus Kirchen- oder Universitätsarchiven zu suchen.

Provenienztyp

Viele Archive verwahren Bestände sehr unterschiedlicher Provenienz – zum Teil auch Bestände aus Provenienzen, die man nach dem reinen Sprengelprinzip in einem Archiv eines anderen Archivtyps erwarten würde. Das Bundesarchiv verwahrt beispielsweise auch Bestände von Parteien und politischen Gruppen (zum Beispiel der Deutschnationalen Volkspartei oder der Parteien und Massenorganisationen der DDR), von Banken und Wirtschaftsunternehmen (zum Beispiel der Bank der deutschen Arbeit oder des Flick-Konzerns) oder Nachlässe von Privatpersonen. Um auch diese Bestände auffindbar zu machen, wird ergänzend zum Archivtyp auch nach einer ähnlichen Systematik, die im Detail noch auszuarbeiten ist, als weiterer Sucheinstieg für die einzelnen Findmittel

auch der Provenienztyp angeboten.

Inhaltliche Suche

Für Territorien, Epochen und Archive bzw. Archiv- und Provenienztypen wird der Findmittelkatalog normierte Sucheinstiege bieten, die beim Aufbau des Grunddatenbestands einheitlich vergeben werden. Die thematische Erschließung der Findmittel wird jedoch die unterschiedliche Beschreibungstradition von Archiven und Bibliotheken widerspiegeln – so können die Bibliotheken aufgrund gedruckter Findmittel keine detaillierte Beschreibung der Bestände liefern, andererseits erfordert eine bibliothekarische Sacherschließung mit normierten Schlagwörtern eine Vertrautheit mit bibliothekarischen Regelwerken, die von Archiven nicht erwartet werden kann. Zu Anfang werden daher die gedruckten Findmittel mit normierten Schlagworten auf der Grundlage der SWD versehen, wie sie etwa die Jahresberichte für Deutsche Geschichte oder der WebGuide von Clio-online verwenden – die von den Archiven gemeldeten Ressourcen werden hingegen mit ausführlichen Beschreibungen mit Freitext versehen. Im Laufe der Zeit ist eine Konvergenz der Suchmöglichkeiten zu erwarten, da die teilnehmenden Archive die Möglichkeit erhalten, auch die Datensätze ihrer gedruckten Findmittel um entsprechende Beschreibungen zu ergänzen.

Mehrwert für die Nutzer/innen

Der Clio-online-Findmittelkatalog wird als Modul des Fachportals erstmals einen gezielten Überblick über die hybride Welt der publizierten Findmittel ermöglichen. Elektronische und gedruckte Bestandsübersichten, Findbücher und Inventare sowie von Verlagen lizenzierte archivische Findmittel und digitalisierte Archivbestände werden in einer einheitlichen Struktur entsprechend den Recherchebedürfnissen von Historikern/innen präsentiert – Forschende können sich gezielt darüber informieren, welche Hilfsmittel Archive zur ortsunabhängigen Vorbereitung von Archivaufenthalten bereitstellen, durch die Gesamtschau wird zudem die Ermittlung von Findmitteln und Beständen zu einzelnen historischer Territorien, einzelnen Epochen oder zu bestimmten Themen der Forschung möglich.

Ein erster Schwerpunkt beim Aufbau des Findmittelkatalogs sind neu publizierte Findmittel, die unabhängig von ihrer Erscheinungsform durch Zusammenarbeit der Projektträger mit den deutschen Archiven möglichst schnell nachgewiesen und der wissenschaftlichen Öffentlichkeit bekannt gemacht werden sollen. Von besonderem Vorteil ist dabei die Einbindung des Findmittelkatalogs in die Kommunikationskanäle von Clio-online und H-Soz-u-Kult, die in kurzer Zeit einen großen Teil der relevanten wissenschaftlichen community erreichen – neu erschlossene Bestände von Archiven bzw. neu erworbene archivische Ressourcen von Bibliotheken können so bereits nach kurzer Zeit von der Forschung rezipiert werden.

Sofern die Mittel dafür von der DFG bewilligt werden, wird bis Ende 2007 darüber hinaus ein umfassender Grundbestand der wichtigsten bereits publizierten elektronischen und gedruckten Findmittel der vergangenen Jahrzehnte aufgebaut – der Arbeitsschwerpunkt wird dabei zunächst bei den Archiven auf Bundes- und Landesebene liegen, in einem zweiten Schritt werden sukzessive auch die Findmittel anderer Archivtypen in das Nachweissystem einbezogen. Clio-online strebt auch dabei eine enge Zusammenarbeit mit den Archiven an, die zur Mitarbeit eingeladen werden und die Möglichkeit erhalten, über einen passwortgeschützten Zugang zu dem beim Bundesarchiv gehosteten Eingabe- und Pflegemodul ihre Online-Findmittel nachzuweisen sowie bibliografische Informationen und zusätzliche Hinweise zu ihren gedruckten Findmitteln einzubringen. Es wird damit möglich, den Wert älterer gedruckter Findbücher durch aktuelle Hinweise zu verbessern, zum Beispiel durch die Information, dass ein Bestand seit Erscheinen des Findbuchs an einen anderen Archivstandort überführt wurde.

Über den Nachweis von Ressourcen hinaus bietet der Findmittelkatalog den Einstieg in weitere Dienstleistungen von Archiven und Bibliotheken – sei es die Möglichkeit, Kontakt zu dem bewahrenden Archiv aufzunehmen, über das eigene Bibliothekssystem auf eine elektronische Ressource zuzugreifen oder darüber informiert zu werden, dass ein gesuchtes gedrucktes Findmittel in der lokalen Bibliothek vorhanden ist oder per Fernleihe bestellt werden kann. Der Findmittelkatalog leistet damit als Modul des größeren Netzwerks Clio-online einen innovativen Beitrag zur

engeren Vernetzung der historischen Forschung mit Archiven und Bibliotheken.

Dr. Sebastian Barteleit, Archivrat, ist Referatsleiter für Bestandserhaltung im Bundesarchiv Berlin. E-Mail: s.barteleit@barch.bund.de

Robert Zepf, M.St. (Oxon.), ist Fachreferent für Geschichte und Kommissarischer Leiter der Wissenschaftlichen Dienste in der Staatsbibliothek zu Berlin – Preußischer Kulturbesitz. E-Mail: robert.zepf@sbb.spk-berlin.de

MANUSCRIPTA MEDIAEVALIA IM DIGITALEN VERBUND

von Robert Giel

Manuscripta mediaevalia ist das zentrale deutsche Informationsangebot für den Nachweis mittelalterlicher Handschriften. Zu insgesamt etwa 60.000 Codices stehen recherchierbare Daten und/oder digitale Katalogimages zur Verfügung. Das Angebot wird kontinuierlich ausgebaut. Zukünftig werden sich die Träger (Bayerische Staatsbibliothek München, Bildarchiv Foto Marburg und Staatsbibliothek zu Berlin) außerdem verstärkt darum bemühen, Manuscripta mediaevalia an weitere fachlich verwandte Angebote anzuknüpfen. Angestrebt wird sowohl eine eindeutige Referenzierungsmöglichkeit für die Handschriftenbeschreibungen in Manuscripta mediaevalia als auch deren Anbindung an ergänzende Informationsquellen im Web. Dabei werden auch weiterhin die Erfordernisse der Handschriftenforschung zentrale Aufmerksamkeit beanspruchen.

Informationsgehalt und Focus

Um eine Ortsbestimmung von *Manuscripta mediaevalia* vornehmen zu können und das Webangebot in sein digitales Umfeld einzuordnen¹, muss eine knappe Zusammenfassung dessen vorangestellt werden, was *Manuscripta* ausmacht. In seinem Beitrag zum Bibliothekartag 2005 über „Angebot und Nutzen von digitalisierten Altbestandsquellen im Internet“ formulierte Thomas Stäcker: „Das zentrale deutsche Portal DFG-geförderter Projekte für den Nachweis mittelalterlicher Handschriften ist

1 <<http://www.manuscripta-mediaevalia.de>> (27.03.2006). Um unvermeidlicher begrifflicher Ermüdung entgegenzuwirken, belasse ich es im Folgenden bei „Manuscripta“.

Manuscripta mediaevalia.“² Damit ist die Aufgabe knapp umrissen. Der Gegenstand von *Manuscripta* sind Handschriften und ihre Erschließung.

Eine aktuelle Zählung der mittelalterlichen Handschriften in Deutschland kommt auf eine Summe von 60.000 Codices. Im entsprechenden Förderprogramm der Deutschen Forschungsgemeinschaft (DFG) werden die einzelnen Bestände von Bibliotheken und Archiven seit etwa 1960 nach einheitlichen Kriterien sukzessive katalogisiert und publiziert. Als Ergebnis verfügen wir mittlerweile über beinahe 200 gedruckte Kataloge mit Beschreibungen von etwa 24.000 Handschriften und Handschriftenfragmenten, die hinsichtlich ihrer Erschließungstiefe auch im europäischen Rahmen Vorbildcharakter beanspruchen dürfen. In erster Linie für diese Katalogisierungstradition stellt *Manuscripta* das Forum dar und aus dieser Tradition heraus bildet es den chronologisch ersten Teil des digitalen Nachweissystems der DFG für den Altbestand in deutschen Bibliotheken.

Darüber hinaus betreibt *Manuscripta* die Retrokonversion wichtiger Handschriftenkataloge und –inventare, die bereits vor dem Einsetzen der Fördermaßnahmen der DFG entstanden sind, und häufig den einzigen Zugang zu den betreffenden Beständen bieten. Hinzu kommt weiter eine Auswahl moderner Kataloge des Auslands, die für die Handschriftenforschung in Deutschland von besonderem Wert sind. Neben den von der Österreichischen Akademie der Wissenschaften erarbeiteten Katalogen sind dies besonders diejenigen der Österreichischen Nationalbibliothek (ÖNB) Wien, der Universitätsbibliotheken von Basel und Uppsala sowie der Biblioteka Jagiellonska in Krakau. Ergänzt wird das Material durch Angaben aus maschinen- und handschriftlichen Aufzeichnungen, etwa aus dem Handschriftenarchiv der Berlin-Brandenburgischen Akademie der Wissenschaften (BBAW) oder aus bibliotheksinternen Auskunftsmitteln. Last, but not least präsentiert *Manuscripta* die noch unpublizierten Beschreibungen aus laufenden Katalogisierungsvorhaben.

Insgesamt vermittelt *Manuscripta* damit den Überblick über eine etwa

2 Stäcker, Thomas, Das ist doch alles im Netz! – Angebot und Nutzen von digitalisierten Altbestandsquellen im Internet (Vortrag Düsseldorf, Bibliothekartag, 15.03.2005), <http://www.opus-bayern.de/bib-info/volltexte/2005/73/pdf/staecker_duedo-2005.pdf> (27.03.2006).

hundertjährige Katalogisierungstradition. Dementsprechend bietet es den Nutzern/innen keinen homogenen Datenbestand, sondern Handschriftenbeschreibungen in unterschiedlicher Erschließungstiefe und Darstellungsform, insgesamt über 60.000 Stück. Alle Beschreibungen werden entweder als retrievalbare Datenbankdokumente angeboten oder als Images des jeweiligen gedruckten Kataloges oder – wie bei der Masse der Codices – in Verknüpfung dieser beiden Ebenen miteinander. Ergänzend sei auch kurz bemerkt, was *Manuscripta* nicht leistet: *Manuscripta* bietet und vermittelt kein Spezialangebot zu einer einzelnen bestimmten wissenschaftlichen Disziplin. Zu den Nutzern von *Manuscripta* zählen außer Historikern/innen mindestens ebenso Philologen/innen, Rechts- und Kunsthistoriker/innen, Philosophen/innen, Theologen/innen und nicht zuletzt die Katalogbearbeiter/innen selbst. So gibt es auch bei *Manuscripta* zwar die obligatorische Linkliste, doch beschränkt sich diese ganz bewusst auf das Arbeitsgebiet der Handschriftenforschung. Das Objekt Handschrift ist der Dreh- und Angelpunkt von *Manuscripta* und von ihm müssen hier die weiteren Überlegungen ausgehen.

Referenzierung auf *Manuscripta*

Bei der Frage nach den Perspektiven im digitalen Verbund sollte *Manuscripta* als einer von vielen Knotenpunkten verstanden werden: Welche Wege führen zu diesem Angebot hin und welche gehen von dort aus wohin? Grundsätzlich gibt es natürlich die Möglichkeit, auf *Manuscripta* insgesamt zu verlinken. Auch Teilbereiche, wie die digitalisierten Handschriftenkataloge und deren einzelne Druckseiten, sind als statische HTML-Seiten individuell anzusteuern. Dieser Schritt mag jedoch von Fall zu Fall unbefriedigend bleiben. So kann es für einen Codex zwar eine Beschreibung in einem Datenbankdokument geben, diese Beschreibung aber nicht gleichzeitig als Katalogimage vorliegen, auf das Bezug genommen werden könnte. Es sei in diesem Zusammenhang daran erinnert, dass *Manuscripta* zunehmend laufende Katalogisierungsunternehmen präsentiert, für die naturgemäß noch gar keine gedruckten Publikationen zur Verfügung stehen, und dies in manchen Fällen auch nie tun werden.

Ebenso ist es denkbar, dass eine Datenbankangabe sich zwar einer äl-

teren gedruckten Publikation bedient, dieser gegenüber aber in der Zwischenzeit aktualisiert und erweitert wurde. Genau dies geschieht seit einigen Jahren beispielsweise im Projekt der Inventarisierung der deutschsprachigen Handschriften der Staatsbibliothek zu Berlin (SBB).³ Zwar steht das einschlägige gedruckte Verzeichnis von Hermann Degering aus den Jahren 1925-1932 vollständig imagedigitalisiert im Netz. Den ungleich höheren Informationswert jedoch besitzen die aus ihm generierten, mittlerweile aktualisierten und entsprechend umfangreicheren Beschreibungen in der Datenbank. Ein Verweis sollte sich also naturgemäß immer auf diese Dokumente richten, zumal dort ohnehin der interne Link auf das Katalogimage von Degering zur Verfügung steht. Wünschenswert ist also die Möglichkeit der Verlinkung auf einzelne Dokumente im Datenbestand von *Manuscripta*. So kann es für die Betreiber fachlich verwandter Webangebote interessant sein, ergänzend auf die aktuelle Beschreibung einer bestimmten Handschrift zu verweisen. Bearbeiter/innen elektronischer Editionen mögen die benutzten Codices zitieren oder es könnte von digitalen Repertorien zu deren Belegstellen verknüpft werden. Derzeit wird geprüft, auf welchem Weg die hierfür notwendige eindeutige Referenzierung der Beschreibungsangaben erfolgen kann.

Eine zusätzliche Option, den Weg zu *Manuscripta* zu finden, wäre eine Öffnung für Web-Suchmaschinen wie Google. Grundsätzlich wäre es etwa denkbar, den gesamten Datenbestand in statische HTML-Seiten zu exportieren, um diese durch die Webcrawler der Suchmaschinen erfassen und indexieren zu lassen. Links auf diesen Seiten würden die Nutzer/innen bei der Auswahl eines Google-Treffers dann direkt in den Datenbestand selbst führen. Ein weiteres Stichwort lautet OAI-Fähigkeit. Einem Harvester könnten mit Hilfe dieses Protokolls Metadaten übergeben werden, die an anderer Stelle für den Aufbau eines Index genutzt würden. Diese Konstruktion macht natürlich nur dann Sinn, wenn die Informationen im Index mit weiteren Angaben anderer Datenprovider gemischt und gemeinsam für eine Recherche zur Verfügung gestellt werden. Ein solcher Weg wird gegenwärtig vom Consortium of European Research Libraries

3 <http://www.manuscripta-mediaevalia.de/hs/projekt_riecke.htm> (27.03.2006).

(CERL) beschriftet, das auf dem Sektor des alten Buches bislang vornehmlich durch die „Hand-Press-Book-Database“ in Erscheinung getreten ist.⁴ Unter seiner Ägide wird eine Software unter Verwendung des OAI-Protokolls entwickelt, die die verteilte Suche in elektronischen Handschriftennachweissystemen Europas und der USA ermöglichen soll. *Manuscripta* ist im Beirat des Unternehmens vertreten.

Verknüpfung von *Manuscripta* mit verwandten Angeboten

Für die Nutzer und Nutzerinnen interessanter als der Weg hin zu *Manuscripta* ist aber sicher die Frage nach den Verknüpfungsmöglichkeiten, die sich von dort aus bieten – immer ausgehend vom zentralen Gegenstand des Angebots, der Handschriftenerschließung. Es liegt daher nahe, an erster Stelle weitere Datenquellen dieses Arbeitsgebiets in den Blick zu nehmen. Die – zumindest unter dem Gesichtspunkt des einheitlichen Recherchezugangs – am nächsten liegende Variante besteht natürlich in der Übernahme bereits existierender Daten. Hier ist auf die jüngste Entwicklung auf dem Gebiet der Austauschformate zu verweisen. Zukünftig soll die Datenübernahme über das mit EU-Mitteln entwickelte MASTER-Format (heute Teil von TEI-P5) erfolgen können, das zwischenzeitlich auf die Bedürfnisse der deutschen Handschriftenerschließung hin abgestimmt wurde.⁵ Entsprechend formatierte Daten können dann in den Datenbestand von *Manuscripta* importiert bzw. Beschreibungen von dort in das MASTER-Format exportiert werden.

In vielen Fällen wird eine „Nachnutzung“ externer Katalogisierungsinformationen aber auf einen Verweis in diese externen Datenangebote beschränkt bleiben. Vorbedingung hier wie bei anderen Verknüpfungszielen ist naturgemäß eine Klärung der Rechte und eine Genehmigung der Seiteninhaber sowie eine gewisse Dauerhaftigkeit des verlinkten Angebots. Im Vordergrund stehen solche Katalogisierungsergebnisse, die bereits in *Manuscripta* vorhandene Informationen ergänzen können. Als Beispiel

4 <<http://www.cerl.org/HPB/hpb.htm>> (27.03.2006).

5 Vgl. hierzu die Projektseite der Herzog-August-Bibliothek Wolfenbüttel: <<http://www.hab.de/forschung/projekte/master.htm>> (27.03.2006).

mag der nur im Netz verfügbare Nachtrag zu Hermann Julius Hermanns Katalog der deutschen romanischen Handschriften der ÖNB Wien dienen.⁶ Hier finden sich zu über 50 Codices, für die *Manuscripta* bereits die eher kargen Informationen des älteren Tabulae-Katalogs enthält, erstmals eingehende Beschreibungen des Einbands und des Buchschmucks sowie eine fundierte kunsthistorische Einordnung. Hier ist eine Verknüpfung mehrerer unterschiedlicher Beschreibungen einer einzelnen Handschrift vorstellbar. Ergänzende elektronische Nachweise sollten in *Manuscripta* zumindest in die topografische Systematik der digitalisierten Handschriftenkataloge eingestellt werden. Ich denke hier an lokale Seiten wie diejenigen von Den Haag oder Lyon sowie das Datenangebot der Österreichischen Akademie zu den Handschriften von Klosterneuburg⁷, das seinerseits an die Signaturenfolge der bereits in *Manuscripta* vorhandenen Beschreibungen anschließt.

In Ergänzung der Beschreibungsinhalte war es immer erklärtes Ziel der Betreiber, möglichst zahlreiche Images aus den beschriebenen Codices selbst zur Verfügung zu stellen. Dabei sollte das Bild grundsätzlich der Erschließung folgen, also nicht für sich selbst stehen, sondern die Beschreibung erläutern und verifizierbar machen. Die etwa 90 derzeit im Angebot befindlichen digitalisierten Handschriften zeugen nicht eben vom bisherigen Erfolg dieser Idee. Hierzu hat in erster Linie der allgemein gestiegene Anspruch an die Bildqualität beigetragen. Die Farbdigitalisierung vom Original in hoher Auflösung ist für herausragende Stücke oder geschlossene Sammlungen geringer oder mittlerer Größe sicherlich eine geeignete Herangehensweise. Auf die in *Manuscripta* vertretenen Objektmengen ist sie jedoch nicht ohne sehr erheblichen finanziellen und zeitlichen Aufwand zu übertragen.

Es soll daher an dieser Stelle noch einmal für einen ergänzenden An-

6 <<http://www.onb.ac.at/sammlungen/hschrift/kataloge/ergaenzungen/ergaenzungen.htm>> (27.03.2006).

7 Zu Den Haag vgl.: <<http://www.kb.nl/kb/manuscripts/index.html>> (27.03.2006); zu Lyon: <[http://sged.bm-lyon.fr/Edip.BML/\(csthnlrkd0virt450luil145\)/Pages/Redirector.aspx?Page=MainFrame](http://sged.bm-lyon.fr/Edip.BML/(csthnlrkd0virt450luil145)/Pages/Redirector.aspx?Page=MainFrame)> (27.03.2006); zu Klosterneuburg: <<http://www.oeaw.ac.at/ksbm/kln/index.htm>> (27.03.2006).

satz geworben werden, auch wenn dieser auf den ersten Blick (hier durchaus in doppelter Wortbedeutung) weniger attraktiv zu sein scheint. Unter Nutzung der in vielen Bibliotheken vorhandenen Sicherheitsverfilmung, zu der bereits ein jüngerer Überblick zu Mengen und Qualitäten existiert, könnten Handschriften in erheblicher Zahl und kurzer Zeit in Graustufen digitalisiert und den entsprechenden Erschließungsangaben zur Seite gestellt werden. Für die überwiegende Zahl der Fragen zum Textbestand einer Handschrift, für Schriftvergleiche oder einfach für den häufig notwendigen, dann aber auch oft ausreichenden ersten optischen Eindruck würden diese vollkommen genügen. Kunsthistorische und kodikologische Fragestellungen sind naturgemäß auf diesem Weg nicht zu beantworten – in diesem Sinne könnten herausragende Stücke natürlich ausgespart bleiben –, aber speziell für diese Anliegen ist der Gang zum Original ohnehin oft unumgänglich. Auf diesem Weg ist unter vertretbarem Einsatz in kurzer Zeit die Einbindung großer Bildmengen denkbar. Ungeachtet der Idee, Images direkt bei *Manuscripta* einzustellen, wird es zukünftig verstärkt darum gehen, auf Bilder digitalisierter Handschriften in externen Angeboten zu verlinken. Auch hier haben natürlich Abbildungen von solchen Beständen Priorität, deren Erschließung bereits über *Manuscripta* zugänglich ist. Hier ist etwa an die digitalisierten Handschriften der Stadtbibliothek Olmütz zu denken⁸, deren Beschreibungen bereits zur Verfügung stehen.

Neben Kataloginformationen und Handschriftenimages bieten sich weitere Hilfsmittel aus dem Bereich der Handschriftenkunde für einen gezielten Verweis aus den Beschreibungen heraus an. Hierzu zählen bibliografische Angaben zu einzelnen Codices sowie Bildquellen zur Einband- und Wasserzeichenbestimmung. So verfügen einige Bibliotheken mit umfangreichem Handschriftenbestand über eine teilweise weit in die Vergangenheit zurückreichende separate Dokumentation zu ihren Objekten. Diese wurden und werden etwa an der Bayerischen Staatsbibliothek (BSB) oder der Herzog-August-Bibliothek (HAB) Wolfenbüttel bereits in

8 <<http://dig.vkol.cz/>> (27.03.2006).

digitale Form überführt.⁹ Die signaturbezogene Verlinkung auf die bibliographischen Angaben der zu einer Handschrift erschienenen Forschungsliteratur ist für die Wolfenbütteler Dokumentation bereits umgesetzt, für die Handschriften der BSB befindet sich diese Funktionalität derzeit in Arbeit. Die Nutzer/innen gelangen also aus der Beschreibung einer Handschrift in *Manuscripta* heraus direkt auf die lokale Bibliografie zu dieser Handschrift. Ergänzend käme hier noch die Dokumentation der Universitätsbibliothek Augsburg in Betracht, die Forschungsliteratur zu den Handschriften und Inkunabeln der Sammlung Oettingen-Wallerstein nachweist.¹⁰

Zum Stichwort Einbandforschung: Seit dem Jahr 2003 erleichtert die DFG-geförderte Einbanddatenbank (getragen von der Württembergischen Landesbibliothek (WLB) Stuttgart, HAB Wolfenbüttel, BSB München und der SBB Berlin) den an Handschriften, Inkunabeln und Frühdrucken Interessierten die Bestimmung von Einbandstempeln und –werkstätten.¹¹ Im Umkehrschluss wäre es nun wünschenswert, die Befunde bei der jeweiligen Handschrift durch die Illustration des Stempels in der Einbanddatenbank zu unterfüttern. Anknüpfungspunkt für Handschriften, deren Einbände bereits durchgerieben wurden, wäre dort die Kategorie der „buchbinderischen Einheit“, für noch nicht erfasste Codices könnte die so genannte „Zitiernummer“ den Bezugspunkt abgeben. Mit der Datenbank zur Piccardschen Wasserzeichensammlung beim Hauptstaatsarchiv Stuttgart steht auch das zentrale Auskunftsmittel für die Identifizierung und Datierung der in Handschriften verwendeten Papiere seit einiger Zeit online zur Verfügung.¹² Analog zu den Einbandstempeln könnte ein Link hierhin die nüchterne Angabe von Repertoriennummer und/oder Motiv in der Beschreibung illustrieren. Bei den nicht über die bereits gedruckten Bände zugänglichen Zeichen würde die Verlinkung sie überhaupt erst

9 Für die BSB vgl. <<http://www.bsb-muenchen.de/Forschungsdokumentation.172.0.html>> (27.03.2006); für die HAB vgl. <<http://www.hab.de/forschung/projekte/handschriftendokumentation.htm>> (27.03.2006).

10 <http://www.bibliothek.uni-augsburg.de/sondersammlungen/oettingen_wallerstein/fordok/> (27.03.2006).

11 <<http://www.hist-einband.de/>> (27.03.2006).

12 <http://www.landesarchiv-bw.de/sixcms/detail.php?template=hp_artikel&id=5171&sprache=de> (27.03.2006).

vergleichbar machen.

Ein zentrales Ziel der Handschriftenkatalogisierung ist die Identifikation und Einordnung der in einem Codex vertretenen Texte. Wo es eine solche gibt, wird in der Beschreibung Bezug auf die entsprechende Edition genommen. Da mittlerweile eine ganze Reihe von Textcorpora auch elektronisch im Netz verfügbar ist, könnte diese Angabe unmittelbar an die korrespondierende Stelle der digitalen Edition angeknüpft werden. Bei einer Teilüberlieferung wäre je nach Strukturierungsgrad der angesteuerten Edition im besten Fall sogar auf einen bestimmten Textabschnitt zu verweisen. Zu diesem Zweck müsste das Datenangebot aber frei im Netz verfügbar sein. Bezogen auf historische Texte liegt es nahe, an die digitalen *Monumenta Germaniae Historica* (dMGH) anzuknüpfen.¹³ Von der Beschreibung eines Textes in einem Codex direkt zum Volltext der Edition in den MGH zu gelangen bzw. umgekehrt von der Edition aus zur Beschreibung und/oder den Images der dieser Edition zu Grunde liegenden Handschrift/en zu verknüpfen, wäre zweifellos wünschenswert. Die gegenseitige Nutzbarmachung bereits existierender Datenangebote würde fraglos einen Mehrwert dieser Einzelangebote bewirken und so verstärkt die bereits eingesetzten Mittel rechtfertigen. Ein Zusammenhang, der nicht zuletzt mit Blick auf potentiell notwendige Fördermittel zu betonen ist.

Realisierung

Abschließend seien noch ein paar Gedanken zur Gewichtung der genannten Möglichkeiten umrissen. *Manuscripta* hat zwei grundsätzliche Rechercheinteressen zu befriedigen: Entweder ein Nutzer/eine Nutzerin benötigt möglichst umfangreiche Informationen zu einer bestimmten Handschrift oder er/sie sucht möglichst viele relevante Handschriften zu einer bestimmten Fragestellung.

Unter dieser Voraussetzung sei als Beispiel noch einmal die Idee einer Anbindung an Google aufgegriffen. Über die bekannteste Web-Suchmaschine gefunden zu werden, klingt erst einmal hochattraktiv. Genau genommen dient diese Option aber weder dem einen noch dem

13 <<http://www.dmgd.de/>> (27.03.2006).

anderen der genannten Suchinteressen. Sinnvoll scheint sie lediglich für Anfragen zu sein, bei denen Handschriften nicht ursprünglich im Blick waren, für die Treffer aus *Manuscripta* dann aber doch relevant sind. Anders als bei großen Bilddatenbanken oder verstreuten Datenbeständen mit neuzeitlichem Material scheint es aber unsicher, ob solche „Zufallstreffer“ in Google *Manuscripta* wesentlichen neuen Nutzerkreisen erschließen. Zudem muss berücksichtigt werden, dass eine Realisierung der angesprochenen Verbesserungsmöglichkeiten grundsätzlich auf das gesamte – und eben oftmals inhomogene – Datenmaterial von *Manuscripta* bezogen sein sollte. Dies wirkt sich beispielsweise auf die angesprochene Verknüpfung von Handschriftenbeschreibungen und elektronischen Editionen aus. Da die Masse der Daten in *Manuscripta* aus den Registern der Kataloge stammt, enthalten sie in der Regel keine Angaben zu benutzten Editionen. Daher liegt es an dieser Stelle nahe, in einem ersten Schritt die Edition auf die für ihre Erstellung verwendeten und in *Manuscripta* beschriebenen Handschriften zu beziehen, um erst in einem zweiten Schritt auf der Grundlage dieser Verknüpfungen umgekehrt eine Verlinkung aus der Textbeschreibung einer Handschrift zur Edition anzulegen. Anpassungsbedarf ergäbe sich auch bei den Wasserzeichen: Abbildungen aus den bereits gedruckten und bei der Handschriftenbearbeitung in der Vergangenheit verwendeten Piccard-Bänden wären nicht ohne weiteres über die bislang übliche Piccard-Zitiernummer anzusprechen, sondern diese müssten im Vorlauf auf eine interne Zählung im Stuttgarter Angebot umgesetzt werden.

Schließlich unterliegen zumindest alle im Volltext erfassten Beschreibungen einem Urheberrecht ihrer Verfasser/innen. Diese Katalogbeschreibungen sind eigenständige wissenschaftliche Leistungen und ihre Ergänzung – so sinnvoll sie im Einzelfall auch sein mag – bedarf der Zustimmung ihrer Erzeuger.

Diese einschränkenden Bemerkungen sind kein Ausdruck vornehmer Selbstbescheidung. Vielmehr müssen durchgängig technische Möglichkeiten, inhaltliche Notwendigkeit und entstehende Kosten sinnvoll aufeinander abgestimmt werden. *Manuscripta* ist zwar fest bei seinen drei Trägerinstitutionen (BSB München, Bildarchiv Foto Marburg und SBB)

verankert und damit institutionell langfristig gesichert. Entwicklungen wie die oben umrissenen müssten jedoch zu einem wesentlichen Teil als Projektarbeit unter Beteiligung weiterer Partner organisiert werden. Dabei sollte es um die Entwicklung von Routinen gehen, die nach Projektende im Regelbetrieb weiter langfristig einsetzbar sind. Welche Zusammenstellung von Projektideen und –teilnehmern sich tatsächlich ergeben mag, ist natürlich von den individuellen Interessen und Zielen möglicher Partner abhängig. Wo in diesem Verbund *Manuscripta* stehen könnte, ist nun aber hoffentlich deutlicher erkennbar.

Dr. Robert Giel ist Mitarbeiter der Handschriftenabteilung der Staatsbibliothek zu Berlin und Mitglied der Zentralredaktion von Manuscripta mediaevalia. E-Mail: robert.giel@sbb.spk-berlin.de

DIE REGESTA IMPERII ONLINE

von Andreas Kuczera

Als Inventar aller urkundlichen und historiografischen Quellen der römisch-deutschen Könige von den Karolingern bis zu Maximilian I. sowie der Päpste des frühen und hohen Mittelalters gehören die Regesta Imperii zu den großen Quellenwerken zur deutschen und europäischen Geschichte. Mit dem Abschluss der im Juni 2001 in Kooperation mit der Bayerischen Staatsbibliothek begonnenen Digitalisierung, Erschließung und Bereitstellung der Regesta Imperii im Sommer 2006 stehen alle bis Anfang 2006 gedruckten Bände als Volltext im Internet bereit. Im vorliegenden Artikel wird nach einer Bestandsaufnahme ein Ausblick auf die nächsten Schritte der Regesta Imperii Online gewagt, insbesondere die Verlinkungs- und Editionsöglichkeiten.

Einleitung

Die Regesta Imperii (RI) entstanden auf Initiative des Frankfurter Stadtbibliothekars Johann Friedrich Böhmmer (1795-1863), der im Jahre 1829 mit dem Sammeln und Publizieren von Urkunden der deutschen Kaiser und Könige begann. Ziel war es, durch die Kombination der Urkunden eines Herrschers mit den Informationen aus den zeitgenössischen „erzählenden“ Quellen (Annalen, Chroniken usw.) eine lückenlose Dokumentation zur Person (Itinerar) und des politischen Handelns zu erstellen. Ursprünglich als Vorarbeit zu den „Monumenta Germaniae Historica“ gedacht, entwickelten sich die RI nach dem Tode Böhmers mit einem erweiterten Regestenkonzept, das eine exakte Wiedergabe von Form und Inhalt sowohl der Urkunden als auch der historiografischen Nachrichten in Kurzform vorsah,

zu einem selbstständigen Grundlagenwerk fort, welches sich in der Mediävistik längst als unentbehrlich erwiesen hat. Die Akademien in Mainz, Berlin und Wien arbeiten derzeit an 15 Teilprojekten.

Neben der allgemeinen politischen Geschichte decken die RI insbesondere die Bedürfnisse der Forschung im Bereich der Rechts- und Verfassungsgeschichte, und darüber hinaus auf verschiedenen Feldern der Sozial-, Wirtschafts- und Personengeschichte ab. Die Schwerpunkte variieren dabei naturgemäß bezogen auf die jeweils behandelte Epoche. Der Ansatz, das gesamte abendländische Mittelalter von 751 bis 1519 quellenmäßig aus der Perspektive der Zentralgewalt(en) zu erschließen, erfordert in hohem Maße Konzeptionsstabilität. Dies gilt erst recht mit Blick auf die räumlichen Bezugfelder, die das gesamte römisch-deutsche Imperium und somit weite Teile Zentraleuropas umfassen. West-europäische Bindungen (insbesondere in der Karolingerzeit) stehen neben osteuropäischen (etwa in der Zeit der luxemburgischen Herrscher), ebenso wechseln geografisch gesehen nördliche Schwerpunkte (Ottonen, Salier) und südliche (Staufer, Habsburger) einander ab, alles jeweils mehr oder minder lose verklammert durch die Bindungen an das römische Papsttum.

In den Regesta Imperii werden die aus Archiven und Bibliotheken zusammen getragenen Quellen unter Beachtung der an der Zentralgewalt orientierten Perspektive, jedoch unter Berücksichtigung der epochalen Spezifika nach klar definierten wissenschaftlichen Kriterien aufbereitet dargeboten. Die enge Verzahnung – im Bereich Edition – mit anderen großen mediävistischen Quellenunternehmungen (Monumenta Germaniae Historica und Jahrbücher der Deutschen Geschichte für den Bereich Früh- und Hochmittelalter, Deutsche Reichstagsakten für den Bereich Spätmittelalter) haben die RI zum idealen Einstieg für quellenfundierte Forschungen werden lassen. Die hohe Akzeptanz, die das Unternehmen in Forschung und Lehre sowie in Archiven und Bibliotheken findet, beruht nicht zuletzt auf dieser Drehscheibenfunktion.

Im Gegensatz zu vielen anderen Akademieunternehmungen haben die RI bereits früh begonnen, sich in Richtung elektronische Publikationsformen zu orientieren. Mit der seit Mitte der 1990er Jahre betriebenen Entwicklung einer CD-ROM-Fassung, für die als Prototyp das am weitesten

„elektronisch“ fortgeschrittene Teilprojekt der „Regesten Kaiser Friedrichs III.“ ausgewählt wurde, konnten bereits erste Erfahrungen mit einer elektronischen Aufbereitung und den hier realisierbaren Suchmöglichkeiten gewonnen werden.

Entwicklungsstand

Seit Juni 2001 erfolgt in Kooperation mit der Bayrischen Staatsbibliothek unter dem Kennwort „Regesta Imperii Online“ die Digitalisierung, Erschließung und Bereitstellung der RI 1831-2004 und künftiger Bände im Rahmen des Deutsche Forschungsgemeinschaft-Förderungsbereichs Retrospektive Digitalisierung von Bibliotheksbeständen. Mit dem Auslaufen der letzten Förderungsperiode im Sommer dieses Jahres geht der mit dem Projekt verbundene Personalaufwand allein auf die Regestenkommission über.

Stand der Arbeiten – Regesten

Gerade am Beispiel der Regesten Kaiser Friedrichs III. lassen sich die Vorteile digitaler Publikationsmethoden veranschaulichen. Als dezentral angelegtes Projekt, bei dem jeweils einzelne Archivbestände aufgearbeitet und anschließend heftweise publiziert werden, zeigt das Projekt, dass die Benutzung des Gesamtwerkes mit der Anzahl der Hefte zunehmend schwieriger wird. Hier bieten digitale Publikationsmethoden neue Möglichkeiten, die über die Probleme gedruckter Faszikel hinweghelfen können.

Hat man nach der Eingabe der Suchkriterien den „Suchen“-Button gedrückt, erscheint nach kurzer Zeit im unteren Frame eine Liste mit den Treffern in Kurzanzeige. In dieser Übersicht kann eine Vorauswahl getroffen werden. Zur Vollansicht eines Regests kommt man entweder über einen Klick auf die Nummer des Treffers, wobei die Vollansicht im oberen Rahmen geöffnet wird, oder durch einen Klick auf das Doppelfenster-Icon, wobei hier nun die Vollansicht in einem neuen Browserfenster geöffnet wird.

Die Vollansicht des Regests ist dreigeteilt. Im oberen Frame werden Abteilung Band und Nummer sowie Verweise auf den vorherigen und den

nächsten Treffer genannt. Die linke Spalte zeigt das Regest im Zusammenhang des eigenen Bandes an, der rechte Frame bietet den Volltext des Regests wobei die Formatierung jener des Bandes nachempfunden ist. Unter jedem Regest findet sich der Hinweis: Sie haben Nachträge, Hinweise, Ergänzungen. Dort können die Nutzer/innen in einem Webformular Hinweise zu den Regesten an die Kommission mailen.

Stand der Arbeiten – Register

Über die Volltextsuche hinaus gibt es für die Regesten Kaiser Friedrichs III. bereits ein per Hand über alle Hefte der Abteilung kumuliertes Register. Über einen Klick auf eine als Link gekennzeichnete Registernummer erhält man im unteren Frame zunächst eine Kurzansicht, wie wir sie bereits aus der Volltextsuchemaske kennen. Für die Zukunft ist es aber angedacht, das Register von den Mitarbeitern/innen selbst erweitern und ausbauen zu lassen. So ist es denkbar, dass die zu den einzelnen Heften zu erstellenden Register nicht mehr in Word, sondern über eine Eingabemaske gleich in die kumulierte Fassung im Netz eingegeben werden. Anschließend könnten über Exportfunktionen die zum entsprechenden Heft gehörigen Einträge formatiert ausgegeben und als Register der Druckvorlage hinzugefügt werden.

Stand der Arbeiten – RI-Opac

Wichtiger Bestandteil eines Regests sind Verweise auf gedruckte Quellen und Literatur. Allerdings sind die Literaturangaben oft mit für die Benutzer/innen kryptischen Abkürzungen wiedergegeben, die nur mühsam über Umwege in der Literaturrecherche rekonstruiert werden können, oder das Quellen- und Literaturverzeichnis fehlt gänzlich. Beide Fälle kommen sowohl in den gedruckten Bänden des 19. als auch des 20. Jahrhunderts vor. Mit der direkten Verknüpfung des Kurzzitats mit der Literaturdatenbank wird für die Benutzer/innen ein deutlicher Mehrwert erzielt. Basis für die Direktverlinkung ist ein in der Datenbank für jeden Literaturtitel gespeicherter Kurztitel, der aus einem eindeutigen Kurzzitat – in der Regel bestehend aus Verfassernamen und wichtigsten Titelbestandteil – gebildet wird.

Über die eigentliche Aufgabe, nämlich die Bereitstellung der in den

Regestenbänden zitierten Literatur, entwickelt sich der RI-Opac mit seinen mittlerweile über 700.000 Datensätzen zu einem immer mächtigeren Werkzeug, das ohne jegliche thematische Einschränkung zeitlich übergreifend, vor allem jedoch Titel zur mittelalterlichen Geschichte des gesamten europäischen Raumes nachweist.

Besondere Merkmale sind die tiefe Erschließung unselbstständigen Schrifttums (Aufsätze aus Zeitschriften und Beiträge zu Fest- und Sammelchriften). Neben „großen“ Zeitschriften wie der Historischen Zeitschrift, dem Deutschen Archiv, den Mitteilungen des Österreichischen Instituts für Geschichtsforschung, *Moyen Age*, *English Historical Review* und vielen anderen werden auch kleinste regionale Zeitschriften und Sammelwerke systematisch von Erscheinungsbeginn an erfasst.

Zur sachlichen Erschließung der Inhalte bietet die Literaturdatenbank einen systematischen sowie einen alphabetischen Thesaurus, der die Verschlagwortung nach geografischen, chronologischen und thematischen Gesichtspunkten erschließt. Einen Überblick über Verfassernamen und die Titelstichwörter der verzeichneten Werke geben alphabetische Indexe. Über die in den RI-Bänden zitierte Literatur hinaus wird der RI-Opac ständig von Fachkollegen/innen aktualisiert und um neu erschienene Literatur erweitert.

Perspektiven

Interne Vernetzung

Momentan dient der Opac gleichzeitig sowohl als Basis für die Literaturlinks (Literaturangaben in den Regesten) als auch als künftige zentrale Arbeitsdatenbank für die über den gesamten deutschen Sprachraum verteilten Regestenbearbeiter/innen. Via Internet sollen die Mitarbeiter/innen die Literaturangaben der neuen Regestenbände bei der Erstellung mit Kurztiteln auszeichnen, womit sich die bisher angewandte nachträgliche Aufbereitung erübrigen würde. Für die Regestenbearbeiter/innen ergeben sich bei der Erstellung neuer Bände durch die Nutzungsmöglichkeiten der Online-Regesten zahlreiche Vorteile. Als Beispiele seien nur die leichtere Vereinheitlichung von Formelteilen oder der chronologische Zugang und die

Suchmöglichkeit nach Quellenschreibweisen genannt, die vor allem disloziert strukturierten Projekten die einheitliche Arbeit erleichtern. Hier deuten sich erste Möglichkeiten an, die sich ergeben, wenn digitale Quellen nicht über die Retrodigitalisierung bereits bestehender Quellenbände erstellt, sondern von vorne herein digital strukturiert eingegeben werden. So ist es zum Beispiel denkbar, Literaturhinweise oder Register über zentrale Nachweisinstrumente (RI-Opac) zu erstellen und bei deren Ausbau und Pflege die Quellenbearbeiter/innen miteinzubeziehen oder Verlinkungen zu anderen Quellenwerken im Internet vorzunehmen, die das betreffende Stück auf andere Weise erschließen, wie Urkundeneditionen, Siegelabbildungen usw.

Neben der durch Tiefenerschließung zu erwartenden inhaltlichen Aufwertung des gedruckten Regests wird zunehmend auch die Erweiterung des Internetangebotes nach außen an Bedeutung gewinnen. Textuell zentrierte Strukturen – wie die momentan vorhandene Regestendatenbank – werden durch die Einbindung von Bildmaterialien – wie den Faksimiles der behandelten Urkunden –, die Einbindung von Siegelabbildungen oder hilfswissenschaftlichen Rahmeninformationen – wie Herrscheritinerar, Kartenmaterial usw. – ergänzt. Es wird interessant sein, zu beobachten, inwieweit sich solche Ergänzungen der bisherigen Publikationsformen auch auf die inhaltliche Ausgestaltung der Werke auswirken werden. Leider steht dem einfachen Einscannen der Urkundenabbildungen und der anschließenden Einstellung ins Netz bisher die noch recht heterogene Rechtslage bezüglich der Veröffentlichungsrechte der Bilder entgegen. Letztendlich müßte mit jedem Archiv einzeln eine Vereinbarung erzielt werden, in der die Bereitstellung der Urkundenfaksimiles im Internet genehmigt wird.

Eine wünschenswerte Alternative zur Bereitstellung der Urkundenfaksimiles durch die Regesta Imperii wäre natürlich die Bereitstellung durch die entsprechenden Archive im Internet. Auf diese Angebote könnte dann vom einzelnen Regest aus verwiesen werden.

URI für Regesten

Auf die Verlinkung der in den Regesten vorhandenen Literaturhinweise

mit dem Opac wurde im vorherigen Abschnitt bereits eingegangen. An einem ähnlichen Verfahren arbeiten wir momentan für die Verlinkung der Volltextregesten. Ein URI (*Uniform Resource Identifier*) bietet einen eindeutigen Link auf ein Regest und erleichtert damit die direkte Verlinkung. Ein Aufruf des Links oder die Eingabe in die Adresszeile des Browsers führen direkt zum entsprechenden Regest.

Hier ein Beispiel:

```
<http://www.regeste-imperii.de/regesten/
  regshow.php?uri=1474-02-14_001_000_013_
  003_000_000000125>.
```

Der URI lehnt sich dabei an die Regesten-ID an, die innerhalb der „Regesta Imperii Online“ zur eindeutigen Identifikation der Regesten Anwendung findet. Hierbei handelt es sich um eindeutige Verweise, bei denen ein URI genau auf ein Regest verweist.

Versionierung

Ziel der RI war es von Anfang an, neben der Bereitstellung bereits gedruckten Materials Nachträge und Verbesserungen zügig der interessierten Öffentlichkeit zur Verfügung zu stellen. Wenn Nachträge in die Onlineversion der Regesten eingearbeitet werden unterscheidet sich diese aber von der gedruckten Fassung. Eine Lösung dieses Problems wäre eine Versionsverwaltung für die Online-Regesten, die jedoch den Nachteil mit sich brächte, dass die URIs komplizierter und die Online-Regesten auch unübersichtlicher werden. Deshalb verfolgen wir momentan einen Ansatz, bei dem das gedruckte Regest im Mittelpunkt bleibt, aber eben durch Nachträge, Ergänzungen und Verbesserungen bereichert wird. Die mit einem Datum versehenen Nachträge werden unterhalb des ursprünglichen Regests angeordnet. Damit bleibt die Entwicklung des Regests transparent.

Wer darf was?

Nachträge sind eine begrüßenswerte Ergänzung der Regesten. Sie bringen aber auf der anderen Seite natürlich auch Probleme mit sich. So muss geklärt werden, wer was wo einbringen darf und welche Kontrollinstanzen über die Ergänzungen entscheiden. Kommen die Nachträge von den Bear-

beitern/innen selbst, reduziert sich der Aufwand auf die Einarbeitung und die anschließende Veröffentlichung. Kommen die Nachträge von außen, sind personelle Ressourcen für die Kontrollinstanzen nötig, die das Material sichten und gegebenenfalls freigeben.

Im übertragenen Sinne und kleinerem Maßstab befinden wir uns vor einer ähnlichen Problematik wie die Wikipedia, bei der durch die offene Struktur in kürzester Zeit eine Vielzahl von qualitativ hochwertigen Artikeln zusammengekommen sind, meistens aber flankiert von ebenso negativen Beispielen schlecht recherchierter Artikel.

Fazit

Letztendlich befinden wir uns in einem Spannungsfeld zwischen den Möglichkeiten der digitalen Quellenedition, bei der sich die Einbindung der Nutzer ja geradezu aufdrängt und dem Anspruch, dass die Online-Regesten ebenso wie die Druckversionen wissenschaftlichen Kriterien genügen und zitierfähig bleiben müssen. Der Wechsel vom Druckmedium hin zur digitalen Edition bringt neue Möglichkeiten der Quellenerschließung und vor allem der Quellenvernetzung mit sich, die aber gleichzeitig wieder mit einem höheren Aufwand an finanziellen Ressourcen verbunden sind. Bei der aktuellen Knappheit der Mittel wird es eine Aufgabe sein, die um digitale Komponenten erweiterten Geschichtswissenschaften in diesem Spannungsfeld weiterzuentwickeln.

Dr. Andreas Kuczera ist Wissenschaftlicher Mitarbeiter der Deutschen Kommission für die Bearbeitung der Regesta Imperii e.V. bei der Akademie der Wissenschaften und der Literatur Mainz. Dort ist er unter anderem mit der fachinformatischen Weiterentwicklung des Online-Angebots der Regesta Imperii betraut. E-Mail: Andreas.Kuczera@geschichte.uni-giessen.de

DIE PRESSEARCHIVE VON HWWA UND ZBW – RETRODIGITALISIERUNG DER ALTBESTÄNDE VON 1900 BIS 1930

von Thomas Sergej Huck und Max-Michael Wannags

Das Hamburgische Welt-Wirtschafts-Archiv (HWWA) wurde 1908 als Zentralstelle des Kolonialinstituts in Hamburg gegründet. Nach dem Ersten Weltkrieg erhielt es seinen heutigen Namen. Seine Aufgabe bestand zunächst darin, durch den Ausbau der Presseauschnittsarchive und die Einrichtung einer Bibliothek Informationen über die wirtschaftliche und soziale Entwicklung insbesondere in den überseeischen Ländern für Wissenschaft, Wirtschaft und Politik bereitzustellen. Heute besitzt das HWWA eines der größten öffentlich zugänglichen Pressearchive im wirtschafts- und sozialwissenschaftlichen Bereich in Europa. Das Wirtschafts-Archiv des Instituts für Weltwirtschaft, heute Teil der Kieler Zentralbibliothek für Wirtschaftswissenschaften (ZBW), entstand 1914 als erste Informationsabteilung des neu gegründeten Instituts für Seeverkehr und Weltwirtschaft. Hier wurden ebenfalls aktuelle Presseartikel über das Wirtschaftsgeschehen in aller Welt bereitgehalten. Um den Zugriff für die Benutzer/innen zu erleichtern, wird im Rahmen des Projektes der Deutschen Forschungsgemeinschaft (DFG) „Retrospektive Digitalisierung von historischen Presseartikeln der Archive des HWWA und des IfW (heute ZBW)“ das historische Pressematerial des HWWA und des Instituts für Weltwirtschaft aus der Zeit von circa 1826 bis 1930/5 nun digitalisiert, aufbereitet und über eine Datenbank im Internet zur Verfügung gestellt.¹

¹ Überarbeitete Fassung eines auf der .hist 2006-Tagung „Geschichte im Netz – Praxis, Chancen, Visionen“ vom 22. bis 24. Februar 2006 an der Humboldt-Universität zu Berlin gehaltenen Vortrages. Besonderer Dank für kritische Anmerkungen und Durchsicht des Textes gebührt Johanna Becker und Kirsten Jeude.

Konzeption und Strategie der Digitalisierung

Angetrieben von den technischen Möglichkeiten und wissenschaftlichen Erfolgen während der Hochzeit der Industriellen Revolution und erfasst von einem positivistischen Wissenschaftsgeist, der die Grenzen menschlicher Erkenntnisse und Möglichkeiten in weite Ferne rücken ließ, wurden in den Jahren um die Wende vom 19. zum 20. Jahrhundert in Europa mehrere Institute mit in concreto zwar spezifischen, letztlich aber doch unter nur einem Oberthema rubrizierbaren Zielsetzungen gegründet, nämlich denen des Aufbaus eines Archivs der Welt oder eines Weltengedächtnisses, eines „mundaneum“, wie dies bei dem 1895 in Brüssel von Paul Otlet und Henri La Fontaine gegründeten Institut International de Bibliografie² heißt, einer Bibliothek im Sinne eines alles und jeden umfassenden Speichers des (Welten-)Wissens mit der alleinigen teleologischen Perspektive und dem Zweck der Steigerung der Effizienz sowohl in der Produktion als auch in der Wissenschaft.

Worum es ging, zeigte schon 1897 der Kaiserliche Bibliothekar Christlieb Gotthold Hottinger (1848-1914) aus Berlin-Steglitz, der einen „Bücher-Zettel-Katalog und ein Bio-Ikono-Bibliographisches Sammelwerk“³ aufbauen wollte. Es dürfte sich kaum um einen Zufall handeln, dass ein Teil des Titels zweier seiner Publikationen die „bibliotheca universalis“ des Conrad Gessner von 1548 gewissermaßen paraphrasieren – plante dieser doch das ihm zugängliche Weltenwissen in enzyklopädischer Weise mit Querverweisen untereinander verschränkt dauerhaft aufzube-

-
- ² Rückblickend zum Anspruch des Instituts: Otlet, Paul, Sur les possibilités pour les entités administratives d'avoir à tout moment leur situation présentée documentairement. Rapport au 4ième congrès international des sciences administratives; 6ième section, Madrid 1930. Zur Person und zum Wirken des Henri La Fontaine: Lubelski-Bernard, Nadine, Un partisan inconditionnel de la paix, in: Goffin, Henry (Hg.), Tracé(s) d'une vie. Henri La Fontaine. Un prix nobel de la paix 1854-1943, Mons (Editions Mundaneum) 2002, S. 71-87.
- ³ Bio-Ikono-Bibliographia universalis. Angeregt von Prof. Dr. Christlieb Gotthold Hottinger, Kaiserlicher Hofbibliothekar, Südende-Berlin, 7. Oktober 1901 sowie ders., Ein Bücher-Zettel-Katalog und ein Bio-Ikono-Bibliographisches Sammelwerk, ib. 1911 (Erw. Neuauflage des Rundschreibens von 1901).

wahren und so zugänglich zu halten.⁴ Weitere Institute mit dem Ziel der restlosen Erfassung des Weltenwissens (wie das ähnlich Markus Krajewski kürzlich nannte) ließen sich beispielsweise unter Rückbezug auf Wilhelm Ostwalds Münchener „Institut zur Organisierung der geistigen Arbeit“ von 1911 jederzeit belegen⁵, doch hier seien mit besonderer Berücksichtigung der allen diesen Instituten eigentümlichen Arbeitsinstrumente im Zusammenhang mit der Logik der Zugänglichmachung dieses Wissens das am 20. Oktober 1908 als Zentralstelle des Kolonialinstituts gegründete Hamburgische Welt-Wirtschafts-Archiv⁶ (HWWA) genannt sowie das Wirtschaftsarchiv des Instituts für Weltwirtschaft⁷, das im Jahr 1914 als erste Informationsabteilung des neugegründeten Instituts für Seeverkehr und Weltwirtschaft ins Leben gerufen wurde. In allen diesen, teils könig- und kaiserlichen, teils gänzlich privaten Instituten wurde die vergangene, gegenwärtige und – zumindest war das geplant, auch: – zukünftige Welt unter Berücksichtigung der jeweils unterschiedlichen, eben institutsspezifischen Funktionsaspekte in eine Liste von Begriffen gepresst, in einen

-
- 4 Umfassend jüngst: Zedelmaier, Helmut, *Bibliotheca universalis und bibliotheca selecta: das Problem der Ordnung des gelehrten Wissens in der frühen Neuzeit*, Köln (Beihefte zum Archiv für Kulturgeschichte 33) 1992, bes. S. 147-173. Lesenswert ferner Haber, Peter, *Der wiedererwachte Traum von der „Bibliotheca Universalis“*. Das totale Wissen im digitalen Zeitalter, in: *Neue Zürcher Zeitung*, 24.01.2000 sowie im vorliegenden Kontext besonders: Krajewski, Markus, *Zettelwirtschaft. Die Geburt der Kartei aus dem Geiste der Bibliothek*, Berlin 2002, bes. S. 17ff.
- 5 Ders., *Restlosigkeit. Wilhelm Ostwalds Welt-Bildungen*, in: Pompe, Hedwig; Leander Scholz (Hgg.), *Archivprozesse. Die Kommunikation der Aufbewahrung*, Köln 2002, S. 173-185.
- 6 Leveknecht, Helmut, *90 Jahre HWWA. Von der Zentralstelle des Hamburgischen Kolonialinstituts bis zur Stiftung des HWWA. Eine Chronik*, Hamburg 1998. Jüngst, mit Schwerpunkt auf der Frühgeschichte des Instituts: Becker, Johanna, *Die Gründung des Deutschen Kolonialinstituts in Hamburg. Zur Vorgeschichte der Hamburgischen Universität*, Hamburg 2005.
- 7 Zu den Anfängen des Instituts jüngst Glaeßer, Hans-Georg, *Das Institut für Seeverkehr und Weltwirtschaft. Von den Anfängen eines Kieler Forschungsinstituts*, in: Elvert, Jürgen u.a. (Hgg.), *Kiel, die Deutschen und die See*, Stuttgart 1992, S. 155-168 mit weiterer Literatur S. 167f. sowie aus zeitgenössischer Perspektive des Gründers: Harms, Bernhard, *Das Königliche Institut für Seeverkehr und Weltwirtschaft an der Christian-Albrechts-Universität zu Kiel*. Kaiser Wilhelm Stiftung. Kiel, 2. Aufl., 1917.

Thesaurus sozusagen, gemäß dem die zeitgenössische nationale und internationale Buch- und Presselandschaft lektoriert wurde. Zu einem Begriff passende Beiträge – das konnten Kurzmeldungen ebenso sein wie längere Aufsätze, ganze Bücher, Werbebroschüren oder irgendwelche sonstigen gedruckten Emanationen menschlichen Geistes – wurden dann separiert und in einer zu diesem Begriff angelegten Themenmappe gesammelt. Und dies, wie sich zumindest im Falle des Wirtschaftsarchivs des Instituts für Weltwirtschaft (heute ZBW) und des Hamburgischen Weltwirtschaftsarchivs zeigen lässt, ohne Unterbrechung bis in die Gegenwart, jedenfalls bis zum 31. Dezember 2005, zu welchem Tag die Pressedokumentation des HWWA endgültig geschlossen worden ist.

Um welche Mengen es sich bei den Pressearchiven in Hamburg und in Kiel handelt, ist nicht ganz klar; es kursieren Mengengerüste, die sich zwischen 17 und 20 Millionen bewegen, wobei ebenfalls offen bleiben muss, ob es sich dabei dann um Artikel, Dokumente oder um Seiten handelt. Zählbar ist das alles ohne weitere Anstrengungen jedenfalls nicht, und die absoluten Zahlen selbst sind auch nicht entscheidend bei diesem weltweit einzigartigen Sammlungsgut der Pressedokumentationsabteilungen beider Häuser. Deren Singularität ergibt sich nicht allein aus der schier Masse des Materials, sondern insbesondere aus der selbstverständlich einem gesellschaftlichen Wandel unterliegenden Konstruktion und Logik der benutzten Thesauri als den zentralen Arbeitsinstrumenten bei der Abbildung und Sammlung des Weltenwissens und seiner Erschließung sowohl für eine allgemeine als auch für eine Fachöffentlichkeit – über nunmehr fast ein ganzes Jahrhundert. Diese 17 oder 20 oder wie viel Millionen Seiten oder Dokumente liegen in den Magazinen des HWWA und der Kieler ZBW⁸ teils originaliter, zum größeren Teil aber als seit den 1960er Jahren erstmals angefertigte Sicherungsmedien wie Microfiche oder Rollfilm (bei gleichzeitiger Kassation der Originalvorlagen) vor und werden seit Juni 2004 im Rahmen eines DFG-geförderten Projektes digitalisiert. Wie heutzutage möglich und deshalb üblich, werden dabei zugleich

8 In welcher das Wirtschaftsarchiv des Instituts für Weltwirtschaft zurzeit ebenso wie das HWWA spätestens zum 01.01.2007 im Rahmen einer Fusion aufgehen.

material- und dokumentspezifische Metadaten erhoben, die mit den Digitalisaten gekoppelt und in einer gemeinsamen Datenbank über das Internet recherchierbar verwahrt und so der Öffentlichkeit zugänglich gemacht werden.

In diesem Projekt, das Pilotcharakter für die Digitalisierung des Gesamtbestandes besitzt⁹, wird zunächst nur der Zeitraum bis in die 1930er Jahre bearbeitet, womit rund drei Millionen Dokumente im Umfang von etwa 4,8 Millionen Digitalisaten zur Verarbeitung anstehen (nach der Erschließung entspricht dies wenigstens neun Millionen digitalen Bildern). Um diese Massen bewältigen zu können, werden im Falle der originalen Themenmappen, die jedoch mit rund 120.000 Dokumenten im Umfang von etwa 400.000 digitalen Bildern den kleineren Teil ausmachen, die zu digitalisierenden Materialien vor der Vergabe an einen Scandienstleister seitenweise mit formalen und bedingt auch inhaltlichen Erfassungs- wie Erschließungsdaten versehen. Diese werden seitens des Dienstleisters in den Dateinamen der digitalen Bilder übernommen, womit sich schließlich die angesprochene internetfähig recherchierbare Datenbank füttern lässt. Ein Retrieval über Formalerfassungsdaten unter Beigabe weiterer institutionen- wie personen- und thesaurusspezifischer Stammdaten ist sofort möglich. Die Sammlung und Pflege dieser Stammdaten wurde sowohl in Kiel als auch in Hamburg von Anfang an intensiv betrieben.

Bei der Digitalisierung von Sicherungsmedien wie Microfiches oder Rollfilmen ist dieses Verfahren nicht möglich, da sich die Dokumente auf diesen Medien nicht in der beschriebenen Weise deklarieren und auszeichnen lassen. Digitale Bilder von solchen Medien können also bestenfalls mit der Film- oder Fiche-Signatur numerus currens angesprochen werden. Selbst ein Retrieval über Formalerfassungsdaten ist so kaum sinnvoll

9 Kritisch setzen sich mit dem Projekt auseinander: Buchholtz, Rüdiger, *Retro-Digitalisierung einer Pressedokumentation am Beispiel des HWWA. Konzepte und Probleme*, Hamburg (Master-Arbeit am Institut für Bibliothekswissenschaft, Bereich Fernstudium, der Humboldt-Universität zu Berlin) 2005 und Jeude, Kirsten, *Besondere Herausforderungen bei der retrospektiven Digitalisierung historischen Pressematerials am Beispiel eines Projekts des Hamburgischen Welt-Wirtschafts-Archivs*, Hamburg (Abschlussarbeit im thematischen Schwerpunkt Informations-Management) 2005.

möglich. Dies ist zugleich der Hauptgrund dafür, dass diese Medien von den Digitalisierungsbestrebungen in europäischen Archiven und Bibliotheken weitgehend ausgeschlossen wurden. Das aber könnte sich künftig ändern, denn im Rahmen des Hamburgischen Retrodigitalisierungsprojektes wurde neben dem aus einem handelsüblichen Presse-Clipping-Software-Modul bestehenden Erfassungs- und Erschließungsinstrument für Digitalisate, die von Papiervorlagen erzeugt wurden, parallel ein weiteres Instrument konzipiert, das sich besonders für die Erfassung und Erschließung von Digitalisaten eignet, die auf der Basis von Sicherungsmedien wie Microfiches und Rollfilmen produziert wurden. Nachfolgend werden die Besonderheiten am Beispiel der Inventarisierung von Rollfilm-Digitalisaten dargestellt, während auf die Präsentation von mittlerweile bereits etablierten Strategien und Techniken zur Erschließung von Digitalisaten papierner Provenienz verzichtet werden kann.

Technische Umsetzung der Digitalisierung

Sicherungsmedien wie Rollfilm und Mikrofiche dienen als Mittel der Langzeitarchivierung. Zu den Vorteilen, wie ihrer Reproduzierbarkeit sowie die Platzersparnis an laufenden Regalmetern, sind auch ihre Nachteile zu nennen: vor allem der sequentielle Zugriff. Zur Langzeitarchivierung gehören die entsprechenden Schreib- und Lesegeräte. Durch den technischen Fortschritt und entsprechende Nachfragesteuerung mangelt es zunehmend an im Publikumsverkehr einsetzbaren Lesegeräten. Was nützt ein reproduzierbares Medium, wenn die dazu erforderlichen Geräte fehlen oder nicht genutzt werden? Berücksichtigt man ferner noch die durch die Internetnutzung gewandelte Kundenerwartung, so kristallisiert sich deutlich die Anforderung heraus: Die Sicherungsmedien und verbleibenden Originalarchivalien sind in eine Form zu überführen, die sowohl einen direkten Zugriff als auch künftige Reproduzierbarkeit sowie ortsungebundene Recherche gestatten. Dem Hamburger Retrodigitalisierungsprojekt liegen – neben einem erheblich kleineren Papierarchiv – 2.758 Rollfilme mit jeweils circa 1.500 A3-Bildern zugrunde. Das Papierarchiv umfasst den Themenbereich Personen, die Rollfilme bilden das Institutionen-, Sach- und Warenarchiv ab. Die Ordnung erfolgte nach lektorierten Namens- und

Systemstellen. Das Sammlungsgut ist heterogen, die Dokumentarten reichen von der aufgeklebten Pressemeldung auf einer A4-Seite über mehrseitige Aufsätze bis zu über einhundert und mehr Seiten starken Geschäftsberichten oder Festschriften. Die Presseartikel in mehreren Sprachen entstammen hunderten von Quellen aus unterschiedlichen Zeiträumen mit verschiedenen Fonts. Die Vorlagen liegen häufig nicht einheitlich justiert, sondern schräg bis quer, Übergrößen können abschnittsweise über mehrere Seiten verteilt sein.

Zunächst musste die Frage beantwortet werden, wie möglichst schnell eine möglichst große Zahl an Dokumenten ortsungebunden präsentiert werden kann. Dabei wurde deutlich, dass hierfür auf die alte Mappenstruktur zurückgegriffen werden muss. Es wurde also ein Index gebildet, der die alten Mappenüberschriften repräsentiert. Die Disparität der Vorlagen gestattet keine einfache Volltexterstellung. Der Volltext ist auch bei bestem Ergebnis zu fehlerbehaftet, als dass er die Bildvorlage ersetzen könnte. Die Mappen werden also Bilddateien enthalten. Gleichwohl wurden in die Erfassungsinstrumente von Anbeginn an OCR-Module integriert. Obschon sie lauffähig sind, werden sie momentan aufgrund der problematischen Texterkennungsquoten nicht standardmäßig eingesetzt, könnten dies jedoch im Falle verbesserter Texterkennungsraten. Dieser Weg ist also nicht verbaut, sondern wird aus genannten Gründen momentan lediglich nicht beschränkt.

Eine weitere Frage befasste sich damit, wie viele Informationen der Bilddatei übergeben werden müssen, um sie eindeutig in der Mappenstruktur zu verorten. Das war die Idee zum Rollfilmslider.¹⁰ Das Prinzip des Sliders sieht vor, am lokalen PC einen Rollfilm wie an einem Lesegerät durchzumustern, wobei im Gegensatz zum Lesegerät aber jedem Bild eindeutige Merkmale zugeordnet werden können. Dazu werden alle 1.500 A3-Bilder eines Rollfilms ausgelesen, zu einer großen Bitmap zusammen-

¹⁰ Darstellung der Funktionsweise dieser Software siehe auch in: Huck, Thomas Sergej; Wannags, Max-Michael, Inhaltliche und technische Darstellung der retrospektiven Digitalisierung der historischen Pressearchive der ZBW und des HWWA, in: IIE aktuell. Mitteilungen aus der Abteilung Überregionale Bibliographische Dienste, Nr. 29, Januar 2006, S. 7-12.

gestellt, die wiederum Einzelbilder auf dem PC-Bildschirm anzeigt. Eine zur Bitmap korrespondierende Binärdatei verzeichnet zur späteren Auswertung pixelgenau alle Änderungen. Aus Sicherheitsgründen, um einem Datenverlust auf dem lokalen PC vorzubeugen, wird diese Binärdatei auf einem zentralen Server gespeichert.

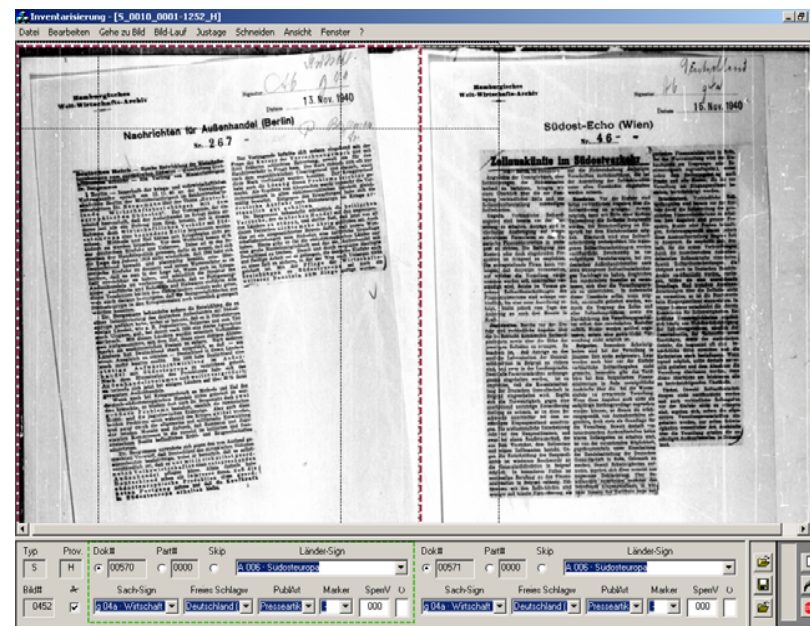


Abbildung 1: Ansicht der Rollfilmslider-Oberfläche

Bei Durchsicht der Bilder liest das Bearbeitungspersonal zuerst die zentralen Ordnungsmerkmale der Mappe aus dem angezeigten Bild aus und ordnet sie einer Auswahlliste im Editierbereich des Sliders zu. Dieses Merkmal wird jetzt fortlaufend automatisch den nachfolgenden Bildern zugeordnet, bis aufgrund eines Themenwechsels eine erneute manuelle Zuordnung erfolgt. Gleichzeitig nimmt bei jedem Bildwechsel ein Zähler um einen Betrag zu, wahlweise umstellbar auf Dokument- bzw. Seitenzählung. Eine wesentliche Weiterentwicklung des Sliders bestand in dem

Hinzufügen von Schneidrechtecken. Das Abbild des Rollfilmbildes auf dem PC-Monitor zeigt in der Regel zwei A4-Seiten. Innerhalb dieses Abbildes kann ein Thema zwischen den beiden A4-Seiten wechseln. Es erwies sich daher als zweckmäßig, von vornherein auf eine A4-Seite als kleinste Einheit zu setzen und die Editierfelder des Sliders paarig auszulegen. Mit den skalierbaren Schneidrechtecken, die standardmäßig einen A4-Rahmen abdecken, werden die Seiten der Vorlage auf dem Bitmap zum Schneiden markiert. Auch hier gilt das Prinzip, dass die Positionen der Rechtecke bei den Folgebildern solange erhalten bleiben, bis sie manuell geändert werden. Bei Übergrößen, bei denen ein Dokument das gesamte Rollfilmbild beansprucht, können die Schneidrechtecke abgewählt und ein Editierfeld ausgeblendet werden.

Der größte Durchsatz beim Bildersichten wird erreicht, wenn keine individuellen Einträge erfolgen, sondern nur geprüft werden muss, ob ein neues Mappenkriterium vorliegt oder die folgenden Bilder Seiten eines übergeordneten Dokumentes darstellen. Trotz dieser Einschränkung erfordert die Arbeit mit dem Slider hohe Konzentration. Allein die HWWA-Mappennotation kann bis zu drei Merkmale enthalten, die abhängig von dem zugrunde liegenden Archiv wechseln können und deshalb einzeln verwaltet werden müssen. Hinzu kommen anklickbare Indikatoren, ob und bis wann eine urheberrechtsbedingte Sperrung¹¹ erforderlich ist und um

11 Nach dem Scannen und Aufspielen auf den Projekt-Server sind unmittelbar jene Materialien verfügbar, die als urheberrechtlich unbedenklich gelten. Dies sind anonyme und pseudonyme Werke, deren Schutzfrist bereits siebenzig Jahre nach der Veröffentlichung abgelaufen ist sowie mit einem Urheber/einer Urheberin bezeichnete Werke, deren urheberrechtlicher Schutz abgelaufen ist, solche also, die von Autoren/innen verfasst sind, die länger als 70 Jahre (plus maximal 364 Tage) verstorben sind (vgl. §§ 64, 69 UrhG). Zudem werden nach dem Prinzip einer „moving wall“ jährlich zusätzlich jene Werke online verfügbar gemacht, die oben genannte Bedingungen des UrhG erfüllen, um schutzfrei öffentlich wiedergegeben zu werden. Umgesetzt wird das Prinzip der „moving wall“ mittels eines Sperrvermerks, der an jedem Dokument angebracht wird. Es werden drei Arten von Sperrvermerken unterschieden: 1. Das Werk ist schutzfrei. 2. Das Werk ist noch geschützt plus Todesjahr des Autors/der Autorin sowie 3. das Werk ist geschützt, das Todesjahr des Autors/der Autorin ist nicht zu ermitteln. Dokumente der ersten Gruppe sind unmittelbar online verfügbar. Dokumente der zweiten Gruppe werden erst in dem Jahr verfügbar, in dem die Schutzfrist abgelaufen ist.

welche Dokumentart es sich handelt. Es können durch Ausblenden („Skippen“) beider Editierfelder auch Vorlagen übergangen werden. Um Platz zu sparen wurden manchmal beim Verfilmen der Originale doppelseitig beklebte Vorlagen parallel belichtet. Die Leserichtung läuft dann nicht von links nach rechts, sondern von oben nach unten. Weil dieser Fall gar nicht so selten auftritt, musste für den Slider ein Verfahren entwickelt werden, das die Zählung von parallelen Streifen erlaubt. Das Softwarekonzept erwies sich als flexibel genug, noch weitere Abweichungen der Leserichtung zu integrieren. Oft wurden auch einzelne Zeitungsspalten als Überlängen quer über das A3-Format verfilmt. Da die gängigen Internet-Browser selten die Möglichkeit einer Bildrotation bieten, und den Betrachtern/innen eine unbequeme Kopfhaltung erspart bleiben sollte, erhielt der Slider die Möglichkeit, das aktuell markierte Rechteck mit einer Rotationsanweisung von dreimal jeweils 90 Grad zu versehen.

Die Mappennotation ist das bestimmende Merkmal der Pressearchive. Durch diese Klammer wurden die vielfältigen Konvolute fast einhundert Jahre zusammengehalten und strukturiert.¹² Deshalb wird ihr im Rahmen des Projektes gesonderte Pflege zuteil.

Bei Personen ist es der Name, der, verbunden mit den Lebensdaten und dem Wirkungsfeld, einen eindeutigen Schlüssel darstellt. Aufwändiger wird es bei den Institutionen, die nicht unter ihrem Namen, sondern unter einer Notation abgelegt wurden, um Namenswechsel, Rechtsformänderungen, gegebenenfalls auch Tochtergesellschaften in einer Mappe zu gruppieren.

Dies wird vom System anhand des eingegeben Todesjahres automatisch mittels eines Prüfalgorithmus berechnet. Die Dokumente werden automatisch zum 1. Januar des entsprechenden Jahres freigegeben. Dokumente der dritten Gruppe können erst dann freigegeben werden, wenn der urheberrechtliche Status ermittelt wurde bzw. soviel Zeit vergangen ist, dass man sicher davon ausgehen kann, dass der Urheber/die Urheberin länger als 70 Jahre tot ist. Bis dahin bleiben sie gesperrt und können nur vor Ort im HWWA oder in der ZBW eingesehen werden. Ein Verweis auf diese Dokumente ist in der Datenbank enthalten.

12 Zu dem (unter dem Begriff „Hamburger System“ bekannt gewordenen) Ordnungs- und Kategoriensystem des HWWA: Bench, Bernhard, Formale Erfassung – Ausgangspunkt eines jeden Informationssystems, in: Auskunft. Mitteilungsblatt Hamburger Bibliotheken 3 (1983), S. 213-221 und Kreuzfeldt, Hans; Scherwath, Wolfgang, Inhaltliche Erschließung wirtschaftsrelevanter Literatur, in: ebd., S. 221-238.

ren. Die Sachnotation weist manche historischen Brüche auf und bildet nur in Verbindung mit einer Ländernotation einen Schlüssel. Für die Warengruppen war ursprünglich gar keine Notation sondern nur eine Schlagwortliste vorgesehen. Das Alphabet bildete in diesem Fall ein willkürliches Ordnungssystem, ist aber, wie beim Sacharchiv, durch Kombination mit einer Ländernotation doch noch zu einer eindeutigen Systematik fortentwickelt worden. Um diese vielfältigen Informationen zusammenzuhalten und zu bearbeiten, wurde für jede Mappennotation ein eigener Normdatensatz mit eindeutigem Zahlenschlüssel angelegt, der schließlich in den Dateinamen der neu generierten Bilddatei übernommen wird. Mit dem Normdatensatz ergeben sich dokumentunabhängige Möglichkeiten, die inhaltliche Erschließung des Mappeninhalts mit erläuternden Hinweisen und neuen Relationen angemessen zu ergänzen oder auch Themencluster, beispielsweise zu aktuellen Anlässen, zu bilden.

The screenshot shows a web application interface for a person's profile. The main area displays details for 'Ballin, Albert' (1857-1918). The details include:

- Beruf: Reederei, HAPAG-Vorstand (1888), Generaldirektor (1896)
- Institution: Hamburg Amerikanische Packetfahrt Actien Gesellschaft (Hamburg)
- Wirkungsbereich: 2DE Deutschland Wirtschaft
- Nationalität: 2DE Deutschland
- Pressemappe: 1910-
- Web-Links: de.wikipedia.org/wiki/Albert_Ballin
- Identnummern: 129654 IFS: 952
- über die Person: 163 Grafische Textdateien (davon 50 gesperrt)

 On the left, there is a search bar and a list of related persons, including:

- Baerentzen, Peter, 1830-1927
- Baeyer, Adolf von, 1835-1917
- Baehrot, Walter, 1826-1877
- Bahr, Hermann, 1863-1934
- Bahr, Abe, 1864-1940
- Baker, George F., 1840-1931
- Balabanoff, Angelica, 1878-1965
- Balbo, Italo, 1896-1940
- Baldern, Oliver, 1899-1958
- Baldern, Stanley Baldwin, 1867-1947
- Balfour, Arthur James, 1848-1930
- Ballin, Albert, 1857-1918
- Bamberger, Ludwig, 1823-1899
- Bann, Nina Ellinger, 1866-1928
- Bann, Ewald, 1883-1953
- Baraber, Albert Ernest Augustin, 1864-1917
- Barboza, Roy, 1849-1923
- Barbours, Henry, 1873-1935
- Barry, Nicolaus, 1871-1935
- Barthelet, Edouard, 1901-1979
- Barth, Ernst, 1870-1938
- Barthel, Julius, 7-1938

 At the bottom, there is a table of image files:

0002 (21 S.)	JPEG
0003 (2 S.)	JPEG
0004 (3 S.)	JPEG
0005 (2 S.)	JPEG
0006	JPEG
0007	JPEG
0008 (3 S.)	JPEG
0009	JPEG
0010	JPEG
0011	JPEG
0012	JPEG

 The interface is supported by DFG (Deutsche Forschungsgemeinschaft).

Abbildung 2: Einbindung der Normdaten in die Präsentation der Personenmappe

Soweit wie möglich wurden bereits standardisierte Normsätze aus dem Bibliotheksbereich übernommen oder Normsätze weitgehend nach deren Regeln angepasst. Dies bietet künftig die Möglichkeit über den Normsatz und dessen Schlüssel zu anderen öffentlichen Datensammlungen zu verlinken, wie das bereits vorbildhaft in einem Projekt der Wikipedia mit der Deutschen Bibliothek realisiert wurde, wobei Personeneinträge der Wikipedia auf Katalogeinträge der DDB verweisen und demnächst auch auf den Bestand Personenarchiv von HWWA und ZBW. Die Normdatensätze werden zentral gepflegt und daraus periodisch Schlüssel Listen generiert. Diese werden beim Aufruf der lokalen Slidersoftware automatisch geladen und stehen als Auswahllisten für den Inventarisierungsvorgang zur Verfügung. Beim Auftreten bislang undokumentierter Merkmale aus der Vorlage kann so relativ schnell ein neuer Schlüssel generiert und verteilt werden. Wenn die Bearbeitung eines Rollfilms abgeschlossen ist und die Plausibilitätsprüfung keine fehlenden Einträge mehr aufdeckt, können auf dem lokalen PC die Bilder aus den Rollfilmdateien entsprechend den gespeicherten Koordinaten gedreht, geschnitten und als eigene Dateien generiert werden. Anschließend werden sie auf einen zentralen Server verschoben.

Auf einem Rollfilm lässt sich kein direkter Zugriff auf ein gesuchtes Objekt realisieren. Es müssen jedes Mal alle Merkmale des Ordnungssystems und damit alle Dokumente nacheinander durchsucht werden um zu einem möglichen Treffer zu gelangen. Dagegen enthält jedes Bild des geschnittenen Filmes im Dateinamen alle Merkmale zur präzisen Bestimmung. Die Datei kann in einer zentralen Verzeichnisstruktur unter einem eindeutigen Mappenschlüssel gruppiert werden. Überwiegend ist eine regionale Beschreibung im Dateinamen beigefügt, manchmal noch ein kennzeichnendes Schlagwort. Es ist aus den Merkmalen lokalisierbar, von welcher Art das abgebildete Dokument ist, ob es einzeln oder als Teil vorliegt.

Zusammenführung aller erschlossenen Metadaten

Zusätzlich herauszuheben ist, dass weder die mit dem Rollfilmslider erzeugten Metadaten noch die digitalen Bilder in einer spezifischen Daten-

bank-Software dauerhaft archiviert werden müssen. Sie müssen lediglich auf Festplatten in einer vorab nach den Erfordernissen der Formalerfassung festgelegten Hierarchie gespeichert werden. In dieser Form können sie in ein Langzeitsicherungskonzept mittels magneto-optischen Speichermedien integriert werden. Möglich wird dies durch eine einfache Namenskongruenz zwischen der Metadatenverwaltungsdatei im XML-Format einerseits und den Namen der digitalen Bilddatei andererseits. Zugleich entspricht der Aufbau der Metadaten-datei weitgehend dem Dublin-Core-Standard¹³, sodass außerdem auch die Richtlinien für die internationale Austauschbarkeit dieses elementaren Datentyps erfüllt werden. Dies gilt schlussendlich auch für das digitale Bildformat selbst, für das allerdings – anders als die herrschende Lehre es empfiehlt – nicht das Format TIF sondern JPG gewählt werden musste, da anderenfalls unvertretbare Datenmassen entstanden wären.¹⁴ Dieselben, eben dargelegten Konventionen gelten natürlich auch für die Arbeit mit den Digitalisaten auf Basis originaler Papiervorlagen, die mit der erwähnten Presse-Clipping-Software durchgeführt wird. Diese Bilddateien lagern freilich temporär in dieser Spezialdatenbank. Doch dies lediglich bis zur Fertigstellung der zugehörigen Metadatenätze. Danach werden die Bilder wie die korrespondierenden Metadatenätze ebenfalls exportiert.

Somit liegt für sämtliche Digitalisate formal eine einheitliche Datenorganisation vor und inhaltlich eine gleichartige Erfassungs- und Erschlie-

- 13 Andresen, Leif, Standardisation of Dublin Core in Europe, in: Zeitschrift für Bibliothekswesen und Bibliographie Heft 1, 47 (2000), S. 39-45 sowie (für das HWWA-ZBW-Retro-Projekt von besonderem Interesse) Weber, Jutta, Nachlässe und Autographen im WWW. Dublin Core in Museen, Archiven und Bibliotheken, in: ebd., S. 63-69.
- 14 Allerdings wurde die bei dem JPG-Format erforderliche Parametrisierung im Hinblick auf die Faktoren Farbraum, Komprimierung und Qualität ausgiebig in Bezug auf alle möglichen Ausgabemedien getestet; was den (Offset-) Druck angeht, sind die erforderlichen Parameter schließlich empirisch durch das Erstellen verschiedener Proofs ermittelt worden. In diesem Zusammenhang vgl. auch den Aufsatz von Pollmeier, Klaus, Vom Baryt zum Bit. Zur Konservierung analoger und digitaler fotografischer Bilder, in: Digitale Bildverarbeitung, eine Erweiterung oder radikale Veränderung der Fotografie? Dokumentation des Symposiums am 12. und 13.11.2004 im Museum Volkswang Essen, Stuttgart 2005, S. 56-63.

ungsstruktur, was jetzt schon über die Homepage des Projektes eingesehen werden kann:









Digitalisierung der Archive der Pressedokumentationen von HWWA und ZBW				
Zu den Dossiers				Über die Archive
 Personen	Im Rahmen des DFG-Projektes "Retrospektive Digitalisierung von historischen Presseartikeln der Archive des HWWA und der ZBW (vormals Wirtschaftsarchiv des IFW)" wird das historische Material der Pressedokumentationen beider Institute von der Zeit ihrer Gründung bis in die Mitte der 1930er Jahre digitalisiert.			 Wirtschaftsarchiv
 Firmen	Hierbei handelt es sich vorwiegend um Presseauschnitte und Aufsätze zu einzelnen Personen und zu einzelnen Waren und Warengruppen, sowie um Artikel, Geschäftsberichte und Festschriften zu Firmen und anderen Körperschaften, ferner um Ausschnitte aus nationalen und internationalen Zeitungen und Zeitschriften zu Sachthemen aus dem Bereich der gesamten Weltwirtschaft.			 HWWA HAMBURG Pressedokumentation
 Warengruppen	Diese Materialien sind zur Zeit nur vor Ort zu benutzen. Um den Zugriff für die Benutzer zu erleichtern, werden diese Materialien nun digitalisiert, aufbereitet und über diese Datenbank im Internet zur Verfügung gestellt.			 Forschungsstelle für die Ostsee-Deutschen
 Sachgruppen	Insgesamt handelt es sich um ca. 4,8 Millionen Presseartikel aus dem oben genannten Zeitraum. Bei Materialwünschen jüngerer Datums erfolgt der Zugang wie bisher. Es ist jedoch geplant, in späteren Projekten auch dieses Material digital zugänglich zu machen.			In eigener Sache
	Veröffentlichungen	Gefördert durch 	Impressum	Stand: 16.08.2006

Abbildung 3: Projekthomepage¹⁵

Zusammenfassung

Derzeit werden in einem DFG-geförderten Projekt Altbestände der Archive der Pressedokumentation des Hamburgischen Welt-Wirtschafts-Archivs und der Zentralbibliothek für Wirtschaftswissenschaften in Kiel (mit den übernommenen Beständen des Wirtschaftsarchivs des Instituts für Weltwirtschaft) aus der Zeit von etwa 1826 bis Mitte der 1930er Jahre digitalisiert, über das Internet sichtbar gemacht und parallel dazu sukzessive die Dokumente einzeln formal erfasst und inhaltlich erschlossen. Bei den einzelnen Beständen handelt es sich im Wesentlichen um diverse Zei-

¹⁵ Vgl. <<http://webopac.hwwa.de/digiview/>>.

tungsausschnittssammlungen zu den verschiedensten Themengebieten (Personen, Institutionen, Sachthemen, Waren, Produkte usw.), aber auch um weiteres Material wie Geschäftsberichte, Aufsätze, Festschriften, Korrespondenzen usw., die für den aktuellen Digitalisierungszeitraum aus rund drei Millionen meist mehrseitigen Einzeldokumenten bestehen. Gegenwärtig dürfte kein weiteres Retrodigitalisierungsprojekt existieren, das derartig ausgesprochen heterogenes Archivgut von teils fragilem (Erhaltungs-)Zustand in dieser großen Masse verarbeitet. Vor diesem Hintergrund werden Strategien der Organisation, Erschließung und Präsentation der anfallenden Digitalisate mitsamt zugehörigen Meta-Daten sowie Aspekte und Perspektiven zum Problem der Datensicherung (Langzeitarchivierung) erörtert und eine spezielle Software für die Erfassung und Erschließung von Digitalisaten vorgestellt, die auf der Grundlage von Sicherungsmedien wie Rollfilm oder Mikrofiche erzeugt wurden.

Langfristig können mit den Methoden und Arbeitsgängen des gegenwärtigen Projektes die gesamten Archivbestände der Pressedokumentationen beider Institute bearbeitet werden, womit einerseits auf Seiten der Nutzer/innen ein bequemer und inhaltlich in neue Tiefen reichender Zugang geschaffen wäre und andererseits mit der dann möglichen Zurückhaltung des Altmaterials, das nur noch in Sonderfällen dem Magazin entnommen werden müsste, ein gewisser Beitrag zu seiner Erhaltung geleistet würde, sofern dieses Material bzw. die von den Originalen erzeugten Sicherungsmedien wie Film und Fiche adäquat gelagert werden. Nicht zu übersehen freilich und deshalb an dieser Stelle gesondert zu erwähnen ist schließlich das mit der Digitalisierung überhaupt erst entstandene Problem der Langzeitarchivierung digitaler Daten, das gegenwärtig weder in dem HWWA-ZBW-Retrodigitalisierungsprojekt noch in irgendeinem anderen vergleichbaren Projekt als wirklich abschließend gelöst betrachtet werden kann.

Dr. Thomas S. Huck ist Projektmanager für Retrodigitalisierung am Hamburgischen Welt-Wirtschafts-Archiv. E-Mail: thuck@t-online.de

Dipl. Bibl. Dipl. Ing. Max-Michael Wannags ist Leiter des Bereichs Informationssysteme am Hamburgischen Welt-Wirtschafts-Archiv. E-Mail: wannags@hwwa.de

DER VIRTUELLE LESESAAAL
DES ARCHIVS FÜR ZEITGESCHICHTE (ETH ZÜRICH) –
ONLINE-PRÄSENTATION UND RETRIEVAL VON ARCHIVGUT

von Daniel Nerlich

AfZ Online Archives ist ein Webfrontend mit integrierter Retrievaltechnologie, das Erschließungsinformationen sowie Text-, Bild- und Tonquellen des Archivs für Zeitgeschichte der ETH Zürich in digitaler Form vermittelt. Die Kombination von probabilistischer Volltextrecherche und strukturgestützter Suche ermöglicht Anwendern/innen eine effiziente Doppelstrategie. Per Mausclick werden Treffer in derselben Ansicht mit der Bestandestektonik synchronisiert angezeigt, wo direkt weitergeblättert werden kann. Informations anbietende Stellen können mit dem Einsatz des Retrieval System neben der Komfortsteigerung für die Benutzung auch arbeitsökonomische Gewinne erzielen. Erschließungsinformationen müssen nicht in einem zusätzlichen Arbeitsprozess für die Darstellung auf dem Internet aufbereitet werden. Sie lassen sich on the fly und damit auf dem aktuellsten Stand und aus unterschiedlichen Datenkollektionen generieren.

„Was können Archive mit dem Internet anfangen?“¹ Wer heute diese von Karsten Uhde 1996 im Mitteilungsblatt des deutschen Archivwesens publizierte Frage wiederholt, provoziert eher Mitleidsbekundungen als eine ernsthafte Diskussion. Dass dies noch vor zehn Jahren völlig anders war, belegt die Tatsache, dass damals im deutschsprachigen Raum kaum ein Archiv das neue Medium nutzte. Uhde präsentierte vor diesem Hintergrund

1 Uhde, Karsten, Archive und Internet, in: Der Archivar 49 (1996), Heft 2, S. 205-216.

ein Dreistufenmodell, wonach Archive zunächst Grundinformationen über die Institution, in einem zweiten Schritt Beständeübersichten und in einem dritten schließlich Detailinformationen im Internet anbieten könnten. Während er die Stufen eins und zwei klar befürwortete, sollte die Präsentation von Findmitteln „in einzelnen Archiven exemplarisch erprobt und auf der Grundlage dieser Ergebnisse gründlich diskutiert werden“.² AfZ Online Archives³, der virtuelle Lesesaal des Archivs für Zeitgeschichte (AfZ) der ETH Zürich⁴, unterstreicht heute ein konsequentes Ja auch zu diesem dritten Schritt der Darstellung im Netz. Das Webfrontend mit integrierter Retrievaltechnologie, das Metadaten und digitale Reproduktionen von Text-, Bild- und Tonquellen vermittelt, versteht sich nämlich als Basis und Service für die historische Forschung, deren Informationsnachfrage sich in den vergangenen Jahren längst irreversibel und weit stärker ins Internet verlagert hat als das Angebot der Informationsproduzenten/innen.

Das AfZ ist ein Dokumentations- und Forschungszentrum zur schweizerischen Zeitgeschichte im europäischen und globalen Kontext. Sein Engagement gilt der Sicherung bedeutender privater institutioneller Quellenbestände und Nachlässe in der Schweiz.⁵ Als Spezialarchiv fördert es mit seinen Dienstleistungen die historische Forschung und legt damit besonderen Wert auf einen hohen Erschließungsgrad und die Zugänglichkeit seiner Bestände. Dadurch war es schon früh motiviert, im Sinne Uhdes in der Schweiz eine Vorreiterrolle bei der Findmittelpräsentation im Internet zu übernehmen. Dazu gehört auch die Erweiterung des Quellenangebots durch größere Pilotprojekte, welche das AfZ seit 1997 unter Einsatz eines hybriden Dokumentenmanagements durchführt. Neben Sicherungs-, Schutz- oder Ersatzverfilmungen sieht dieses Verfahren die Digitalisierung von Originalen mit dem Ziel der Verfügbarkeit der Einzeldokumente als elektronische Kopien vor. Ausgesuchte Quellenmaterialien werden so nicht

2 Ebd., S. 216.

3 Vgl. <<http://onlinearchives.ethz.ch>> (26.09.2006).

4 Vgl. <<http://www.afz.ethz.ch>> (26.09.2006).

5 Zu den Sammlungsgebieten des Archivs für Zeitgeschichte siehe <<http://www.afz.ethz.ch/fsdokustellen.html>> (26.09.2006).

mehr nur indirekt über Metainformationen, sondern direkt zugänglich, was den Nutzungskomfort entscheidend erhöht. Zu den Highlights bereits digitalisierter Bestände gehören rund 500.000 Presseartikel zur Schweizer Wirtschaftspolitik der Jahre 1943 bis 1974, Tondokumente zu rund 140 Kolloquien mit Zeitzeugen/innen seit 1973 sowie eine Plakatsammlung, welche die wirtschaftspolitischen Abstimmungskämpfe der Nachkriegszeit aus Sicht der Schweizer Wirtschaft illustriert. Im Folgenden werden zunächst die Bedeutung dieser drei exemplarischen Bestände für die Forschung zur schweizerischen Zeitgeschichte und die Erfahrungen mit deren Digitalisierung als Grundlage für die Internetaufbereitung umrissen. Im Sinne eines ersten Werkstattberichts werden anschließend die Features der Präsentation im Rahmen von AfZ Online Archives vorgestellt, dessen Angebot in den nächsten Jahren kontinuierlich ausgebaut wird.

Die Gesellschaft zur Förderung der schweizerischen Wirtschaft (wf) wurde 1942 mit dem Zweck gegründet, das staatliche Vollmachtenregime der Kriegszeit abzubauen und die Interessen der Privatwirtschaft im Sinne einer marktwirtschaftlichen Ordnung zu vertreten. In der tagespolitischen Lobbyarbeit stellte die wf im Bereich der Public Relations Referenten/innen für Vorträge oder Radio- und Fernsehsendungen und gab eigene Publikationen wie die monatlich erschienenen Wirtschaftspolitischen Mitteilungen oder Pressequerschnitte (Für Sie gelesen) heraus. Die Gesellschaft hielt dabei engen Kontakt zu Bundesparlamentariern/innen sowie Verbands- und Parteienvertretern/innen, die sie dokumentierte und beriet, wofür ihr Dokumentationsarchiv die Basis bildete. In wirtschaftsrelevanten Abstimmungskampagnen war die wf schließlich für Konzeption, Organisation und Finanzierung verantwortlich. Im Jahr 2003 gab die Gesellschaft ihre Selbstständigkeit auf und fusionierte mit dem Schweizerischen Handels- und Industrie-Verein zum neuen Dachverband namens *economiesuisse*.

Die historische Zeitungsausschnittsammlung der Zürcher Hauptgeschäftsstelle der wf ist eine der größten und wichtigsten Dokumentationen zu Politik und Wirtschaft in der Schweiz aus dem Zeitraum 1943 bis 1993. Sie beruht auf der Auswertung von verschiedenen Pressediensten, bis zu 100 Tages- und Wochenblättern, von Fachpublikationen und wirtschaftspolitischen Informationsschriften. Im AfZ wurde diese Sammlung nach

historischen Kriterien systematisiert, wobei rund 3.500 Sachdossiers entstanden. Für die Ersatzverfilmung des ersten Teils der Zeitungsausschnitte bis ins Jahr 1974 im Umfang von rund 1.100 Archivschachteln wurde eine 50-fache Verkleinerung auf einem 16mm-Film gewählt, einem Speichermedium, dessen lange Haltbarkeit geprüft und bewährt ist. Mit einem Kodak Digital Science Scanner / Microimager 990 wurde für die Verarbeitung ein Gerät eingesetzt, welches neben dem Film für die Komfortsteigerung von Recherche und Reproduktion TIFF-G4-Dateien produzierte. Erkauft wurde dies mit erheblichem Aufwand bei der Arbeitsvorbereitung, da dem Durchlaufgerät die einzelnen Dokumente ohne Büro- und Bostitch-Klammern zugeführt werden mussten. Dieser handarbeitsintensive Vorgang konnte allerdings mit der ohnehin notwendigen Drehung der chronologischen Ordnung des Materials zusammengelegt werden. Ein Archiven immer höchst willkommener Effekt war zudem die Freilegung von Lagerraumkapazität: Die rund 500.000 Aufnahmen dieses ersten Sammlungsteils beanspruchen statt der ursprünglichen 120 laufenden Meter noch rund 50 GB Speicherkapazität. Die verfolgte hybride Strategie gewährleistet schließlich bei künftigen Migrationsproblemen jederzeit den Rückgriff auf das Sicherheitsnetz der erstellten Mikrofilme – ein Muss angesichts der hier unbestrittenen Unwägbarkeiten digitaler Formatentwicklungen.

Seit 1973 führt das AfZ eine Kolloquienreihe mit Zeugen/innen der Schweizer Zeitgeschichte durch und sichert mündliche Quellen im Sinne der Oral History. Ein eigenentwickeltes Konzept gewährleistet eine gleich bleibende Qualität der Kolloquien, in deren Zentrum jeweils der Rückblick eines oder mehrerer Zeitzeugen/innen steht. Gewünscht wird ein bewusst subjektives Zeugnis: Was war aus der persönlichen Sicht prägend? Welches waren die wichtigsten Stationen, die Hauptprobleme und Lösungsversuche während der beruflichen Tätigkeit? Auf diese Weise entstanden einzigartige, in der Regel inhaltlich konzise Tondokumente, darunter zu inzwischen verstorbenen Zeitzeugen/innen, von denen sonst nur in Ausnahmen mündliche oder schriftliche Erinnerungen erhalten geblieben sind. Der Bestand umfasst über 140 Tonbänder mit Zeitzeugnissen, die von der Forschung und für Ausstellungen nun immer öfter beigezogen werden. Die

Veranstaltungen wurden vom AfZ bis in die jüngste Gegenwart auf Spulentonbändern festgehalten und werden heute digital aufgezeichnet. Um den Anforderungen schneller Verfügbarkeit und gezielter elektronischer Recherche zu genügen, aber auch zu Sicherungszwecken wurde in den vergangenen zwei Jahren die gesamte Sammlung digitalisiert. Analoge Tonbänder weisen erfahrungsgemäß eine gute Haltbarkeit auf, doch können Zerfallsprozesse trotz fachgemäßer Lagerung relativ unvermittelt einsetzen. Im Vordergrund stand – zwecks Schonung der Originale – die Herstellung von mittelfristig haltbaren Sicherungsformaten und Benutzerkopien. Bei Tondokumenten ist die Frage nach digitalen Langzeitformaten noch offener als bei Texten, wo sich gewisse Standards bereits etabliert haben. Als Ergebnis eingehender Evaluationen und Tests wurden die Originalbänder sukzessive auf Digital Audio Tapes (DAT) überspielt und die Tonsignale gleichzeitig als WAV-Audiodateien gespeichert. Von dieser digitalen Kopie werden für die Benutzung und zum beschleunigten Datenaustausch komprimierte MP3-Kopien erstellt. Die vielfältigen Einsatzmöglichkeiten sowie das einfache Handling dieser Digitalisate und nicht zuletzt die Aussicht auf ein Spoken Document Retrieval sind echte Gegenwerte, welche die vorübergehende Bindung von Arbeitskraft für solche Konversionsprozesse in langfristiger Perspektive sicherlich aufwiegen.

Rund 200 Plakate der wf aus der zweiten Hälfte des 20. Jahrhunderts illustrieren die für die Privatwirtschaft geführten Abstimmungskämpfe in der Schweiz, eine Besonderheit des direktdemokratischen politischen Systems. Die großformatigen Dokumente fungierten in der Zeitspanne ihres Aushangs in erster Linie als wirtschaftspolitische Werbeträger. Da ihre Botschaft in einer Bildsprache abgefasst sein musste, die allgemein verständlich war und auf Sehgewohnheiten basierte, machen sie zum einen politische Denk- und Rezeptionsmuster ihrer jeweiligen Entstehungszeit sichtbar. Andererseits sind sie unersetzliche Zeugen der Gesamtentwicklung von Plakatkunst und -grafik in der Schweiz. Der kleine, aber ergiebige Bildquellenbestand wurde nach der Einzelschließung wie das Dokumentationsarchiv der wf hybrid bearbeitet, wobei in diesem Fall aus praktischen und finanziellen Überlegungen die Digitalisierung im Anschluss an die bei Großformaten klar preisgünstigere Mikroverfilmung

erfolgte. Zunächst wurden alle Plakate mit einer Zeuschel OK 200 auf 35mm-Ilford-Farbdias sicherheitsverfilmt. Der Scanvorgang lieferte dann unkomprimierte TIFF-Dateien für die Langzeitarchivierung und für die Benutzung JPG-Formate. Wie im Fall der Tonbänder stieg mit der schnellen Verfügbarkeit von Reproduktionen die Nutzung des Bestandes spürbar an, was zusammen mit der Schonung der teilweise wertvollen Originale auch diesen Digitalisierungsentscheid legitimiert.

Die Digitalisierung von Kollektionen ist aber nur eine erste wichtige Etappe auf dem Weg zu deren Online-Präsentation. Es folgen Herausforderungen vor allem für die Informatikabteilungen, welche die Langzeitarchivierung der anfallenden Dateien und deren Verbindung und kombinierte Anzeige mit den zugehörigen Erschließungsinformationen sicherstellen müssen. Im Vorfeld der Aufschaltung der wf-Dokumentation, die in dieser Frage als Pilotprojekt diente, musste ein Webclient programmiert werden, der es ermöglichte, die auf der Ebene der Metadaten absuchbaren digitalisierten Dossiers aufzurufen, um sie dann bequem „durchblättern“ zu können. Kernstück dieses vom Archiv für Zeitgeschichte entwickelten Systems war der Aufbau dynamischer HTML-Seiten mittels Hypertext Preprocessor (PHP) in Kombination mit MySQL. Ein automatisches Update fügte jeder Verzeichnungseinheit der zentralen Archivdatenbank einen Link an, welcher den Befehl zum Aufruf der Digitalisate dieser Verzeichnungseinheit aus einem Ordner auf dem Dokumenten-Server enthält. Der Befehl wird im PHP-Programmskript des Webservers übernommen, das über eine MySQL-Datenbankabfrage alle im bezeichneten Ordner befindlichen Digitalisate zusammenstellt und zur Anzeige übergibt, womit den User/innen schließlich eine Gesamtschau aller Dokumente der digitalisierten Dossiers angeboten wird. Dieses für die Anzeige von Textdokumenten konzipierte „Kern-DMS“ wurde dann für die Anzeige von Media-Assets (Fotos, Tondokumente und Filme) weiterentwickelt, womit die Grundlagen für den korrekten Aufruf digitalisierten Archivguts geschaffen waren.

Um nicht nur die Darstellung, sondern auch die in diesem Stadium noch unbefriedigende Datenabfrage zu optimieren, wurde zusammen mit Retrievalspezialisten/innen der Firma xylix.software (Zürich) zwischen

Oktober 2003 und Ende 2005 AfZ Online Archives realisiert.⁶ Die konsequent auf XML und einem integrierten Information Retrieval System basierende Entwicklung bezweckt die Anzeige und effiziente Abfrage sämtlicher archivischen Erschließungsdaten und digitalen Reproduktionen in einem einfach zu bedienenden Webfrontend. Einstiegsrecherchen unterstützt eine Volltextsuche in der Beständeübersicht mit Metadaten zu den über 400 Nachlässen, institutionellen Archiven und weiteren Quellen aus den Sammlungsgebieten des AfZ. Die tiefer gehende probabilistische Datenbanksuche, mit der zur Zeit rund ein Drittel der Bestände auf Stufe der Verzeichnungseinheiten abgefragt werden kann⁷, ist mit zusätzlichen Verfeinerungsmöglichkeiten wie die Eingrenzung von Laufzeiten ausgestattet. Soweit Archivgut bereits in digitaler Form vorliegt und freigegeben ist, kann es über AfZ Online Archives direkt angesteuert und eingesehen werden. Damit wird der virtuelle Katalog- und Lesesaal Realität.

Suchergebnisse des Information Retrieval System, die nach errechneter Relevanz sortiert werden, können darüber hinaus stets in der Bestandes-tektonik angezeigt werden, und zwar nach Belieben auf allen Ebenen der Hierarchie. Diese Kombination von probabilistischer Volltextrecherche und strukturgestützter Suche ermöglicht ihren Anwendern/innen eine effiziente Doppelstrategie. Die synchronisierte Trefferanzeige erfolgt per Mausklick ohne einen Wechsel der Ansicht, womit direkt in der Tektonik weitergeblättert werden kann. Eine Merkliste erlaubt schließlich die Zusammenstellung individueller Suchresultate im Vorfeld von Bestandeskonsultationen. Vorteile bietet das System aber nicht nur für die Benutzung. Einen arbeitsökonomischen Gewinn zieht das AfZ seinerseits daraus, dass Erschließungsinformationen nicht in zusätzlichen Arbeitsprozessen für die Darstellung im Internet aufbereitet werden müssen. Sie lassen sich on the fly und damit auf dem aktuellsten Stand aus der operativen Datenbank generieren. Mit seinen Partnern testet das Archiv bereits heute eine Folgeversion des Systems, welche etwa die Möglichkeit der Sucheingrenzung

6 Zur Konzeption von AfZ Online Archives siehe <<http://www.afz.ethz.ch/fsprojekte.html>> (26.09.2006).

7 Vgl. zum aktuellen Stand der freigeschalteten Bestände den Hilfetext unter <<http://onlinearchives.ethz.ch/>> (26.09.2006).

auf Teilkollektionen und als letzten Schritt die Recherche und Trefferanzeige innerhalb der Digitalisate selbst vorsieht.



Abbildung: Synchronisierte Suche in AfZ Online Archives. Die Trefferanzeige im Kontext der Tektonik ermöglicht das Weitersuchen in der Bestandesstruktur.

Angesichts der massiven Nachfrageverschiebung seitens der Forschung auf Online-Angebote von Archiven ist absehbar, dass es nur eine Frage der Zeit sein wird, bis die von Uhde skizzierte dritte Stufe einer detaillierten Bestände- und Dokumentenvermittlung im Internet nicht mehr nur von einzelnen Pionierinstitutionen erreicht, sondern in der Archivlandschaft zum Standard wird. Ob sich dieser Standard aber bis auf die Ebene digitaler Reproduktion erstrecken wird, hängt von der zurzeit offenen Entwicklung insbesondere des Urheberrechts ab, auf die an dieser Stelle nicht eingegangen werden kann. Mindestens ebenso entscheidend ist in der Praxis aber ein ganz anderer

Umstand: „Die totale Digitalisierung ist nicht finanzierbar.“⁸ Dieser lapidaren Feststellung Hermann Leskiens muss man natürlich zustimmen. Das gläserne Archiv ist auch unter dem pekuniären Aspekt eine Utopie. Täglich leichter wird es allerdings für Archive und verwandte Institutionen – vom Argument der Bestandserhaltung einmal abgesehen – das Postulat Leskiens zu erfüllen, dass Digitalisierungen zwingend einen Mehrwert erbringen müssen. Verzweifelten und zweifelnden Budgetverantwortlichen sei zum Schluss empfohlen: Checken Sie mal Ihre Benutzungsstatistiken!

Dr. Daniel Nerlich ist stellvertretender Leiter des Archivs für Zeitgeschichte der ETH Zürich. Zu seinen Arbeitsschwerpunkten gehören die Sicherung von privaten Archivbeständen zur schweizerischen Wirtschafts- und Außenhandelspolitik sowie die Förderung des Einsatzes von Informationstechnologien bei der Vermittlung des Archivguts. E-Mail: nerlich@history.gess.ethz.ch

QUELLENARBEIT IM PROJEKT „PREUßEN ALS KULTURSTAAT“ – STRUKTURIERTE INFORMATIONSERFASSUNG MIT DEM „ARCHIV- EDITOR“

von Alexander Czmil und Bärbel Holtz

Historiker/innen leben von archivalischen Funden, mit denen sie Entwicklungen oder Ereignisse rekonstruieren, erklären sowie zweifelsfrei belegen können. Die Fülle an Fakten, Zusammenhängen und Fundorten jedoch überfordert ein individuelles Gedächtnis und interessante, gerade erst aus den Aktenbergen zutage gebrachte Einzelheiten drohen sogleich erneut in Vergessenheit zu geraten. Bei umfassendem Aktenstudium stehen Historiker/innen somit vor dem Problem, in ihren eigenen Arbeitsmaterialien die unterschiedlichsten Sachverhalte einer Akte effizient und strukturiert festzuhalten, sie dabei stets an ihrer Quelle zu verorten und darüber hinaus möglichst für die eigene Fragestellung abfragbar aufzubereiten. Hierfür eröffnet die EDV neue Wege. Im Rahmen der TELOTA-Initiative der Berlin-Brandenburgischen Akademie der Wissenschaften wurde mit der Entwicklung eines Softwaretools begonnen, das genau dieses strukturierte archivarische Arbeiten abbilden und unterstützen soll: auf der einen Seite durch das systematische Erfassen von Akteneinträgen im Zusammenhang mit semantischen Informationen, auf der anderen Seite durch eine flexible Suche, welche die Möglichkeit bietet, beliebige Anfragen entsprechend der Fragestellung an die vorher erfassten Daten zu stellen. So kann auf einmal erfasste Informationen jederzeit zugegriffen werden. Dieses Softwaretool, der „Archiv-Editor“, soll in seiner Funktionalität präsentiert und diskutiert werden.

8 Leskien, Hermann, Die retrospektive Digitalisierung löst und impliziert Probleme, vgl. <<http://www.uni-muenster.de/Forum-Bestandserhaltung/konversion/digi-leskien.shtml>> (26.09.2006).

Einleitung

Das Akademievorhaben „Preußen als Kulturstaat“¹ erforscht das Verhältnis von Staatsbildung, Kultur und Zivilgesellschaft im 19. und frühen 20. Jahrhundert. Ziel ist es, das Zusammenwirken wie Gegeneinander gesellschaftlicher und bürokratischer Kräfte dieses sozial, kulturell und regional unterschiedlich strukturierten europäischen Machtstaats auszuleuchten. Umfängliche archivalische Überlieferungen enthalten hierfür eine unverzichtbare Quellenbasis wie sie vor allem aus den zentralstaatlichen Akten Preußens hervorgehen.

Das Geheime Staatsarchiv Preußischer Kulturbesitz in Berlin-Dahlem ist das kulturhistorische Gedächtnis des preußischen Staates. Es bewahrt etwa 35 laufende Kilometer Archivalien unterschiedlichster Herkunft auf: staatliche Akten, private Nachlässe und Sammlungen. Jede einzelne Akte ist ein Unikat, die in ihrer Zusammensetzung kein zweites Mal existiert. Entstanden sind die meisten dieser Akten durch das Funktionieren von Verwaltung und Gesetzgebung auf den verschiedensten Ebenen, durch das Reflektieren zeitgenössischer Zustände in Denkschriften oder Korrespondenzen, durch das Kommentieren vergangener Ereignisse und Prozesse. Ihr damaliger Entstehungszusammenhang und genaues Wissen um den früheren Behördenaufbau liefern den heutigen Archivbenutzern/innen die erforderliche Orientierung, um in der Fülle des Aktenmaterials einzelne Schriftstücke oder komplette Vorgänge ausfindig zu machen – gleich, ob es beispielsweise die Sachakte einer Abteilung des Kultusministeriums ist, in der sämtliche Angelegenheiten zur staatlichen Kirchenpolitik im Vormärz, egal welcher Konfession, streng chronologisch abgeheftet wurden; ob es um die Personalakte eines Ministerialrats geht, die seine Laufbahn vom Assessor, über den Landrat und Regierungsrat in den unterschiedlichsten Regionen Preußens bis hin zu seiner Berufung ins Ministerium wiedergibt; oder ob es um den privaten Briefwechsel eines Kultusministers geht, den er mit Gelehrten oder Politikern zu Einzelheiten seines Ressorts geführt hat. Diese Vielfalt an Themen, dieses Springen zwischen Akteuren

1 Vgl. <<http://www.bbaw.de/bbaw/Forschung/Forschungsprojekte/actaborussica/de/Startseite>> (31.03.2006).

und Betroffenen und die Authentizität der Schriftstücke machen das Lesen von Akten so spannend und interessant, das strukturierte Erfassen der darin enthaltenen Fakten und Zusammenhänge aber auch schwierig und oft aufwändig.

Akten sind nicht gegliedert wie Bücher, die das Ergebnis geistiger Arbeit an einer Fragestellung sind und den Gedankengang ihres Verfassers oder ihrer Verfasserin strukturiert wiedergeben. Akten sind das Produkt von Verwaltungshandeln. Diese Archivalien vereinen in sich nicht selten völlig verschiedene Vorgänge, die freilich wichtige Details und Zusammenhänge für mehrere Fragestellungen freigeben. Finden sich nun in einer solchen Akte sowohl wichtige Schriftstücke über das staatliche Vorgehen gegenüber der Landeskirche und der sich separierenden religiösen Minderheiten als auch bislang unbekannt biografische Angaben zu deren führenden Figuren, so sind beides wichtige Funde, die es für die spätere wissenschaftliche Auswertung festzuhalten gilt. Historiker/innen leben von solchen archivalischen Funden, mit denen sie Entwicklungen oder Ereignisse rekonstruieren, erklären sowie zweifelsfrei belegen können. Die Fülle an Fakten, Zusammenhängen und Fundorten jedoch überfordert jedes individuelle Gedächtnis und interessante, gerade erst aus den Aktenbergen zutage gebrachte Einzelheiten drohen sogleich erneut in Vergessenheit zu geraten. Bei umfassendem Aktenstudium stehen Historiker/innen somit vor dem Problem, in ihren eigenen Arbeitsmaterialien die unterschiedlichsten Sachverhalte einer Akte effizient und strukturiert festzuhalten, sie dabei stets an ihrer Quelle zu verorten und darüber hinaus möglichst für die eigene Fragestellung abfragbar aufzubereiten. Hierfür wurden im Rahmen des „Telota – Projekt des Monats“² neue Wege eröffnet.

Telota – The electronic life of the academy

Seit Februar 2005 stellt die Berlin-Brandenburgische Akademie der Wissenschaften (BBAW) durch das „Telota“³ – Projekt des Monats“⁴ (TPM) im

2 Vgl. hierzu Neumann, Gerald, Wörterbücher als digitale Ressourcen für Mensch und Maschine, in diesem Band.

3 Vgl. <<http://www.bbaw.de/initiativen/telota/index.html>> (31.03.2006).

4 Vgl. <<http://www.bbaw.de/pom.html>> (31.03.2006).

Rahmen der Initiative „Telota“ monatlich ein neues, digitales Projekt der Akademievorhaben vor. Durch diese Initiative wird das Angebot an Informationen zu geisteswissenschaftlichen Themen an der Akademie erheblich erweitert und verbessert. Ein weiterer Schwerpunkt von TPM liegt auf der Unterstützung der Akademievorhaben durch den Einsatz von Standards und elektronischen Hilfsmitteln, welche auf den Bedarf der einzelnen Projekte abgestimmt sind. Unter diesem Aspekt wurde unter anderem die Emanzipation von proprietärer kostenpflichtiger Software als Ziel gesetzt.

Um die bisherige Arbeitsweise der Mitarbeiter/innen des Vorhabens „Preußen als Kulturstaat“ zu unterstützen und einheitlich zu strukturieren, wurde der „Telota-Archiv Editor“⁵ entworfen, mit dessen Hilfe aus den Akten exzerpierte Vorgänge so erfasst werden können, dass eine umfassende Suche und systematische wissenschaftliche Auswertung der Daten aller Mitarbeiter/innen möglich ist. Des Weiteren können die Daten, die in einer an den Editor angebotenen lokalen Version der nativen XML-Datenbank *eXist*⁶ gespeichert werden, in diverse Formate wie HTML, PDF oder XML exportiert werden.

Die eigentliche Arbeitsweise der Projektmitarbeiter/innen ändert sich trotz der neuen Software nur unwesentlich. Anstelle des von den jeweiligen Mitarbeiter/innen bevorzugten Texteditors, in den meisten Fällen Microsoft Word, in dem die Aktenvorgänge bisher erfasst wurde, starten die Mitarbeiter/innen nun den „Archiv Editor“. Nach der Benutzeranmeldung zeigt sich die Arbeitsoberfläche des „Archiv Editors“. Die Erfassung der Aktenvorgänge im „Archiv Editor“ erfolgt „personenorientiert“. Auf der linken Seite befindet sich im oberen Teil der Bereich, in dem die aktuell bearbeitete Person oder Quelle angezeigt wird. Direkt darunter ist eine Liste aller in der Datenbank befindlichen Personen und, hinter dem zweiten Karteireiter, Quellen. Auf der rechten Seite befindet sich der

5 Die Homepage des Editors inklusive einer Kurzanleitung zu Download und Installation befindet sich unter <<http://pom.bbaw.de/ae>> (31.03.2006).

6 Der Terminus „native XML-Datenbank“ bezeichnet die besondere Eigenschaft einer Datenbank spezielle technische Voraussetzung für die genuine Verarbeitung von XML-Dokumenten zu besitzen. Bei der vom TPM eingesetzten Datenbank *eXist* handelt es sich um ein Open Source-Projekt der TU Darmstadt, vgl. <<http://www.exist-db.org>> (31.03.2006).

Bereich zur Anzeige aller zu einer Person oder Quelle gehörenden Ereignisse. In diesem Fall ist ein Ereignis ein aus der Akte in den Editor übertragener Aktenvorgang, der relevante Informationen zu einer Person enthält.

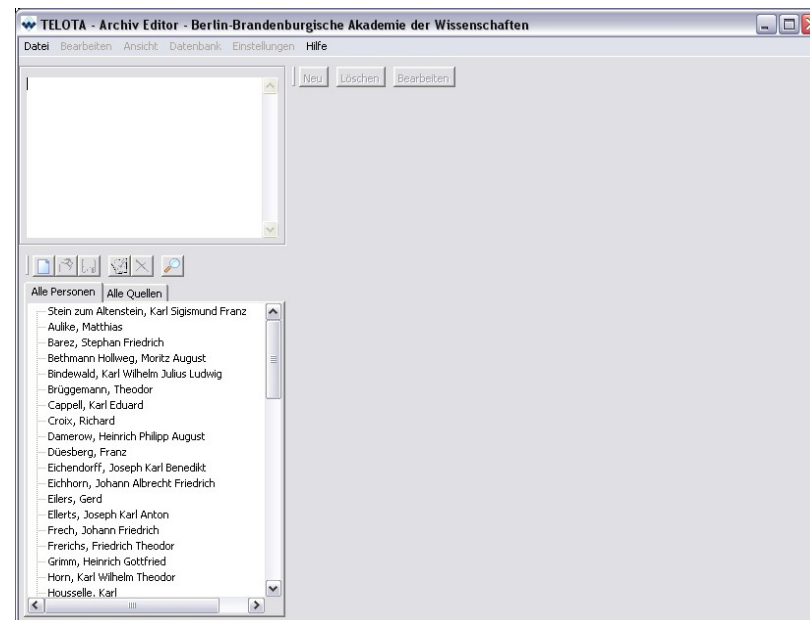


Abbildung 1: Der „Archiv Editor“ nach dem Laden und der Benutzerauthentifizierung

Aus der Liste aller Personen kann man die gewünschte Person aussuchen und per Doppelklick oder durch Klick auf das „Öffnen“-Icon in das Editorfenster laden. Im linken oberen Teil stehen nun alle Details zu der ausgewählten Person. Auf der rechten Seite haben sich die Ereignisse geöffnet, welche in verschiedene, frei konfigurierbare, Kategorien unterteilt sind. Jedes Ereignis besteht aus einem Datum, eventuell aus einer Unterkategorie, der eigentlichen exzerpierten Information, der Quelle, einem Hinweis auf den Mitarbeiter/die Mitarbeiterin, der/die dieses Ereignis aufgenommen

men hat und einer systemweit eindeutigen Identifikationsnummer zu Referenzierung.

Über der Liste aller Personen befinden sich Buttons zum Erstellen eines neuen Personen- oder Quelldatensatzes sowie zum Öffnen, Speichern, Bearbeiten, Löschen und für die Suchfunktion. Ein Klick auf den „Neu“-Button öffnet, je nach Karteireiter, entweder einen Dialog zur Erfassung einer neuen Person oder einer neuen Quelle. Bei der Person kann sowohl die Liste der Adelstitel, als auch die Liste der Adelsprädikate frei definiert werden.

The screenshot shows a dialog box titled 'Person' with the following fields and options:

- Nachname: Aulike
- Vorname: Matthias
- Adelstitel: Graf
- Adelsprädikat: (empty)
- Rufname: Fürst, Fürstin, Graf, Gräfin, Großherzog
- Geschlecht: männlich
- Bemerkung: (empty text area)
- Buttons: Speichern, Abbrechen

Abbildung 2: Bearbeiten der Basisdaten einer Person

Rechts neben dem Knopf zum Erstellen von neuen Personen oder Quellen

befinden sich die Knöpfe zum Öffnen eines Datensatzes und zum Speichern aller neuen oder modifizierten Datensätze in der Datenbank. Als nächstes folgen die Knöpfe zum Bearbeiten des aktuell geladenen Personen- oder Quelldatensatzes und zum Löschen des in der Liste ausgewählten Personen- oder Quelldatensatzes. Der letzte Knopf führt zur Suchoberfläche.

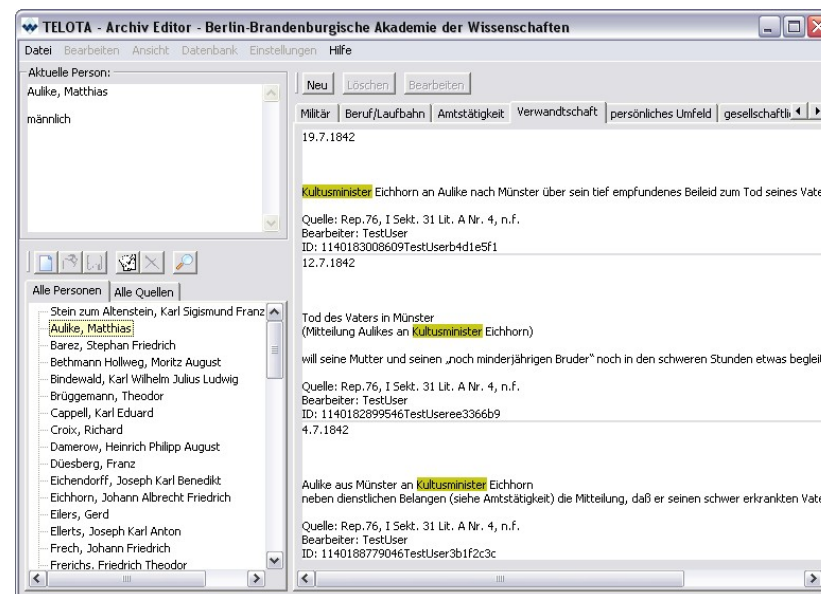


Abbildung 3: Die Daten zur Person Matthias Aulicke wurden in den Editor geladen.

Sind Teile des Ereignistextes mit sogenannten Eigenschaften verbunden, erscheinen diese grün hinterlegt. Typ und Inhalt der Eigenschaft können über einen Klick mit der rechten Maustaste auf den grünen Bereich in Erfahrung gebracht werden. Eigenschaften sind Informationen, die andere Informationen aus dem Ereignistext näher definieren. So können zum Beispiel unterschiedliche Schreibweisen von Namen oder Abkürzungen

durch ein kontrolliertes Vokabular leicht für die Suche zugänglich gemacht werden. Zudem ist es so einfacher, ähnliche Ereignisse zu finden. Die Eigenschaften werden über Listen definiert, die von den Mitarbeitern/innen frei erweiterbar sind. Die Ereignisansicht von Quellen ähnelt der von Personen mit dem Unterschied, dass die Kategorien fehlen.

Abbildung 4: Bearbeiten eines Ereignisses

Ereignisse anlegen und bearbeiten

Im oberen Teil auf der rechten Seite in der Hauptansicht des Editors befin-

den sich die Buttons zum Anlegen von neuen sowie zum Löschen und Bearbeiten von alten Ereignissen. Wie bereits erwähnt, besteht ein Ereignis aus diversen Komponenten, die über das Ereigniserfassungsfenster eingegeben und modifiziert werden können. An erster Stelle ist dies das Datum, das der einheitlichen Erfassung wegen aus Auswahllisten zusammengestellt werden kann. Dabei ist es möglich, nicht genau bestimmbare Datumszeiträume anzugeben wie zum Beispiel „vor 1850“ oder „Sommer 1867“.

Jedes Ereignis ist einer Kategorie, einer Art Schlagwort, zugeordnet. Die Kategorien in der Auswahlliste stimmen mit den Karteireitern im rechten Bereich des Editors überein. Die rechte Auswahlliste enthält die sogenannten Unterkategorien, die es ermöglichen, ein Ereignis weiter zu spezifizieren. So kann man einem Ereignis aus der Kategorie „Person“ zum Beispiel die Unterkategorie „Geburt“ zuweisen. Kategorien und Unterkategorien sind über eine Konfigurationsdatei beliebig änder- und erweiterbar.

Der eigentliche Text, der aus der Akte übertragen wird, steht im Ereignistextfeld. Diesem Text können aus den frei konfigurierbaren Eigenschaftsauswahllisten normierte Informationen mitgegeben werden, indem die gewünschte Eigenschaft aus der Liste ausgewählt wird. Anschließend kann man einen Teil des Textes markieren und diesem per „Zuweisen“ die Information zuweisen oder man kann per „Einfügen und Zuweisen“ den normierten Text aus der Liste an die Stelle des Cursors in das Ereignistextfeld übernehmen.

Zu jedem erfassten Ereignis muss eine Quelle angegeben werden, welche man aus einer Liste bestehender Quellen auswählen kann. Soll eine neue Quelle angelegt werden, kann diese über die Hauptansicht oder direkt aus dem Ereignisdialog heraus hinzugefügt werden. Zu einer vollständigen Quellenangabe gehört selbstverständlich auch die Angabe über die genaue Fundstelle innerhalb der Quelle, also das Blatt der Akte sowie die Angabe recto oder verso.

Export

Der Menüpunkt „Datei“ enthält im gegenwärtigem Entwicklungsstadium

lediglich den einen Unterpunkt „Exportieren als...“, der es ermöglicht, den aktuell bearbeiteten Personendatensatz in verschiedenen Formaten zu speichern. Neben den Formaten XML und Text werden auch HTML und PDF angeboten. Letztere können in der Form der Ausgabe beliebig beeinflusst werden. Die dafür zuständigen Dateien *html.xml* und *pdf.xml* befinden sich im Unterverzeichnis *stylesheets* des „Archiv Editors“. Die nötigen Kenntnisse vorausgesetzt, ist man in der Gestaltung der Ausgabe somit völlig frei.

Suche

Die Suchfunktion basiert auf der vom *World Wide Web Consortium*⁷ (W3C) vorgeschlagenen *XQuery*-Technologie⁸ zur Abfrage von XML-Datenbanken. Im augenblicklichen Entwicklungsstadium des „Archiv Editors“ muss man die *XQuery*-Anfrage noch manuell zusammenstellen.⁹ In den folgenden Versionen wird eine komfortable Suchoberfläche die einfache Suche in den Texten oder in den Eigenschaften unterstützen. Die Anzeige der Suchergebnisse ist unterteilt in Personen und Quellen. So werden zum Beispiel der Name der Person sowie das Ereignis angezeigt, welches den Suchkriterien entspricht. Ein Doppelklick auf ein Ereignis öffnet den dazugehörigen Personen- oder Quellendatensatz in der oben beschriebenen Hauptansicht des Editors.

Konfiguration

Neben den bereits erwähnten Dateien zur Beeinflussung des HTML und PDF-Exports existieren weitere Möglichkeiten, um die Darstellung und Funktionen des Editors den eigenen Bedürfnissen anzupassen. Zum einen ist dies die Möglichkeit, die Kategorien – mithin die Karteireiter – und die Unterkategorien für die Ereignisse festzulegen, indem man die entsprechenden Werte in der Datei *config.properties* anpasst. Zum anderen kann man alle Eigenschaftsauswahllisten selbst definieren. Dazu legt man

7 Vgl. <<http://www.w3.org/>> (31.03.2006).

8 Vgl. <<http://www.w3.org/XML/Query/>> (31.03.2006).

9 Eine Beispielabfrage nach allen Ereignissen, die den Begriff „Kultusminister“ enthalten, sähe in *XQuery*-Notation folgendermaßen aus: `//event[.&=„Kultusminister“]`.

einfach eine Textdatei mit der Endung *txt* an, welche alle potentiellen Werte zu einer Eigenschaft enthält. Die Liste muss im Verzeichnis *properties* gespeichert sein.

Fazit

Neben den umfangreichen Funktionen, die der „Archiv Editor“ bereits zur strukturierten Erfassung von Aktenvorgängen bietet, ist selbstverständlich noch ein großer Teil an Wünschen offen. So ist eine Benutzerverwaltung mit integriertem Rechtemanagement für die verteilte Arbeit unabdingbar. Ebenso ist eine zentrale Instanz der Datenbank erforderlich, die über das Internet mit den Arbeitsergebnissen der einzelnen Projektmitarbeiter/innen gespeist werden kann. Im Gegenzug können die einzelnen Mitarbeiter/innen ihre lokale, im „Archiv Editor“ integrierte XML-Datenbank mit dem aktuellen Stand der zentralen Datenbank aktualisieren.

Auch die Suchoberfläche bedarf einer Erweiterung, so dass die Suche nicht mehr als *XQuery* eingegeben werden muss, sondern über grafische Bedienelemente kombiniert werden kann.

Des Weiteren soll der „Archiv Editor“ in Zukunft nicht nur auf Windows-Systemen, sondern auch auf Macintosh und Linux einsetzbar sein. Sobald eine stabile Version des „Archiv Editors“ vorliegt, wird diese der interessierten (Fach-)Öffentlichkeit als Open Source-Produkt zur Verwendung und aktiven Weiterentwicklung angeboten.

Alexander Czmiel ist Wissenschaftlicher Mitarbeiter an der Berlin-Brandenburgischen Akademie der Wissenschaften. Er arbeitet im Projekt „Telota – Projekt des Monats“. E-Mail: czmiel@bbaw.de

Dr. Bärbel Holtz ist Wissenschaftliche Mitarbeiterin an der Berlin-Brandenburgischen Akademie der Wissenschaften und Arbeitsstellenleiterin des Akademienvorhabens „Preußen als Kulturstaat“. E-Mail: holtz@bbaw.de

QUELLEN IM NETZ

QUELLEN IM NETZ, VORSTELLUNG UND ZUSAMMENFASSUNG DER SEKTION

von Patrick Sahle und Georg Vogeler

Die Vorbemerkungen und die Zusammenfassung zur Sektion versuchen die Beiträge in einen weiteren Kontext zu stellen. Zusammen bilden die einzelnen Präsentationen den aktuellen Entwicklungsstand im Bereich digitaler Quellenerschließung und Edition ab. Von den Organisatoren ist dabei ganz bewusst aus einer größeren Zahl von vorgeschlagenen Vorträgen ein möglichst breites Spektrum zusammengestellt worden, das verschiedene Epochen, Materialarten und Methodiken umgreift. Dadurch sollte in der Zusammenschau auch die Frage nach der Entwicklung einer ‚gemeinsamen‘ Methodologie aufgeworfen werden. Der Vergleich der verschiedenen Ansätze zeigt, dass es im Übergang zu digitalen Arbeits- und Publikationsformen nicht (nur) zu einer weiteren Spezialisierung kommt, sondern auch die Inhalte, Techniken und Methoden zunehmend konvergieren. Schließlich ist der Frage nachzugehen, wie sich die Veränderungen auf der ‚Angebotsseite‘ auch auf die ‚Benutzung‘ und damit auf die Heuristik historischer Forschung auswirken.

Fragestellungen

Das Internet ist laut Urteil des Hanseatischen Oberlandesgerichts Hamburg vom 24.02.2005 eine „bekannte Nutzungsart“ seit 1993 (andere Urteile lauten auf 1995).¹ Das Internet ist also mit uns seit inzwischen 13 Jahren.

¹ Siehe hierzu den Beitrag in „Kommunikation und Recht“ 2006, Heft 1, S. 46-48. Das Urteil online auf den Seiten des Oberlandesgerichts unter <<http://www.jurpc.de/rechtspr/20050121.htm>> (06.08.2006).

Eine der ersten deutschen Übersichten über digitale Quellen zum Mittelalter – die heute natürlich allenfalls medien- und wissenschaftshistorisch von Interesse sein kann – datiert aus dem Jahr 1995.² Mehr als zehn Jahre sind für ein so dynamisches Medium wie das Internet eine lange Zeit, wie es andererseits für die Geschichtswissenschaft eine sehr kurze Zeit ist. Aus beidem ergibt sich das Paradox: Wir sind schon weit, stehen aber immer noch am Anfang. Umso mehr lohnt eine weitere Bestandsaufnahme aktueller Entwicklungsstände. Um es direkt in einer zeithistorisch wissenschaftsgeschichtlichen Perspektive zu verorten: Die anfängliche Phase vereinzelter kleiner innovativer Experimente geht langsam ihrem Ende entgegen. Auf den „Inseln“ stehen jetzt „Leuchttürme“: Wir sehen inzwischen immer mehr ausgereifte, stabile, umfassende und gut benutzbare Quellensammlungen, die unmittelbar die praktische Arbeit mit den Quellen beeinflussen können und Vorbildcharakter gewinnen. Bei den folgenden Beiträgen geht es also nicht um eine Bestandsaufnahme, welche Quellen im Netz verfügbar sind, sondern darum wie sie im Netz *anders* sind als man es von den gedruckten Quellenwerken her kennt.

Die Quellenerschließung bedient sich heute *selbstverständlich* digitaler Werkzeuge. Wir sind allerdings nach wie vor weit davon entfernt, dass auch die Ergebnisse der Quellenerschließung genauso selbstverständlich in digitalen Medien publiziert werden. Wir haben es immer noch eher mit einzelnen Vorreiterprojekten zu tun, die in ganz besonderer Weise sowohl inhaltlich als auch methodisch jene Pfade anlegen, denen eines Tages die dann alltägliche Praxis der Quellenerschließung folgen können. Wir sind immer noch weit davon entfernt, eine abgeschlossene Methodik der digitalen Quellenerschließung einfach als gegeben annehmen zu können, auf deren Grundlage wir uns endlich vor allem um die Inhalte kümmern könnten.

In der Vorbereitung dieser Sektion bestand das (Luxus-)Problem, aus einer ganzen Reihe höchst interessanter Themenvorschläge auswählen zu müssen. Leicht hätte man sich zum Beispiel auf eine bestimmte Quellen-

² Derzeit archiviert unter <http://www.uni-koeln.de/~ahz26/dateien/maalt/mahomepageunix.html> (06.08.2006).

gattung beschränken können. Die mittelalterlichen Urkunden scheinen derzeit eines der Felder zu sein, auf dem die breiteste Entwicklung in inhaltlicher und methodischer Hinsicht zu beobachten ist. Wir haben uns gegen eine solche Fokussierung entschieden – nicht nur, weil Ende Februar 2007 eine ganze Tagung zu diesem Komplex in München stattfinden wird. So haben wir versucht, die ganze Bandbreite der Entwicklungen abzudecken: unterschiedliche Quellengattungen und Materialtypen sind in die Überlegungen ebenso einzubeziehen wie die methodischen Ansätze und technischen Lösungswege. Hier geht es uns um eine Bestandsaufnahme zum Stand der Kunst: Was ist heute möglich? Was ist gewissermaßen schon etablierte „best practice“? Was sind die gängigen Paradigmen an denen man sich orientiert? Aber auch: wo liegen die noch nicht gelösten Probleme? Und: wo sind die Perspektiven für die nächste Zukunft?

Zu diesen Fragen liefern die Beiträge der Sektion „Quellen im Netz“ jeweils ganz eigene interessante Antworten. Wir glauben jedoch, einige Fragen erhellen zu können, die aus der Zusammenschau so unterschiedlicher Ansätze klar werden. So hoffen wir, dass die Sektion als Ganzes mehr ist als die Summe der Teile, wenn man die Beiträge unter zwei Aspekten betrachtet:

1. Das, was sich unter dem Schlagwort „**Konvergenz**“ versammeln lässt: Konvergieren die einzelnen Quellenerschließungsprojekte in *inhaltlicher* Hinsicht? Führen sie zu umfassenden Sammlungen gleichartiger Quellen? Dazu ist auch die Frage zu stellen: Konvergieren sie in *technischer* Hinsicht? Lassen sie sich in der Praxis wirklich vernetzen und integrieren, weil man sich inzwischen auf gemeinsame technische Paradigmen und gemeinsame Standards verständigen kann? Konvergieren die digitalen Unternehmungen darüber hinaus in *methodologischer* Hinsicht? Bildet sich also eine neue übergreifende Methodologie der Quellenerschließung aus? Eine, die dann auch so unterschiedliche Materialien umfassen kann, wie mittelalterliche Urkunden, mittelalterliche Weltkarten, neuzeitliche Akten und Protokolle und modernes Filmmaterial? Und konvergieren damit auch die Aktivitäten der verschiedenen Akteure und ihrer Praktiken? Wird aus den Grunderschließungsleistungen der Archivare und Bibliothekare, den tiefer gehenden Editions-

arbeiten, den Analysearbeiten und den Publikationsarbeiten der Fachwissenschaftler ein großes Ganzes?

2. Wir gehen zunehmend zu grundsätzlichen Fragen über: Es verändern sich die Inhalte, es verändern sich die Arbeitsweisen und es verändern sich die Methoden. Die Methoden der Quellenerschließung stehen dann aber unweigerlich auch in engem Zusammenhang mit unserer Haltung *zu* den Quellen, unserer Sicht *auf* die Quellen und damit auch mit unserer Benutzung der Quellen. Der zweite Aspekt, der die folgenden Beiträge überspannen könnte ist damit jener der **Heuristik**: Verändern sich durch die neuen Angebote, durch die neuen Darstellungsweisen (fast könnte man sagen: die neuen Existenzweisen) und durch die neuen Nutzungsmöglichkeiten nicht auch unsere Fragestellungen und Arbeitsweisen? Wenn Geschichtswissenschaft bedeutet, Antworten auf immer wieder neue Fragen zu geben, dann müssten die Auswirkungen auf die Geschichte insgesamt offensichtlich werden.

Ergebnisse

Die sechs Beiträge der Sektion „Quellen im Netz“³ haben sehr unterschiedliche Dokumentarten vorgestellt: Urkunden, Filme, Landkarten, Protokolle, Güterverzeichnisse. Für diese Quellentypen werden spezifische digitale Erschließungs- und Präsentationsformen verwendet: Eine Einzelbildannotation in einem Film steht für eine andere Methodik als das semantisch gedeutete Inhaltsverzeichnis einer Landkarte in Form einer Topic Map. Das diachrone Modell eines hochkomplexen Quellentextes steht der Datenmenge minutiöser Dokumentation in politischen Akten des 20. Jahrhundert gegenüber.

In der Zusammenschau kann mit dem Bereich der mittelalterlichen Geschichte begonnen werden, dem sich vier Beiträge der Sektion mit Blick auf die Erschließung und Verarbeitung von mediävistischen Quellenarten gewidmet haben. Heutige nationale Grenzen, die immer noch die Aufbewahrung und Erschließung der Überlieferung bestimmen und begrenzen, kannte das Mittelalter noch nicht. Karl Heinz hat sich mit dem Monasteri-

3 Vgl. zu den Beiträgen der Sektion die Aufsätze in diesem Band.

um-Projekt⁴ diesen – und anderen – Grenzüberwindungen gewidmet. Das Projekt dokumentiert nicht nur die Verflechtungen kirchlicher Urkundenüberlieferung im östlichen Mitteleuropa, sondern ist auch ein Beispiel für die Integration traditionell getrennter Erschließungsebenen: Bilder von Urkunden werden kombiniert mit archivischen Informationen, mit Regesten und Editionen. Außerdem beleuchtet es Chancen und Probleme verteilter, kollaborativer und inkrementeller Arbeiten.

Michael Gervers hat mit seinem Beitrag „Towards the implementation of advanced research and result presentation facilities for a published online collection of medieval Latin charters“ am Beispiel des DEEDS-Projekts⁵ umfangreiche Möglichkeiten für die Nutzung von Dokumenten mit tiefer semantischer Codierung aufgezeigt. Sie helfen, traditionelle, aber immer noch zentrale quellenkritische Fragen, wie die nach der Datierung und Echtheit eines Dokuments, in neuer Weise zu beantworten. Sie ermöglichen darüber hinaus aber auch inhaltliche Recherchen und Darstellungen, etwa zu Fragen nach religiösen Einstellungen, nach der Relevanz von römischen Entscheidungen in der Peripherie oder nach Rechtsentwicklungen.

Matthias Perstling hat demonstriert, wie elektronische Medien die Präsentation von komplex strukturierten Quellen ermöglichen, deren Mehrschichtigkeit in traditionellen Formen gar nicht adäquat abgebildet werden kann. Am Beispiel des Steirisch-landesherrlichen Marchfütterurbars von 1414/1426 hat er die editorische Problematik einer seriellen, kontinuierlich veränderten Quelle geschildert und dazu das Modell einer „Online-Faksimile-Edition“ vorgestellt.

An der Quelle *als Bild* sind auch die Konzepte orientiert, die Günther Görz bei der Präsentation „Kognitiver Karten“ des Mittelalters, das heißt mittelalterlicher Weltkarten, vorgestellt hat. Das Kartenmaterial, das als Text-Bild-Kombinationen die Handlungsräume narrativ vermittelter Geschichte und Geschichten darstellt und sich mit dem jeweiligen Wissensstand und durch Wissenstransfer verändert, ist mit einem systematischen

4 <<http://www.monasterium.net>> (20.02.2006).

5 <http://www.utoronto.ca/deeds/research/deeds_project.html> (21.02.2006).

vergleichenden Stellenkatalog erschlossen worden. In den circa 250 Weltkarten des Projektes (von der Hereford- und Ebstorfkarte über Portulane und Ptolemäuskarten bis zum Behaim-Globus) können so Bilder, Legenden und Bild-Text-Kombinationen ge- und untersucht werden. Der Beitrag hat damit die Frage aufgeworfen, ob die zur Erschließung angewandte formale Ontologie auf der Basis einer für den Behaim-Globus erstellten Konzepthierarchie Erkenntnisse ermöglicht, die mit herkömmlichen Mitteln nicht gewonnen werden konnten.

Alle diese Beiträge geben Zeugnis davon, dass mittelalterliches Quellenmaterial mit Hilfe des Internets aus den modernen Grenzen wieder in überregionale Zusammenhänge gebracht werden kann (wie sie im Mittelalter gegeben waren), dass vielschichtige Quellenstrukturen abbildbar und in Suchanfragen integrierbar werden und dass Werkzeuge entwickelt werden können, die komplexe Zusammenhänge im Quellenmaterial sichtbar machen. Die Frage, ob für das Quellenmaterial aus anderen Epochen die gleichen Fragestellungen und Beobachtungen anwendbar sind, helfen die folgenden Beiträge zu beantworten.

Die Internetpräsentation einer zentralen zeithistorischen Quelle stellte Jörg Filthaut vor. Dabei ist die Neuedition der „Kabinettsprotokolle der Bundesregierung“⁶ nicht nur inhaltlich mit der gedruckten Edition der „Akten der Reichskanzlei“ der Weimarer Republik verbunden, sondern wird auch im Rahmen der Retrodigitalisierung dieser Edition in einer gemeinsamen Softwareplattform präsentiert. Das bisherige Nutzer-Feedback und bestehende Anfragen zeigen, dass es der Forschung in zunehmendem Maße möglich sein wird, große Mengen an Quellen in der Zusammenschau auszuwerten. Damit wird das Material insbesondere quantifizierenden Fragestellungen zugänglich. Als Perspektive ergibt sich daraus, dass es auch für die zeithistorische Forschung wünschenswert wäre, wenn elektronische Findmittel, wie zum Beispiel der Karlsruher Virtuelle Katalog, beim Suchergebnis als erste Treffer die Links zu den digitalen Angeboten liefern würden und erst dann die Buchform nachweisen würden. Eine weitere Perspektive des Referats war die Vernetzung

6 <<http://www.bundesarchiv.de/kabinettsprotokolle/>> (06.08.2006).

zwischen Edition und elektronischen Exzerpten der Nutzer, mit der die Edition endgültig zu einer offenen und dynamisch wachsenden Arbeitsumgebung wird.

Christopher Carlson richtete den Blick schließlich auf eine eher ungewöhnliche Quellengattung, nämlich historische Filmeditionen. Die Diskussion über den Quellenwert von Filmmaterial hat zu editorischen Ansprüchen geführt, die von elektronischen Editionen zu erfüllen sind: Filmische Quellen müssen technisch und inhaltlich sorgsam ausgewählt, Provenienz und Dignität müssen überprüft und nachgewiesen werden. Die „Film-Edition“ muss das Material mit einer umfassenden quellenkritischen Analyse und einer entsprechenden Dokumentation absichern. Mithilfe von Bewegtbild-Datenbanken können die Filme endlich sehr viel besser bzw. tiefer erschlossen und annotiert werden, als dies früher möglich gewesen wäre. Neben dem qualitativen Gesichtspunkt entstehen auch neue *quantitative* Perspektiven: Umfangreiche, punktgenaue Annotationen sind möglich, die die Menge und das Niveau von textkritischem Apparat und inhaltlichem Kommentar klassischer Texteditionen erreichen können – ohne die Benutzbarkeit des Materials zu behindern. Christopher Carlson hat exemplarisch die vorhandenen Optionen anhand vielschichtiger Beispiele aktueller Filmeditionsarbeit aufgezeigt. Neuere technische Plattformen unterstützen dabei die Multimedialität der Wissensrepräsentation.

Welche Schlüsse können nun aus diesen Referaten für die beiden Grundfragen der Sektion gezogen werden? Einige weiterführende Überlegungen sollen abgeleitet und im Folgenden vorgestellt werden.

Verschwinden die Unterscheidungen zwischen Archivierung, Erschließung, Edition und wissenschaftlicher Publikation in der digitalen Welt?

Wir haben gefragt, ob die unterschiedlichen Aktivitäten und Praktiken der Akteure konvergieren. Die Ausgangserfahrung ist hier, dass Bibliothekare zum Beispiel unter Quellendigitalisierung gemeinhin eine Imagedigitalisierung von Buchseiten, Archivare digitale Abbildungen und Inhaltserschließung der Originale und Fachwissenschaftler elektronische Volltexteditionen verstehen. Die Referate der Sektion zeigen aber, wie schwach diese Grenzen ausgeprägt sind: Archivare edieren Texte ebenso

umfassend, wie Forscher Grundlagenmaterial online zur Verfügung stellen. Die Referenten und Referate belegen also, dass die Grenzen digitaler Quellenrepräsentationen schon heute immer weniger den institutionellen Traditionen von Archiv, Bibliothek und Universität folgen. Insbesondere die Vorstellung des Monasterium-Projektes⁷ hat gezeigt, wie sehr diese Formen in der digitalen Welt zusammenfließen: Ist ein umfangreiches Bildarchiv mit Regesten und Transkriptionen eine Faksimileausgabe, ein Regestenwerk, ein Repertorium oder eine Edition? Die mediale Festlegung der traditionellen Trennungen etablierter Publikationsformen verschwinden in einer digitalen Umgebung – ohne dass deshalb die inhaltliche Unterscheidung zwischen Faksimile, Regest und Edition verschwindet. Diese werden vielmehr als inhaltliche Module übergreifender Informationssysteme erkennbar.

Die Quellen im Netz konvergieren so in einem informationswissenschaftlichen Modell, das die bislang getrennten Formen der Quellenrepräsentation einander zuordnet und die historische Forschung mit ihren Quellen verbindet. An anderer Stelle ist dazu von einem „digitalen Informationsraum“ mit den Achsen „Benutzbarkeit, Funktionalität, Qualität“, „Menge“ und „Erschließungstiefe“ gesprochen worden.⁸ Man kann damit einer Vernetzung zwischen den unterschiedlichen Formen der Quellenrepräsentation den Vorzug gegenüber einer Prozesskette geben.⁹ Die mit der Bewahrung und Bearbeitung historischer Quellen beschäftigten Akteure in Archiv, Museum, Bibliothek und wissenschaftlicher Forschung können die erprobten Arbeitsgänge unter einer modularen Perspektive abwickeln und ihre Ergebnisse zum einzelnen Quellenobjekt verknüpfen.

Denn das Quellenobjekt ist, wenn man von modernen elektronischen

7 <<http://www.monasterium.net>> (20.02.2006).

8 Vgl. Patrick Sahle, Urkunden-Editionen im Internet – Einführung und Überblick, in: Archiv für Diplomatik 52 (2006). Eine ähnliche Sicht bei Georg Vogeler: Towards a standard of encoding medieval charters with XML, in: Literary and Linguistic Computing 20 (2005), S. 269-280.

9 Ohne den Begriff zu gebrauchen, beschreibt Karsten Uhde, ein Pionier der Verwendung des Internets für die Archive, in seinem Beitrag: Urkunden im Internet - Neue Präsentationsformen alter Archivalien, in: Archiv für Diplomatik 45 (1999), S. 441-464, die Prozesskette „Archiv->Regest->Edition->Forschung“.

Verwaltungsakten absieht, selbst noch nicht digital: Die Handschrift, die Filmrolle, das Exemplar eines Buches, der Globus werden von den dafür zuständigen Einrichtungen aufbewahrt und für wissenschaftliche Auswertung zur Verfügung gestellt. Die ehemals getrennten Repräsentations- und Erschließungsformen sind Abstraktionen von diesem realen Objekt, die digital reproduziert werden können:

1. eine Beschreibung des Quellenobjektes (zum Beispiel die archivische Formalbeschreibung)
2. der Text / das Bild als Abstraktion des Quellenobjektes, die gerne – und nicht immer ganz richtig – mit dem eigentlichen Quellenobjekt gleichgesetzt werden, insbesondere da die Quellenobjekte Texte und Bilder in Kopie überliefern können
3. der „historische Ort“, das heißt die im Bild oder Text enthaltene Sachinformation. Diese ist Ergebnis einer historischen Interpretation unterschiedlicher Methodik: Sprachliche Methoden machen aus lateinischen Urkunden eine Übersetzung. Eine typologische Einordnung macht aus einem Stück Papier ein Konzept. Inhaltliche Erforschung macht aus einer Wortliste ein Verzeichnis von geografisch identifizierbaren Orten. Strukturanalyse macht aus einem Punkt auf einer Karte eine Stadt, aus einer Textphrase eine Arenga.

Die Quellen im Netz sind unterschiedliche Aspekte eines konkreten Quellenobjektes, die unterschiedlich eng mit historischer Deutung verknüpft, aber auf jeden Fall Teil der Darstellung von Geschichte im Internet sind. Oder anders formuliert: Quellen im Netz sind Verbindungsknoten zwischen virtueller Geschichte und der realen Welt. Der jetzigen wie der vergangenen.

Können wir noch etwas Konkretes aus den Beiträgen dieser Sektion über die Konvergenz lernen? Es sind zwei Fragen, die wir dabei ansprechen möchten, zu denen die einzelnen Beiträge zwar nicht explizit Stellung bezogen haben, die implizit aber doch immer eine Rolle spielen:

1. Welche Datenformate werden verwendet?
2. Über welche Kanäle stellt man die Daten der Öffentlichkeit zur Verfügung?

Es ist inzwischen Konsens, „proprietäre“ Datenformate zu vermeiden.

Offen gelegte und weit verbreitete Datenformate wie ASCII, tiff oder jpeg sind in der Diskussion längst Selbstverständlichkeiten. XML ist als Metagrammatik, die es uns erlaubt Daten dokumentnah und zugleich in einiger Komplexität zu strukturieren, zu einem ähnlichen Paradigma geworden, obwohl stark strukturierte Daten sehr wohl auch mit Hilfe von SQL offen archivierbar sind. Aber XML beschreibt semistrukturierte Daten und ist damit für historische Informationen besonders geeignet. XML hat noch einen weiteren Vorteil: XML-Daten werden gerne nach öffentlichen Standardsprachen gebildet. Das Textgestaltungsformat OpenDocument (ODF) ist zwar noch zu jung, um sich schon etablieren zu können, und muss deshalb hinter dem Nicht-XML-Format pdf zurückstehen.¹⁰ Aber im geisteswissenschaftlichen Bereich sind andere XML-Schemata verbreitet und akzeptiert: Die Guidelines der Text Encoding Initiative (TEI)¹¹ oder die Encoded Archival Description (EAD).¹² Weitere sind in Entwicklung begriffen (wie zum Beispiel ein Standard zur Auszeichnung von Urkunden: Charters Encoding Initiative (CEI))¹³ oder stoßen in neue, noch völlig unreflektierte Bereich vor, wie die Historical Events Markup Language (HEML), die sich auf eine faktenorientierte Übersichtsdarstellung konzentriert.¹⁴ Historikerinnen und Historiker sollten sich der Diskussion um solche Standards stellen, denn sonst drohen sie, in die Modelle von Sprachwissenschaftlern und Bibliothekaren gezwungen zu werden. Ein Bereich könnten hier vielleicht auch die von Manfred Thaller 2003 beschriebenen „Digital Autonomous Cultural Objects“ (DACO) sein, deren spezifisch historische Eigenschaften festzulegen wären.¹⁵ Auf jeden Fall

10 <http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=office> (20.02.2006).

11 <<http://www.tei-c.org>> (20.02.2006).

12 Encoded Archival Description (EAD). Official EAD Version 2002 Web Site <<http://www.loc.gov/ead/>>, (20.02.2006).

13 <<http://www.cei.lmu.de>> (20.02.2006).

14 Bruce Robertson: Introduction to Hempl, 2000-2004 <<http://heml.mta.ca/heml-cocoon/description>> (20.02.2006).

15 Vgl. Thaller, Manfred, 'Historische Fachinformatik'. Ein Kölner Modell, in: Burckhardt, Daniel; Hohls, Rüdiger; Ziegeldorf, Vera (Hgg.), Geschichte und Neue Medien in Forschung, Archiven, Bibliotheken und Museen. Tagungsband .hist 2003 (Historisches Forum 7,1), Berlin 2005, S. 45-72.

stehen auch jetzt schon akzeptierte und offen dokumentierte Standards zur Verfügung, so dass zumindest der Datenaustausch problemlos gestaltet werden könnte.

Auch der Kanal, über den diese digitalen Daten ausgetauscht werden, ist eigentlich kaum noch strittig: Das Internet hat die Diskussionen über die Haltbarkeit von CD-ROMs oder über den Wechsel der Diskettenformate beendet – und die Diskussion über die Langfristarchivierung den Institutionen – den Bibliotheken, Archiven und Rechenzentren – zugewiesen. Das Internet ist der gegebene Kanal für die Verbreitung der Daten. Für die Frage der Konvergenz bedeutet das, dass die Daten über die dem WWW zu Grunde liegenden Protokolle zugänglich wären. Die Möglichkeiten, zwar in offenen Formaten zu archivieren, in proprietären oder zumindest nur schwer weiterzuverarbeitenden Formaten wie kennwortgeschützten PDF-Dateien zu veröffentlichen oder sie hinter einem Zugriffsschutz zu verbergen, werden aber ebenfalls genutzt. Dagegen wäre es zu bevorzugen, die digitalen Quellenrepräsentationen nicht nur in offenen Datenformaten zu speichern, sondern die Daten auch clientgesteuerten Lesevorgängen zugänglich und schematisch ansprechbar zu machen, so dass Metadatenprovider, gemeinsame Portale oder weiter verarbeitende Dienste auf die Daten zugreifen können (zum Beispiel nach den Standards der Open Archives Initiative - Protocol for Metadata Harvesting¹⁶). Die einzelnen Objekte sollten, wie aus der Druckwelt gewohnt, mit Hilfe von kanonischen Referenzen in der Art von „MGH DD F.I. 218“ ansprechbar sein.¹⁷ Erst dann erscheint die Vision vom „KVK für Quellen“, das heißt einem Institutionen übergreifenden Nachweisinstrument ebenso realistisch wie eine angebotsübergreifende Volltextsuche.¹⁸ Gerade letzteres ist ein zentrales

16 <<http://www.openarchives.org>> (20.02.2006).

17 Vgl. den unpublizierten Beitrag von Bernhard Assmann auf der Tagung „A DTD for Medieval Charters“, München 2004, vgl. Vogeler, Georg, Ein Standard für die Digitalisierung mittelalterlicher Urkunden. Bericht zum Workshop (München 5./6. April 2004), in: Archiv für Diplomatik 50 (2004), S. 23-34.

18 Vgl. dazu die Aktivitäten des Projektes DING: Heller, Markus; Vogeler, Georg, Modern information retrieval technology for historical documents, in: Humanities, Computers and Cultural Heritage : Proceedings of the XVIth international conference of the Association for History and Computing, Amsterdam 2005,

Desiderat für die historische Interpretation der Quellen.

Wie verändert sich das historische Arbeiten durch die Digitalisierung von Quellencorpora?

Die digitalen Repräsentationsformen von Quellen verändern nicht nur die Arbeit an den Quellen selbst, sondern auch die Forschung auf der Grundlage der Quellen.

Inhaltlich

Nur eines der vorgestellten Projekte kann vielleicht von sich behaupten, ein gänzlich neues Forschungsfeld überhaupt erst zu begründen: monasterium.net integriert nämlich einen historischen Raum, der durch nationale Grenzen – und bis zum Fall des eisernen Vorhangs sogar Grenzen ganzer Staatenblöcke – fragmentiert war. Die klösterliche Kultur Ostmitteleuropas ist schließlich keine nationale gewesen: Die Verbindungen zwischen den Zisterzienserklöstern des Raumes beruhten auf Filiationen, nicht auf politischen Grenzen. Das Projekt hat also einen historischen Raum redefiniert, der neu zu erforschen ist.

Methodik

Am stärksten verändern digitale Quellenrepräsentationen die Forschungsmethodik: Dabei ist an den Nutzen in der Forschungspragmatik nur deshalb zu erinnern, damit er nicht als Banalität missverstanden wird: Quellen im Netz erleichtern dem Forscher die Arbeit ungemein, wenn er 7 Tage in der Woche 24 Stunden von überall in Sekundenschnelle auf die Quellen zugreifen und diese sammeln, nachprüfen, weiter verarbeiten und in seine Publikationen einfließen lassen kann.

Quellen im Netz betreffen aber auch die „echte“ Forschungsmethodik, das heißt sie lösen nicht nur pragmatische Probleme der Forschung sondern beeinflussen auch die Wahl der Analyse Kriterien und die Auswertungsmethoden: Die Geschwindigkeit des Zugriffs stärkt eine in der modernen Geschichtsforschung eher untergenutzte Methode: die Statistik. Der Sie-

S. 143-148 und <http://www.cis.uni-muenchen.de/~heller/Classes/ding/> (20.02.2006).

geszug der Kulturgeschichte hat die Gruppe von Historikern, die die verbreitete Abneigung gegen „trockene“ Zahlen teilen, wohl wachsen lassen. Die digitale Repräsentation von Quellen wird nun einen solchen Trend nicht brechen und die statistischen Jahrbücher nicht zur zentralen Quelle der Geschichtsforschung machen. Sie wird viel mehr Häufigkeiten von Quellenaussagen und ihren Anteilen an einer Grundgesamtheit alleine deswegen wieder einen gewissen Stellenwert verschaffen, weil sie leicht zu ermitteln sind und sie also zu einem ebenso leicht überprüfbar Argument werden, wie es das direkte Quellenzitat ist.

Das beruht auf einer Eigenschaft digitaler Quellenrepräsentationen, die auch für qualitative Auswertungen gilt: Die ungeheuren Materialmengen, die einerseits durch elektronische Tools überhaupt edierbar geworden sind, können andererseits mit Suchbefehlen schnell wieder auf ein bearbeitbares und auf die Fragestellung zugeschnittenes Maß reduziert werden. Der elektronische Volltext ermöglicht darüber hinaus analysierende und interpretierende Operationen, die man unter dem Stichwort „Text-Mining“ zusammenfassen könnte: Dabei geht es nicht nur darum, passende Dokumente zu einem Suchwort zu ermitteln („Information retrieval“) oder Fakten aus Texten zu extrahieren, sondern in den Texten enthaltene Informationen explizit zu machen.¹⁹ Es geht also darum, sprachliche Merkmale der Quellen so mit einander in Beziehung zu setzen, dass historische Erkenntnisse daraus gewonnen werden können. Die eine digitale Quelle (zum Beispiel ein historiografischer Text) kann damit die Hinweise liefern, die eine andere digitale Quelle (zum Beispiel eine Karte) verständlich macht.

Bei vielen Projekten zur Digitalisierung von Quellen stehen jedoch nicht die Volltexte am Anfang, sondern Bilder. Die Bilderfeindlichkeit der älteren historischen Forschung hat zunächst etwas mit der philologischen Tradition der Geisteswissenschaften zu tun. Hinzu kamen die hohen Kosten der Bildreproduktion, also historisch-technische mediale Bedingungen.

19 Vgl. Mehler, Alexander, Textmining in: Lemnitzer, Lothar; Lobin, Henning (Hgg.): Texttechnologie. Perspektiven und Anwendungen. Tübingen 2004, S. 329-352.

Es kann nicht verwundern, dass die Interpretation bildlicher Quellen bzw. die Auswertung von bildlichen Eigenschaften der Quellen mit dem Siegeszug der digitalen Reproduktionstechniken einen immensen Aufschwung genommen hat.²⁰ Die Beiträge der Sektion haben gezeigt, dass die digitale Bildrepräsentation aber auch inhaltliche Möglichkeiten bietet, die über die reine Fotografie hinaus gehen: Bilder und Bildbereiche können direkt mit sachlichen Erläuterungen versehen werden, das heißt die digitale Quellenrepräsentation von Bildquellen ermöglicht eine kritische Annotation der Bilder wie sie in Texteditionen schon lange Standard ist, deren Bedarf bei Bildern aber erst durch die Diskussion um die Bilder der Wehrmachtsausstellung deutlich geworden ist.²¹ Digitale Bildedition macht so kritische Bildinterpretation öffentlich und nachvollziehbar. Sie ist eine wirklich multimediale Quellenrepräsentation, die Sachinformation, Text und Bild direkt einander zuordnet.

Fazit

Digitale Quellenrepräsentationen verlieren ihre traditionellen, von institutioneller und medialer Logik vorbestimmten Grenzen. Im Verarbeitungsprozess entstehen neue Grenzen, die nichts mit der sachlichen Logik der Quelle selbst zu tun haben müssen. Diese Grenzen sind unscharf und können durch technische Lösungen überwunden werden, die den Austausch der Ergebnisse erleichtern. Dazu sind viele Technologien schon entwickelt, fachspezifische hat die Historikerzunft aber noch wenige zu bieten. Digitale Quellenrepräsentationen wirken derzeit auf die Forschung vorrangig methodisch: Bilder können kontrolliert als Sachaussagen interpretiert werden. Das Volumen verarbeitbaren Materials nimmt zu. Die Vielschichtigkeit der Quellen wird transparenter. Die Informationen aus unterschiedlichen Quellen vernetzen sich. Inhaltlich scheint die digitale

20 Vgl. z.B. Bruhn, Matthias; Borgmann, Karsten (Hgg.), Sichtbarkeit der Geschichte. Beiträge zu einer Historiografie der Bilder (Historisches Forum 5), Berlin 2005.

21 Vgl. Schmidt-Neuhaus, Dieter: Die Tarnopol-Stellwand der Wanderausstellung 'Vernichtungskrieg - Verbrechen der Wehrmacht 1941-1944'. Eine Falluntersuchung zur Verwendung von Bildquellen, in: Geschichte in Wissenschaft und Unterricht 50 (1999), S. 596-603.

Quellenrepräsentation zunächst einer Globalisierung des historischen Bildes förderlich. Weder aus der Multimedialität der digitalen Quellenrepräsentation noch aus den informationswissenschaftlichen Modellen ihrer Erschließung haben sich bislang durchweg neue Fragestellungen ergeben, die bislang unbekannte historische Forschung provoziert hat. Das aber kann sich noch ändern ...

Patrick Sahle M.A. ist Wissenschaftlicher Mitarbeiter an der Staats- und Universitätsbibliothek Göttingen, Laufende Projekte: Zentrales Verzeichnis Digitalisierter Drucke (zvd), Online-Portal digitalisierter Kulturgüter Niedersachsens (OPAL). Arbeitsschwerpunkte: Digitale Quellenschließung und -edition, Geisteswissenschaften und Neue Medien, Texttheorie. E-Mail: sahle@sub.uni-goettingen.de

Dr. Georg Vogeler ist Wissenschaftlicher Assistent an der Ludwig-Maximilians-Universität, Historisches Seminar, Abteilung Geschichtliche Hilfswissenschaften; Arbeitsschwerpunkte: Amtsbücherkunde, Quellendigitalisierung, Digitalisierung von mittelalterlichen und frühneuzeitlichen Urkunden (CEI), Überlieferung der Urkunden Kaiser Friedrichs II. E-Mail: g.vogeler@lrz.uni-muenchen.de

NEW METHODS FOR THE ANALYSIS OF DIGITIZED MEDIEVAL LATIN CHARTERS

by Michael Gervers and Michael Margolin

Dramatic advances in information technology and significantly increased ease of access to the Internet have produced new and exciting research methods for historians. The DEEDS Project at the University of Toronto in Canada has built a corpus of about 9,000 securely dated medieval English charters which are accessible from the World Wide Web. This paper describes the application of advanced research and presentation methods to an analysis of scribal variations among a group of about 1,300, mostly undated, medieval Latin charters from The Cartulary of the Knights of St. John of Jerusalem in England. Research methods include a wide array of computerized textual analyses: full and segregated content searches, inline phonetic transformations, computational geometric representations, statistical processing of results, and the interactive graphical interpretations of those results.

Records of property transfer are as ancient as writing itself. Consequently, the deed, or conveyance, represents the most continuous form of legal documentation available to the historian. The current objective of the DEEDS Project, founded at the University of Toronto in 1975, is the provision of computerized access to the content of English conveyances of the twelfth and thirteenth centuries.¹ Our main priority at the moment is to

1 Gervers, Michael, The DEEDS Project: Towards the dating and analysis of English private charters of the twelfth and thirteenth centuries, in: *Le Médiéviste et l'Ordinateur* 41 (2002), S. 60-66. The internet version can be consulted at

develop the means to provide dates for private charters, 92 per cent of which, from the time of William the Conqueror in 1066 to the accession of King Edward II in 1307, were issued without chronological indicators.² The questions which can be addressed through the content of our databases are limitless, but to give a few examples of our research directions in recent years we may cite the following: 1) a study of the donations to the twelfth-century Order of the Hospital of St. John of Jerusalem which led to the conclusion that the Order became militarized in response to the fall of Edessa in 1144 and the call for the Second Crusade in 1145³; 2) changing forms of address appearing in grants to the Hospitallers in England during the twelfth and thirteenth centuries, which closely reflect the nature and degree of the Order's presence in, or absence from, the Holy Land⁴; 3) the relative growth of the estates of the military orders in London and its suburbs from the twelfth to fourteenth centuries⁵; and 4) evidence of social unrest in England under the Interdict (1208-1214).⁶ We are presently

<http://lemo.irht.cnrs.fr/41/mo41_07.htm> (12.10.2006). See also, Michael Gervers, "The Deeds Project and the Development of a Computerised Methodology for Dating Undated English Private Charters of the Twelfth and Thirteenth Centuries", in: Michael Gervers (Hg.), *Dating Undated Medieval Charters, Rochester-Woodbridge 2000*, pp. 13-35; Fiallos, Rodolfo "An Overview of the Process of Dating Undated Medieval Charters: Latest Results and Future Developments", in: idem, pp. 37-48.

2 Gervers, Michael, "The Dating of Medieval English Private Charters of the Twelfth and Thirteenth Centuries", in: Brown, Jacqueline; Stoneman, William P. (Hgg.), *A Distinct Voice. Medieval Studies in Honor of Leonard E. Boyle, O.P., Notre Dame (Indiana) 1997*, S. 455-504; idem, "Identifying Irregularities and Establishing Chronology in Medieval Charters", in: Keats-Rohan, Katharine S.B. (Hg.), *Resourcing Sources (Prosopographica et genealogica 7)*, Oxford 2002, S. 164-78.

3 Gervers, Michael, "Donations to the Hospitallers in England in the Wake of the Second Crusade", in: Gervers, Michael (Hg.), *The Second Crusade and the Cistercians*, New York 1992, pp. 155-61.

4 Gervers, Michael, "Changing Forms of Hospitaller Address in English Private Charters of the Twelfth and Thirteenth Centuries", in: Laszlovszky, József; Hunyadi, Zsolt (Hgg.), *The Crusades and the Military Orders: Expanding the Frontiers of Medieval Latin Christianity*, Budapest 2001, pp. 395-405.

5 Gervers, Michael, "The Commandery as an Economic Unit in England", in: Luttrell, Anthony; Pressouyre, Léon, *La Commanderie, Institution des ordres militaires dans l'Occident médiéval*, Paris 2002, pp. 245-50.

6 Gervers, Michael and Nicole Hamonic, "Pro Amore Dei: Diplomatic Evidence of

researching means to identify the medieval English scribe of the late twelfth through thirteenth centuries as an author.

Our Corpus of digitized medieval charters was started a decade ago using the Oracle version 5 database running on a 386-33 PC. At that time we had a set of seven main tables, namely DOCUMENT, PERSON, PROPERTY / COMPENSATION, LEASE, RELATION, LINKAGE and ROLE, with a variety of supporting code tables.⁷ Encoding information was achieved at the data-entry level by means of a linguistically based coding ‘language’. Three types of ‘sentences’ were used to describe connections between people, property, and people and property. Actions, tense, number and mode were similarly coded. It worked very well for the time.

However, today, the DEEDS Corpus, together with metadata, is hosted by the latest Oracle database and is widely available on the World Wide Web.⁸ The Corpus presently includes two groups of Latin charters, mainly from twelfth- and thirteenth-century England, derived either from printed sources or transcribed directly from the manuscripts (Figure 1).

Social Conflict During the Reign of King John”, in: Papers in honour of James A. Brundage. Publication anticipated in 2007.

7 Gervers, Michael; Long, Gillian; McCulloch, Michael, “The DEEDS Database of Mediaeval Charters: design and coding for the RDBMS ORACLE 5”, in: History and Computing, 1 no. 3 (1990), pp. 1-12.

8 <<http://www.utoronto.ca/deeds/research/research.html>> (01.10.2006).

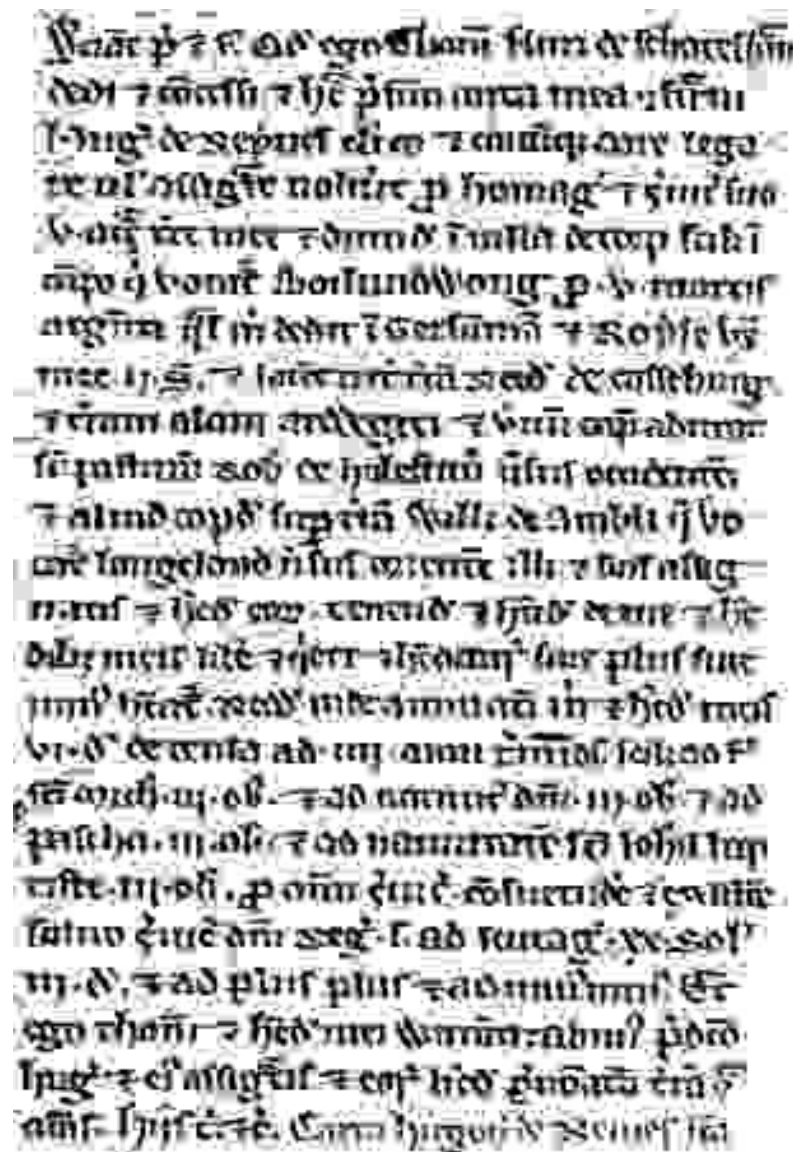


Figure 1: Charter Source: Manuscript

The first group, of about 9,500 charters, has been dated internally or by the editor of the manuscript, using internal evidence (Figure 2).

59.

[John de Haliwell grants to Reginald the Beadle land in Grove Street.]

Sciunt presentes & futuri quod ego Iohannes de Halliwelle coruisarius assensu & consensu Alicie uxoris mee & amicorum meorum dedi & concessi & quietumclamaui³ [&c.] pro me & heredibus meis Reginaldo seruienti Uniuersitatis Oxonie totum ius meum & clamamium [*sic*] quod habui vel quod habere potui [vel]⁴ quocumque modo possit mihi descendere de tota illa terra cum omnibus pertinentiis suis que est inter terram que quondam fuit magistri Hugonis de Seydegerd ex una parte & terram que quondam fuit Iohannis de Chelsee ex altera in Groppecuntelane in parochia beate Marie virginis Oxonie; habendum & tenendum predicto Reginaldo & heredibus suis vel suis assignatis libere & quiete, integre, bene & in pace, faciendo inde capitalibus dominis illius feodi seruicium debitum & consuetum. Pro hac autem donatione [&c.] predictus Reginaldus dedit michi pre

C. 6. 1
Mich.
1260-
Mich.
1261

¹ Peter de Kyllum was rector from 1248 to the end of 1274.

² *Bracton*, iv. 367 (Rolls Series) speaks of an important lawsuit by the rector of St. Mary's who was next in succession after William Hardel; this was Peter de Kyllum. Bracton's language rather suggests that William Hardel had treated the land of the rector as his private property and had

sold it. Bracton says the case was tried before the King at Woodstock, and as it is generally assumed that he refers to nothing later than 1254, our deed would be of 1248-1254.

³ The use of this word shows that Reginald already had a rent or some other share in the tenement.

⁴ Not in MS.

manibus centum solidos sterlingorum in gersummam in magno negocio meo. Ut igitur [&c. sealing], hiis testibus, Adam Feteplace tunc maiore Oxonie, Galfrido Aurifabro, Ricardo filio Nicholai tunc balliuis, Willelmo de Wintonia espeicer, Willelmo le Speicer iuniore, Thoma le Speicer, Willelmo de Eu, Henrico Inge, Ricardo Marescallo, Willelmo Russello cordewanario, Henrico Barbatore, Reginaldo Illuminatore & aliis.¹

¹ This deed and the next refer to a property mentioned in deeds 141 and 142, lying on the east side of Bedel Hall.

Figure 2: Charter Source: Published Cartulary Copy

```
- <sourceDocument dnum="00930059" created="2003-08-19" cartulary="Oriol College - Oxford">
  <content>Sciunt presentes et futuri quod ego Iohannes de Halliwelle coruisarius
  assensu et consensu Alicie uxoris mee et amicorum meorum dedi et concessi et
  quietumclamaui etc pro me et heredibus meis Reginaldo seruienti Uniuersitatis
  Oxonie totum ius meum et clamamium quod habui vel quod habere potui [vel]
  quocumque modo possit mihi descendere de tota illa terra cum omnibus
  pertinentiis suis que est inter terram que quondam fuit magistri Hugonis de
  Seydegerd ex una parte et terram que quondam fuit Iohannis de Chelsee ex
  altera in Groppecuntelane in parochia beate Marie virginis Oxonie habend et
  tenend predicto Reginaldo et heredibus suis vel suis assignatis libere et quiete
  integre bene et in pace faciendo inde capitalibus dominis illius feodi seruicium
  debitum et consuetum Pro hac autem donatione etc predictus Reginaldus dedit
  michi pre manibus centum solidos sterlingorum in gersummam in magno
  negocio meo Ut igitur etc hiis testibus Adam Feteplace tunc maiore Oxonie
  Galfrido Aurifabro Ricardo filio Nicholai tunc balliuis Willelmo de Wintonia
  espeicer Willelmo le Speicer iuniore Thoma le Speicer Willelmo de Eu Henrico
  Inge Ricardo Marescallo Willelmo Russello cordewanario Henrico Barbatore
  Reginaldo Illuminatore et aliis</content>
+ <data>
+ <notes>
+ <markup>
</sourceDocument>
```

Figure 3: Digital Representation of Medieval Charter

The second group, of about 1,300 mostly undated charters, comes from The Cartulary of the Knights of St. John of Jerusalem in England (British Library, Cotton ms. Nero E vi).⁹ In the computer, they are stored independently, but both groups are available for computerized analyses. Each charter is stored in digitized form, together with a photographic image of the original printed version, or of the manuscript itself. The text and any accompanying information is extracted from the original source by scanning, using the Optical Character Recognition Program (OCR) or by transcription directly from the manuscript. We have developed a program which we call "Document Manager" to convert the text of the charter, and all external and analytical data, into an Extensible Markup Language (XML) document which is later sent to the database (Figure 3). Each document stored in the database is the source for the searchable online

⁹ Gervers, Michael, *The Hospitaller Cartulary in the British Library (Cotton MS Nero E VI). A Study of the Manuscript and its Composition, with a Critical Edition of Two Fragments of Earlier Cartularies for Essex*, Toronto 1981; idem, *The Cartulary of the Knights of St. John of Jerusalem in England: Essex*, 2 vols., London 1982-96.

version of that charter in Hyper Text Markup Language (HTML).

There are several ways to access DEEDS digitized data:

1. Remotely: over the Internet, using File Transfer Protocol (FTP), connecting directly to the DEEDS database server, using Dedicated Web Services
2. Locally: from workstations that have DEEDS applications installed on their hard drive.

Charters can be viewed over the Internet using the “Browse” option on the DEEDS Website. This will start the “Document Browser” (Figure 4) program which allows one to select the text of a particular charter, view the original source, and the title page of the printed edition. The source cartulary can be chosen by name or by its internal code. In addition, the “Document Browser” displays the diplomatic parts of the charter and formulae, marked up according to a predefined color legend.

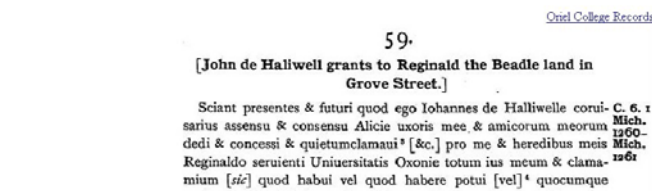
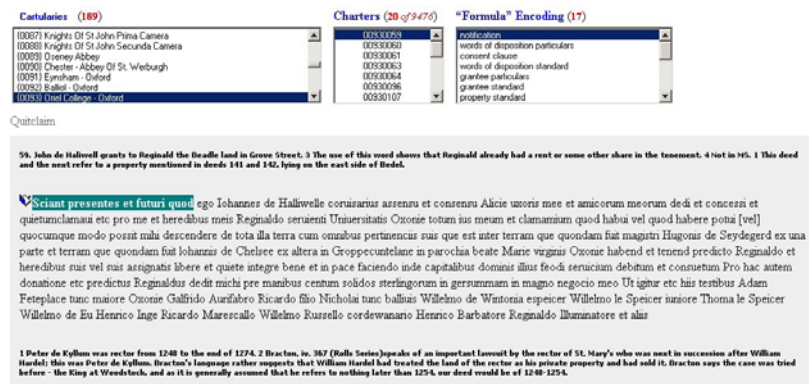


Figure 4: Document Browser

A second online program permits textual queries to be directed to the chosen collection of charters. This can be started from the DEEDS website by selecting the “Search” option. This program supports exact, fuzzy and proximity text patterns. There are also multiple options for further refining the query by specifying meta data restrictions, such as the nature of the charter, the time span, the source of the text, the principal individuals appearing in it, etc. The program spans two panels. The first displays query restrictions while the second renders a search result in one of the three available modes:

1. „Search“,
2. „Context“,
3. „Chart“.

When the “Search” mode is selected, the lower panel will display a list of charters together with their dates. Any charter from this list can be viewed with a highlighted query term (Figure 5). The same view also allows one to find a word or phrase in the current document.



Figure 5: Query Term Search

When the “Context” mode is selected, each query term will be shown within its context. The context boundaries can be expanded from three words before and after the queried pattern, up to a maximum of ten words. These context words can be alphabetized independently, starting from the word nearest to the query term and moving either forwards or backwards

(Figure 6).

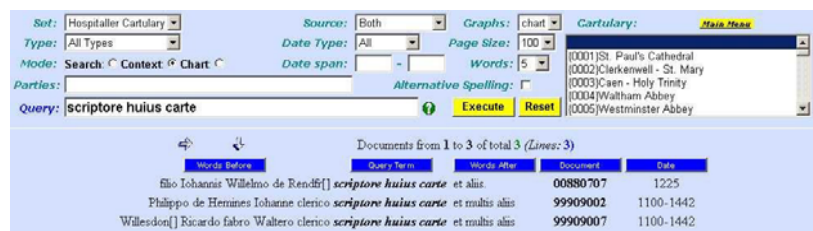


Figure 6: Query Term Context

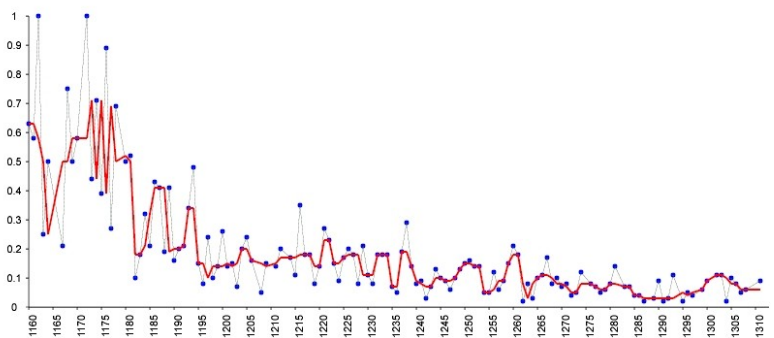
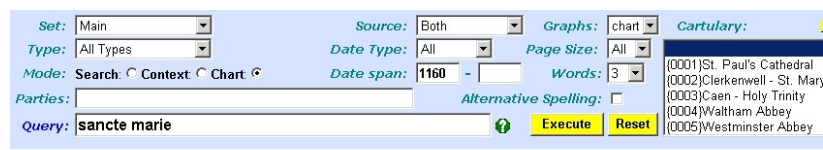


Figure 7: Graphical Output

When the “Chart” mode is selected the lower panel will display a graphical representation of the query using Scalable Computer Graphics (SVG) format. SVG is an XML dialect and therefore is portable across different hardware platforms and operating systems. An additional advantage of SVG is that all graphics can be scaled to any degree without loss of image quality and direct user interaction is allowed (Figure 7). There is also one

more advanced option called “Alternative Spelling”. When this option is enabled some common spelling variations, like double consonants, are removed.

We have recently successfully applied our online facilities to explore scribal references in The Cartulary of the Knights of St. John of Jerusalem in England, in an attempt to identify authorship, or “schools” of writing (Figure 8).

Several more programs have been developed to work with digitized charters but, for now, they are only accessible from the DEEDS departmental workstations. We use these programs for the maintenance of digitized documents and for performing chronological and content-driven textual analysis. However, it is our intention to make those programs also accessible remotely over the Internet.

Document	date	name	Location (titulus)	Phrase	
2	99909001	undated	Adam diaconus de Angemare	Sumpthng	presentis scripti notario
3	00880294	1180 c.	Adam scriptor	Chaureth - Chrishall	scriptore
4	99909008	undated	Edmundus	Chippenham	qui hanc cartam scripsit
5	00870019	1190 c.	Hamelinus clericus	Roydon	qui hanc cartam fecit
6	00880046	1200-1204	Hugo capellanus	Cressing-Witham: Hatfield Peve	presentis carte scriptore
7	00880394	1155 c.	Hugo scriptor	Chaureth - Henham	scriptore
8	99909004	1257-1258	Willelmus capellanus	Hampton	qui hanc cartam composuit
9	99909002	undated	Iohannes clericus	Hampton	scriptore huius carte
10	99909009	undated	Iohannes de Shryneham clericus	Chippenham	qui presens scriptum composuit
11	99909000	undated	Magistro Philippo medico	Ravensthorpe	qui hanc cartam scripsit
12	00870095	1269	Nicholaus clericus	Rainham	huius scripti notario
13	99909003	undated	Petrus capellanus de Codham	Hampton	qui hanc cartam scripsit
14	00880541	1235-1245	Petrus capellanus de Hastede	Gestingthorpe	qui hanc cartam scripsit
15	00870103	1280 c.	Radulphus de Boughton	West Thurrock	presencium scriptore
16	00880698	1245 c.	Reginaldus clericus	Gestingthorpe	qui hanc cartam fecit
17	00880734	1230 c.	Robertus de Stistede	Gestingthorpe	qui hanc cartam scripsit
18	00880940	1255 c.	Simon clericus	Sampford	qui hanc cartam fecit
19	99900274	1219	Stephanus de Esseleia	London	qui hoc scripsit
20	99909006	undated	Thomas clericus	Hampton	qui hanc cartam scripsit
21	00880319	1190-1200	Walterus	Chaureth - Sawbridgeworth	qui cartam scripsit
22	00880523	1230 c.	Walterus clericus	Bumpstead Helion	huius scripti scriptore
23	99909007	undated	Walterus clericus	Hampton	scriptore huius carte
24	00880538	1185 c.	Walterus de Hull clerico magistri	Bumpstead Helion	qui hanc cartam scripsit
25	99909005	undated	Willelmus capellanus	Hampton	qui hanc cartam scripsit
26	00880707	1225	Willelmus de Rendfr[Gestingthorpe	scriptore huius carte
27	00880155	1220-1230	Willelmus de Salseton	Little Maplestead	qui scripsit hanc cartam
28	00880512	1230-1240	Willelmus filius Derkini	Bumpstead Helion	qui hanc cartam scripsit
29	99900283	undated	Willelmus filius Derkini	Herefeld	qui hanc cartam scripsit

Figure 8: Hospitaller Cartularies Scribal References

We use a set of securely dated charters to draw any quantitative conclusions with respect to chronology or content. Obviously, the number of dated charters available for each time span varies, as does the accuracy of the chronological evidence, which varies from the exact day, month and

year to a range of several years. To overcome discrepancies in chronological attributes, we have developed a method of normalizing chronological information in order to facilitate computer-aided analyses.

We convert the attributed date to the Julian calendar and then apply a special computer program to analyze the availability of charters over a given time span (in our case from 1050 to 1359). This program breaks up the total chronological span into periods when approximately equal numbers of dated charters are available. An index of this computer-generated period (henceforth referred to as a “tile”) is subsequently assigned to each charter. The tile value is later used as a proxy for the original date in all statistical computations. The size of the tile can vary from just one day to several years, while the number of charters referenced remains almost unchanged. The tile number assigned to the charter is automatically updated when charters are added to, or removed from, the collections. Both the actual and normalized availability of our first group of charters is shown in Figure 9.

We can apply computer-aided analysis to the charter text in order to evaluate the chronology and content. Since this is an official legal document written or issued by a religious, lay or royal institution, inevitably its vocabulary, structure and content reflect the time of its creation and also carry the “footprint” of the issuer. The core of our approach is that, by analyzing different aspects of the text, a link can be established between the given text and a set of similar charters whose attributes are known. The fact that all digitized charters are well attributed and are linked to a significant amount of meta data allows us to conduct computerized textual analyses from different angles, using the meta data for filtering information. This meta data includes elements such as: the type of legal action documented, information about the parties represented in the charter, details concerning the place of issue and locations, the name of the scribe, etc. One of its most important features is identification of the different diplomatic parts and any content-related formulae in the text of the charter. This information, which is currently available for about 50 per cent of our charters, can be accessed under the “Browse” option on our website menu.

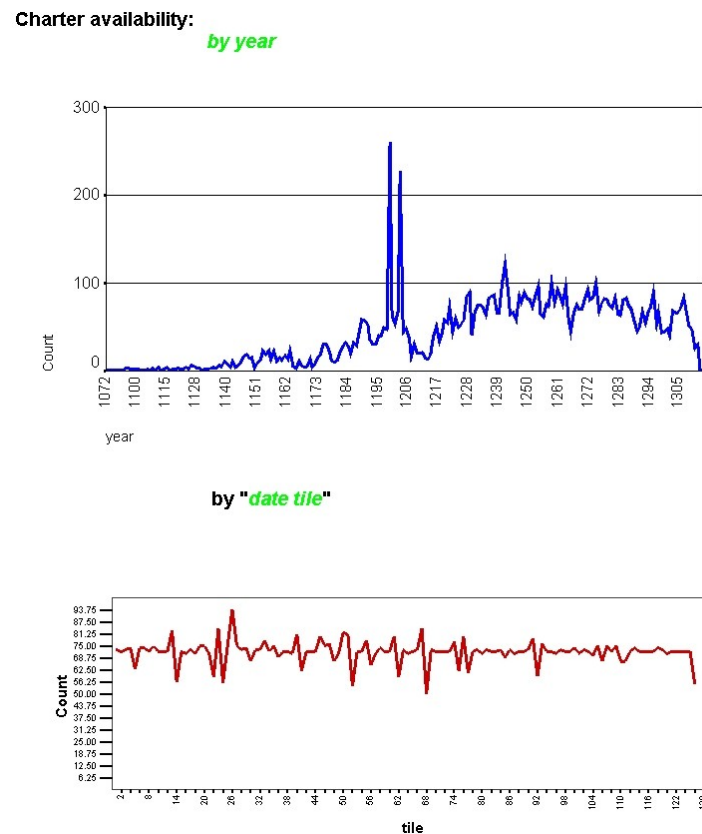


Figure 9: DEEDS Corpus Charters

The examination of vocabulary involves extracting combinations of two or more adjacent words, so-called “word-patterns”, in consecutive order from the text of the charter and then finding occurrences of each pattern in a collection of digitized charters. The total number of valid word-patterns derived from a typical charter varies from hundreds to thousands, depending on the size of the text. All attributes of charters in which a given query term occurs are collected and later processed by the

computer program. Depending on the circumstances, word-patterns can be generated in three different ways:

1. from the original text
2. from the partially normalized text
3. from the fully normalized text.

Normalization of the text (Figure 10) can include lexical transformations, phonetic transformations, or both. Partial normalization involves lexical transformation and includes the replacement of most Roman and Arabic numerals, prepositions, definite articles and measurement units by a one-letter proxy in the text of the document. Full normalization adds phonetic transformation by further replacing words by a phonetic proxy. We tested different phonetic matching algorithms¹⁰ and eventually decided to use the modified version of the Lawrence Philips' Double Metaphone algorithm for the phonetic transformations (Figure 11).¹¹

Sciant presentes et futuri quod ego Iohannes de Halliwelle coruisarius assensu et consensu Alicie uxoris mee et amicorum meorum dedi et concessi et quietumclamaui etc pro me et heredibus meis Reginaldo seruianti Uniuersitatis Oxonie totum ius meum et clamamium quod habui vel quod habere potui [vel] quocumque modo possit mihi descendere de tota illa terra cum omnibus pertinenciis suis que est inter terram que quondam fuit magistri Hugonis de Seydegerd ex una parte et terram que quondam fuit Iohannis de Chelsee ex altera in Groppecuntelane in parochia beate Marie virginis Oxonie habend et tenend predicto Reginaldo et heredibus suis vel suis assignatis libere et quiete integre bene et in pace faciendo inde capitalibus dominis illius feodi seruicium debitum et consuetum Pro hac autem donatione etc predictus Reginaldus dedit michi pre manibus centum solidos sterlingorum in gersummam in magno negocio meo Ut igitur etc his testibus Adam Feteplace tunc maiore Oxonie Galfrido Aurifabro Ricardo filio Nicholai tunc ballius Willelmo de Wintonia espeicer Willelmo le Speicer iuniore Thoma le Speicer Willelmo de Eu Henrico Inge Ricardo Marescallo Willelmo Russello cordewanario Henrico Barbatore Reginaldo Illuminatore et aliis

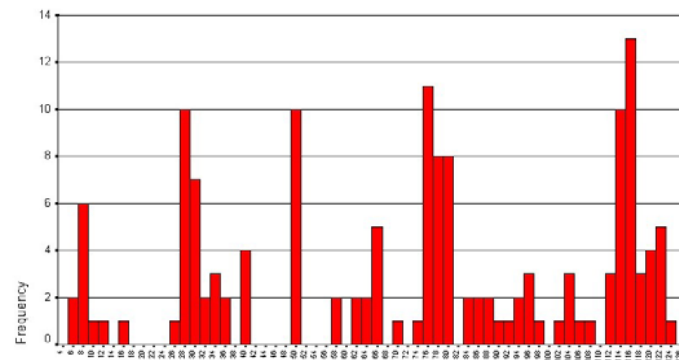
Figure 10: Original Text

- 10 Zobel, Justin; Dart, Philip, *Phonetic String Matching: Lessons from Information Retrieval*, <<http://goanna.cs.rmit.edu.au/~jz/fulltext/sigir96.pdf>> (12.10.2006).
- 11 Philips, Lawrence, *The Double Metaphone Search Algorithm*, *C/C++ Users Journal*, June 2000,, <<http://www.ddj.com/dept/cpp/184401251>> (12.10.2006).

scent prsent cj futr cj eg iohns pr hall corsr asns cj consn als uxrs me cj amkrm merm ded cj consc cj qutmk etk pr me cj herdp mes rejnl sernt unfrs oxn totm ius mem cj clnmm cj hap cj cj hapr pot cj qukmk mod post mih desnd pr tot il ter pr ommps pbrtn sus qu est pr term qu qundm fut majst hukns pr sedjr pr nm part cj term qu qundm fut lohns pr cels pr altr pr grpkn pr pars bet mar virjn oxn hapnd cj tennd prdkt rejnl cj herdp sus cj sus asnts lipr cj qut inkr ben cj pr pas fasnd ind captl domns ils fed serm deptm cj const pr hak cj donxn etk prdkt rejnl dedt mih pr manps nm nt strln pr gersm pr man nekt me cj igrtr etk his testp adm fetpl tunk mar oxn galfr aurfp rikrd fil nkl tunk balfs willm pr wintn espsr willm le spsr iunr tom le spsr willm pr eu henrk inj rikrd marsk willm rusl cordn henrk barpt rejnl ilmnt cj als

Figure 11: Transformed Text

#00930059 Gregorian date: 1260-61, Julian Date: 2181727, Tile: 79



```
1 hit
Total number of hits: 17
Date Files: 3 70 91 60 16 52 69 1 58 76 112 104 80 120 79 85 64
Statistics: 76

1..25 hits
UP size: 1: 75.7143
UP size: 2: 82
UP size: 3: 73.4022
UP size: 4: 75.1224
UP size: 5: 94.0056
UP size: 6: 95
UP size: 7: 94.5
UP size: 8: 90
Statistics: 03.0101

Total number of hits: 111600
```

Figure 12: Word-Pattern Frequencies Over Time

We process the results that are returned by word-pattern queries in separate flows, depending on their nature, amount and quality:

1. When a word-pattern produces more than 25 hits, each such distribution is evaluated independently.
2. When a word-pattern produces between one and 25 hits, results are first tabulated using all those of the same number and later evaluated using one for each word-pattern size.
3. All word-patterns which produce a single hit are accumulated and processed just once at the end of the process.

We accumulate intermediate results at various stages of the processing. All intermediate and final computations are performed by the integrated Statistical Engine (Figure 12).

Frequently, when word-patterns generate multiple hits, they produce a distribution with peaks in multiple clusters. In this case the correct local distribution can be chosen by using results produced by the other two flows (Figure 13). The addition to the program of an industrial strength statistical engine has enabled us greatly to improve the quality of calculations through all stages of word-pattern processing. We apply Robust Statistical Routines and Fibonacci ratios base weighting to achieve our final results. An example of output (Figure 14) shows the accuracy of our fully-automated chronological evaluation of a sample of 100 charters using original word-patterns.

M-Estimators

	Huber's M-Estimator	Tukey's Blweight	Hampel's M-Estimator	Andrews' Wave
V2	77.1410	77.1700	77.2130	77.1694

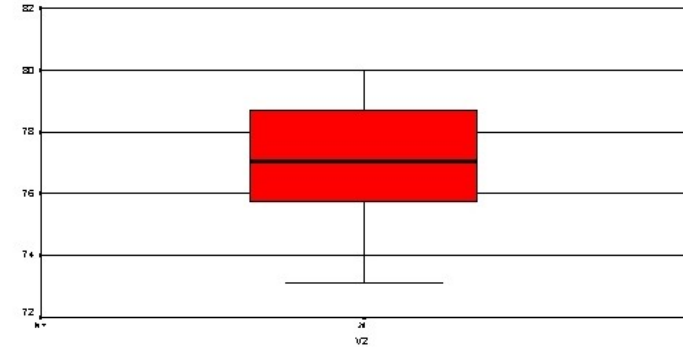


Figure 13: Local Distribution

Average Error

	Huber's M-Estimator	Tukey's Blweight	Hampel's M-Estimator	Andrews' Wave
MIN	3.1655	3.1831	3.2411	3.1932
AVG	5.54	5.27	5.46	5.27

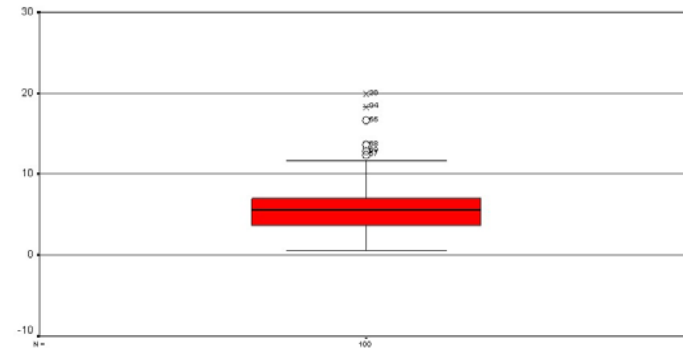


Figure 14: Charter Chronological Evaluations Test

Using computer-generated geometrical content footprints during charter analysis appears to be a promising way of comparing charter content. According to this method, original text is converted to digital form by replacing each letter, including spaces, by their arithmetic ASCII value. After that, the program computes the frequencies of each value. At the next stage, a Computational Geometry program treats frequency/character value pairs as objects of virtual two-dimensional space. Later the program extracts a series of convex polygons (called “convex hulls”) starting with the largest, using Onion Peeling Algorithms.¹² The smallest hull generated by the program is the one eventually used for a content representation of the text. Generally speaking, an overlap of hulls generated from different charters reflects the similarity of their content (Figure 15). The current implementation of our program for content footprint evaluation uses the latest release of the Computational Geometry Algorithms Library (CGAL).¹³

12 Poulus, Marios; Papavlasopoulos, Sozon; Chrissilopoulos, Vasilios, A Text Categorization Technique based on a Numerical Conversion of a Symbolic Expression and Onion Layers Algorithm, in: *Journal of Digital Information*, 6/1 (2004), article no. 276.

13 Computational Geometry Algorithms Library (CGAL), <<http://www.cgal.org>> (01.10.2006).

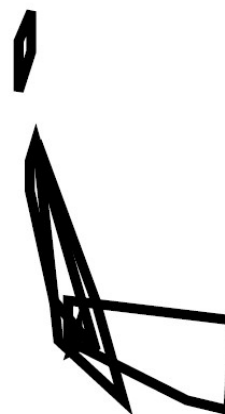


Figure 15: Geometrical Content Footprint

To sum up, this paper has introduced the DEEDS Project approach to the textual analysis and management of a collection of digitized Medieval Latin charters. All our digitized documents are encoded using XML and, although we use our own naming rules for XML elements and attributes, the structure of encoded charter documents can easily be exported to such encoding systems as the Text Encoding Initiative (TEI) and the emerging Charter Encoding Initiative (CEI).¹⁴ We have developed a different technical approach to encoding the metadata and physical placement of the Markup Language tags into the text. At the storage level, we store separately an original text of the charter, encoding information and meta data information. Generally speaking we use dynamic encoding as an alternative to the more commonly used static system. This means that our system is capable of generating a variety of statically encoded documents in response to a specific client request. By employing dynamic encoding we have also eliminated the problem of overlapping markup elements.¹⁵

14 Vogeler, Georg, Towards a standard of encoding medieval charters with XML, in: *Literary and Linguistic Computing* 20 (2005), S. 269-280.

15 Gervers, Michael, “Application of Computerized Analyses in Dating Procedures for Medieval Charters”, in: *Le Médiéviste et l’Ordinateur* 42 (2003), S. 7-25. Co-

We have also discussed how we apply new methods to vocabulary and content analyses of Medieval Latin texts using “word-patterns” and Computational Geometry. Tight integration of a Statistical Engine into the process has allowed us to improve significantly the accuracy and reliability of our output results, and the use of analytical online tools like our “Browse” and “Search” programs can greatly facilitate the research. The application of Scaleable Vector Graphics as an alternative to commonly used Bitmap pictures has improved considerably the quality of graphical information available to Internet users.

For the future, we see great potential in the further development of word-pattern frequency analyses by using multiple mutually independent data flows. As we are always dealing with inexact statistical data, the search for alternative statistical methods remains one of our main priorities. We are currently working on the integration of the vocabulary and content evaluation tools into our website and we continuously seek to improve its appearance and functionality.

Michael Gervers is Professor of History at the University of Toronto. His research interests include medieval diplomatics, codicology, palaeography, chronology, charter analysis, textual editing and the dating of medieval documents. E-Mail: gervers@chass.utoronto.ca

Michael Margolin is Research Programmer at the DEEDS Project, University of Toronto. His research interests are quantitative analyses and digitization of medieval charters for the purpose of establishing chronological and authorship attribution. E-Mail: m.margolin@utoronto.ca

DAS MONASTERIUM-PROJEKT – EIN BEISPIEL FÜR KOLLABORATIVE UND INKREMENTELLE EDITIONEN

von Karl Heinz

Es ist die Absicht der folgenden, auf einem Vortrag auf der Berliner .hist 2006 Tagung basierenden Darstellung, das Monasterium-Projekt in seinen grundsätzlichen Strukturen vorzustellen und die Verwertungsmöglichkeiten der Projektergebnisse aufzuzeigen. Im Einzelnen kommen dabei die grundsätzlichen Ziele des Monasterium-Projekts zur Sprache, wobei die Darstellung der projektinternen Arbeitsweise eine Art Werkstattbericht geben wird. Neben diesen Grundlagen wird die Vereinigung des Monasterium-Konsortiums vorgestellt und in einem weiteren Punkt die verbesserten und neuen Forschungsmöglichkeiten für die Geschichtswissenschaft verdeutlicht. Zuletzt wird noch ein Ausblick auf die kommenden zwei bis drei Jahre versucht, indem die Perspektiven des Monasterium-Projekts erörtert werden.

Einleitung

Die Erfahrungen der letzten Jahrzehnte haben gezeigt, dass klösterliche Urkundenbestände – sei es in Österreich oder in anderen Staaten Europas – zwar vorhanden, für eine intensive wissenschaftliche Nutzung aber nur in ungenügendem Ausmaß vorbereitet sind. Die Gründe dafür sind mannigfaltig. In vielen Klosterarchiven sind Behelfe wie Findbücher oder Repertorien Mangelware. Wenn diese vorhanden sind, dann verfügen sie über ein ansehnliches Alter. Die Personalstruktur und die budgetäre Lage der Klöster sind ebenfalls nicht dazu geeignet, einen regelmäßigen Benutzerdienst zu installieren. Ein Besuch in Klosterarchiven muss meist von

langer Hand vorbereitet werden, sofern das Archiv nicht überhaupt für die öffentliche Nutzung gesperrt ist. Hinzu kommt, dass oft gut ausgebildetes Personal fehlt.

Ziele des Monasterium-Projekts

Angesichts der großen historischen Bedeutung des in Frage stehenden Archivmaterials möchte das Monasterium-Projekt hier Abhilfe schaffen; im Einzelnen hat es sich folgende Ziele gesetzt: Grundsätzliche Absicht ist die Digitalisierung mittelalterlicher Urkundenbestände geistlicher Provenienz sowie die wissenschaftliche Aufbereitung derselben unter Schaffung einer umfassenden und recherchierbaren Datenbank. Die einzelnen Digitalisate werden mit den entsprechenden Metadaten verknüpft und über die Monasterium-Seite¹ im Internet publiziert. Das sichtbare Ergebnis der Bemühungen ist die Schaffung eines kostenfreien und möglichst benutzerfreundlichen Zugangs zu dieser erwähnten Form des historisch-kulturellen Erbes. In einer ersten Projektstufe waren es die Bestände der Stifte und Klöster des Landes Niederösterreich, welche – gleichsam als ein Pilotprojekt – zwischen den Jahren 2002 und 2005 aufgearbeitet worden sind.

Abwicklung und Finanzierung

Zur organisatorischen und fachlichen Abwicklung des Projekts wurde im Jänner 2002 das Institut zur Erschließung und Erforschung kirchlicher Quellen (IEEkQ) als Trägerverein des Projekts mit Sitz in St. Pölten gegründet.²

¹ <<http://www.monasterium.net>>.

² <<http://www.monasterium.net/at/ieekq.html>>.



Abbildung 1: Startseite von Monasterium.net

Die erste Stufe des Monasterium-Projektes ist auf drei Jahre (Mitte 2002 bis Mitte 2005) anberaumt worden, wobei sich die finanziellen Mittel einerseits aus den Beiträgen aus öffentlicher Subventionierung andererseits aus den Zahlungen der Ordensstifte Niederösterreichs und Wiens zusammensetzten. Zu den Geldgebern der öffentlichen Hand gehören das Bundesministerium für Bildung, Wissenschaft und Kultur (BMBWK), das Kulturredesamt des Landes Niederösterreich, die Stadt Wien, die Landeshauptstadt St. Pölten, die Stadtgemeinde Ybbs, die Stadtgemeinde Geras, die Marktgemeinde Schönbühel-Aggsbach, und die Marktgemeinde Heiligenkreuz. Neben diesen Körperschaften steuerten auch die noch bestehenden Stifte und Klöster Niederösterreichs ihren Anteil bei. Es handelte sich um die Benediktinerabtei Altenburg (NÖ), das Prämonstratenser-Chorherrenstift Geras (NÖ), das Benediktinerstift Göttweig (NÖ), das Zisterzienserstift Heiligenkreuz (NÖ), das Augustiner-Chorherrenstift Herzogenburg (NÖ), das Augustiner-Chorherrenstift Klosterneuburg (NÖ), das Zisterzienserstift Lilienfeld (NÖ), das Benediktinerstift Melk (NÖ), die Benediktinerabtei Schotten (W), das Benediktinerstift Seitenstetten (NÖ) und das Zisterzienserstift Zwettl (NÖ).

Arbeitsweise

Integration und Vernetzung bereits bestehender Informationen

In einem ersten Arbeitsschritt wurden bereits bestehende Urkundeneditionen, sei es im Volltext oder als Regest, eingescannt und mittels einer Optical Character Recognition (OCR) Software in eine Textdatei konvertiert. Es dauerte ungefähr ein halbes Jahr (Juli bis Dezember 2002), bis das verwendete Programm *ABBYY FineReader*³ durch eine Vielzahl an Tests die Retrokonversion mit einer Fehlerquote von unter einem Prozent bewerkstelligen konnte. Die so gewonnenen Texte wurden einem gründlichen Lektorat unterzogen, um die verbliebenen Fehlesungen zu eliminieren. Es hat sich gezeigt, dass ein großer Teil der Lesefehler, die nach der Konversion verblieben sind, auf Unschärfen und Unsauberkeiten im Druck zurückzuführen waren und immer ähnliche Buchstabenkombinationen betroffen haben. Nach diesen Fehlern konnte maschinell gesucht werden, was eine große Zeitersparnis bedeutete.

Kodierung

Der nächste Schritt betraf die Vorbereitung der Texte für das World Wide Web. Um die Texte auch im Internet darstellbar zu machen, wurden sie mit HTML-Markierungen versehen. Diese Markierungen beschränkten sich auf die Kennzeichnung von Originalschreibweisen (<i>) und von Superskripten (<sup>). Diese Auszeichnungsarbeit konnte zum Teil automatisiert unter Zuhilfenahme der Software Lotus Word Pro durchgeführt werden, wobei allerdings auch hier ein sorgfältiges Lektorat unentbehrlich war. Die so korrigierten und präparierten Texte werden nun in eine Datenbank gespielt, wobei die Inhalte von Hand in die entsprechenden Felder der Datenbankmaske kopiert wurden.

Indizierung

Besonders wichtig für die Benutzbarkeit des im Monasterium-Projekt bereitgestellten Materials ist, dass die Urkunden auch über einen Index erschließbar sind. Sinn der Indizierung ist es, die Recherchemöglichkeiten nicht nur auf die Volltextsuche zu beschränken, sondern ein Hilfsmittel

beizusteuern, das genauere Abfragen ermöglicht. Jede einzelne Urkunde wird mit einem Indexbestand versehen, wobei die Indexbegriffe sechs Hauptgruppen zugeordnet werden. Diese Gruppen umfassen die Orts- und Personennamen, Sachbegriffe, Stifte und Klöster, Pfarrkirchen, Pfarren und Topografisches. Jeder Indexeintrag besteht aus einem Haupt- und einem Unterindexfeld. Im Hauptfeld wird der Begriff zunächst in der Originalschreibweise der Urkunde erfasst. Im Unterfeld erfolgt dann eine Spezifizierung des Begriffes. Im Fall von Orten werden diese identifiziert und die Ortsnamen aufgelöst. Die Zuordnung der Orte wird durch die Zuweisung zu einem bestimmten Gerichtsbezirk eindeutig ausgewiesen. Personennamen werden der modernen Schreibung entsprechend standardisiert. Allgemein bekannte Persönlichkeiten, wie Kaiser/Könige, Herzöge, Päpste, Bischöfe, Äbte etc. werden zusätzlich mit ihrer Funktion (Titel) und ihren Eckdaten versehen. Positionen wie Aussteller/innen, Empfänger/innen, Siegler/innen oder Zeugen/innen werden im Speziellen als solche gekennzeichnet, um nach ihnen gezielt suchen zu können. Dabei kann auch nach bestimmten personellen Kombinationen recherchiert werden, womit beispielsweise bestimmte Siegelpartnerschaften per Knopfdruck nachweisbar werden. Juristische Sachbegriffe werden im Unterindexfeld in ihrer Schreibung normalisiert und in kürzester Form erklärt.

Der große Vorteil der Indexrecherche liegt nun darin, dass man nach Begriffen suchen kann, die im Volltext nicht vorkommen. So ist es beispielsweise möglich, nach Orten mit ihrer heutigen Bezeichnung und nach Personen in ihrer modernen Namensschreibung zu suchen. Das Ergebnis bietet den Benutzern/innen gleichzeitig auch eine Mindestinformation über den gesuchten Begriff. Die Möglichkeit nach Originaltextstellen zu suchen bleibt natürlich unbenommen.

Neubearbeitungen

Eine Reihe von großen und wichtigen Beständen ist bis heute entweder ungenügend oder überhaupt nicht dokumentiert. In diesen Fällen ist das Monasterium-Projekt bestrebt, Neubearbeitungen in Form von Vollregesten zu erarbeiten. Klöster, die auf diese Art und Weise bearbeitet werden, sind das Benediktinerstift Melk, das Augustiner Chorherrenstift Kloster-

3 <http://www.abbyy.com/finereader_ocr/>.

neuburg sowie das Zisterziensertift Zwettl.

Bilddigitalisierung

Ein weiterer Projektschwerpunkt besteht in der Erstellung von farbigen Urkundendigitalisaten, welche mit Hilfe eines transportablen Scanners erstellt werden, da es ausgeschlossen war, die zu erfassenden Urkunden aus den Archiven zu entfernen. Die Urkunden wurden vor Ort in den jeweiligen Stiftsarchiven mit einer Auflösung von 400 dpi gescannt. Die Farbdigitalisate wurden als Masterfiles im TIFF-Format abgespeichert und dienten als Grundlage für alle weiteren Verwendungen. Aus diesen Masterfiles wurden zwei Qualitäten von JPEG-Bildern – eine hoch aufgelöste Form für die interne Verwendung und eine Variante mit niedrigerer Auflösung für die Darstellung im Netz – sowie Voransichtsbilder generiert. Diese Bilder wurden dann mit dem entsprechenden Datensatz verbunden, um sie gleichzeitig mit dem Text darstellbar zu machen.

Präsentation

Von Beginn an war ein Fokus des Projekts darauf gerichtet, mit seinem Angebot ein möglichst breites Publikum zu erreichen. Es erwies sich in diesem Zusammenhang als ausgesprochen förderlich, nach Absolvierung der oben angeführten Arbeitsschritte als letzte Stufe eine öffentliche Präsentation in den jeweils betroffenen Klöstern zu veranstalten. Bei diesen Gelegenheiten werden einerseits der Urkundenbestand vorgestellt als auch der praktische Umgang mit der Datenbank erläutert.

Das Monasterium-Konsortium

Schon kurz nach Projektbeginn führte der Einblick in die niederösterreichischen Urkundenbestände zu der Erkenntnis, dass die Aufarbeitung derselben nur ein erster kleiner Schritt sein konnte, dem noch mehrere Projektphasen zu folgen hätten. Es zeigte sich deutlich, dass die geistlichen Institutionen in ihren Beziehungen und Kontakten nicht nur die Grenzen des Landes, sondern auch die heutigen Staatsgrenzen Österreichs weit hinter sich ließen und Teil eines gut funktionierenden mitteleuropäischen Netzwerkes waren.

Aus diesem Umstand heraus wurde das Monasterium-Konsortium als

eine lose Arbeitsgemeinschaft von Archiven, Bibliotheken, wissenschaftlichen Instituten und sonstigen Kulturerbe-Institutionen ins Leben gerufen. Alle diese Einrichtungen arbeiten gemeinsam an der Verwirklichung einer mitteleuropaweiten Urkunden-Datenbank, wobei dafür folgende vertraglich festgelegte Rahmenbedingungen gelten:

- Entwicklung, Ausbau und Einhaltung gemeinsamer technischer und wissenschaftlicher Standards
- Nutzung von Synergien und gemeinsamen Schnittstellen in den Bereichen Öffentlichkeitsarbeit, Projektmanagement, Marketing, Digitalisierung, Know-how etc.
- Mehrsprachigkeit: gleichberechtigte Verwendung der Sprachen der Partnerstaaten – mehrsprachige Darbietung von Ressourcen (Texte, Indizes, Regesten)
- Wahrung der Eigenständigkeit jeder einzelnen teilnehmenden Institution: Nutzung der Monasterium-Plattform soweit dies nötig ist unter Wahrung jeglicher Urheber- und Verwertungsrechte an den unter Monasterium.net bereitgestellten Bild- und Textmaterialien.

Im Interesse der Wahrung einer möglichst großen Flexibilität jeder teilnehmenden Institution gibt es die Möglichkeit, auf verschiedenen Ebenen am Monasterium-Konsortium teilzunehmen:

- Koordinator: Organisation und Vernetzung des Gesamtprojektes, Bereitstellung technischer und organisatorischer Infrastruktur (derzeit: Institut zur Erschließung und Erforschung kirchlicher Quellen, St. Pölten/A)
- Arbeitsgruppe (aktiv): selbstständige wissenschaftliche und technische Erschließung von Beständen sowie Bereitstellung und Verwaltung von Daten (Texte/Bilder)
- Content Provider (passiv): Bereitstellung von Beständen zur digitalen Erfassung (zum Beispiel Klosterarchive)
- Consulter: Mitarbeit an der wissenschaftlichen und technischen Weiterentwicklung des Gesamtprojektes oder dessen Teilbereichen

Allgemeine Nutzungsmöglichkeiten

Durch die virtuelle Bereitstellung großer Urkundenbestände (gegenwärtig

etwa 20.000 Stück) ergibt sich ein bequemer und vor allem auch gebührenfreier Zugang für alle Interessierten (Wissenschaftler/innen, Heimatforscher/innen, Studierende, Lehrer/innen usw.). Erstmals ist es möglich geworden, Urkundenbestände bestandsübergreifend zu durchsuchen. Die durch die Indizierung erreichte Aufarbeitung bzw. Normalisierung der wichtigsten Urkundeninhalte macht dieses Material auch für historische Laien benutzbar und erschließt im Zusammenhang mit der schulischen und universitären Lehre neue Vermittlungsmethoden. Die Daten stehen allen Schulen rund um die Uhr zur Verfügung. Die Stücke dienen als wertvolle Ergänzung zum im Geschichtsunterricht vermittelten Lehrstoff und veranschaulichen das bis dato zumeist nur theoretisch Gelernte, wobei die Daten in vielfältiger Weise auch interdisziplinär eingesetzt werden können.

Neben den erweiterten Möglichkeiten des Recherchierens bietet das Projekt auch eine breite Palette anderer Nutzungsfelder an, die den Benutzern/innen als Hilfsmittel dienen. Diese betreffen die historischen Hilfswissenschaften (Paläografie, Chronologie etc.), die Ortsnamenidentifizierung (zum Beispiel *Orbis latinus*) und Wörterbücher ebenso wie Links zu nützlichen Online-Nachschlagewerken bzw. Bibliothekskatalogen. Damit stehen den Interessierten weiterführende Möglichkeiten offen, die es ermöglichen, die Suchergebnisse bis zu einem gewissen Grad selbst zu interpretieren.

Neue Forschungsansätze

Eine digitale Quellensammlung wie sie die Monasterium-Datenbank darstellt, bietet der historischen Forschung in ihren Fragestellungen Möglichkeiten, welche bislang entweder nicht oder nur unter zeitlichem und finanziellem Aufwand durchführbar waren.

Die Bereiche, in welchen neue Erkenntnisse gewonnen bzw. bestehende verifiziert werden können sind:

Personengeschichte

Einzelne Persönlichkeiten lassen sich durch die Recherchierbarkeit der Daten sowohl im Volltext- als auch im Indexmodus über die Jahrzehnte hinweg verfolgen und nachweisen. Neben biografischen Angaben lassen

sich nun auch Aussagen zum geografischen Wirkungsbereich und sozialen Umfeld einzelner Persönlichkeiten machen, indem Aussteller/innen, Siegler/innen und Zeugen/innen von Urkunden innerhalb eines bestimmten Zeitraums zueinander in Beziehung gesetzt werden, wodurch Partnerschaften im diplomatischen Sinn gleichsam auf Knopfdruck erkennbar werden.

Orts- und Lokalgeschichte

Gerade der klösterliche Urkundenbestand birgt in vielen Fällen sehr frühe Nennungen von einzelnen Siedlungen, Gegenden oder sonstigen topografischen Gegebenheiten, was für die Lokal- und Regionalgeschichte von nicht hoch genug einzuschätzender Bedeutung ist. In diesem Zusammenhang lassen sich im Rahmen zielgerichteter Abfragen auch besitzgeschichtliche Zusammenhänge rekonstruieren, wobei durchaus überraschende Ergebnisse erzielt werden können, wie das folgende Beispiel verdeutlichen mag:

Der winzige Ort Tallesbrunn (306 Einwohner/innen) liegt nicht weit von der österreichisch-slowakischen Grenze im östlichen Niederösterreich. Das fiktive Projekt, eine umfassende Ortschronik von Tallesbrunn zu verfassen, soll die Erleichterungen verdeutlichen, welche den Forschern/innen durch die Nutzung der Monasterium-Datenbank geboten werden können. Die Eingabe des Ortsnamens in die Indexsuche ergibt für den Zeitraum zwischen 1115 und 1359 nicht weniger als 23 Treffer in zehn verschiedenen Urkunden von vier verschiedenen Institutionen. Dabei handelt es sich um die Stifte Heiligenkreuz, Herzogenburg, Lilienfeld und das Schottenkloster in Wien. Neben der für die geringe Größe des Ortes nicht unbedeutenden Trefferquote ist vor allem der Umstand erwähnenswert, dass die betreffenden Urkunden in Archiven von ziemlich weit voneinander entfernten Stiften zu finden sind, welche in drei verschiedenen Landesvierteln Niederösterreichs liegen. Es ist sehr zweifelhaft, ob der fiktive Forscher bzw. die fiktive Forscherin auf den Gedanken gekommen wäre, für die Geschichte von Tallesbrunn auch das Archiv von Lilienfeld zu konsultieren, welches immerhin an die 160 km entfernt liegt. Unbestreitbar ist auch die immense Kosten- und Zeitersparnis, welche durch die Verwendung einer zentralen Datenbank erzielt werden kann, denn für die Recherche in vier Klosterarchiven sind bei herkömmlicher Arbeitsweise

wohl einige Wochen bis Monate zu veranschlagen, sofern der Zutritt in diese Archive überhaupt gewährt werden würde.

Institutionsgeschichte

Naturgemäß sind die Urkundenbestände der Stifte und Klöster die Hauptquellen für deren eigene geschichtliche Entwicklung. Allerdings gewähren sie auch Einblick in die Abläufe von territorial übergeordneten Einrichtungen, wie Pfarren, Diözesen und Erzdiözesen.

Bi- und multilaterale Beziehungen im mitteleuropäischen Raum

Gerade die Klosterurkunden bieten ein reichhaltiges Betätigungsfeld bei der Erschließung von bilateralen bzw. multilateralen Beziehungen von Staaten, Ländern, Regionen und Institutionen in Mitteleuropa. Die Grenzen, welche heute die Landkarte dieses Raumes prägen, sind, historisch gesehen, entweder sehr neuen Datums (beispielsweise die Grenzen zwischen Ungarn und der Slowakei, zwischen Tschechien und der Slowakei oder Ungarn und Siebenbürgen) oder hatten ursprünglich nicht diese trennende Bedeutung, wie in der allerjüngsten Geschichte (Grenzen Österreichs zu den ehemals sozialistischen Nachbarstaaten). Geistliche Institutionen dokumentieren diesen Umstand sehr deutlich, da die einzelnen Klöster untereinander europaweit vernetzt waren und miteinander im regen kulturellen Informationsaustausch standen. Die Vernetzung manifestierte sich aber nicht nur innerhalb eines Ordens (die Verbindung der einzelnen Häuser war und ist besonders bei den Zisterziensern aufgrund ihrer Ordensverfassung sehr intensiv ausgeprägt), sondern lässt sich auch zwischen den Vertretern verschiedener Orden bzw. im Verhältnis zu weltlichen Persönlichkeiten oder Institutionen nachweisen.

Internationale Kontakte einzelner Institutionen lassen sich durch gezielte Abfragen ebenfalls in kürzester Zeit eruieren. Ein sehr plakatives Beispiel bietet die Überlieferung des Zisterzienserstiftes Heiligenkreuz im Wienerwald, das in seinem Urkundenbestand für den Zeitraum zwischen 1208 und 1327 um die 50 ungarischen Königsurkunden aus der Arpaden- und Anjouzeit aufbewahrt. Einen ähnlich großen Bestand bilden die Urkunden des Bischofs und des Domkapitels von Győr/Raab (1203-1361). Abgesehen davon, dass diese Urkunden in Ungarn bis dato weitgehend

unbekannt waren, zeigen sie eine starke besitzrechtliche Verflechtung von Heiligenkreuz im Herrschaftsbereich der Stephanskronen.

Von allergrößtem Interesse wäre es angesichts des bedeutenden ungarischen Bestandes in Heiligenkreuz, nach möglichen Spuren des Stiftes in ungarischen Archiven, in diesem speziellen Fall namentlich in der königlichen Kanzlei und im Diözesanarchiv von Győr/Raab zu suchen und damit die Beziehungen des Zisterzienserstiftes noch besser zu dokumentieren. Erklärte Absicht des Monasterium-Konsortiums ist es, eine solche, bislang noch nicht realisierbare Recherche zu ermöglichen, indem die Bestände auch der ungarischen Archive in die Monasterium-Datenbank integriert werden sollen.

Hilfswissenschaftlicher Bereich (Diplomatik, Paläografie, Sphragistik)

Bedingt durch die Tatsache, dass das Monasterium-Projekt die geistlichen Urkunden des Mittelalters mitteleuropaweit so gut wie flächendeckend zusammenzuführen beabsichtigt, können systematische Untersuchungen im Bereich der historischen Hilfswissenschaften durchgeführt werden. Besonders profitieren könnte davon die Lehre im Zusammenhang mit der archivwissenschaftlichen Ausbildung, speziell im Bereich der Paläografie, da, bedingt durch die flächendeckende Erfassung der Bestände, sowohl generelle Entwicklungen als auch regionale Besonderheiten in der Schriftentwicklung vom 9. bis zum 15. Jahrhundert nachvollziehbar und anschaulich gemacht werden können.

Weitere Disziplinen

Neben der Geschichtswissenschaft und den mit ihr assoziierten Disziplinen können auch weitere Fächer Nutzen aus der virtuellen Datenbank ziehen. Die elektronische Aufbereitung und Auszeichnung ermöglicht es der Germanistik und der Klassischen Philologie, unterschiedliche Schreibweisen und Formulierungen in deutschen und lateinischen Texten über die Jahrhunderte hinweg nachzuvollziehen. Ebenso kann das Eindringen bestimmter juristischer Normen in die geübte Rechtspraxis anhand der Urkundentexte geografisch und zeitlich abgestuft festgemacht werden und somit den Rechtswissenschaften eine Erkenntnismöglichkeit geboten werden.

Ausblick

Neben der oben skizzierten Bestandserweiterung im Rahmen des Monasterium-Konsortiums stellt die Entwicklung eines auf XML basierenden, kollaborativen Redaktionssystems (EditMOM)⁴, das eine umfassende Online-Bearbeitung ausgewählter Bestände ermöglichen soll, die vordringlichste Herausforderung der folgenden zwei Jahre dar. Dabei handelt es sich um ein System, das in der Lage sein soll, einem möglichst breiten Publikum über die Verwendung eines einfachen Webbrowsers die Mitarbeit am Monasterium-Projekt zu ermöglichen. Damit können bisher für die Datenerfassung nicht genutzte Kapazitäten außerhalb des eigentlichen Projektbereichs erschlossen werden. Voraussetzung für die kollaborative Mitarbeit ist nach der erfolgten Registrierung die Einhaltung der von der projektkoordinierenden Stelle vorgegebenen Editions- bzw. Auszeichnungsstandards. Die kollaborativen Betätigungsbereiche werden sich über die einfache Erstellung von Transkriptionen über die Indizierung bis hin zur diplomatisch-paläografischen Tiefenerschließung erstrecken. Die Ergebnisse der Mitarbeit werden zunächst durch ein Moderationssystem einer Qualitätsprüfung unterzogen und nach deren Freigabe in die Datenbank integriert und online gestellt. Potentielle Interessenten/innen sind in diesem Zusammenhang Studierende der Geschichte bzw. Archivare/innen in Ausbildungskursen (zum Beispiel des Instituts für Österreichische Geschichtsforschung) oder Heimatforscher/innen.

Bei der Auszeichnung der Urkundentexte lehnt sich das Projekt an die Vorgaben des an der Ludwig-Maximilian-Universität in München entwickelten Standards der Charters Encoding Initiative (CEI) an.⁵

Zusammenfassung

Zum Schluss seien an dieser Stelle nochmals die wesentlichsten Anliegen des Monasterium-Projekts zusammengefasst. Grundsätzliches Ziel des Projekts ist die Schaffung einer mitteleuropäischen Datenbank geistlicher

4 <<http://pcghw51.geschichte.uni-muenchen.de:8080/cocoon/EditMOM/index.html>>.

5 <<http://www.cei.lmu.de/>>.

Urkunden (und Handschriften) im Rahmen des Monasterium-Konsortiums unter Überwindung der räumlichen Zersplitterung vieler Bestände und der trennenden zwischenstaatlichen Grenzen. Der erfasste Urkundenbestand wird unter Verwendung des CEI Auszeichnungsstandards in hilfswissenschaftlicher und inhaltlicher Hinsicht möglichst tief erschlossen.

Der auf diese Art erschlossene Datenbestand soll einerseits für die Forschung, andererseits für möglichst breite Bevölkerungsschichten kostenlos bereitgestellt werden, die durch das Editionstool EditMOM zur aktiven Mitarbeit animiert werden sollen, wodurch nicht zuletzt die Urkunden als wichtige Grundlagen der Historiografie verstärkt im Bewusstsein der Bevölkerung verankert werden.

Dr. Karl Heinz, MAS, ist Geschäftsführer des Instituts zur Erschließung und Erforschung kirchlicher Quellen (IEEkQ) und in der Projektkoordination des Monasterium-Projekts tätig. E-Mail: karl.heinz@monasterium.net

DARSTELLUNG MEHRSCICHTIGER, KOMPLEX-STRUKTURIERTER QUELLEN – DIE COMPUTERGESTÜTZTE DYNAMISCHE EDITION

von Matthias P. Perstling

Die Darstellung komplexer Strukturen in historischen Quellen galt lange Zeit als eine der Hauptschwierigkeiten bei der Erstellung von kritischen Editionen. Besonders die Wiedergabe von mehrschichtigen und mehrdimensionalen Quellenstrukturen in einer traditionell gedruckten und somit zweidimensionalen Form konnte die Erwartungshaltung vieler Historiker/innen meist nicht befriedigen. Mit Hilfe eines elaborierten Informationssystemes lässt sich ein Informationsgewinn bei der Repräsentation komplex-strukturierter Quellen durch die Überwindung der Linearität des gedruckten Textes erzielen, indem fortgeschrittene Datenbank- und Hypertext-Techniken zusammenwirken. Überdies wird den Benutzern/innen des Editionssystemes neben den unterschiedlichsten Zugriffsmöglichkeiten zur schnelleren Auffindung der gewünschten Information auch die Möglichkeit zur Unterscheidung zwischen dem Inhalt der Quelle und dem Wissen des Editors bzw. der Editorin darüber geboten. Die methodischen Voraussetzungen und die funktionelle, sowie technische Umsetzung eines solchen Editionssystemes sollen anhand eines Beispiels aufgezeigt werden.

Einleitung

Seit Karl Lachmann (1793-1851) seine Methode zur Erstellung einer kritischen Textedition präzisier¹, verfolgt die Editionswissenschaft den

1 Vgl. Fuhrmann, Horst, Über Ziel und Aussehen von Texteditionen, in: Monumenta Germaniae Historica (Hg.), Mittelalterliche Textüberlieferungen und ihre kritische Aufarbeitung, München 1976, S. 12-27, hier S. 12f.; Stackmann,

Zweck, einen Text in möglicher Reinheit und Vollständigkeit, übersichtlich und lesbar darzustellen.² Dies gilt sowohl für „philologische“ als auch für „historische“ Editionen³, wenn auch die unterschiedlichen Zielsetzungen der jeweiligen Disziplin zu leichten Abweichungen in den Darstellungen führen. Dementsprechend versucht eine historische Edition das Dokument so zu erschließen, dass es Auskunft über politische, soziale, kirchliche, geistige und andere Fragen zu geben vermag. Zudem bietet sie Information über die Rezeption des Textes, sowie über bestehende parallele Überlieferungen. Das ist auch der Unterschied zu den philologischen Editionen, die den verschiedenen Schreibvarianten, und somit auch dem kritischen Variantenapparat, mehr Bedeutung beimessen.

Ungeachtet dessen gelten bzw. galten die Methode sowie auch die Technik der kritischen Edition in beiden Fächern als ausgereift⁴, und so kam es in den letzten 150 Jahren kaum zu Veränderungen der Techniken. Umso mehr stimmt der Umstand traurig, dass eine Vielzahl an interessanten historischen Quellen noch nicht kritisch ediert wurde bzw. die Versuche dazu fehlschlagen. Es gibt viele verschiedene Gründe für ein solches Scheitern, wie zum Beispiel, dass die Finanzierung von Editionsprojekten ein oft unüberwindbares Hindernis darstellt oder dass manche Quellen für

-
- Karl, Mittelalterliche Texte als Aufgabe, in: Foerste, William; Borek, Karl-Heinz (Hgg.), Festschrift für Jost Trier, Köln 1964, S. 241-267, hier S. 243.
- 2 Rehbein, Malte, Die digitale Textedition, in: Ebeling, Hans-Heinrich; Thaller, Manfred (Hgg.), Digitale Archive. Die Erschließung und Digitalisierung des Stadtarchivs Duderstadt, Göttingen 1999, S. 103-124, hier S. 103.
- 3 Bei der Definition von Edition folgt der Autor der Meinung von Patrick Sahle: „Edition ist die erschließende Wiedergabe von historischen Dokumenten.“ Sahle, Patrick, Vom editorischen Fachwissen zur digitalen Edition: Der Editionsprozess zwischen Quellenbeschreibung und Benutzeroberfläche, in: Thaller, Manfred (Hg.), Fundus – Forum für Geschichte und ihre Quellen, Beiheft 2: Quellen und Quelleneditionen im neuen Medienzeitalter, Göttingen 2003, S. 75-102, hier S. 76.
- 4 Dem Autor ist die Diskussion um die Umsetzung der „New Philology“ durchaus bewusst, jedoch blieb diese stets eine Methodendebatte der Philologen, die sowohl Fürsprecher/innen als auch energische Gegner/innen hervorrief und bei der noch keine Richtungsentscheidung gefallen ist. Vgl. Stackmann, Karl, Die Edition – Königsweg der Philologie?, in: Bergmann, Rolf; Gärtner, Kurt (Hgg.), Methoden und Probleme der Edition mittelalterlicher deutscher Texte, Tübingen 1993, S. 1-18, bes. S. 4.

bestimmte Personengruppen als nicht edierungswürdig gelten. Ein Hauptgrund ist aber mit Sicherheit auch der, dass viele Quellen zu *komplex* sind, um sie mit Hilfe der traditionellen Methode zur Erstellung einer kritischen Textedition in Buchform zu edieren. In diesem Zusammenhang bedeutet *komplex*, dass mehrere „*nicht-lineare*“ Ebenen übereinander geschichtet sind. Im Gegensatz dazu gilt ein Text als *linear*, wenn er nur für sich selbst steht. Als idealen *linearen* Text müsste man sich somit einen Autografen vorstellen, der von einer Person geschrieben wurde, keine Verweise zu anderen Texten hat, in sich abgeschlossen ist und nicht nach außen auf andere Texte wirkte. Je weniger *linear* und somit *komplexer* ein Text ist, desto schwieriger ist natürlich auch seine Darstellung.⁵

Die Komplexität einer historischen Quelle kann sich aber auf verschiedenen Ebenen abspielen und ist auch vom Texttyp der Quelle geprägt. So wird zum einen eine Quelle als *komplex* zu bezeichnen sein, wenn viele unterschiedliche Überlieferungen daraus entstanden sind oder der Text aus mehreren, verschiedenen Werken zusammengesetzt wurde, auch wenn es sich dabei um *den* von nur einer Person geschriebenen Originaltext der Quelle handelt.⁶ Zum anderen wird aber auch ein Autograf ohne Verweise zu anderen Texten, bzw. ohne weitere Rezeption als komplex gelten müssen, wenn die Quelle voll von Streichungen, Glossen und Nachträgen von verschiedenen Schreiberhänden ist, oder eine mehrschichtige innere Struktur besitzt. Als Exempel für diese Kategorie dienen Handschriften wie etwa Nekrologe, Verbrüderungsbücher oder spätmittelalterliche Urbare.

5 Vgl. Rehbein, Textedition (wie Anm. 2), S. 104. Die Problematik linearer, komplexer Texte stellt auch Manfred Thaller dar und bietet eine technische Beschreibung. Vgl. Thaller, Manfred, A Draft Proposal for a Standard for the Coding of Machine Readable Sources, in: Greenstein, Daniel I. (Hg.), *Modelling Historical Data: Towards a Standard for Encoding and Exchanging Machine-Readable Texts* (Halbgraue Reihe zur Historischen Fachinformatik A 11), St. Katharinen 1991, S. 19-64.

6 Es sei hier als Beispiel auf die kritische Edition des „*liber pontificalis*“ durch Theodor Mommsen (MGH *Gesta Pontificum Romanorum* 1,1, Berlin 1898.) bzw. auf die nie vollendete Edition der „*Kapitulariensammlung des Benedictus Levita*“ (Vgl. Schmitz, Gerhard, „Unvollendet“ – „Eingestampft“ – „Kassiert“. Nie Erschienenes und Missglücktes, in: *Zur Geschichte und Arbeit der Monumenta Germaniae Historica*, München 1996, S. 64-73) verwiesen.

Neben dieser in unterschiedlicher Weise ausgeprägten Komplexität von historischen Quellen gibt es noch weitere Eigenheiten solcher Dokumente, die bei einer Edition zu berücksichtigen sind und zu Problemen führen können: Zum einen sind historische Informationen generell *unscharf*, das bedeutet, dass sie sich nicht ohne weiteres eindeutigen Kategorien zuordnen lassen.⁷ Zum anderen sind historische Informationen *kontextsensitiv*, das heißt, dass historische Angaben in aller Regel nicht unabhängig voneinander interpretierbar sind.⁸

Methodische Überlegungen – Vom linearen zum dynamischen System

Muss sich die Wissenschaft also mit der Tatsache abfinden, dass es nie zu befriedigenden kritischen Editionen von historischen Quellen mit einer komplexeren Struktur kommen wird? Oder gibt es auch andere Möglichkeiten der Präsentation von Quellen, die über die traditionelle Methode der Wiedergabe in einem zweidimensionalen System – in Buchform – hinausgehen?

Um diese Fragen zu beantworten, sollte der eingangs erwähnte Zweck einer kritischen Textedition – den Text in möglicher Reinheit und Vollständigkeit übersichtlich und lesbar darzustellen – etwas genauer betrachtet werden. So bedeutet für Ingo H. Kropač eine Quelle zu edieren vielmehr „die Informationsstruktur, den Informationsgehalt, die Informationsdichte und die Überlieferungsstruktur einer Quelle auf eine Weise zugänglich zu

7 Vgl. Thaller, Manfred, Ungefähre Exaktheit. Theoretische Grundlagen und praktische Möglichkeiten einer Formulierung historischer Quellen als Produkte ‚unscharfer‘ Systeme, in: Nagl-Docekal, Herta; Wimmer, Franz (Hgg.), *Neue Ansätze in der Geschichtswissenschaft* (Conceptus Studien 1), Wien 1984, S. 77-100, bes. S. 90-96.

8 Vgl. Thaller, Manfred, Datenbasen als Editionformen?, in: Schwob, Anton; Kranich-Hofbauer, Karin; Suntinger, Diethard (Hgg.), *Historische Edition und Computer. Möglichkeiten und Probleme interdisziplinärer Textverarbeitung und Textbearbeitung*, Graz 1989, S. 215-241, bes. S. 230; Levermann, Wolfgang, Kontextsensitive Datenverwaltung (Halbgraue Reihe zur Historischen Fachinformatik B 8), St. Katharinen 1991; Thaller, Manfred, Gibt es eine fachspezifische Datenverarbeitung in den historischen Wissenschaften? Quellenbanken in der Geschichtswissenschaft, in: Kaufhold, Karl Heinrich; Schneider, Jürgen (Hgg.), *Geschichtswissenschaft und elektronische Datenverarbeitung* (Beiträge zur Wirtschafts- und Sozialgeschichte 36), Wiesbaden 1988, S. 45-83.

machen, die so vollständig als möglich sein soll und den Benutzer einer Edition in die Lage versetzt, diese analytisch zu verwenden, ohne die eigentliche Quelle einsehen zu müssen.“⁹ In dieser bewusst allgemein gehaltenen Definition tritt zum ersten Mal der Begriff der *Information* in Erscheinung. Betrachtet man eine Edition informationswissenschaftlich und abstrahiert die durch sie entstehenden Prozesse, so stellt eine Edition im Grunde genommen nichts anderes als eine Transformation von Information von einem Medium in ein anderes dar. Dabei ergibt sich das Problem, dass eine Transformation immer zu einem Informationsverlust führt. Im speziellen Fall ist der Informationsverlust eine Verringerung der Authentizität der Quelle.¹⁰ Die Editoren/innen versuchen freilich, den Informationsverlust auf ein Minimum zu reduzieren, aber sie werden definitiv daran scheitern, wenn ein komplex-strukturierter Text ohne eine adäquate Darstellung der internen Quellenstruktur ediert wird. Genau deshalb muss die Repräsentation einer mehrschichtigen Quelle das zweidimensionale System eines gedruckten Buches überwinden. Infolgedessen bietet ein computergestütztes Editionssystem¹¹, das unterschiedliche Ebenen abbil-

9 Kropač, Ingo H., Theorien, Methoden und Strategien für multimediale Archive und Editionen, in: van Eickels, Klaus; Weichselbaumer, Ruth; Bennowitz, Ingrid (Hgg.), *Mediävistik und Neue Medien*, Ostfildern 2004, S. 306.

10 Vgl. Kropač, Ingo H., Gain et perte d'information. Problèmes fondamentaux posés par l'édition informatisée de données historiques, in: Genet, Jean-Philippe (Hg.), *Standardisation et échange des bases de données historiques. Actes de la troisième Table Ronde internationale tenue au L.I.S.H. (C.N.R.S.), Paris 1988*, S. 49-57; Brunnhofer, Regina; Kropač, Ingo H., Digital Archives in a Virtual World, in: Royal Netherlands Academy of Arts and Sciences (Hg.), *Humanities, Computers and Cultural Heritage. Proceedings of the XVI international conference of the Association for History and Computing, 14-17 September 2005, Amsterdam 2005*, S. 83-89; Sahle, Patrick, Bemerkungen zur Edition historischer Quellen zwischen Informationsverlust und Informationsgewinn, vgl. <<http://ba.tuxomania.net/kerpen/edition/infos/text-sahle.html>> (13.02.2006).

11 Der Begriff „computergestützt“ wird im Folgenden noch genauer spezifiziert und vom Autor im Sinne von Kropač (Vgl. Kropač, Theorien (wie Anm. 9), S. 307.) verwendet. „Computergestützte Edition“ bezeichnet somit einen Teilbereich der „digitalen Edition“, wie sie von Peter Robinson (Vgl. Robinson, Peter, *What is a Critical Digital Edition?*, in: *Variants – Journal of the European Society for Textual Scholarship* 1 (2002), S. 43-62, bes. S. 51-57.) und Patrick Sahle (Vgl. Sahle, Patrick, *Digitales Archiv – Digitale Edition. Anmerkungen zur Begriffsklärung*, vgl. <http://www.germanistik.ch/publikation.php?id=Digitales_Archiv_und_digitale_Edition> [13.02.2006].) gesehen und definiert wird.

den kann, die einzige Möglichkeit, die verschiedenen Schichten angemessen darzustellen.

Wenn zuvor konstatiert wurde, dass eine Edition immer zu einem Informationsverlust führt, dann wurde nur eine Seite berücksichtigt. Denn wenn eine Edition auch einen Informationsverlust bedeutet, so erhält man im Gegenzug dazu auch einen Informationsgewinn. Dieser Informationsgewinn entsteht, da in die Edition das Wissen des Editors bzw. der Editorin über die Quelle und seine bzw. ihre Interpretation einfließen.¹² Aber genau dieses Wissen stellt sich auch als nicht-linear heraus und muss noch zur Elementarinformation der Edition hinzugefügt werden. Folglich resultiert aus jedem Editionsversuch eine mehrdimensionale Repräsentation, auch wenn der Originaltext nahezu linear ist. Daraus ergibt sich, dass eine kritische Edition zugleich komplexer *und* weniger komplex als die Originalquelle ist sowie mehr *und* weniger Information in einem bietet. Das Ziel ist die exakte Information zu erhalten, welche die Benutzer/innen zu einem bestimmten Zeitpunkt für einen speziellen Zweck benötigen. Um dieses Ziel zu erreichen und befriedigende Ergebnisse zu erhalten, muss die Präsentationsform dynamisch sein und Hypermedia¹³ im eigentlichen Sinn bieten.

Aus den oben genannten Anforderungen wird die Notwendigkeit eines Systems offenkundig, das den Text in einer auf Computern basierenden, dynamischen Form wiedergibt.¹⁴ Da dies offensichtlich mehr beinhalten

Archiv_und_digitale_Edition> [13.02.2006].) gesehen und definiert wird.

12 Kropač, Gain (wie Anm. 10), S. 51f.; Sahle, Bemerkungen (wie Anm.10), S. 1.; Köhn, Tilo, Erfahrungen mit einer Digitalen Edition – am Beispiel des Traktats Heinrichs von Antwerpen, vgl. <<http://golm.rz.uni-potsdam.de/hva/vortrag.htm>> (13.02.2006).

13 In diesem Zusammenhang zieht der Autor „Hypermedia“ dem Begriff „Hypertext“ vor, da es bei digitalen Editionssystemen auch zur Darstellung von Faksimiles in digitalen Formaten (Rastergrafik) kommt und somit der Übergang von logischen Verbindungen zwischen ausschließlich Textpassagen zu solchen zwischen unterschiedlichen Medien gelingt. Zur Grundidee des Hypertextes vgl. Nelson, Theodor H., *Getting it Out of Our System*, in: Schecter, George (Hg.), *Information Retrieval. A Critical Review*, Washington 1967, S. 191-210, sowie auch Freisler, Stefan, *Hypertext – Eine Begriffsbestimmung*, in: *Deutsche Sprache* 22 (1994), S. 19-49.

14 Manfred Thaller hatte bereits 1988 die Vision, dass „Datenbanken [...] daher in

muss, als nur eine klassisch gedruckte Edition in das Internet zu stellen¹⁵ und diese somit zu digitalisieren¹⁶, bedienen wir uns einer elaborierten Methode der digitalen Edition: der Methode der *Integrierten Computergestützten Edition*¹⁷, kurz ICE genannt. In diesem Kontext bedeutet „computergestützte Edition“, dass „der editorische Forschungsprozess zumindest teilweise auf formalen Verfahren beruht, das Ergebnis (die Publikation) mit formalen Methoden analysiert werden kann, und die Realisierung auf einem Informationssystem beruht, das ein verbessertes Verständnis der editorischen Entscheidungen sowie der edierten Entitäten gewährleistet.“¹⁸ Des Weiteren steht der Terminus „integriert“ in der Methode dafür, dass „durch iterative Forschungsprozesse folgende Ziele erreicht werden: Transparenz, Nachvollziehbarkeit und Korrekturmöglichkeit; Modularität

mittlerer Zukunft eine Editionsform werden [können]“, die „auf Grund einer dynamischen und integrierten Darstellung sämtlicher Textüberlieferungen allmählich zu einem Instrument werden“, um eine dynamische Edition zu erzeugen. Thaller, Datenbasen (wie Anm. 8), S. 233f. Jedoch meint er dazu auch: „Diese Art einer dynamischen Edition setzt allerdings Konzeptionen voraus, die, insbesondere wegen der anzuwendenden internen Textdarstellung, den Stand derzeit in der Datentechnik üblicher Instrumente sprengen. Ihre Realisierung setzt daher eine bewusste Anstrengung zur Formulierung (und offensichtlich auch zur Realisierung) von Datenstrukturen voraus, die derartige mehrschichtige Textdarstellungen in einschlägige Datenbanksoftware integrierbar machen.“ Thaller, Datenbasen (wie Anm. 8), S. 234.

- 15 Weder Editionstexte als PDF-Dateien, noch reine Transkriptionen im Web mit irgendeiner Auszeichnungssprache, sei es auch XML, gefertigt, genügen dem Anspruch, dass es sich dabei um eine computergestützte Edition handelt. Vgl. Kropač, Theorien (wie Anm. 9), S. 307.
- 16 Vgl. Sahle, Archiv (wie Anm. 11), S. 6.
- 17 Zu den Grundlagen der ICE-Methode vgl. Kropač, Ingo H., Ad fontes. Von Wesen und Bedeutung der Integrierten Maschinellen Edition, in: Ebner, Herwig; Haselsteiner, Horst; Wiesflecker-Friedhuber, Ingeborg (Hgg.), Geschichtsforschung in Graz, Graz 1990, S. 465-482; Kropač, Ingo H.; Botzem, Susanne, As You Like It. Archiving, Editing and Analysing Medieval Documents, in: Smets, Josef (Hg.), Histoire et Informatique. Ve Congrès “History & Computing”, 4.-7. 9. 1990 à Montpellier, Montpellier 1992, S. 301-313.; Kropač, Ingo H.; Botzem, Susanne; Kurschel, Henriette, Das ICE-Projekt, in: Informatik Forum 8,4 (1995), S. 159-164.; Kropač, Theorien (wie Anm. 9), S. 295-316.; vgl. dazu auch die Homepage der ICE: Kropač, Ingo H.; Kropač, Susanne; Vasold, Gunter; Boshof, Heidrun, Integrierte Computergestützte Edition (ICE), vgl. <<http://www.fcr-online.com/ice>> (10.02.2006).
- 18 Kropač, Theorien (wie Anm. 9), S. 308.

in der Systemorganisation; verbesserte Verfügbarkeit und ‚Benutzbarkeit‘ und offene Systemtechnologien¹⁹ in typologischer und quantitativer Hinsicht.“²⁰

Diese ausgereifte Methode, die auf einem offenen und erweiterbaren Konzept aufbaut, wurde in den *Fontes Civitatis Ratisponensis* (FCR), einem Forschungsprojekt und Editionsunternehmen an der Karl-Franzens-Universität Graz²¹, in die Praxis umgesetzt und eingehend getestet.²² Zur Wiedergabe komplex-strukturierter Quellen benötigt die Methode der ICE allerdings eine Erweiterung: Wie bereits oben erwähnt ist für die Darstellung von mehrschichtigen und mehrdimensionalen Quellen ein dynamisches System erforderlich, das auf die gestellten Anforderungen in unterschiedlicher Weise eingehen kann.

Computergestützte dynamische Edition: „ICE-dynamisch“

Um ein dynamisches System zu betreiben, ist eine Datenbasis, wie zum Beispiel eine Datenbank, unerlässlich. Die in dieser Datenbasis enthaltenen

-
- 19 Diese offenen Systemtechnologien müssen natürlich internationalen Standards entsprechen, sowie Unabhängigkeit, Erweiterbarkeit und Interoperabilität der Daten garantieren.
- 20 Kropač, Theorien (wie Anm. 9), S. 311.
- 21 Die *Fontes Civitates Ratisponensis* (FCR) bestehen bereits seit 1990 an der Karl-Franzens-Universität Graz unter der Leitung von Ao. Univ.-Prof. Dr. Ingo H. Kropač. Nachdem das Forschungsprojekt ursprünglich am Forschungsinstitut für Historische Grundwissenschaften beheimatet war, ist es in der Zwischenzeit dem Fachbereich „Historische Fachinformatik und Dokumentation“ des Instituts für Geschichte eingegliedert worden.
- 22 Vgl. dazu Kropač, Ingo H.; Kropač, Susanne, Prolegomena zu einer städtischen Diplomatie des Spätmittelalters, das Beispiel Regensburg, in: Prevenier, Walter; de Hemptinne, Thérès (Hgg.), La diplomatie urbaine en Europe au moyen âge. Actes du congrès de la Commission internationale de Diplomatique, Gand, 25.-29. 8. 1998 (Studies in Urban Social, Economic and Political History of Medieval and Early Modern Low Countries 9), Louvain 2000, S. 229-265.; Boshof, Heidrun, Die *Fontes Civitatis Ratisponensis*: Geschichtsquellen der Reichstadt Regensburg online, in: van Eickels, Klaus; Weichselbaumer, Ruth; Bennewitz, Ingrid (Hgg.), Mediävistik und Neue Medien, Ostfildern 2004, S. 279-294.; Vasold, Gunter, Edition à la carte? Usability, Interfacing und Datenmigration für webbasierte Editionssysteme, in: van Eickels, Klaus; Weichselbaumer, Ruth; Bennewitz, Ingrid (Hgg.), Mediävistik und Neue Medien, Ostfildern 2004, S. 261-278. Die Seiten der FCR sind erreichbar unter <<http://www.fcr-online.com>> (11.02.2006).

Daten müssen in einer gut strukturierten Form aufgenommen werden, damit die darin gespeicherte Information leicht verarbeitet werden kann. Dabei ist von immenser Wichtigkeit, dass sich die Datenstruktur möglichst nahe an die Quellenstruktur anlehnt und diese wiederzugeben versucht, ohne auf das Aussehen der späteren Edition Rücksicht zu nehmen. Mit anderen Worten: Es muss dadurch zu einer strikten Trennung von (Informations-)Inhalt und (Publikations-)Form der Daten kommen.²³ Auf dieselbe Art werden die unterschiedlichen Schichten oder Varianten einer komplex-strukturierten Quelle behandelt: Sie müssen inhaltlich getrennt und ausdrücklich gekennzeichnet werden.

Der dynamische Teil des Systems muss so gestaltet sein, dass die Benutzer/innen ihre „eigene“ Edition verfassen können. Die speziellen Anliegen der Benutzer/innen an die Quelle sollen erkannt und in eine nützliche Repräsentation umgesetzt werden. Diese Anforderungen verlangen eine ausgeklügelte Benutzeroberfläche mit unterschiedlichsten Zugriffsmöglichkeiten auf das System und die darin gespeicherten Daten. In diesem Zusammenhang müssen allerdings auch einige Gegebenheiten beachtet werden: Zuallererst muss sichergestellt sein, dass der edierte Text immer zitierbar bleibt, auch wenn die Textversion auf dynamische Weise erstellt wird. Zudem muss für die Benutzer/innen der Edition der Inhalt der Quelle immer so genau wie möglich vom Wissen des Editors bzw. der Editorin über die Quelle unterscheidbar bleiben.

Diese Erweiterung der Methode der ICE für komplex-strukturierte Quellen erscheint möglicherweise ein wenig abstrakt und nicht präzise definiert, doch das war vollends beabsichtigt (im vollen Bewusstsein über deren Ungenauigkeit). Patrick Sahle stellte in seinen Thesen über digitale Editionen fest, dass „unterschiedliche Quellen unterschiedliche Editionsverfahren erfordern“²⁴ und der Autor stimmt mit ihm darin vollkommen überein. Genau deshalb wäre *eine* explizit ausgearbeitete Methode für alle

23 Sahle, Patrick, Digitale Editionstechniken und historische Quellen, in: Jenks, Stuard; Marra, Stephanie (Hgg.), Internet-Handbuch Geschichte, Köln 2001, S. 153-166, hier S. 163.; Sahle, Fachwissen (wie Anm. 3), S. 76f.

24 Sahle, Patrick, Digitale Edition (Historischer Quellen) – Einige Thesen, vgl. <<http://www.uni-koeln.de/~ahz26/dateien/thesen.htm>> (10.2.2006), S. 3.

Arten von komplex-strukturierten Quellen sinnlos. Je spezialisierter eine Methode ist, desto mehr Ausnahmen und Modifikationen werden für die verschiedenen Beispiele benötigt. Das oben beschriebene Modell soll vielmehr einen Rahmen bilden, in dem Editionen von vielen verschiedenartigen Quellen mit komplexer Struktur realisiert werden können, die diesen Grundvoraussetzungen entsprechen.

Die Methode der „computergestützten dynamischen Edition“ („ICE-dynamisch“) stellt somit eine Konzeption dar, die Manfred Thaller in seinen Thesen von 1988²⁵ als Voraussetzung für eine dynamische Edition forderte. Seine damalige Vision, dass „Datenbanken [...] daher in mittlerer Zukunft eine Editionsform werden [können]“, die „auf Grund einer dynamischen und integrierten Darstellung sämtlicher Textüberlieferungen allmählich zu einem Instrument werden [... um eine dynamische Edition zu erzeugen]“²⁶, scheint auf diese Weise in Erfüllung gegangen zu sein.

Faksimile als Unterstützung

Eine weitere Möglichkeit im Zuge von digitalen Editionssystemen, auf die hier noch nicht eingegangen wurde, bietet die Darstellung von Faksimiles. Da jedoch ein Faksimile noch keine textkritische Leistung darstellt, kann eine reine Faksimile-Ausgabe niemals ein Editionsersatz sein.²⁷ Das Faksimile kann jedoch als unterstützendes Medium einer Edition dienen, wenn es dabei hilft, den Entstehungszusammenhang von Textfassungen auch in seiner optischen Dimension deutlicher hervorzuheben.²⁸ Immanenter wird die Notwendigkeit von zusätzlichen Faksimiles zum edierten Text noch bei den Fontes Civitatis Ratisponensis, bzw. bei der ICE-Methode: Ziel dabei ist es, „Texte mit wissenschaftlich nachprüfbar Verfahren zu erstellen, sie zu präsentieren und zugänglich zu machen [..., wobei] als oberste Prämisse Transparenz [gilt], um den Benutzer in die Lage zu versetzen, alle Entscheidungen und Interpretationen des Bearbeiters rekonstruieren,

25 Thaller, Datenbasen (wie Anm. 8), S. 234.

26 Ebd., S. 233f.

27 Plachta, Bodo, Editionswissenschaft. Eine Einführung in Methode und Praxis der Edition neuerer Texte, Stuttgart 1997.

28 Ebd.

überprüfen und gegebenenfalls auch korrigieren zu können.“²⁹ Folglich kann durch die Präsentation von Faksimile und Text nebeneinander ein verbessertes Verständnis der editorischen Entscheidungen, sowie Nachvollziehbarkeit und eventuell auch Korrekturmöglichkeit gewährleistet werden.³⁰

„ICE-dynamisch“ konkret: Das Editionssystem des „Marchfutterurbars“

Als Beispiel für ein digitales Editionssystem einer komplex-strukturierten Quelle basierend auf der vorgestellten Methode soll auf die Darstellung des „Steirisch-landesfürstlichen Marchfutterurbars von 1414/1426“³¹ (MFU) verwiesen werden. Diese spätmittelalterliche Handschrift aus dem Steiermärkischen Landesarchiv in Graz ist ein Urbar, in dem die Haferabgaben der Holden in der Mittelsteiermark an das landesfürstliche, berittene Heeresaufgebot über zwölf Jahre hinweg von verschiedenen Schreibern penibel verzeichnet wurden. Die besondere Bedeutung der Quelle für die Regionalgeschichte ergibt sich daraus, dass zum ersten Mal in der Geschichte eine so enorme Anzahl an Abgabepflichtigen³² in einem großen Gebiet der Steiermark namentlich aufgenommen wurde. Der Umstand, dass die Quelle so viele Streichungen und Nachträge enthält – aufgrund der häufigen Wiederverwendung der Handschrift – führte zu ihrer mehrschichtigen Struktur und folglich auch zum Fehlen einer adäquaten Edition.³³ Nur ein Versuch, das Urbar zu edieren, „gelang“ im Jahr 1910³⁴ jedoch in einer

29 Boshof, *Fontes* (wie Anm. 22), S. 279.

30 Vgl. Kropač, *Theorien* (wie Anm. 9), S. 308, 311.

31 Steirisch-landesfürstliches Marchfutterurbar von 1414/1426: Steiermärkisches Landesarchiv (StLA), Laa. A. Antiquum II, Stockurbar 43/64. Vgl. Perstling, Matthias P., *Das steirisch-landesfürstliche Marchfutterurbar von 1414/1426. Teiledition und Interpretation*, Dipl.-Arb., Graz 2003.

32 Es sind darin über 10.000 Personen namentlich genannt.

33 Vgl. Spreitzhofer, Karl, *Steirische Quellen zur Geschichte des Marchfutters in der frühen Neuzeit*, in: *Mitteilungen des Steiermärkischen Landesarchivs* 27 (1977), S. 51–65, hier S. 53.

34 Dopsch, Alfons; Mell, Alfred (Hgg.), *Die Landesfürstlichen Gesamturbare der Steiermark aus dem Mittelalter* (Österreichische Urbare, I. Abt., Bd.2), Wien 1910, S. 311–590.

besonders unbefriedigenden Weise: Die „Edition“ erfolgte in Tabellenform, ohne auf die prosopografische Relevanz der Quelle einzugehen, indem einfach die Namen der Holden weggelassen wurden.

Funktionelle Aspekte des Editionssystems

Das Ziel einer zufrieden stellenden Edition des Marchfutterurbars wäre demnach, für jeden bestimmten Zeitpunkt den jeweils damals aktuellen Zustand der Quelle wiedergeben zu können und alle genannten Personen mit ihrer entsprechenden Abgabe in einer klar ersichtlichen und anschaulichen Form in die Repräsentation zu integrieren. Verbunden mit diesem Bestreben muss den Benutzern/innen die Möglichkeit gegeben werden, jede einzelne Schicht des Textes zu isolieren. Außerdem soll es ihnen möglich sein, die exakte Information, die sie für ihren expliziten Zweck benötigen, zu erhalten.

Mit diesem Editionssystem wird somit den Benutzern/innen ein Werkzeug in die Hand gegeben, mit dem sie ihre „eigene“ Edition zusammenstellen können. Der Output aus dem System ist in folgedessen variabel, **dynamisch erzeugt** und kann auf verschiedenartige Weise dargestellt werden. Um die Möglichkeiten des Systems auch nutzen zu können, ist eine ausgereifte Benutzeroberfläche essentiell, die eine Vielzahl an unterschiedlichen Optionen des Systemzugangs und -zugriffs auf die Datenbasis bietet. Den Kern des Systems bildet eine Datenbank, welche die Strukturdatenbank enthält und den Output erzeugt.

Wie bereits erwähnt, bietet eine digitale Edition die Möglichkeit, Faksimiles der Quelle zu zeigen. Selbst die beste kritische Textedition kann eine Handschrift nicht „visualisieren“, und der Informationsverlust in Bezug auf die Abnahme an Authentizität der Quelle kann durch die Abbildung der Digitalisate reduziert werden. Im Editionssystem des Marchfutterurbars sind unterschiedliche Darstellungsarten der Bilder enthalten: Zum einen sind die Faksimiles, die jeweils eine Seite des Urbars zeigen, als Thumbnails ersichtlich und die Benutzer/innen können so digital durch die Handschrift blättern. Die Größe der Bilder ist veränderbar und so kann eine wechselnde Anzahl an Thumbnails pro Bildschirmseite angezeigt werden. Durch Klick auf die Bilder können diese auch in einen Viewer geladen

werden, in dem Vergrößerungen bis zur doppelten Originalgröße möglich sind (siehe Abbildung 1). Zusätzlich gibt es auch eine Ansicht, in der das Faksimile und der Editionstext nebeneinander am Bildschirm erscheinen, wodurch sie sehr leicht miteinander verglichen werden können. Dies ist zugleich auch die Standardeinstellung des Systems.

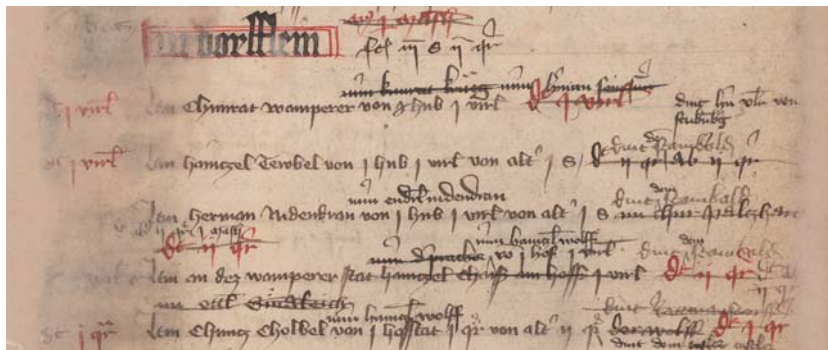


Abbildung 1: Textpassage aus dem MFU, fol. 17r (StLA)

Der gängigste Einstieg in eine Edition eines historischen Dokuments führt die Benutzer/innen auch in einem digitalen System zum eigentlichen Text, wobei dieser selten von Anfang bis Ende durchgelesen wird. Besonders Urbare wie das Marchfutterurbar, aber auch alle anderen wirtschaftsgeschichtlichen Quellen, werden konsultiert, um spezifische Informationen zu erhalten. Um das Information-Retrieval zu erleichtern, sind in diesem System zusätzliche Module wie Glossar, Personen- und Ortskatalog installiert, sowie einfache Volltextsuche und erweiterte Suchmöglichkeiten integriert. Somit kann die von Benutzern/innen aller Fachbereiche benötigte Information gefunden werden. Doch auf welche Weise wird sie dargestellt? Für die unterschiedlichen Arten der Repräsentation muss die traditionelle Form des edierten Texts als erstes erwähnt werden (siehe Abbildung 2). Diese Form ist für das Marchfutterurbar jedoch fast nicht benutzbar, da durch die mehrschichtige Struktur, als Folge der zahlreichen Streichungen und Nachträge, der dazugehörige kritische Anmerkungsappa-

rat immens anwächst und sehr viel Information im Variantenapparat verschwindet. Dennoch wird diese Form (der Vollständigkeit wegen) im System angeboten und kann auf zwei unterschiedliche Weisen dargestellt werden: entweder als ein der Seitenmetapher folgender Text neben dem Faksimile, oder als allein stehender Text, der der internen Struktur des Marchfutterurbars – nach Pfarreien und Ortschaften geordnet – entspricht, ohne dabei Seitenumbrüche des Originals zu berücksichtigen.

3.16		Im Dorfflein ¹ <i>facit 3 s 2 qr.</i> ^a	
3.16.1	<u>Dedit 1 virl.</u> ^b	Item Chunrat Wamperer ^c <i>nun Herman</i> ^d von ½ hub 1 virl. ^e	<i>Dint hern Vlr(eich) von Stubenberg.</i> ^f
3.16.2	<u>Dedit 1 virl.</u> ^g	Item Hainczel Tewbel von 1 hub 1 virl von alter 1 s. ^h	<i>Dint dem</i> ⁱ <i>Stainbald.</i> ^j
3.16.3	[<u>Dedit ...</u>] <u>qr [...]</u> <u>Dedit</u> <u>aber 2 qr</u> <u>1 mash.</u> ^k	Item Herman Nidendran <i>nun Enderl</i> <i>Nidendran</i> ¹ von 1 hub 1 virl von alter 1 s. ^m	<i>Dint dem</i> ⁿ <i>Stainbald.</i> ^o
3.16.4	<u>Dedit [...]</u> <u>vir l c.</u> ^p	Item an dez Wamperer stat <i>nun Hainczel Wolff</i> ^q von 1 hof 1 virf. ^s	<i>Dint dem</i> ^t <i>Stainbald.</i> ^u
3.16.5	<u>Dedit 1 qr.</u> ^v	Item Chuncz Cholbel <i>nun Hainczel Wolff</i> ^w von 1 hofstat 1 qr von alter 2 qr. ^x	<i>Dint dem</i> ^y <i>Tastler.</i> ^z
<p>3.16 ^a Von Hand 7 nachgetragen. ^b Von Hand 4 nachgetragen. ^c Danach über der Zeile von Hand 3 nachgetragen und von Hand 5 gestrichen: nun Kunrat Krüg. ^d Über der Zeile von Hand 5 nachgetragen. Danach über der Zeile von Hand 5 nachgetragen und von Hand 6 gestrichen: sein sun. ^e Danach der Nachtrag von Hand 2, der von Hand 4 wieder gestrichen wurde: <u>Dedit 1 virl.</u> ^f Von Hand 5 nachgetragen.. ^g Von Hand 4 nachgetragen. ^h Danach von Hand 2 nachgetragen und von Hand 4 gestrichen: Dedit 2 qr. Aber 2 qr. ⁱ Über der Zeile von Hand 5 nachgetragen. ^j Von Hand 2 nachgetragen. ^k Von Hand 4 nachgetragen. ^l Über der Zeile von Hand 4 nachgetragen. ^m Danach von Hand 2 nachgetragen und von Hand 4 gestrichen: Nu Chunrat Palczhart. <u>Dedit 2 qr.</u> ⁿ Über der Zeile von Hand 5 nachgetragen. ^o Von Hand 2 nachgetragen. ^p Von Hand 4 nachgetragen. ^q Über der Zeile von Hand 5 nachgetragen, statt</p>			

dessen der Nachtrag von Hand 4 über der Zeile gestrichen: nun der Pracher.
^r Über der Zeile von Hand 4 nachgetragen, statt dessen gestrichen: Hainczel
 Chaiss ain hoff 1 virl. ^s Danach der Nachtrag von Hand 2, der von Hand 4
 wieder gestrichen wurde: Dedit 2 qr. Dedit aber 2 qr. Danach von Hand 3
 nachgetragen und von Hand 4 gestrichen: Nun Ottel Sindleich. ^t Über der
 Zeile von Hand 5 nachgetragen. ^u Von Hand 2 nachgetragen. ^v Von
 Hand 4 nachgetragen. ^w Über der Zeile von Hand 5 nachgetragen.
^x Danach von Hand 2 nachgetragen und von Hand 4 gestrichen: Der Wolff.
Dedit 1 qr. ^y Danach von derselben Hand (Hand 5) gestrichen Tasler.
^z Unter der Zeile geschrieben, statt dessen der Eintrag von Hand 2 über der
 Zeile gestrichen: Dint Rattmanstorffer.

¹ Dörfel (R), KG Ponigl, OG Thannhausen, GB u. BH Weiz.

Abbildung 2: Traditionelle Editionsform der oben gezeigten Textpassage
 (MFU, fol. 17r); die Unterstreichungen stehen für Rubrizierungen.

Eine Erleichterung und Hilfe zur Unterscheidung der diversen Varianten
 und Schichten wird den Benutzern/innen mit einer anderen Repräsentati-
 onsform geboten: Die verschiedenen Schreiberhände sind dabei ausge-
 zeichnet und farblich hinterlegt, ebenso auch alle Arten von Streichungen
 und Nachträgen – jede Farbe steht für eine Hand und die Streichungen in
 der Handschrift werden auch im edierten Text ausgestrichen. Streng ge-
 nommen handelt es sich bei dieser Form des Textes um eine Transkription
 und keine kritische Edition, doch dieser Zwischenschritt wird den Benut-
 zern/innen aus genau einem Grund zugänglich gemacht: um die Transpa-
 renz und Nachvollziehbarkeit der Entscheidungen des Editors bzw. der
 Editorin zu gewährleisten.

Die eigentliche Innovation der computergestützten dynamischen Edi-
 tion des Marchfutterurbars ist jedoch die **Repräsentation der separierten
 Schichten**, bei der es sich um eine kritische Edition im eigentlichen Sinn
 handelt, da bei dieser Darstellungsform bereits das Wissen des Editors
 bzw. der Editorin über die Quelle in die Texte eingearbeitet worden ist. In
 diesem Urbar lassen sich fünf chronologisch-unterscheidbare Ebenen
 finden:

- Der Zustand von vor 1414 (auf einer verschollenen Vorlage basie-

- rend),
- der Zustand von 1414 mit den jeweiligen Abgaben,
 - der Zustand zwischen den Jahren 1414 und 1426, einer Periode, in der
 etliche Änderungen vorgenommen wurden,
 - der Zustand von 1426 mit den damaligen Abgaben,
 - und schließlich der Zustand nach 1426 mit den abschließenden Modi-
 fikationen.

Diese fünf Schichten sind als Variante „a“, „b“, „b-1“, „b-2“ und „c“
 bezeichnet und können entweder einzeln oder vergleichend untereinander
 ausgegeben werden (siehe Abbildung 3). Die Benutzer/innen können
 dadurch mühelos die unterschiedlichen Zustände und Veränderungen im
 Text wahrnehmen, sowie auch die entsprechende Abgabe eines Holdes zu
 einem bestimmten Zeitpunkt. Des Weiteren ist es natürlich auch möglich,
 zum Beispiel nur Variante „b“ und „c“ eines Eintrages zu vergleichen und
 sich diese neben dem Faksimile anzeigen zu lassen.

3.16a	Im Dorfflein ¹	
3.16c	Im Dorfflein ¹ facit 3 s 2 qr.	
3.16.1a	Item Chunrat Wamperer von ½ hub 1 viril.	
3.16.1b	Item Chunrat Wamperer von ½ hub 1 viril. <u>Dedit 1 viril.</u>	
3.16.1b'	Item Chunrat Wamperer nun Kunrat Krüg von ½ hub 1 viril. <u>Dedit 1 viril.</u>	
3.16.1b''	<u>Dedit 1 viril.</u> Item Chunrat Wamperer nun Herman sein sun von ½ hub 1 viril.	Dint hern Vlr(eich) von Stubenberg.
3.16.1c	<u>Dedit 1 viril.</u> Item Chunrat Wamperer nun Herman von ½ hub 1 viril.	Dint hern Vlr(eich) von Stubenberg.
3.16.2a	Item Hainczel Tewbel von 1 hub 1 viril von alter 1 s.	
3.16.2b	Item Hainczel Tewbel von 1 hub 1 viril von alter 1 s. Dedit 2 qr. Aber 2 qr.	Dint Stainbald.

3.16.2c	<u>Dedit 1 viri.</u>	Item Hainczel Tewbel von 1 hub 1 viri von alter 1 s.	Dint dem Stainbald.
3.16.3a		Item Herman Nidendran von 1 hub 1 viri von alter 1 s.	
3.16.3b		Item Herman Nidendran von 1 hub 1 viri von alter 1 s. Nu Chunrat Palczhart. <u>Dedit 2 qr.</u>	Dint Stainbald.
3.16.3b'	[Dedit ...] qr [...] Dedit aber 2 qr 1 mash.	Item Herman Nidendran nun Enderl Nidendran von 1 hub 1 viri von alter 1 s.	Dint Stainbald.
3.16.3c	[Dedit ...] qr [...] Dedit aber 2 qr 1 mash.	Item Herman Nidendran nun Enderl Nidendran von 1 hub 1 viri von alter 1 s.	Dint dem Stainbald.
3.16.4a		Item an dez Wamperer stat Hainczel Chais ain hoff 1 viri.	
3.16.4b		Item an dez Wamperer stat Hainczel Chais ain hoff 1 viri. <u>Dedit 2 qr.</u> Dedit aber 2 qr.	Dint Stainbald.
3.16.4b'		Item an dez Wamperer stat Hainczel Chais ain hoff 1 viri. <u>Dedit 2 qr.</u> Dedit aber 2 qr. Nun Ottel Sindleich.	Dint Stainbald.
3.16.4b''	Dedit [...] viri c.	Item an dez Wamperer stat nun der Pracher von 1 hof 1 viri. <u>Dedit 2 qr.</u> Dedit aber 2 qr.	Dint Stainbald.
3.16.4c	Dedit [...] viri c.	Item an dez Wamperer stat nun Hainczel Wolff von 1 hof 1 viri.	Dint dem Stainbald.
3.16.5a		Item Chuncz Cholbel von 1 hofstat 1 qr von alter 2 qr.	
3.16.5b		Item Chuncz Cholbel von 1 hofstat 1 qr von alter 2 qr. Der Wolff. <u>Dedit 1 qr.</u>	Dint Rattmanstorffer.

3.16.5c	<u>Dedit 1 qr.</u>	Item Chuncz Cholbel nun Hainczel Wolff von 1 hofstat 1 qr von alter 2 qr.	Dint dem ^a Tastler.
^a <i>Danach gestrichen</i> Tasler.			
¹ Dörfel (R), KG Ponigl, OG Thannhausen, GB u. BH Weiz.			

Abbildung 3: Repräsentation der separierten Schichten des oben genannten Textes (MFU, fol. 17r); die Unterstreichungen stehen für Rubrizierungen, Variante „a“, „b“, etc. für die verschiedenen Schichten.

Eine weitere Möglichkeit, die Linearität und Zweidimensionalität der gedruckten Edition zu durchbrechen und folglich eine neue Art des Zugriffs zu erstellen, bietet die **erweiterte Suche**. In diesem Kontext bedeutet diese Suchfunktion nicht nur die Verwendung Boolescher Operatoren, sondern sie beinhaltet vielmehr die Option unterschiedliche Einträge, die in einem Themengebiet zueinander in Beziehung stehen, auszugeben, ohne von der Seitenmetapher der Handschrift abhängig zu sein. So können sich die Benutzer/innen zum Beispiel alle Holden eines speziellen Grundherren, deren Einträge und Abgaben simultan anzeigen lassen, ohne darauf Rücksicht nehmen zu müssen, dass zwischen den Einträgen in der Originalquelle womöglich Dutzende Seiten liegen. Das erleichtert natürlich die Arbeit mit der Quelle und kommt den Historikern/innen zu Hilfe, deren Interessen in der Regionalgeschichte, Genealogie und Prosopografie liegen.

Jenen Benutzern/innen, die Informationen über eine bestimmte Region suchen, bietet das System einen zusätzlichen Zugang. Da ein sprachlich gestalteter Ortskatalog exakte Daten zu den Orts- und Flurnamen liefert, jedoch nur schwer einen Überblick über ein geografisches Gebiet geben kann, wird auch eine **digitale Landkarte** angeboten, in der alle Orte eingezeichnet und mit den betreffenden Textstellen in der Edition verlinkt sind. Somit können Zusammenhänge zwischen zwei benachbarten Orten viel leichter aufgedeckt werden, auch wenn deren Einträge in der Handschrift durch viele Seiten getrennt sind. Eine Anwendungsmöglichkeit

dafür wäre zum Beispiel die viel einfachere und präzisere Lokalisierung einer Gegend, in der es Ernteauffälle gab, wenn dies in mehreren Textpassagen der Handschrift erwähnt wird. Das kann für Wirtschaftshistoriker/innen von großer Relevanz sein, aber es gibt in dem System noch andere speziell für sie geschaffene Repräsentationen der Quelle: Die Abgaben der einzelnen Holden können in Tabellenform dargestellt werden und es ist des Weiteren möglich, mit diesen Daten zu rechnen bzw. sie statistisch auszuwerten und zu analysieren. Auch für Spezialisten/innen der Paläografie und der Kodikologie wird ein eigener Zugang bereitgestellt. Sie können das System über den paläografischen und den kodikologischen Kommentar erreichen, in dem zum Beispiel die verschiedenen Schreiberhände, die Wasserzeichen oder die Lagenformel der Handschrift beschrieben sind. Natürlich gelangt man von den Beispielen immer via Hyperlink wahlweise zu dem Faksimile oder dem edierten Text.

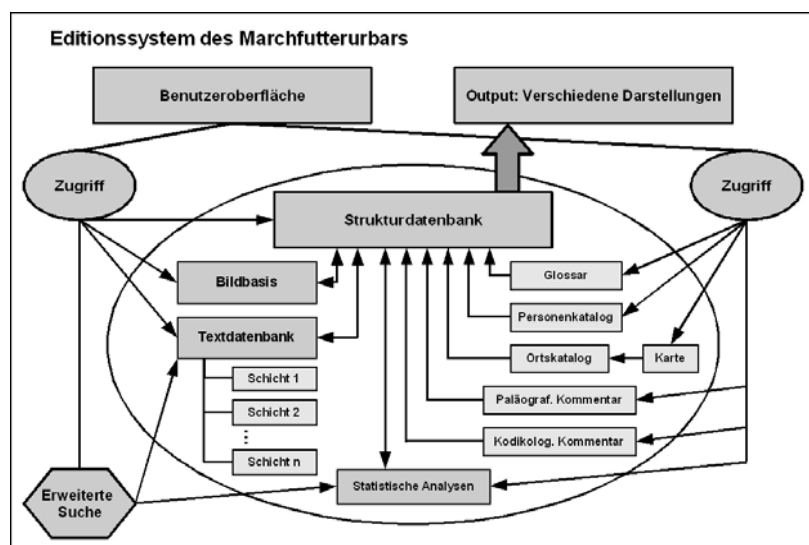


Abbildung 4: Funktionalitätsschema des Gesamtsystems

Aus all den zuvor genannten Überlegungen ergibt sich ein Gesamtsystem

für eine computergestützte dynamische Edition des Marchfutterurbars, das eine mehrdimensionale Quellenrepräsentation mit unterschiedlichen Zugriffsmöglichkeiten auf die Datenbanken beinhaltet und dessen Funktionalität in Abbildung 4 in abstrahierter Weise dargestellt ist.

Technische Aspekte des Editionssystems

Das Gesamtsystem ist als Client-Server-Lösung konzipiert und zielt darauf ab, dass zur Benutzung der digitalen Edition ein Standardbrowser³⁵ genügt. Um das zu gewährleisten, erhalten die Clients vom Server nur Dateien im XHTML-Format³⁶ und Bilder als JPEGs.³⁷ Die Dynamisierung des Systems entsteht dadurch, dass die auf dem Server enthaltenen XML-Datenbanken durch XSL-Transformationen³⁸ in die von den Benutzern/innen gewünschte Form verändert und spezifiziert werden. In Teilbereichen kommen auch php-Lösungen³⁹ mit MySQL-Datenbankverbindungen⁴⁰ zum Einsatz.

Im Inneren des Systems existieren Datenbanken, die Daten in unterschiedlichen Zuständen enthalten. Die transkribierten Rohtext-Daten werden – gemäß der Anforderung, offene Systemtechnologien zu verwenden

35 Der Browser muss derzeit lediglich Java und JavaScript verarbeiten und mit Frames umgehen können.

36 XHTML: eXtensible HyperText Markup Language, vgl. <<http://www.w3.org/TR/xhtml1>> (12.02.2006).

37 JPEG: Joint Photographic Experts Group, vgl. <<http://www.jpeg.org/index.html?langsel=de>> (12.02.2006).

38 XSL: eXtensible Stylesheet Language; XSL ist eine Familie von Sprachen zur Erzeugung von Stylesheets (Layouts) für XML-Dokumente und gliedert sich in zwei Untergruppen (Dialekte): XSL-FO (XSL Formatting Objects) und XSLT (XSL Transformation). XSLT ist eine in XML kodierte, deklarative Programmiersprache zur Transformation von XML-Dokumenten und wird gemeinsam mit XPath (XML Path Language) angewendet. XPath ist eine Pfadbeschreibungssprache, die zur Navigation innerhalb von XML-Dokumenten dient und die Teile eines Dokuments adressieren kann. Vgl. <<http://www.w3.org/TR/xslt20>> (13.02.2006); <<http://www.w3.org/TR/xpath>> (13.02.2006); Bogner, Frank, XSLT 2.0. Das umfassende Handbuch, Bonn 2005.

39 php: Serverseitige Skriptsprache, die auf WWW-Servern zur Generierung von dynamischen Inhalten dient. Vgl. <<http://www.php.net>> (12.02.2006).

40 MySQL ist ein SQL-Datenbankverwaltungssystem (SQL = Structured Query Language, strukturierte Abfrage-Sprache). Vgl. <<http://www.mysql.com>> (12.02.2006).

den, die internationalen Standards entsprechen – in der Extensible Markup Language (XML)⁴¹ ausgezeichnet, deren Struktur mittels einer XML Schema-Datei⁴² definiert ist. Diese Strukturierung folgt der Seitenmetapher und den einzelnen Einträgen, die jeweils eine eindeutige Identifikationsnummer erhalten. Zugleich wird der Text auch nach den verschiedenen Schreiberhänden, Nachträgen und Streichungen ausgezeichnet.⁴³ Aus diesen XML-Datenbasen werden mittels XSLT⁴⁴ neue XML-Dokumente generiert, die – wiederum auf eine XML Schema-Datei aufbauend – das Schichtenmodell der Handschrift darstellen. Diese Generierung erfolgt (fast) vollautomatisch den Transformationsroutinen des Editors bzw. der Editorin entsprechend.⁴⁵ Lediglich die Fußnoten müssen noch in den so entstandenen Editionstext eingefügt werden. Auf diese XML-Datenbasen, welche die Varianten des Editionstextes widerspiegeln, greifen die Benutzer/innen des Editionssystems zu und können sich ihre „eigene“ Edition zusammenstellen. Auch diese Modifikationen werden intern über XSLT verarbeitet und es entstehen dynamisch erzeugte XML-Dokumente. Diese Dokumente werden jedoch nicht im reinen XML-Format vom Server an den Client geschickt, sondern zuvor noch über eine weitere XSL-Transformation in XHTML-Dateien umgewandelt.

Neben diesen Textdatenbasen und der Bilddatenbank ist im System

41 XML: eXtensible Markup Language (erweiterbare Auszeichnungssprache), vgl. <<http://www.w3.org/xml>> (12.02.2006).

42 XML Schema ist eine komplexe Sprache zum Definieren von XML-Dokumentstrukturen, vgl. <<http://www.w3.org/TR/xmlschema-0>> (12.02.2006); Skulschus, Marco; Wiederstein, Marcus, XML Schema, Bonn 2004.

43 Die spätere XML-Notation zur Unterscheidung der einzelnen Schreiberhände und Varianten wurde bereits 1987 von Manfred Thaller in einem seiner Aufsätze vorweg genommen: vgl. Thaller, Manfred, *Secundum Manus*. Zur Datenverarbeitung mehrschichtiger Editionen, in: Härtel, Reinhard (Hg.), *Geschichte und ihre Quellen*. Festschrift für Friedrich Hausmann zum 70. Geburtstag, Graz 1987, S. 629-637.

44 Um XML Schema verarbeiten zu können, wird die relativ neue Version XSLT 2.0 verwendet.

45 Somit ist auch die Bedingung erfüllt, dass es sich hierbei um eine computergestützte Edition handelt, da ein formales Verfahren zur Erzeugung des Editionstextes angewendet wird. Vgl. dazu Kropač, *Theorien* (wie Anm. 9), S. 308.

noch eine weitere Datenbasis notwendig, welche die Verknüpfungen zwischen den unterschiedlichen Entitäten herstellt, aber auch als Schnittstelle zwischen den verschiedenen Zugriffsmöglichkeiten und den Primärdaten fungiert. In diese Strukturdatenbasis wird die gesamte abstrahierbare Information des Urbar eingetragene und mit XML ausgezeichnet. Dabei handelt es sich um alle in der Quelle genannten Orte, Personen, deren Abgaben und die dazugehörigen Grundherrschaften.⁴⁶ Die Strukturdatenbasis folgt im Gegensatz zur Textdatenbasis jedoch nicht der Seitenmetapher der Handschrift, sondern der logischen Struktur des Marchfutterurbar⁴⁷, wodurch eine „Loslösung“ der einzelnen Entitäten von den Seiten möglich wird. Als Referenz zwischen den unterschiedlichen Datenbasen dienen immer die eindeutigen Identifikationsnummern der Einträge, die sich an der Gliederung der Strukturdatenbasis orientieren.

Um die in der Strukturdatenbasis gespeicherte Information in einer umfangreicheren und anspruchsvolleren Weise auswerten zu können ist eine Datenbank notwendig, die mit komplex-strukturierten Daten aus historischen Quellen umgehen kann, sowie die Verwendung von Ontologien und weitreichende statistische Analysen ermöglicht. Genau für diese Bedürfnisse wurde das Softwarepaket *kleio*⁴⁸ geschaffen, das auf diesem

46 Die Struktur dieser Datenbank wurde hier nur sehr rudimentär erläutert und vereinfacht dargestellt. In Wirklichkeit ist sie bedeutend umfangreicher und enthält 175 (!) unterschiedliche Felder, die durch 50 XML-Elemente aufgebaut werden, was aufgrund einer modularen Anordnung ermöglicht wird.

47 Die logische Struktur des Marchfutterurbar ist in Pfarren, Ortschaften und „abgabepflichtige Immobilien“ gegliedert.

48 Das Programmpaket *kleio* wurde von Manfred Thaller in Zusammenarbeit mit mehreren europäischen Forschungsinstitutionen entwickelt. Für die datenbankorientierten Komponenten von *kleio* wurde ein kontextsensitives Datenmodell entwickelt, das seinerseits auf dem Konzept semantischer Netze basiert. Das Programm bietet (datentechnische) Hilfsmittel, der extremen Variabilität historischen Quellenmaterials und dessen inhärenten Unschärfen effizient zu begegnen. Da sich in historischen Quellen eine Reihe von Informationen, die durch die klassischen Datentypen nicht fassbar oder nicht unmittelbar interpretierbar sind, finden lassen, stellt *kleio* Mittel bereit, solche Informationen in ihren Merkmalen zu definieren und so neben den Quellen auch Wissen über sie in so genannten „logischen Objekten“ zu verwalten. Dazu zählen etwa Definitionsumgebungen für erweiterte Datentypkonzepte, der Ausgleich orthografischer Varianten, die Lemmatisierung des Lateinischen, numerische und

Gebiet seit Jahren zur Zufriedenheit vieler Historiker/innen arbeitet.

Im Editionssystem des Marchfütterurbars wird *kleio* eingesetzt, um mit den vielen unterschiedlichen, nichtdezimalen Maß- und Währungsangaben, die in der Quelle vorkommen, rechnen zu können. Außerdem wird dadurch auch eine Schnittstelle zum Statistikprogramm „SPSS“⁴⁹ angeboten. Neben dieser Verwendung werden noch andere Eigenschaften des Softwarepakets zur Erstellung von topografischen und prosopografischen Katalogen genutzt: *kleio* unterstützt den „Ausgleich orthografischer Varianten“ und „definierbare Verbindungen zwischen Entitäten mehrerer Datenbanken (Record Linkage)“.⁵⁰ Die Generierung der Kataloge geht dabei so vor sich, dass die unterschiedlichen Schreibweisen von Orts- oder Personennennungen im Urbar mit Hilfe des leicht variierten Soundex-Algorithmus⁵¹, der in *kleio* implementiert ist, ausgeglichen und miteinander verglichen werden. Anschließend werden die identifizierten Individuen in normalisierter Schreibweise in eine sekundäre Datenbank aufgenommen und über Record Linkage⁵² mit der Primärdatenbank, der Strukturdatenbank, verknüpft.

nicht numerische Klassifikationssysteme, definierbare Verbindungen zwischen Entitäten mehrerer Datenbanken (Record Linkage), die Verwaltung nichtdezimaler Maß- und Währungsangaben, die auch temporären Schwankungen unterworfen sein können und auch die Interpretation historischer Datierungsangaben. Daneben enthält *kleio* auch hochspezialisierte Komponenten wie etwa Interfaces zu Statistikprogrammen, zum Satzprogramm "TeX" und zu HTML. Vgl. Thaller, Manfred, „kleio“. A Database System (Halbgraue Reihe zur Historischen Fachinformation B 11), St. Katharinen 1993; Vgl. dazu auch Kropač; Kropač, Prolegomena (wie Anm. 22), wo der Umfang von *kleio* zusammengefasst dargestellt wird.

49 SPSS: Superior Performance Software System (früher stand SPSS für „Statistical Package for the Social Sciences“), vgl. <<http://www.spss.com>> (12.02.2006).

50 Kropač; Kropač, Prolegomena (wie Anm. 22).

51 „Soundex“ ist ein quasi-phonetischer Algorithmus zur Indizierung von Wörtern und Phrasen nach ihrem Klang. Vgl. <<http://www.archives.gov/publications/general-info-leaflets/55.html>> (12.02.2006).

52 Zu Vorgehensweise und Einsatz von „kleio“ für Record Linkage-Prozeduren siehe beispielsweise Woollard, Matthew; Denley, Peter, Source-Oriented Data-Processing for Historians: a Tutorial for Kleio (Halbgraue Reihe zur Historischen Fachinformation A 1), St. Katharinen 1993, S. 217-264.

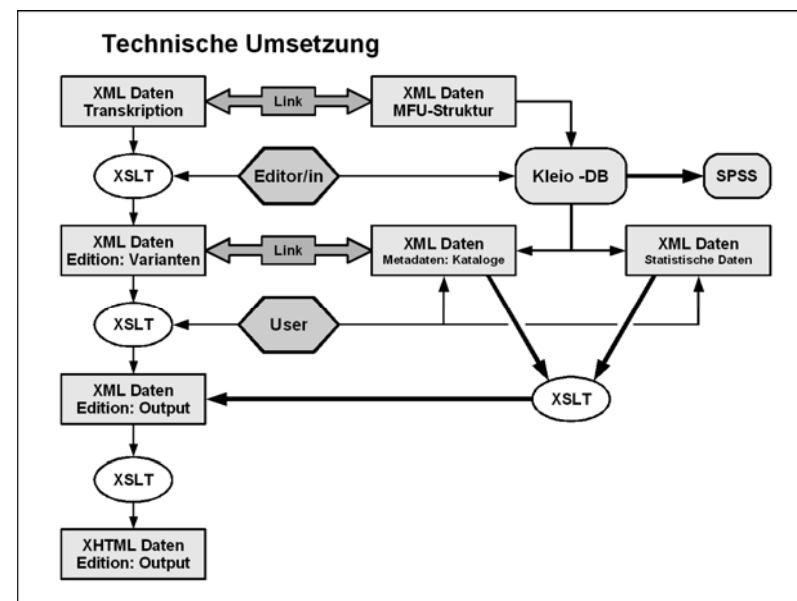


Abbildung 5: Technische Umsetzung des Gesamtsystems

Somit arbeitet das Gesamtsystem mit einer doppelt hybriden Form an Daten: Zum einen mit den XML-Daten in den Textdatenbanken und zum anderen mit der Strukturdatenbasis, in der XML-Daten und *kleio*-Datenbanken zusammenspielen.⁵³ Dabei werden die XML-Daten in die *kleio*-Datenbank eingelesen, wo sie analysiert, gespeichert und angereichert werden. Der von *kleio* generierte Output wird schließlich wieder in XML-(Meta-)Daten umgewandelt, die mit der Strukturdatenbasis und der

53 Vgl. dazu Thaller, Manfred, Texts, Databases, „kleio“: a Note on the Architecture of Computer Systems for the Humanities, in: Buzzetti, Dino; Pancaldi, Giuliano; Short, Harold (Hgg.), Augmenting Comprehension. Digital Tools and the History of Ideas (Office for Humanities Communication Publication 17), London 2004, S. 49-76; Ders., Vom verschwindenden Unterschied zwischen Datenbanken und Texten: Konsequenzen neuerer www-Technologie am Beispiel von museumsnahen Datenbanken, 2000, <http://www.museumtheuern.de/edvtage/g_mat/g12_thal.pdf> (13.02.2006).

Textdatenbasis verlinkt sind und über die verschiedenen Zugangsarten für die Benutzer/innen erreichbar sind. Mittels dieser Metadaten erfolgt die weitere Dynamisierung des Gesamtsystems, da sie das abstrakte Bindeglied unterschiedlicher Einträge auf einer Metaebene darstellen (siehe Abbildung 5).

Zusammenfassung

Ziel des Autors ist die kritische Edition eines speziellen Quellentypus, die aufgrund der Komplexität der Handschrift bisher noch nicht realisiert wurde. Vor allem das Abbilden der mehrschichtigen Struktur führte bis dato zum Scheitern einer gelungenen Edition. Der einzig brauchbare Weg, um die Informationsstruktur, den Informationsgehalt, die Informationsdichte und die Überlieferungsstruktur des Marchfutterurbars darzustellen, ist der Aufbau eines computergestützten dynamischen Editionssystems nach der Methode von „ICE-dynamisch“. Dieses System bietet verschiedene Zugriffe auf die Quelle, ermöglicht den Benutzern/innen leichter zwischen den unterschiedlichen Schichten der Quelle zu unterscheiden und befähigt sie, eine Repräsentation der Handschrift nach ihren eigenen Anforderungen zu erstellen. Die Einschränkungen einer kritischen Edition in gedruckter Form durch deren Linearität können so überwunden werden und die Mehrdimensionalität liefert einen Mehrwert an Information. Das Editionssystem des Marchfutterurbars offenbart sich daher als eine realisierbare Konzeption für eine dynamische Edition, wie sie Manfred Thaller⁵⁴ bereits 1988 angeregt und gefordert hatte.

*Matthias P. Perstling arbeitet am Institut für Geschichte an der Karl-Franzens-Universität Graz. Er ist Forschungsassistent an dem vom österreichischen Wissenschaftsfond und der Stadt Regensburg finanzierten Projekt *Fontes Civitatis Ratisponensis*. E-Mail: matthias.perstling@uni-graz.at*

KOGNITIVE KARTEN DES MITTELALTERS – DIGITALE ERSCHLIESSUNG MITTELALTERLICHER WELTKARTEN

von Günther Görz

Mittelalterliche Weltkarten sind in erster Linie als kognitive Karten zu verstehen. Um den Bestand und das Veränderungspotential des Text-Bild-Materials zu erfassen und zu erschließen, wird im Rahmen eines interdisziplinären Projekts in Erlangen ein systematischer vergleichender Stellenkatalog mittelalterlicher und frühneuzeitlicher Weltkarten erarbeitet. In einem ersten Schritt wurde eine multimediale Datenbank aus hochauflösenden digitalen Bildern circa 300 repräsentativer Weltkarten aufgebaut, die durch zugeordnete Metadaten erschlossen sind. Die Datenbank wird ergänzt durch spezielle Software-Werkzeuge zur Bildbearbeitung und zum Bildvergleich. Mit diesen Mitteln wird der Stellenkatalog erarbeitet, der alle Positionen umfassen soll, die auf den Mappaemundi des 13. bis 16. Jahrhunderts mit Bildern, Legenden und Bild-Text-Kombinationen verzeichnet sind. Als begriffliche Grundlage hierfür dient eine formale Ontologie, die zur Zeit in der beschreibungslogischen „Web Ontology Language“ OWL-DL formuliert wird; sie erweitert eine zunächst für den Behaim-Globus erstellte Konzeptionshierarchie und wird in das „Conceptual Reference Model“ der ICOM-CIDOC (International Committee for Documentation of the International Council of Museums) eingebettet. Diese Art der Formalisierung ermöglicht komplexe Anfragen, die weit über die Möglichkeiten herkömmlicher Datenbanken hinausgehen. Fernziel des Projekts ist die Ausarbeitung der in den Karten dargestellten kognitiven Beziehungen und deren Wandel auf der Grundlage des Stellenkatalogs.

⁵⁴ Thaller, Datenbasen (wie Anm. 8), S. 233f.

Einleitung: Projektdefinition und Ziele

Unter dem Titel „Kognitive Karten des Mittelalters“ wurde ein langfristig angelegtes interdisziplinäres Forschungsvorhaben der Erlanger germanistischen Mediävistik und der Informatik in Zusammenarbeit mit dem Graduiertenkolleg 516 „Kulturtransfer im europäischen Mittelalter“ initiiert. Die Zielvorgabe des Projekts definierten Hartmut Kugler und der Autor:

Die kosmografische und universalkartografische Überlieferung des Spätmittelalters beansprucht zunächst keine auf exakte Messungen und empirische Nachweise gestützte Geltung. Sie zielt vielmehr auf Orientierung im Sinne einer „Weltanschauung“. Dabei dominiert die Autorität von Schriftzeugnissen. Auf ein stark normierendes Verteilungsschema, wie es zum Beispiel besonders wirksam die *Imago mundi* des Honorius Augustodunensis (12. Jahrhundert) etabliert hat, scheinen die Autoren des 13. bis 16. Jahrhunderts nicht verzichten zu können. Selbst in Texten, die sich dezidiert als Reiseberichte und Abenteuerfahrten (Marco Polo, Mandeville) präsentieren, bleibt offenbar ein von der christlichen Enzyklopädik präformiertes „Weltbild im Kopf“ vorgeordnet, eine kognitive Karte, die den linear fortschreitenden Erzähltext organisieren half. Ungeklärt ist bislang, welche Gestaltungsfreiheiten und Variationsbreiten dabei möglich waren. Konnte sich jeder Verfasser eines Reiseberichts oder eines in ferne Erdgegenden führenden Abenteuerromans seine eigene kognitive Karte zurecht machen oder hielt er sich dabei an ein für „objektiv“ gültig gehaltenes Schema? Um diese Frage bearbeiten zu können, ist es hilfreich, die kartografischen Zeugnisse der Zeit systematisch heranzuziehen und auszuwerten. Die *Mappaemundi*, auch sie eine Frucht der christlichen Enzyklopädik, sind in erster Linie kognitive Karten. In ihnen wirken Parameter der *Ars memorativa* und der scholastischen Summenwerke. Ihre grafischen Muster und ihre Bild-Text-Kombinationen visualisieren die Handlungsräume narrativ dargestellter Geschichte und Geschichten. Eine *géographie absolue*, eine nur aus sich selbst heraus begründete Erdkunde, bieten sie nicht. Im Zeitraum zwischen dem ausgehenden 13. und der Wende zum 16. Jahrhundert (Herefordkarte circa 1290, Behaim-Globus 1492) ist aber die verbindliche Setzung und Zuordnung der einzelnen Positionen zu einem

Problem geworden, dem die Hersteller von Weltkarten sich zu stellen versuchten. Der Kanon der topografischen Positionen blieb nicht unverändert, neue Positionen traten hinzu, die es im Kartenschema unterzubringen und mit den alten ins Verhältnis zu setzen galt und die auch zu Veränderungen des Schemas zwangen. Diese Veränderungen lassen sich nur zum Teil evolutiv „aus sich heraus“ erklären. Vielmehr ist es aussichtsreich, sie als Resultate von Transferprozessen zu begreifen und zu studieren. Zum Beispiel führt von der Herefordkarte (in England um 1290 angefertigt) zum sogenannten „Katalanischen Weltatlas“ (um 1375 wohl auf Mallorca angefertigt) keine lineare historische Entwicklung, obwohl viele Positionen auf beiden Karten verzeichnet sind. Ihre Differenz erklärt sich wesentlich aus dem Hinzutreten arabisch-jüdischer Konventionen des Kartografierens, von denen der Hersteller der Herefordkarte nichts wissen konnte. Damit läßt sich die Fragestellung präzisieren und ein konkretes Arbeitsvorhaben ableiten:

- Welche Kontinuitäten, welche Varianten, welche Innovationen sind in der Universalkartografie des 13. bis 16. Jahrhunderts zu verzeichnen?
- Wo sind Schemabrüche und Diskrepanzen festzustellen?
- Wo ist mit der Fortschreibung einer Tradition, wo ist mit Transferprozessen zu rechnen?

Damit dem nachgegangen werden kann, wird ein vergleichender Stellenkatalog aller Positionen erarbeitet, die auf den *Mappaemundi* des 13. bis 16. Jahrhunderts (von der Hereford- und Ebstorfkarte bis zum Behaim-Globus) mit Bildern, Legenden und Bild-Text-Kombinationen verzeichnet sind. Der Stellenkatalog wird auch die kartografischen Schemata und damit die Position der einzelnen Bilder und Legenden sowie ihre kartenbildlichen Relationen zueinander über Bilddateien aufnehmen.

Parallel dazu werden spätmittelalterliche Sachtexte sowie Reise- und Abenteuerbücher auf die Positionen ihrer „erzählten Geografie“ durchzusehen sein und müssen daraufhin überprüft werden, wieweit sich ihre Positionierung mit dem aus den *Mappaemundi* gewonnenen Stellenkatalog vereinbaren läßt oder signifikant davon abweicht.

Die neuen von der Informatik erarbeiteten Methoden der digitalen Bilderfassung, Bildbearbeitung, formalen Erschließung und mehrdimen-

sional zugreifbaren Speicherung von Bildern bieten erstmals mit der Nutzung eines breitbandigen Netzes die Chance, bisher in der Regel schwierig zugängliche Originalkarten als digitale Reproduktionen in der für die Forschung erforderlichen Qualität an den Arbeitsplatz des Wissenschaftlers oder der Wissenschaftlerin zu bringen. Daher wird als zentrales Instrument der Forschung eine multimediale Datenbank repräsentativer mittelalterlicher und frühneuzeitlicher Weltkarten auf der Basis der einschlägigen Forschungsliteratur aufgebaut.¹ Sie besteht aus hochauflösenden digitalen Bildern, die durch zugeordnete Metadaten erschlossen sind, wodurch vielfältige Such- und Zugriffsmöglichkeiten eröffnet werden. Auf der Grundlage dieses Datenmaterials wird dann unter Zuhilfenahme einer Hierarchie einschlägiger Konzepte – das heißt komplexer Objekttypen, die in einer formalen Ontologie organisiert sind – der Stellenkatalog in der Form vernetzter Objektbeschreibungen ausgearbeitet. Über diesen Beschreibungen soll eine kognitiv orientierte Darstellungsebene eingerichtet werden, in der solche Beschreibungen nach visuellen, epistemischen, mnemotechnischen und narrativen Kriterien organisiert und unter Berücksichtigung der diachronen Dimension verknüpft werden können. Diese begriffshistorische Dimension wurde bisher in keinem der bekannten Wissensrepräsentationsschemata berücksichtigt.

Spezielle Benutzungsschnittstellen, wie sie zur Zeit im Kontext der digitalen Bibliotheken auf der Internet-Technologie basiert entwickelt werden, sollen das Hinzufügen beliebiger Annotationen und Dokumente und auch die direkte Einbeziehung der einschlägigen Sekundärliteratur gestatten. Den Mittelpunkt der verteilten Systemarchitektur bildet eine wissensbasierte Suchkomponente, die mit grafischer Unterstützung auch komplexe logische Anfragen über dem gesamten Datenbestand ausführen kann.

1 U.a. Destombes, Marcel, *Mappemondes A.D. 1200-1500*, Monumenta Cartographica Vetustioris Aevi, Amsterdam 1964 und Harley, J.B.; Woodward, David (Hgg.), *The History of Cartography*, Bd. 1, Chicago 1987.

Die Weltkarten-Datenbank

Mit dem Entwurf und der Implementierung einer Datenbank für repräsentative mittelalterliche Weltkarten wurde im Wintersemester 2003/2004 der erste Schritt zur Realisierung des „mappae“-Projekts begonnen. Sie ist gegenwärtig der zentrale Teil einer dem Projekt gewidmeten Website, in der ein Katalog dieser Karten in der Form von Objekt- und Bildbeschreibungen mit zugeordneten Kartenbildern zugreifbar ist. Dieser Datenbestand wird ergänzt durch eine Forschungsbibliografie mit zurzeit mehr als 600 Einträgen sowie durch einschlägige Texte, die online in der Form einer „Digitalen Bibliothek“ vorgehalten werden.

Bei der Auswahl der in den Katalog aufgenommenen Karten wurden alle drei wesentlichen Typen von Weltkarten berücksichtigt – Mappae-mundi, Ptolemäus Weltkarten und Portulane.² Dabei wurden Karten aus dem 9. bis 16. Jahrhundert mit einem Schwerpunkt auf dem 13. bis 16. Jahrhundert berücksichtigt. Maßgebend für die Auswahl war dabei die Verfügbarkeit von Daten, was die Beschreibungen und vor allem qualitativ hinreichende Bildvorlagen betrifft. Bisher wurden in den Katalog über 300 Karten aufgenommen, wobei für die weitaus überwiegende Zahl auch Bilder erfasst sind. Im Hinblick auf die Qualität der Bilder wurde, einer Einteilung der Reproduktionen von Thaller³ in die vier Klassen „illustrierend“, „lesbar“, „paläografisch“ und „enhanceable“ entsprechend, vor allem auf Vorlagen aus den beiden letzten Klassen Wert gelegt, da sie als einzige für eine wissenschaftliche Bearbeitung hinreichende Detailtreue bieten. In einigen Fällen wurden vorläufig auch Bilder geringerer Qualität

2 Als Vorlage für die Auswahl dienten einschlägige Aufstellungen und Klassifikationsschemata bei Arentzen, Jörg-Geerd, *Imago Mundi Cartographica – Studien zur Bildlichkeit mittelalterlicher Welt- und Ökumenekarten unter besonderer Berücksichtigung des Zusammenwirkens von Text und Bild* (Münstersche Mittelalter-Schriften 53), München 1984; von den Brincken, Anna-Dorothee, *Kartographische Quellen. Welt-, See- und Regionalkarten* (Typologie des sources du moyen âge occidental 51) Turnhout, Belgium 1988; Destombes, *Mappemondes* (wie Anm. 1), Harley; Woodward, *History* (wie Anm. 1) und Uhden, Richard, *Zur Herkunft und Systematik der mittelalterlichen Weltkarten*, in: *Geographische Zeitschrift* 37 (1931), S. 321-340.

3 Thaller, Manfred, *From the Digitized to the Digital Library*, in: *DLib Magazine*, Februar 2001.

aufgenommen, die aber im Lauf der Zeit durch bessere Reproduktionen ersetzt werden sollen.

Für die Beschreibungen der Karten wurde ein Metadatenschema erarbeitet, wobei für die Auswahl der Attribute aus Kompatibilitätsgründen auf das sehr einfache Merkmalsystem des deutschen prometheus-Verbundprojekts⁴ aufbauend Vorschläge der Visual Resources Association⁵ und weitere aus dem Bereich der Kunstgeschichte und der Museen wie zum Beispiel AMICO (Art Museum Image Consortium)⁶ berücksichtigt und um kartenspezifische Merkmale ergänzt wurden.

Auf dieser Grundlage wurde ein Datenbankschema definiert und in dem relationalen Datenbanksystem MySQL implementiert. Da üblicherweise nur eine Teilmenge aller Attribute für die Suche verwendet wird, wurde für die einfache Suche ein Satz von Suchattributen ausgezeichnet; daneben sollen aber wie bei den Online-Bibliothekskatalogen (OPAC) in einem Expertenmodus auch Suchanfragen über alle Attribute möglich sein. Eine wichtige Entwurfsentscheidung besteht in der Trennung der Metadaten für Originale, Faksimiles (Bildvorlagen) und Medienobjekte (digitale Bilder). Dadurch können Redundanzen in den Daten weitgehend vermieden werden. Um eine größtmögliche Vereinheitlichung der Attributwerte zu erreichen, wurden zum einen für viele Attribute Typbeschränkungen festgelegt – beispielsweise für Zeitangaben nach dem ISO-Standard –, und andererseits auch Terminologielisten integriert, die in der Benutzungsoberfläche als aufklappbare Menüs erscheinen. Dies soll künftig um direkte Zugriffsmöglichkeiten auf standardisierte Thesauri erweitert werden, zum Beispiel die Normdaten aus dem Bibliothekswesen für Personen und Körperschaften oder die Thesauri der Getty Foundation für Künstlernamen und geografische Bezeichnungen. Da die meisten Bilddaten urheberrechtlich geschützt sind, wurde für den Zugriff auf die Datenbank ein gestuftes System von Zugriffsrechten eingeführt. Auf ausgezeichneten Client-Rechnern im universitären Intranetz ist ein uneingeschränkter Zugriff

4 <<http://www.prometheus.de>> (15.02.2006).

5 <<http://www.vrweb.org/vracore3.htm>> (15.02.2006).

6 <<http://www.amico.org/AMICOLibrary/dataspec.html>> (15.02.2006).

möglich, der auf Antrag für Forschungszwecke auf registrierte externe Benutzer erweitert werden kann, wohingegen im Internet nur gemeinfreie Daten sichtbar sind. Es wird erwogen, mittelfristig die Bilder zusätzlich mit digitalen Wasserzeichen zu schützen.

Abbildung 1: Die Suchmaske

Des Weiteren wurden einige Software-Werkzeuge für die Eingabe, Korrektur und Suche implementiert. Da das gesamte System über jeden leistungsfähigen Web-Browser zu bedienen sein soll, wurden mittels PHP entsprechende Masken geschaffen, sodass Eingabe und Korrektur von

Metadaten ohne weitere technische Vorkenntnisse möglich sind. Da bei der ersten Befüllung der Datenbank größere Datenmengen anfielen, wurde zusätzlich eine Schnittstelle zur Übernahme von Daten aus separat erstellten Excel-Tabellen und speziell für Bibliografien auch aus BibTeX-Dateien eingerichtet. Was die Bilddaten betrifft, musste zunächst eine Auswahl der am Besten erreichbaren Reproduktionen aus der Literatur vorgenommen werden, da nahezu keine frei zugänglichen digitalen Bilder in hinreichender Qualität verfügbar sind. Eine Ausnahme bilden die in Erlangen erstellten digitalen Bilder der Ebstorfer Weltkarte und die Bilder des Behaim-Globus, die in einem früheren Projekt von Groß-Diapositiven, die das Germanische Nationalmuseum Nürnberg dankenswerterweise für Forschungszwecke zur Verfügung gestellt hatte, digitalisiert worden waren. Die ausgewählten Vorlagen wurden mit einer Scan-Auflösung von 400 oder 600 dpi eingescannt, mit Photoshop farblich nachbearbeitet, gegebenenfalls aus Teilscreens zusammengesetzt und dann als Masterdateien im TIFF-Format abgelegt. In die Datenbank werden für die Präsentation daraus erzeugte JPEG-Dateien in drei Auflösungen – volle Größe, Bildschirmgröße und „Thumbnail“ – aufgenommen. Für eine spätere Projektphase ist geplant, in einigen wenigen Fällen direkt von Originalen digitale Fotografien anzufertigen.

Suchanfragen können direkt in eine im Web-Browser aufrufbare Suchmaske eingegeben werden (Abbildung 1). Die Ergebnisse werden als „Thumbnails“ mit einer Kurzbezeichnung präsentiert und können dann in den beiden anderen Größen angezeigt werden, ebenso ihre Metadaten (Abbildung 2). Ist die Treffermenge zu groß, kann die Suche mit weiteren Angaben schrittweise eingeschränkt werden.

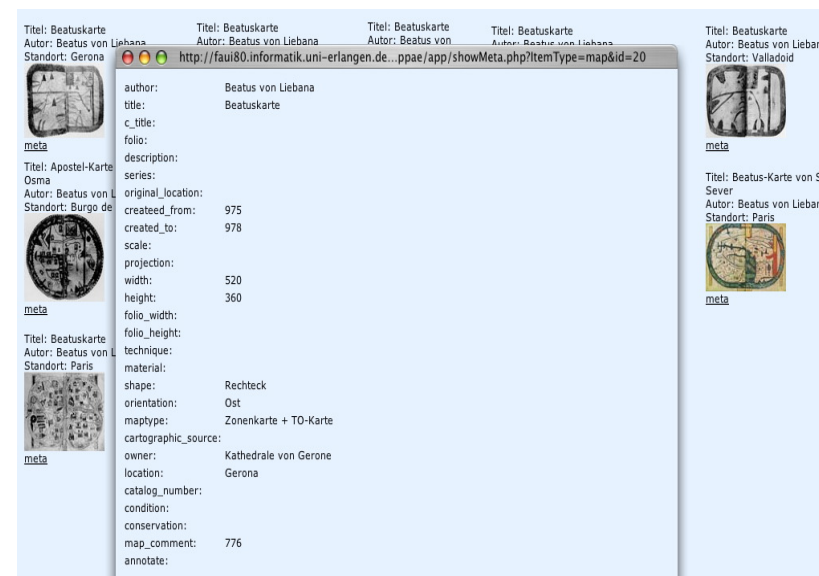


Abbildung 2: Suchmaske mit Ergebnisanzeige (Metadaten)

Für die Anzeige der Bilder in der höchsten Auflösungsstufe wird das in jedem modernen Web-Browser laufende Anzeigeprogramm DIGILIB benutzt, das am Max-Planck-Institut für Wissenschaftsgeschichte Berlin und der Universität Bern entwickelt wurde und als OpenSource-Software verfügbar ist.⁷ Es bietet den Benutzern und Benutzerinnen eine Zoom-Funktion mit der Möglichkeit, einen Bildausschnitt anzuwählen. Bei jeder Suchanfrage liefert der Server immer nur den jeweils angesteuerten aus dem Gesamtbild errechneten Bildausschnitt, was der Performanz sehr zugute kommt. Große Bilder können auf dem Server auch gekachelt abgespeichert werden. Weiterhin besteht die Möglichkeit, im Bild Marken zu setzen und Bilder oder Bildausschnitte durch URLs zu referenzieren (Abbildung 3). Diese URLs können benutzerseitig gespeichert oder versandt und zu einem späteren Zeitpunkt wieder aufgerufen werden, was eine

⁷ <<http://digilib.berlios.de/>> (15.02.2006).

Neuberechnung des betreffenden modifizierten Bilds bewirkt.

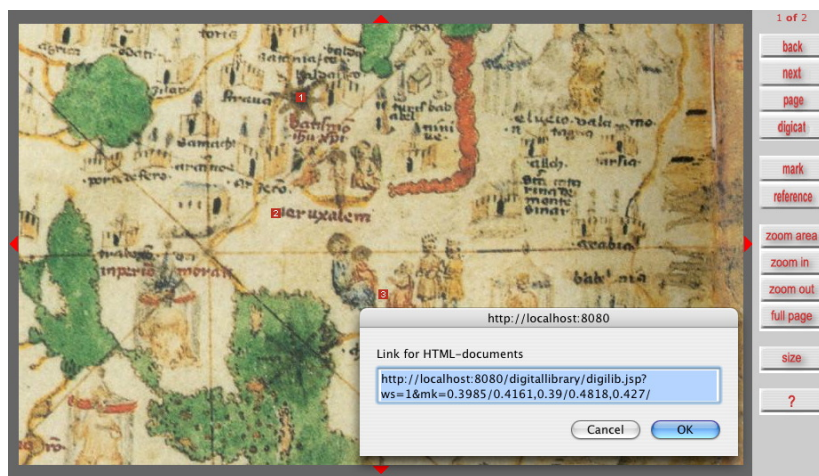


Abbildung 3: DIGILIB-Anzeige mit Marken und generierter URL

Damit ist der status quo der Kartendatenbank skizziert. Künftige Erweiterungen sollen ermöglichen, zu den Marken in DIGILIB Annotationen anzulegen und einige Bildbearbeitungsfunktionen, zum Beispiel zur Manipulation von Helligkeit, Kontrast und Farben, bereitzustellen. Um Bilder in systematischer Weise statisch vergleichen zu können, wurde auf DIGILIB aufbauend das Werkzeug „COMPAGO“⁸ entwickelt, das unter anderem zwei Bilder in nebeneinander liegenden Fenstern anzeigen kann. In unserem Fall würde hiermit der Vergleich von Darstellungen gleicher oder ähnlicher Motive erheblich erleichtert – neben geografischen Objekten verschiedenste Miniaturen wie beispielsweise Stadtsilhouetten, Menschen, Tiere, Monster, spezifische Szenen und allgemein Darstellungen des Fremden oder von Elementen der Heilsgeschichte. Daher soll auch COMPAGO nach Möglichkeit verfügbar gemacht werden.

8 Graßhoff, Gerd; Liess, Hans-Christoph; Nickelsen, Kärin (Hgg.), COMPAGO – Der systematische Biodvergleich. Handbuch (Bern Studies in the History and Philosophy of Science. Educational materials 3), Bern 2001.

MapViewer: Dynamischer Vergleich und 3D-Darstellung von Karten

Zum dynamischen Vergleich und zur 3D-Darstellung von Karten wurde im Projekt das Programm MapViewer⁹ implementiert, das als eigenständige Komponente auf dem lokalen Arbeitsplatzrechner oder aber auf dem Projektserver ausgeführt werden kann. Ihm liegt die Methode der Bildregistrierung zugrunde, ein Verfahren, um zwei Bilder, die analogen Inhalt in unterschiedlicher Darstellung abbilden, aneinander anzupassen.¹⁰ Alle Registrierungsmethoden nutzen das folgende Schema: Zuerst müssen Merkmale im Ausgangs- und Vergleichsbild erkannt beziehungsweise festgelegt werden, die dann aneinander anzupassen sind (Abbildung 4). Dazu ist ein Transformationsmodell zu bestimmen, wofür globale oder lokale Abbildungsmodelle benutzt werden können. Mit seiner Hilfe wird das Ausgangsbild dann transformiert. In beiden Typen von Abbildungsmodellen werden Rotations-, Skalierungs- und Translationsoperationen benutzt, wobei im globalen Fall eine affine Transformation und im lokalen Fall verzerrende Komponenten mit sogenannten Radial Basis Functions eingesetzt werden. Neben der starren Registrierung durch affine Transformation ist auch eine flexible Form mithilfe von Thin-Plate-Splines möglich. Das Ergebnis kann Überlagerung des transformierten Ausgangsbilds mit dem Vergleichsbild („Alpha-Blending“, Abbildung 5) oder durch Verzerrungsgitter (Abbildung 6) veranschaulicht werden. Letztere sind ein Beiprodukt der Registrierung, wobei zwei Möglichkeiten bestehen: Entweder legt man ein gleichmäßiges Gitter über die Ausgangskarte oder man benutzt das Projektionsgitter der Vergleichskarte und verzerrt das Gitter dann entsprechend.

9 Hofmann, Andrea, Vergleich und 3D-Darstellung von alten Landkarten durch Bildregistrierung (Diplomarbeit im Fach Informatik am Institut für Informatik 8, Universität Erlangen-Nürnberg), Erlangen 2005.

10 Gottesfeld-Brown, Lisa, A Survey of Image Registration Techniques, in: ACM Computing Surveys 24 (1992), S. 325-376; Zitová, Barbara; Flusser, Jan, Image Registration Methods. A Survey, in: Image and Vision Computing 21 (2003), S. 977-1000.

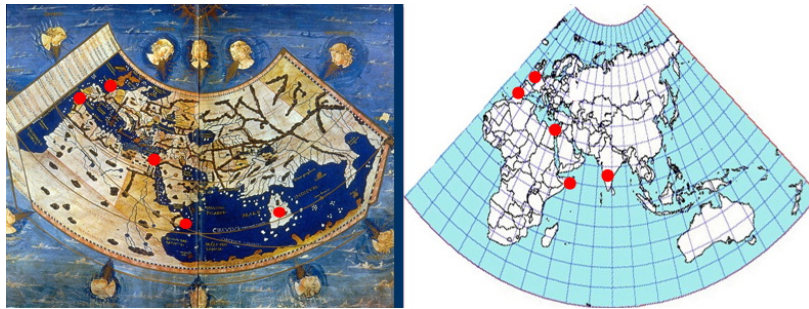


Abbildung 4: Bildregistrierung – gesetzte Passpunkte

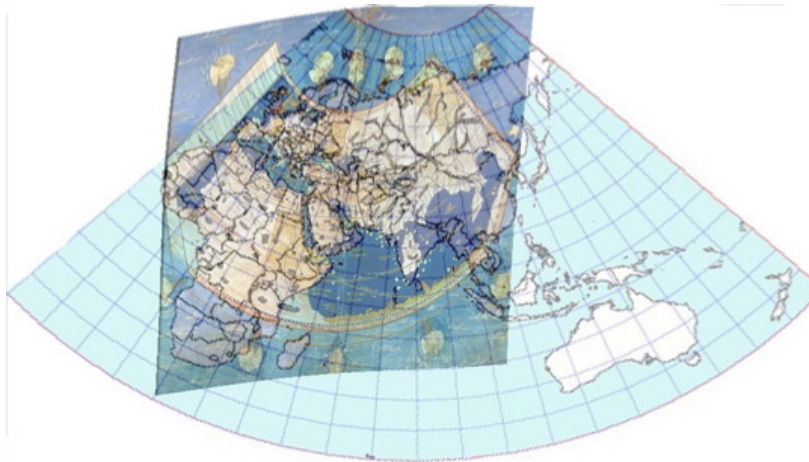


Abbildung 5: Bildregistrierung – „Alpha-Blending“



Abbildung 6: Bildregistrierung – Verzerrungsgitter

Überraschende Einblicke können sich ergeben, wenn historische Karten auf eine Kugel projiziert werden: Simek¹¹ und andere haben darauf hingewiesen, dass etwa in mittelalterlichen Weltkarten, die nach dem T-O-Schema¹² konstruiert sind, durchaus bemerkenswerte geografische Aspekte deutlich werden, wenn man sie auf eine Kugelschale projiziert. Die Analogie zur Ansicht aus einem geostationären Satelliten über dem Mittelpunkt, zumeist Jerusalem, ist nicht von der Hand zu weisen. Auch die Abbildung von Ptolemäuskarten auf eine Kugel eröffnet interessante Gesichtspunkte (Abbildung 7).

Alle Kartenprojektionen bilden ein geografisches Koordinatensystem mit Längen- und Breitengraden in das kartesische der zweidimensionalen Karte ab. Zur 3D-Projektion müssen also unter Berücksichtigung der Grenzen für Länge und Breite die Koordinaten der ebenen Karte in Kugelkoordinaten umgerechnet werden, um diese dann als Textur auf die Kugel abzubilden. Die modernen mathematischen Projektionstypen werden in azimutale, konische und zylindrische eingeteilt.¹³ Liegt eine Karte eines

11 Simek, Rudolf, Erde und Kosmos im Mittelalter, München 1992.

12 „Terrarum Orbis“, Bezeichnung für kreisförmige („O“) geostete Weltkarten, deren obere Hälfte vom asiatischen Kontinent eingenommen wird, die linke untere Hälfte von Europa und die rechte von Afrika. Getrennt werden die Kontinente durch „T“-förmig dargestellte Wassermassen, wobei Don und Nil den Querbalken und das Mittelmeer den Längsbalken bilden.

13 Schröder, Eberhard (Hg.), Kartenentwürfe der Erde. Kartographische Abbildungsverfahren aus mathematischer und historischer Sicht (Mathematische

bekannten Projektionstyps vor, so werden ein Projektionsgitter über die Karte gelegt, die 3D-Geometrie der Gitterpunkte und deren Texturkoordinaten bestimmt und dann die 3D-Projektion mittels der Programm-bibliothek OpenInventor¹⁴ durchgeführt. Mittelalterliche Karten können keinem dieser Projektionstypen zugeordnet werden¹⁵; daher ist zunächst die Registrierung mit einer Referenzkarte bekannter Projektion erforderlich und dann für die registrierte Karte die 3D-Projektion auszuführen.



Abbildung 7: 3D-Projektion einer Ptolemäuskarte

Schülerbücherei 128), Leipzig 1988; Hoschek, Josef, *Mathematische Grundlagen der Kartographie* (BI-Hochschul-Taschenbücher 443/443a), Mannheim 1984; Spallek, Karlheinz, *Kurven und Karten*, Mannheim 1980.

14 <<http://www.sgi.com>> (15.02.2006).

15 Dies trifft auch auf die sogenannten Ptolemäischen Projektionen und Portulane zu; zu historischen Projektionen allgemein siehe: Keuning, Johannes, *The History of Geographical Map Projections until 1600*, in: *Imago Mundi* 12 (1955), S. 1–24.

Besonders instruktiv ist die Animation derart erzeugter virtueller Globen, vor allem das interaktive Drehen und Zoomen. Hierfür ist ein Betrachter für VRML-Dateien erforderlich.¹⁶ VRML, die Virtual Reality Modelling Language, ist eine Beschreibungssprache für 3D-Szenen sowie deren Geometrie, Ausleuchtungen und Animationen, die verschiedene Interaktionsmöglichkeiten bietet. Sie wurde ursprünglich als Erweiterung des OpenInventor-Formats eingeführt; da letzteres nicht von allen VRML-Viewern unterstützt wird, führen wir einen expliziten Export nach VRML durch.

Der Stellenkatalog

Konzeption des Stellenkatalogs

Für den Stellenkatalog muss zuerst eine Systematik für die in ihm abzulegenden Objektbeschreibungen, das heißt ein Klassifikationsschema für die Objekttypen und ihre Eigenschaften, erarbeitet werden. Diese Beschreibungen sind mit Positionen auf den zugehörigen Kartenbildern zu verankern. Hierfür wird von der Markierungstechnik in DIGILIB Gebrauch gemacht, sodass die betreffenden Bildausschnitte mit Marken durch Aufruf der generierten URLs angezeigt werden können. Der Zugriff auf die Einträge des Stellenkatalogs wird durch eine Anfragemaske erfolgen.

Eine derartige, allerdings vereinfachte Möglichkeit ist zum Beispiel in der CD-ROM „Mapamondi“¹⁷ (Katalanischer Weltatlas) der Bibliothèque Nationale de France realisiert; dort können mit diesen Links nur einfache Annotationen in der Form unstrukturierter Texte verbunden werden. Neben der Möglichkeit, benutzerseitig Marken und Annotationen einzutragen, sind vorgegebene Marken verzeichnet, die auf Tabellen mit geografischen Bezeichnungen von Orten in ihrem Umfeld und den Transkriptionen in die moderne Form verweisen (Abbildung 8).

16 VRML-Viewer gibt es als Plugin für verschiedene Browser, zum Beispiel den Cortona VRML-Client von parallelgraphics.org, aber auch als separate Applikationen.

17 Bibliothèque Nationale de France (Hg.), *Mapamondi – une carte du monde au XIVe siècle* (Atlas Catalan). CD-ROM, Paris 1998.

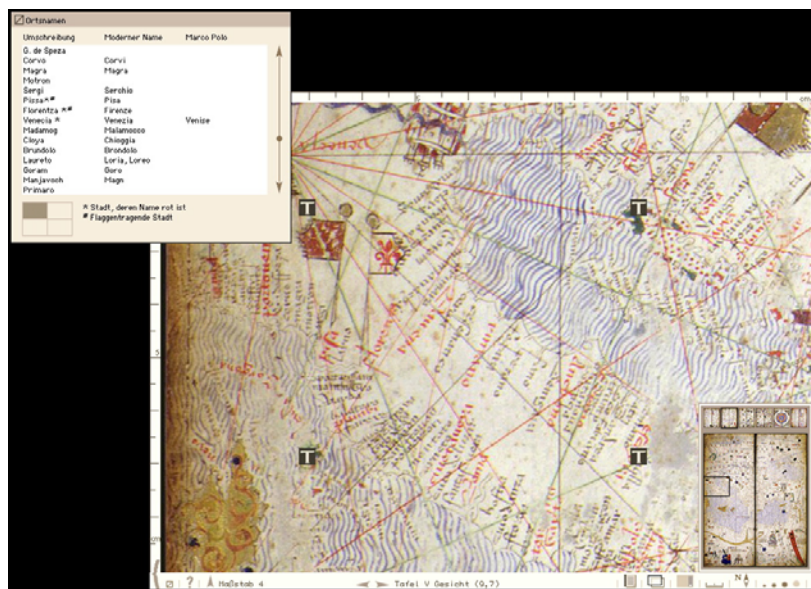


Abbildung 8: CD-ROM „Mapamondi“ der BNF: Bildausschnitt mit Toponymen

Ein anders gearteter Ansatz wird in der von Kline und anderen entwickelten CD-ROM „A Wheel of Memory — The Hereford Mappamundi“¹⁸ verfolgt. Hier ist ein Stellenkatalog in einer relationalen Datenbank abgelegt. Der Datenbank liegt ein System von Kategorien und Subkategorien mittlerer Granularität zugrunde, das dem Benutzer über eine Suchmaske zugänglich ist, in der er entweder selbst eine Eingabe eintippen oder aus vorgegebenen Listen, die als dynamische Menüs angeboten werden, jeweils einen Wert auswählen kann (Abbildung 9). Die Suchergebnisse werden in tabellarischer Form präsentiert und das Bildfenster über dem passenden Bildausschnitt positioniert. Darüber hinaus kann auch in einem Bild der

18 Kline, Naomi Reed, A Wheel of Memory. The Hereford Mappamundi. CD-ROM, Ann Arbor 2001; Dies., Maps of Medieval Thought. The Hereford Paradigm, Woodbridge, Suffolk 2001.

Karte, das aber immer nur einen relativ kleinen Ausschnitt zeigt, grafisch navigiert werden. Es kann stets nur nach einzelnen Stellen oder dargestellten Objekten auf der Karte gesucht werden. Eine weitergehende Verknüpfung der Datenbankeinträge wurde nicht vorgesehen; vielmehr sind zu bestimmten Themen, beispielsweise zur Alexandersage, lediglich feste Bildfolgen vorgegeben, die nicht direkt mit der Datenbank verknüpft sind.



Abbildung 9: CD-ROM „A Wheel of Memory“: Datenbankabfrage

Formale Domänen-Ontologie und Referenzontologie

Im Unterschied zu den beiden genannten Beispielen werden die Einträge in unserem Stellenkatalog aus strukturierten Beschreibungen bestehen. Die grundsätzliche Vorgehensweise zur semantischen Erschließung wurde in einem vorangegangenen Projekt zum Behaim-Globus¹⁹ entwickelt; dieser

19 Görz, Günther; Holst, Norbert, The Digital Behaim Globe (1492), in: Bearman,

Ansatz wird in veränderter Form hier wieder aufgenommen.

Voraussetzung für die Erstellung strukturierter Beschreibungen ist der Entwurf eines objekt-zentrierten Konzeptmodells, das alle Typen von Objekten enthält, die im Stellenkatalog erfasst werden sollen. Dazu gehören geografische Darstellungen, zum Beispiel Inseln oder Stadtsilhouetten, ebenso wie Miniaturen nichtgeografischen Inhalts, zum Beispiel das Paradies oder Tiere, sowie Inschriften verschiedener Art. Von grundsätzlicher Relevanz ist dabei die Unterscheidung verschiedener Beschreibungsebenen: Wir gehen aus von der materiellen, grafischen (1) Darstellung bestimmter (2) konkreter oder abstrakter Gegenstände und Sachverhalte auf einem (3) materiellen Objekt – hier einer Karte –, das entweder (4) direkt digitalisiert oder von einer (5) analogen Reproduktion digital erfasst wurde. Dies muss sich in verschiedenen Klassen deskriptiver Metadaten reflektieren, die im Konzeptmodell zu berücksichtigen sind: zum einen Metadaten für die Karten, ihre Reproduktionen und die davon angefertigten digitalen Bilder mit ihren jeweiligen bibliografischen und technischen Angaben, zum anderen dann für die dargestellten Gegenstände und Sachverhalte sowie ihre grafische und typografisch/textuelle Darstellung.

Dieses Konzeptmodell wird als taxonomische Hierarchie entworfen, das heißt, zu jedem Objekttyp („Konzept“) wird ein „Rahmen“ („Frame“) von Attributen (Merkmalen, Eigenschaften) festgelegt und die Konzepte selbst sowie gegebenenfalls die Attribute werden in einer Vererbungshierarchie angeordnet, mit anderen Worten in der Gestalt einer „formalen Ontologie“. Jede Objektbeschreibung ist nichts anderes als ein für den jeweiligen Objekttyp mit Werten für seine Attribute ausgefüllter Rahmen, eine „Instanz“ (siehe unten). Um möglichst flexible logische Verknüpfungen formulieren zu können, wird als Repräsentationssprache ein entscheidbares Fragment der Quantorenlogik erster Stufe, die sogenannte Beschreibungslogik, gewählt, die für derartige Repräsentationen besonders geeignete Sprachelemente bietet. Für Beschreibungslogiken stehen effi-

David; Trant, Jennifer (Hgg.), *Museum Interactive Multimedia 1997. Cultural Heritage Systems – Design and Interfaces. Selected Papers from ICHIM-97, The Fourth International Conference on Hypermedia and Interactivity in Museums, Archives & Museum Informatics and Louvre, Paris 1997*, S. 157-173.

ziente Inferenzalgorithmen bereit, die vollständige und korrekte Schlüsse aus komplexen, logisch zusammengesetzten Anfragen garantieren.²⁰ Für die Implementation der formalen Ontologie setzen wir die im Rahmen des „Semantic Web“ entwickelte „Web Ontology Language“ OWL-DL²¹ ein, die eine der am weitesten entwickelten Beschreibungslogiken auf der syntaktischen Basis von XML darstellt.

Eine formale Ontologie definiert das terminologische System für einen Gegenstandsbereich und ist normalerweise an einschlägigen Theorien, durch die dieser erschlossen wird, orientiert.²² Im Idealfall stützt sie sich auf den Begründungszusammenhang der jeweiligen Theorien ab und benutzt die diesen Theorien zugrundeliegenden Abstraktionsverfahren. Sofern es sich um axiomatische Theorien handelt, wird die formale Ontologie auch die axiomatische Basis repräsentieren; ob und in welchem Umfang dies im konkreten Fall möglich ist, hängt von der Ausdruckskraft der gewählten Ontologiesprache ab.

Nun greift jede Domänen-Ontologie auf Allgemeinbegriffe für Zeit und Raum, Ereignisse, Akteure, Prozesse und weitere zurück, die für die begriffliche Modellierung aller lebensweltlichen Gegenstandsbereiche von Nutzen sind. Darüber hinaus müssen auch die für derartige Modellierungen benutzten logisch-mathematischen Grundbegriffe wie Klasse, Relation, Zahl, aber auch Mereologie (Teil-Ganzes-Beziehungen) festgelegt werden, die üblicherweise auf einer Meta-Ebene vorgesehen sind. Zu diesem

20 Donini, Francesco; Lenzerini, Maurizio; Nardi, Daniele; Schaerf, Andrea, Reasoning in Description Logics, in Brewka, Gerhard (Hg.), *Foundations of Knowledge Representation*, Stanford 1996, S. 191-236.

21 Smith, Michael K.; Welty, Chris; McGuinness, Deborah L., *OWL Web Ontology Language Guide. W3C Recommendation 10 February 2004*, W3C (World Wide Web Consortium), Geneva 2004.

22 Zu den theoretischen Aspekten formaler Ontologien siehe unter anderem: Menzel, Chris, *Ontology Theory*, in: Euzenat, Jérôme; Gomez-Perez, Asuncion; Guarino, Nicola; Stuckenschmidt, Heiner (Hgg.), *Ontologies and Semantic Interoperability, Proc. ECAI-02 Workshop, (CEUR-WS 64)*, Lyon 2002, Guarino, Nicola, *Formal Ontology and Information Systems*, in: Ders. (Hg.), *Formal Ontology in Information Systems. Proceedings of FOIS-98, Trento, Italy, 6-8 June 1998*, IOS Press, Amsterdam 1998, S. 3-15, zu deren praktischen Aspekten zum Beispiel: Noy, Natalya, *Ontologies*, in: Farghaly, Ali (Hg.), *Handbook for Language Engineers*, Stanford 2003, S. 181-211.

Zweck wurden formale Referenzontologien entwickelt, mit denen dann Domänen-Ontologien begrifflich zu verknüpfen sind, so dass sich die spezifischen Konzepte als Spezialisierungen von allgemeinen Konzepten ergeben. Aus technischer Sicht liegt die besondere Bedeutung von Referenzontologien darin, dass sie die Grundlage für semantische Interoperabilität bieten, zum Beispiel für übergreifende Recherchen.

Für den uns interessierenden Anwendungsbereich erscheint das vom Internationalen Komitee für die Dokumentation des „International Council of Museums“ (ICOM-CIDOC) entwickelte objektorientierte „Conceptual Reference Model“ (CRM)²³ besonders geeignet. Das CRM wurde inzwischen unter dem Titel „Information and documentation – A reference ontology for the interchange of cultural heritage information“ als ISO-Standard 21127 angemeldet. Es war zunächst entwickelt worden, um alle in der Museumsdokumentation relevanten Konzepte zu erfassen, ganz besonders die für den museumsübergreifenden Datenaustausch; es ist aber inzwischen weit darüber hinaus gediehen. Die wichtigsten Klassen des CRM sind in Abbildung 10²⁴ wiedergegeben sowie einige wichtige Beziehungen zwischen Ereignissen (E5 Event), Personen (E39 Actor), Gegenständen (E18 Physical Stuff), Zeiten (E52 Time-Span) und Orten (E53 Place) angedeutet. Eine zentrale Rolle beim Modellierungsansatz spielen die Merkmale (Properties), über denen die Klassen abstrahiert werden, sowie die Beziehungen (134 bei 81 Klassen) zwischen den Klassen. Da an dieser Stelle nicht weiter auf das CRM eingegangen werden kann, sei abschließend hervorgehoben, dass das CRM die den verschiedenen Datenbank-Schemata für vorhandene Dokumentation zugrunde liegende Seman-

23 Siehe: Crofts, Nick; Doerr, Martin; Gill, Tony; Stephen Stead; Stiff, Matthew, Definition of the CIDOC Conceptual Reference Model. Version 4.2, The International Committee for Documentation of the International Council of Museums (ICOM-CIDOC), Paris, June 2005, Produced by the ICOM/CIDOC Documentation Standards Group, continued by the CIDOC CRM Special Interest Group; weitere detaillierte Informationen sind zu finden unter: <http://cidoc.ics.forth.gr/> (15.02.2006). Eine sehr lesenswerte Übersicht mit Beispielen bietet: Ermert, Axel u.a., Das CIDOC Conceptual Reference Model. Eine Hilfe für den Datenaustausch?, Berlin 2004.

24 Ermert, CIDOC (wie Anm. 23).

tik und die Struktur von Dokumenten thematisiert. Die Terminologie, die in den Daten verwendet wird, ist selbst nicht Gegenstand des CRM; es werden aber Schnittstellen zu standardisierten Terminologien wie Thesauri zu Personen- und Ortsnamen empfohlen.

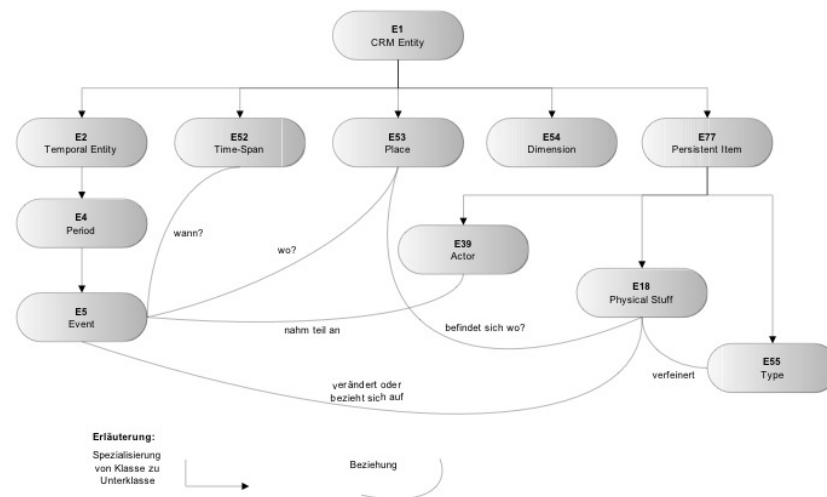


Abbildung 10: Wichtige Klassen (Entities) der CRM-Klassenhierarchie

Vorgehensweise zur Erstellung der formalen Domänen-Ontologie

Als Vorgehensweise zur Erstellung von Domänen-Ontologien hat es sich bewährt, zuerst einen Katalog charakteristischer Fragen aufzustellen. Dieser soll an typischen Beispielen zeigen, was in unserem Fall im Stellenkatalog dargestellt und damit Gegenstand späterer Anfragen werden kann. Die Fragenliste wird zunächst anhand einiger reichhaltig ausgestatteter typischer Weltkarten sowie unter Berücksichtigung der für den Behaim-Globus bereits vorhandenen Konzepthierarchie erstellt und kann später im Hinblick auf weitere Karten erweitert und modifiziert werden. Da der Behaim-Globus Elemente aus den drei wesentlichen Typen von Weltkarten – Mappaemundi, Ptolemäuskarten und Portulane – aufweist, bietet die für

ihn erarbeitete Konzepthierarchie (Abbildung 11) einen guten Ausgangspunkt.

An erster Stelle stehen einfache Fragen, die wir „Navigations-Anfragen“ nennen wollen – sie fragen üblicherweise direkt nach einzelnen Objekten oder Mengen von solchen. Dies ist der einzige Fragetyp, der in Datenbankanwendungen wie der genannten zur Herefordkarte möglich ist. Solche Fragen können in der Praxis entweder durch Ausfüllen einer Suchmaske, die einen leeren Objektrahmen darstellt, oder durch grafische Navigation auf einem Kartenbild gestellt werden. Sie lassen sich typischerweise durch „Was ist ...?“ oder „Wo ist ...?“ wiedergeben. Ein Beispiel hierfür – das wie alle folgenden auf den Behaim-Globus bezogen ist – wäre: „*Wo sind portugiesische Wappen dargestellt?*“

Die zweite Klasse umfaßt begrifflich-deskriptiv zusammengesetzte Fragen, so nach Eigenschaften von Objekten (*Welche Farben haben die Zelte in Afrika?*), nach Objekten und einfachen Objektmengen (*Wie sieht die portugiesische Fahne aus? Zu welcher Klasse gehört ...?*) und nach Objektklassen (*Welche Wappen gibt es? Wie viele Fahnen gibt es? Was sind die Tierkreiszeichen? Welche Tiere sind in Afrika dargestellt? In welcher Region ist ...?*). Weiterhin gehören hierzu komplexe Fragen nach Objekten und Objektmengen (*Wo sind Nürnberger Wappen? Welche Herrschaftszeichen sind benannt?*), nach Topologie und Orientierung (*Welche Regionen sind in ...enthalten? Welche Städte liegen östlich von ...?*) und nach Erklärungen (*Was ist die Ekliptik? Warum ist ...hier?*). Während Erklärungsfragen an sich schon besondere Anforderungen stellen, denn zu ihrer Beantwortung sind neben den in der Konzepthierarchie festgelegten Beziehungen noch weitere Hintergrundinformationen erforderlich, setzen auch schon Fragen nach Topologie und Orientierung eine anspruchsvolle formale Repräsentation räumlicher Objekte bestimmter Granularität, ihrer Beziehungen untereinander samt dazugehörigen Schlussregeln voraus. Das läßt sich leicht vor Augen führen, wenn man überlegt, welche Schlüsse zur Beantwortung einer Frage wie „*Wie heißt die größte Insel zwischen Griechenland und Afrika?*“ notwendig sind.

Von einem solchen Fragenkatalog ausgehend kann die formale Onto-

logie für den Stellenkatalog schrittweise definiert werden:²⁵

- Aufzählung der Objekttypen (zum Beispiel „geografisches Objekt“);
- Unterscheidung von Konzepten und Attributen (zum Beispiel „Insel“ und „geografischer Name“);
- Anordnung der Objekttypen in einer Konzepthierarchie und, soweit sinnvoll, auch der Attribute in einer Rollenhierarchie;
- Bestimmung der Instanzen (Individuenobjekte) für typische Beispiele (zum Beispiel „Island“);
- Bestimmung zusammengesetzter Objekte und ihrer Teile;
- Bestimmung von Eigenschaften und Beschränkungen für Attribute bezüglich Typ und Anzahl (zum Beispiel „Jahr“ ist eine natürliche Zahl zwischen 800 und 1550).

Für die Formulierung derartiger formaler Ontologien sind leistungsfähige Editoren mit grafischen Benutzungsoberflächen wie Protégé²⁶ verfügbar; im konkreten Fall wird Protégé mit einer Erweiterung für OWL eingesetzt (Abbildung 12). Dieser Editor bietet vielfältige Möglichkeiten, wozu unter anderem die Visualisierung des Konzeptgraphen als aktiver, navigierbarer Graph und Konsistenzprüfungen gehören. Letztere sind möglich durch die Kommunikation mit einer OWL-Inferenzmaschine, die als Server-Prozess gestartet werden kann; inkonsistente Eingaben können so festgestellt und in Protégé grafisch hervorgehoben werden.

Auch für die Erfassung der Einträge des Stellenkatalogs ist Protégé bestens geeignet: Es bietet ein eigenes Fenster für die Eingabe von Instanzen, deren Integrität dann gleichfalls in der genannten Weise sofort überprüft werden kann.

25 Siehe dazu für weitere Details: Noy, Ontologies (wie Anm. 22), S. 181-211 und Brachman, Ron J. u.a., Living with CLASSIC: When and How to Use a KL-ONE-like Language, in: Sowa, John (Hg.), Principles of Semantic Networks, Kap. 14, San Mateo 1991, S. 401-456.

26 <<http://protege.stanford.edu/>> (15.02.2006).

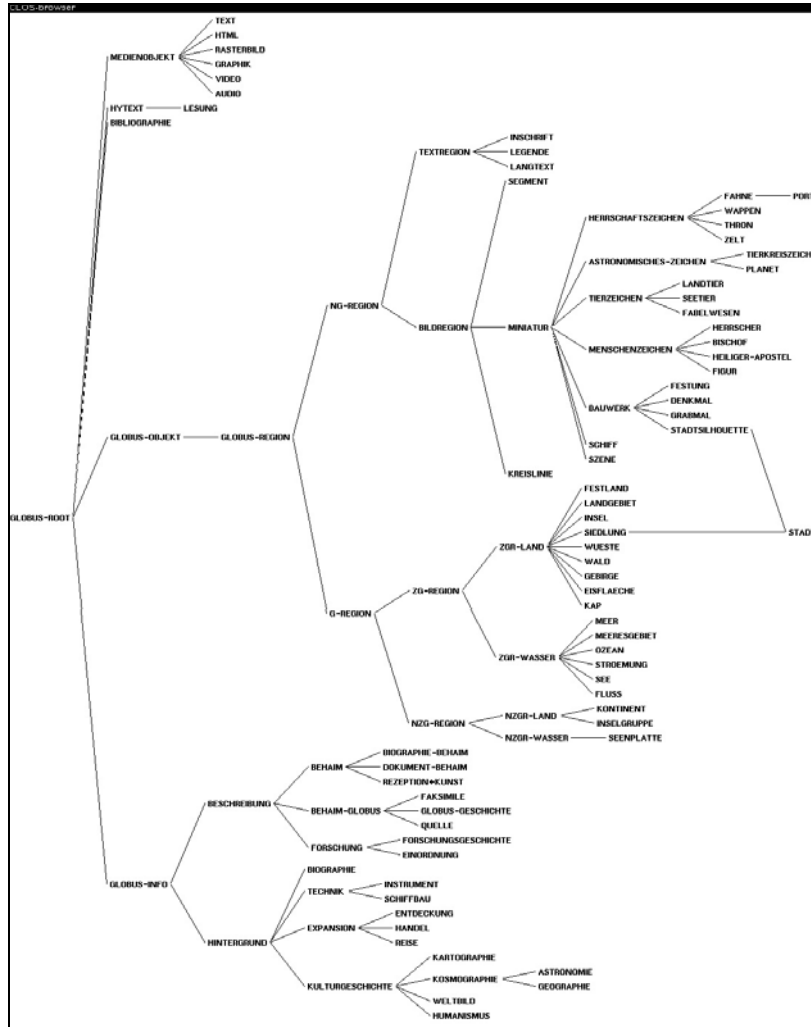


Abbildung 11: Oberer Teil der Konzepthierarchie für den Behaim-Globus

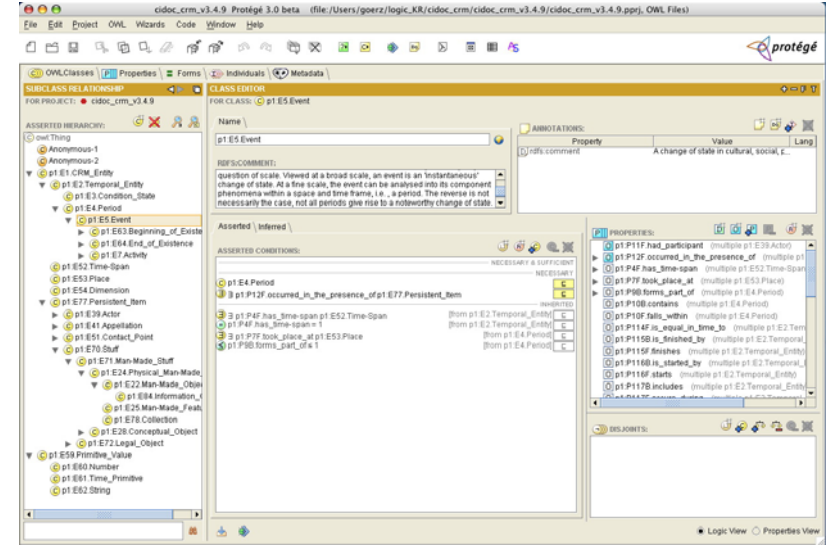


Abbildung 12: Protégé: Ansicht bei der Bearbeitung des CIDOC CRM

Einfache Anfragen

Um die Benutzung des künftigen Stellenkatalogs zu illustrieren, sei auf die im Behaim-Globus-Projekt erarbeitete Lösung verwiesen, die als Vorstufe für den neuen Entwurf dient. Die Konzepthierarchie mit Attributen und Vererbung sowie die Instanzen waren in einer objekt-orientierten Programmiersprache (CLOS in CommonLISP) implementiert worden, die allerdings über keine logische Inferenzkomponente verfügt. Daher waren nur einfache Anfragen nach Objektbeschreibungen im Stellenkatalog möglich, jedoch nicht solche, die die Auswertung logischer Verknüpfungen erfordern, die also über die einfache Navigation im Konzeptgraphen hinausgehen. Zur Formulierung der Anfragen stand eine Anfragemaske bereit, die neben den Feldern für einzutragende Attributwerte auch eine dynamische grafische Darstellung des Konzeptgraphen enthielt (Abbildung 13). Weiterhin konnte auch in den Bildern der Globussegmente navigiert werden. In den Objektbeschreibungen wird auf die bildliche Darstellung der jeweiligen Objekte – Miniaturen oder Textfelder – verwiesen; diese

sind durch einen sie umschließenden Polygonzug im Bildkoordinatensystem dargestellt. Anfrageresultate sind Instanzenbeschreibungen, die in einem jeweils neu generierten Fenster in tabellarischer Form mit Attributnamen und -werten und dem dazugehörigen Bildausschnitt angezeigt werden. Um die Suche gegenüber Schreibvarianten, zum Beispiel bei Toponymen, tolerant zu machen, wurde ein Algorithmus zur approximativen Volltextsuche integriert, der nicht nur zur Analyse der eingegebenen Attributwerte, sondern auch bei der Volltextsuche in den zahlreichen in frühem Neuhochdeutsch formulierten Inschriften eingesetzt wurde.

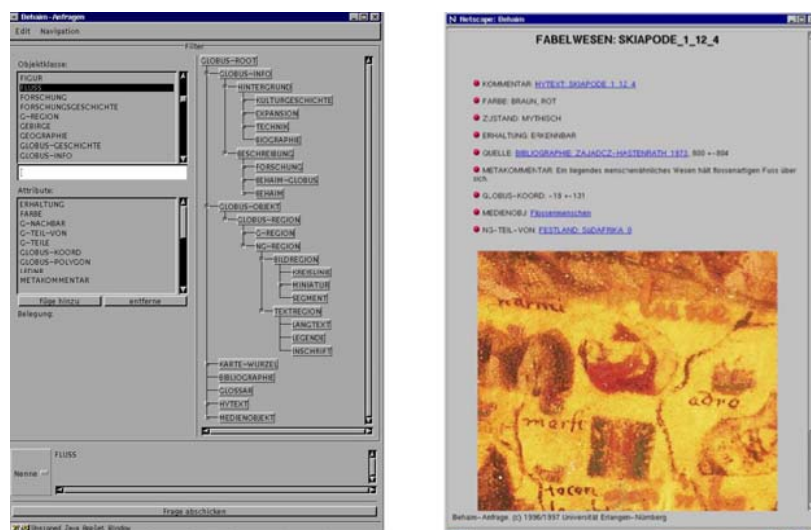


Abbildung 13: Anfragemaske des Behaim-Globus-Systems und Suchergebnis

Beschreibungslogische Wissensrepräsentation und komplexe Anfragen

Als nächster wichtiger Schritt muss die Vernetzung der Objektbeschreibungen gemäss den noch zu erarbeitenden kognitiven Relationen erfolgen. Hierzu ist die Konzepthierarchie um eine weitere Schicht mit geeigneten Konzepten und Attributen zu erweitern. Auch hierfür wird zunächst eine Liste typischer komplexer Anfragen erstellt. Beispiele für solche Fragen

wären

- Was ist das Datum der frühesten Karte, auf der X vorkommt?
- Was ist die chronologische Reihenfolge der Erwähnungen von X?
- Welche Reproduktionen des Kartentyps Y gibt es wo?
- In welchen Veröffentlichungen werden Beatus-Karten behandelt?
- Nenne Beiträge über Portulane, die bei Mittelalter-Konferenzen vorgelesen wurden!

Derartige Anfragen erfordern zu ihrer Bearbeitung nicht nur eine Suche im Konzeptgraphen unter Einbezug von Oberbegriffen, sondern auch die Auswertung logischer Kombinationen. Mit der Einbettung der Domänen-Ontologie in das CIDOC CRM wird eine neue Dimension von Verknüpfungen und damit auch Anfragemöglichkeiten eröffnet: Da im CRM Ereignisse eine besondere Rolle spielen – denn Gegenstände und ihre Beziehungen untereinander können nur durch Ereignisse verändert werden – können auf diesem Weg übergreifende Verbindungen zwischen Einträgen des Stellenkatalogs hergestellt werden. „Ereignisse (E5 Event) verändern Gegenstände (E18 Physical Stuff) in vielfältiger Weise, Personen (E39 Actor) nehmen an Ereignissen in verschiedener Funktion teil, die Ereignisse finden an Orten (E53 Place) und in bestimmten Zeiträumen (E52 Time-Span) statt“.²⁷ Um grundsätzlich zu zeigen, dass eine beschreibungslogische Wissensrepräsentationssprache hinreichend ausdrucksstark ist, um die Anforderungen an die Realisierung des Stellenkatalogs zu erfüllen, wurde bereits im Behaim-Globus-Projekt ein erfolgreicher Versuch durchgeführt. Dazu war das Konzeptmodell in eine gegenüber OWL-DL wesentlich vereinfachte beschreibungslogische Sprache (CLASSIC) übertragen und eine Teilmenge der Objektbeschreibungen aus dem Stellenkatalog in entsprechende Instanzenbeschreibungen umgesetzt worden. In diesem Rahmen konnten komplexe Anfragen gestellt werden, zu deren Beantwortung logische Schlussfolgerungen erforderlich sind, hier allerdings auf räumliche Relationen fokussiert: Topologie, Orientierung und Skalen.

Für Anfragen an in der „Web Ontology Language“ OWL formulierte

²⁷ Ermer, CIDOC (wie Anm. 23).

Wissensbasen wurde die Anfragesprache OWL-QL (OWL Query Language)²⁸ definiert. Da allerdings Eingaben in einer logischen Formelsprache einer fachwissenschaftlichen Öffentlichkeit nicht zumutbar sind, muss zur Erleichterung der Benutzung des Stellenkatalogs noch eine geeignete grafische Benutzungsoberfläche geschaffen werden.

Typische Folgerungen von Aussagen über Konzepte und Instanzen, die von einem beschreibungslogischen System ausgeführt werden können, sind aus formaler Sicht die folgenden:

1. Gegeben sei eine Konzeptbeschreibung:
 - Ist sie konsistent mit der Terminologie?
 - Ist sie spezieller oder allgemeiner als eine andere Beschreibung (Subsumtion)?
 - Ist sie äquivalent (bedeutungsgleich) zu einer anderen Beschreibung?
 - Bezeichnen zwei Klassen dieselbe Menge von Individuen?
 - Was sind die speziellsten eingeführten Beschreibungen, mit denen man sie klassifizieren kann?
 - Auf diese Weise kann eine Strukturierung des Wissensbestands durch automatische Berechnung der Taxonomie erreicht werden.
2. Gegeben sei eine Instanzenbeschreibung:
 - Ist es eine Instanz einer Klasse (Instantiierung) und wenn ja, welcher?
 - Welche sind die Instanzen einer Klasse?

Damit wird die Auswertung kontextsensitiver Beziehungen ermöglicht – also genau dasjenige Leistungsmerkmal, das von dem Stellenkatalog einschließlich der kognitiven Relationen erbracht werden soll. Unter Berücksichtigung der oben genannten Beschreibungsebenen können auch multiple Sichten beziehungsweise Perspektiven auf Objekte dargestellt und ausgewertet werden – etwa auf einen Text als physisches, logisches, rhetorisches oder linguistisches Objekt.

28 Fikes, Richard; Hayes, Pat; Horrocks, Ian, OWL-QL – A Language for Deductive Query Answering on the Semantic Web, KSL-TR-03-14, Knowledge Systems Laboratory, Stanford 2003.

Die Vernetzung der Objektbeschreibungen unter kognitiven Aspekten kann auf diese Weise dynamisch erfolgen. Sind bestimmte Themen als Konzepte in der Hierarchie verankert, können dazu passende Beziehungen festgelegt werden, wobei die Inferenzmaschine die Aufgaben der Konsistenzprüfung und automatischen Klassifikation übernimmt. Dies würde beispielsweise für das Beispiel der Alexandersage auf der Hereford-Karte bedeuten, dass nicht eine separate, von der Datenbank getrennte und fest vorgegebene Folge statischer Bilder erstellt werden muss, sondern dass eine derartige „Themenreise“ anhand der gegebenen Merkmale aus der Datenbasis assoziativ generiert werden kann.

Textbearbeitung und Verwaltung von Links

Für die in den Karten enthaltenen Texte wurde schon beim Behaim-Globus eine Einteilung in drei Klassen vorgesehen: Eigennamen (Toponyme, Personennamen), Formulierungen vom Umfang eines Satzes oder Satzteils und längere kohärente Texte. Die formale Beschreibung der Texte mußte allerdings von Hand vorgenommen werden; außer der fehlertoleranten Zeichenkettensuche waren keinerlei linguistische Analyseverfahren vorgesehen. In die in Instanzen niedergelegten Texte konnten allerdings Hypertext-Links eingetragen werden, die jeweils in einem eigenen Attribut gesammelt und bei Anfragen ausgewertet werden konnten. Inzwischen sind im Hinblick auf komplexe Metadaten neue Darstellungsverfahren entwickelt worden, die davon ausgehen, dass zumindest die längeren Texte mittels einer XML-Sprache repräsentiert und ausgezeichnet sind, wie dies etwa die „Text Encoding Initiative“ (TEI)²⁹ vorschlägt. Derart codierte Texte könnten dann mit einem Schema nach dem „Metadata Encoding and Transmission Standard“ (METS)³⁰ verbunden werden, das unter anderem hinreichende Möglichkeiten zur Darstellung von Strukturinformation und Links bietet. Für die Bearbeitung von in XML codierten Texten und den inhalts-orientierten Zugriff soll der Arboreal XML-Browser³¹ eingesetzt

29 <<http://www.tei-c.org/P4X/>> (15.02.2006).

30 <<http://www.loc.gov/standards/mets/>> (15.02.2006).

31 <<http://archimedes.fas.harvard.edu/arboreal/>> (15.02.2006).

werden, den wir bereits in einem anderen Projekt benutzen. Dieses flexible OpenSource-Werkzeug ermöglicht verschiedene Formen der Suche in solchen Texten, unterstützt den Textvergleich – was von besonderem Interesse für die verschiedenen historischen Lesungen der Inschriften auf dem Behaim-Globus ist –, besitzt Schnittstellen zu Wörterbüchern und Thesauri und linguistischen Analysekomponenten (aktuell zur Wortmorphologie) für verschiedene Sprachen, erlaubt Annotationen und ermöglicht via DIGILIB den Zugriff auf verknüpfte Bilder.

Ausblick

Abschließend seien noch einige Fernziele des Projekts genannt. Hierzu gehören solche, die noch intensiver Grundlagenforschung bedürfen, aber auch Ziele, die eher die praktische Benutzbarkeit eines derartigen Systems thematisieren und daher Entwicklungscharakter haben.

Zur ersten Gruppe gehört zum Beispiel die wichtige Frage nach der Darstellung zeitveränderlicher Konzepte. Es steht außer Zweifel, dass in dem relativ großen Untersuchungszeitraum die Deutung vieler der dargestellten Objekte und Sachverhalte einem Wandel unterliegt. Dies betrifft neben der Darstellung geografischer Gegebenheiten auch viele Darstellungen enzyklopädischen und narrativen Charakters. Es ist eine bisher kaum untersuchte Frage, wie sich dies in einer formalen Sprache zur Wissensrepräsentation abbilden läßt. Auf jeden Fall wird mit der Darstellung der Konzepte auch eine interne Zeitrepräsentation zu verbinden sein. Auf lange Sicht ist es auch unbefriedigend, Einträge des Stellenkatalogs nur in tabellarischer Form zu präsentieren. Daher sollte untersucht werden, inwieweit sich Techniken der automatischen Textgenerierung auf solche Darstellungen anwenden lassen.

Auf der Seite der Benutzbarkeit steht ohne Zweifel die Umsetzung eines an die repräsentationssprachlichen Ausdrucksmöglichkeiten angepassten Navigationskonzepts. Bevor dessen technische Realisierung angegangen werden kann, ist eine Reihe eher grundsätzlicher Probleme zu lösen. Hierzu gehört etwa die Frage, was zu welchem Zeitpunkt und wie erreichbar sein soll, also die des Bezugs zum aktuellen Bearbeitungskontext. Verschiedene Benutzer und Benutzerinnen werden mit dem Stellenkatalog

sehr verschiedene Forschungsfragen bearbeiten wollen; daher wird es notwendig sein, einen gewissen Grad an Personalisierung einzuführen, der unter anderem bestimmte Benutzerpräferenzen, etwa durch kontextsensitive und präferenzgesteuerte Anzeige, berücksichtigt. So sollte allen Benutzern und Benutzerinnen eine ihrem individuellen Interessenprofil angepasste Arbeitsumgebung zur Verfügung stehen, in der sie ihre eigenen assoziativen Verknüpfungen, Lesemarken und Arbeitsnotizen ablegen und bei wiederholten Sitzungen wieder aufrufen können. Unabhängig davon sollte die Anfrageschnittstelle auch einen Zugriff auf verschiedene Glossare und Indexlisten, zum Beispiel Toponyme, Personennamen, Themen und andere Schlüsselwörter bieten. Auch der Zugriff auf die Daten über eine Zeitachse ist ein unbestrittenes Desiderat. Weiterhin sollte den Benutzern und Benutzerinnen nicht nur ein freier Zugang auf alle bisherigen Bearbeitungsschritte mit dem System, das heißt ihre individuelle „Navigationsgeschichte“, möglich sein, sondern von jedem Bearbeitungsschritt ist zu fordern, dass er grundsätzlich umkehrbar ist, sodass jederzeit die Rückkehr in einen früheren Zustand der Bearbeitung möglich ist. Diese Elemente bilden nur den Anfang für eine übergreifende Konzeption von Assistenzfunktionen für die wissenschaftliche Arbeit mit einem solchen Stellenkatalog.

Diskussion: Forschungsfragen, digitale Präsentation und Erschließung

Bei Projekten zur Digitalisierung und formalen Erschließung historischer Quellen wie dem vorgestellten liegt es nahe, zu fragen, inwiefern die gewählte Vorgehensweise zu einer Konvergenz von zur Zeit getrennten Quellenangeboten aus Archiven, Bibliotheken und Wissenschaft führen kann. Dies ist grundsätzlich zu bejahen, auch wenn im aktuellen Fall der mittelalterlichen Weltkarten trotz zahlreicher Digitalisierungsprojekte in diesen Bereichen noch so gut wie keine geeigneten Ressourcen zur Verfügung stehen. Der Schlüssel zu einer Konvergenz liegt in der kompromisslosen Nutzung einschlägiger Standards, wie sie auch in unserem Projekt erfolgt. Die vor allem bei digitalen Bibliotheken und Archiven erfolgte Entwicklung hat zu Standardisierungsvorschlägen geführt, die wie METS

in formaler³² und das CIDOC CRM in inhaltlicher Hinsicht eine Integration heterogener Ressourcen in greifbare Nähe rücken lassen. Sind unter dieser Voraussetzung geeignete digitale Quellen erreichbar, so lassen sie sich ohne weiteres integrieren; allerdings sollte die Verfügbarkeit durch die Verwendung von persistenten URLs abgesichert sein. Auf diese kann eine Zusammenfassung der digitalen Quellen in virtueller Form, erfolgen, um sie dann jeweils an unterschiedlichen Orten unter bestimmten Gesichtspunkten zu erschließen. Um sie allerdings in den Stand einer verteilten erschlossenen „virtuellen Kollektion“ zu bringen, bedarf es noch weiterer Infrastrukturmaßnahmen, etwa in der Form von Servern, die einschlägige Linkfarmen allgemein verfügbar halten. Den jeweiligen Projekten fällt die Aufgabe zu, nicht nur ihre Ergebnisse über Links dort einzubringen, sondern auch Methoden zu entwickeln, um die auf diesem Wege erreichbaren Arbeitsergebnisse unterschiedlicher Partner mit den eigenen zu vergleichen sowie Ergänzungen und Kommentare hinzuzufügen. Dies würde langfristig zu einer neuen Ebene der wissenschaftlichen Kommunikation führen. Die Unterscheidung zwischen Archivierung, Erschließung, Edition und wissenschaftlicher Publikation wird dabei nicht verschwinden, aber die einzelnen Stufen könnten wesentlich dichter vernetzt werden, als es bislang möglich ist.

Das „mappae“-Projekt selbst ist offen angelegt und strebt keine umfassende digitale Edition mittelalterlicher Weltkarten an, aber die Modalitäten der tiefen Erschließung, auch wenn sie nur partiell ist, haben Charakteristika einer Edition. Von einer Edition könnte nur dann gesprochen werden, wenn „Vollständigkeit“ (in irgendeinem Sinn) angestrebt würde. Insofern können die Erschließungsaktivitäten des Projekts im Prinzip Bausteine für Editionen liefern; ihre besondere Leistung besteht in einer mehrdimensionalen Vernetzung sowie einer Systematisierung der Metadaten, gewissermaßen als Editions-Vorstufe.

Eine entscheidend weiter gehende Frage wäre, ob sich Forschungsfragen

32 Zur Anwendung des „Metadata Encoding and Transmission Standard“ METS siehe zum Beispiel Entwurf und Fallstudien der Oxford Digital Library: <http://www.odl.ox.ac.uk/projects/projects_mets.htm> (15.02.2006).

gen und Forschungsergebnisse durch digitale Quellenangebote ändern. Nun werden Digitalisierungstechniken als solche nicht unmittelbar zu besseren Ergebnissen führen. Zwar wird das Datenmaterial in neuer Qualität verfügbar, aber Forschungsfragen sind primär theoriegetrieben und wohl zunächst nur in geringem Maß, wenn überhaupt, abhängig vom Präsentationsmedium.

Allerdings eröffnen bildseitig die durch das digitale Medium erreichbare höhere Präzision und die Anwendung von Methoden der Bildverarbeitung neue Möglichkeiten zur Manipulation digitaler Bilder und der Visualisierung – beispielsweise durch Programme wie „COMPAGO“ und Map-Viewer –, so dass manche der bestehenden Forschungsfragen erst mit diesen Mitteln sinnvoll bearbeitet werden können. Dasselbe gilt für die nunmehr möglich gewordene Vernetzung und Aggregation der Daten, die einen quantitativen Sprung in neue Größenordnungen ermöglicht, zu deren Bearbeitung eine ganze Palette (semi-)automatischer mathematischer Auswertungsverfahren bereit steht. Über die bekannten statistischen Verfahren hinaus wurden in den letzten Jahren mehrfach auch Methoden der Bioinformatik mit großem Erfolg eingesetzt, so etwa Techniken der phylogenetischen Bäume zur Bestimmung von Stammata für Handschriftenvarianten.

Im Hinblick auf die interaktive Bearbeitung von Texten ist neben Volltext- und linguistisch informierter Suche auch an Techniken des Text-Mining zu denken, um die textkritische philologische Arbeit zu unterstützen.

Die Nutzung all solcher Verfahren kann zu neuen Hypothesen führen, die ohne sie nicht möglich gewesen wären und damit auch zu neuen Forschungsfragen anregen.

Der entscheidende Schritt unseres Ansatzes liegt aber im Übergang von der bloßen Werkzeugnutzung zu neuen Strukturierungs- und Beschreibungsverfahren. Auch die theoriegeleitete semantische Erschließung geht von vorhandenen Forschungsfragen aus, doch die Notwendigkeit einer einheitlichen normierten Darstellung erbringt eine neue Qualitätsstufe in der systematischen Auswertung, die ohne sie nicht möglich ist. Die anfängliche Idee des Semantic Web, alle Webseiten mit Metadaten zu annotieren, um damit eine systematische Suche durchführen zu können, hat sich

bekanntlich in einem derart verteilten System ohne zentrale Kontrollinstanz nicht durchsetzen können. Demgegenüber hat sich der Einsatz standardisierter formaler Beschreibungsverfahren innerhalb einzelner Projekte bewährt. Der damit verbundene zweifellos hohe Aufwand erbringt aber mit dem Einsatz der logischen Verknüpfung und des automatischen formalen Schließens einen erheblichen Mehrwert, denn dadurch werden Anfragen an den Datenbestand möglich, die mit traditionellen Erschließungstechniken überhaupt nicht oder zumindest nicht mit den üblicherweise verfügbaren Ressourcen sinnvoll bearbeitet werden konnten. Selbstverständlich werden eine systematische logische Erschließung und automatische Inferenzverfahren allein nicht die Lösung aller Forschungsfragen beschere, aber sie erbringen eine neue Qualität bei der Unterstützung der Interpretation des Datenmaterials. Deduktive Methoden ersetzen nicht die *ars inveniendi*, sie können jedoch neue Bezüge aufzeigen und besitzen damit ein hohes Potential zur Induktion innovativer Forschungsfragen.

Danksagung

An erster Stelle gebührt Hartmut Kugler ein herzlicher Dank für die produktive und stets erfreuliche Zusammenarbeit. Weiterhin ist vielen Studierenden und studentischen Hilfskräften zu danken, insbesondere Andrea Hofmann für MapViewer und Barbara Ries, ohne deren engagierten Einsatz bei der Erfassung und Bearbeitung der Bilder und der Erstellung der Metadaten die Datenbank noch ziemlich leer wäre.

Professor Dr. Günther Görz, Professor für Informatik (Künstliche Intelligenz) an der Friedrich-Alexander-Universität Erlangen-Nürnberg. Arbeitsgebiete: Maschinelle Sprachverarbeitung, insbesondere Dialogsysteme, angewandte Logik und Wissensrepräsentation, digitale Medien. Weitere Interessen: Wissenschaftsgeschichte, Logik und Sprachphilosophie mit Schwerpunkt in der Mediävistik. E-Mail: goerz@informatik.uni-erlangen.de

HISTORISCHE FILMEDITIEN IM INTERNET-ZEITALTER

von Christopher N. Carlson

Die historische Filmeditionsarbeit steht mittlerweile im Spannungsverhältnis zwischen überlieferten Arbeitsweisen und den technischen Möglichkeiten von digitalen und nicht-linearen Bearbeitungsplattformen. Einerseits ist man weiterhin stark mit dem Filmischen (im ursprünglichen Sinne des Wortes) verbunden, was durchaus als limitierender Faktor angesehen werden muss. Andererseits hat man neue Arbeitswerkzeuge an die Hand bekommen, die manches an der historischen Filmeditionsarbeit erleichtern – und anderes überhaupt erst ermöglichen.

Einleitung

Schon seit geraumer Zeit gibt es unter Historikern/innen eine kontroverse Diskussion über den Quellenwert von Filmmaterial. Digitale und nicht-lineare Speicher- und Retrievaltechniken geben dieser Diskussion neue Impulse. Auf jeden Fall wird zu Recht gefordert, dass die Verwertbarkeit von audiovisuellen (AV) Medien in der geschichtlichen Forschung und Lehre durch eine umfassende quellenkritische Analyse mit entsprechender Dokumentation abgesichert werden muss. Auch sind bei AV-Medien besondere urheber- und persönlichkeitsrechtliche Gesichtspunkte zu berücksichtigen.

Die vorgegrifflichen Aspekte der wissenschaftlichen Filmedition herkömmlicher Prägung bestehen unverändert fort. Weiterhin müssen filmische Quellen technisch und inhaltlich sorgsam ausgewählt, und es müssen Provenienz und Dignität überprüft und nachgewiesen werden. Mithilfe von Bewegtbilddatenbanken können jedoch Erschließungen und Annotationen

sehr viel besser durchgeführt werden, als dies früher möglich gewesen wäre. Neben dem qualitativen Gesichtspunkt kommen auch neue *quantitative* Möglichkeiten auf: War man früher auf einen sehr überschaubaren „wissenschaftlichen Handapparat“¹ beschränkt, um den Editionsutzer/die Editionsutzerin nicht mit papierenen Textdokumenten schier zu erschlagen, kann man in nicht-linearen Retrievalumgebungen deutlich mehr Textmaterial unterbringen, ohne die Nutzung weiter zu erschweren. Wenn das Mehr an Material nicht gebraucht wird, kann man es technisch ohne viel Aufwand ausblenden.

Des Weiteren kann man mit entsprechenden Online-Präsentationsmöglichkeiten die erzielten Ergebnisse facetten- und nuancenreicher nachvollziehen. So gibt es beispielsweise Systeme, die technisch imstande sind, nicht nur das digitalisierte Medium, sondern auch die Beschreibungselemente zu segmentieren bzw. zu sequenzieren. Damit können Metadaten wie Schlagwörter oder Klassifikationsstellen, aber auch originalsprachliche Texte oder der im Off gesprochene Kommentar passgenau der jeweiligen Sequenz zugeordnet werden. Die damit einhergehende vermehrte Trennschärfe der Beschreibung ermöglicht es ihrerseits, vergleichsweise komplexe Medienverbundpakete zu schaffen, die zumindest den bisherigen Editions begriff erheblich erweitern dürften.

Im Folgenden sollen exemplarisch die vorhandenen Möglichkeiten anhand verschiedener Beispiele aktueller Filmeditionsarbeit aufgezeigt werden. Neuere technische Plattformen unterstützen dabei die Multimedialität der Wissensrepräsentation. Sicherlich ist es noch zu früh, um eine „goldene Epoche“ der historischen Filmeditionsarbeit auszurufen. Gleichwohl dürfte schon jetzt klar sein, dass die neuen Techniken dieser Arbeit sehr zugute kommen.

Ausgangssituation und Problemstellung

Am 10. Mai 2001 tagte der Arbeitskreis „Editionsprobleme des 20. Jahr-

1 Terveen, Fritz, Vorschläge zur Archivierung und wissenschaftlichen Aufbereitung von historischen Filmdokumenten, in: Geschichte in Wissenschaft und Unterricht 6/3 (1955), S. 169-177.

hunderts“ der Arbeitsgemeinschaft historischer Forschungseinrichtungen (AHF) beim Südwestrundfunk in Stuttgart. Die Veranstaltung wurde von Archivdirektor Dr. Josef Henke, Bundesarchiv Koblenz, moderiert. In einer ersten Stellungnahme würdigte Dr. Edgar Lersch (Historisches Archiv des Südwestrundfunks, Stuttgart) den Stellenwert von Editionen audiovisueller Dokumente im Gesamtangebot vor allem der verschiedenen Text-Editionen. Lersch vertrat den Standpunkt, dass es zwar nennenswerte Mengen an audiovisuellen Materialien gibt, die für eine geschichtswissenschaftliche Nutzung in Betracht kommen, dass dies aber tendenziell eher selten geschieht. Angesichts dessen kann es nicht überraschen, dass es nur wenig Diskussion darüber gegeben hat, welche Anforderungen an die audiovisuell-medial basierte Editionsarbeit zu stellen sind. Die Probleme liegen nicht – wie häufig bei der Edition von Schriftgut – in der Transkription oder in der Übersetzungsqualität, sondern eher im Bereich der Überlieferungsgeschichte, zum Beispiel wenn man für ein und dasselbe Vervielfältigungsstück mehrere sekundäre Bezugsquellen hat. Auch gab es im Zeitalter der linearen Medien erhebliche ungelöste Probleme, audiovisuelle Dokumente mit entsprechenden Text-Annotationen zu versehen. Zugleich führte der relativ hohe Bearbeitungsaufwand angesichts der eher marginalen Rezeption solcher Editionen zwangsläufig zur Frage nach der Kosten-Nutzen-Relation.²

Man kann die entstehende Problematik von Filmeditionen exemplarisch auch anhand eines relativ jungen Falls bei einer Fotoedition beleuchten: der so genannten Wehrmachtausstellung. Zur Erinnerung: Die ursprüngliche Wehrmachtausstellung trug den Titel „Vernichtungskrieg. Verbrechen der Wehrmacht 1941-44“ und lief zwischen 1995 und 1999 als Wanderausstellung in 33 Städten in Deutschland. Aufgrund heftiger Kontroversen, bei denen es auch um die Frage ging, ob einige der in der Ausstellung verwendeten Fotos gefälscht worden seien, wurde die Ausstellung

2 Protokoll über die Tagung des Arbeitskreises „Editionsprobleme des 20. Jahrhunderts“ der Arbeitsgemeinschaft außeruniversitärer historischer Forschungseinrichtungen in der Bundesrepublik Deutschland e.V. (AHF) am 10. Mai 2001 beim SWR in Stuttgart, Quelle: <<http://ahf-muenchen.de/Arbeitskreise/ProbEdiAudioVisDok.shtml>> (16.11.2005).

zunächst geschlossen. Sie wurde später unter neuer Leitung und mit einem neuen Titel wieder eröffnet.

1999 wurde das Bundesarchiv – nicht zuletzt auch, weil es selbst Lieferant einiger der in der Ausstellung benutzten Fotos war – um eine Stellungnahme in einem österreichischen Prozess gebeten. Der Leitende Archivdirektor des Bundesarchivs Koblenz Wolf Buchmann setzte sich mit dem Thema auseinander. Er stellte einen Mangel an hilfswissenschaftlicher Quellenkritik, in der Vergangenheit begangene archivfachliche Fehleinschätzungen sowie Fehler im archivischen Umgang mit der Fotoüberlieferung fest. Bis in die jüngste Vergangenheit wurden in den (Schriftgut-) Archiven mit ihren oftmals großen Fotobeständen noch immer die Provenienzen von Fotografien missachtet und die Einzelstücke in sachthematischen Zusammenhängen oder orts- oder personenbezogenen Treffsammelungen untergebracht. Dies hat dazu geführt, dass der Entstehungszusammenhang und damit häufig auch die Überlieferungsgeschichte nicht mehr rekonstruiert werden konnten. Beide Aspekte sind jedoch sowohl für die Authentizitätssicherung als auch für ihre Interpretation und Auswertung unerlässlich. Diese Angaben müssen auch in jeder Edition aufscheinen.

Anhand der Rekonstruktion einer Fotoserie aus der Ausstellung zeigte Buchmann seinerzeit auf, dass es nur mit Hilfe des Entstehungszusammenhangs und der durch bzw. über ihn recherchierbaren Informationen möglich ist, die Echtheit der Fotos zu beweisen. Dadurch wurde es auch möglich, die ursprüngliche Funktion der Bildsequenz zu rekonstruieren.³

Anlässlich der Auseinandersetzung um die Echtheit von Exponaten der Ausstellung sowie um ihre richtige Interpretation bzw. historische Einordnung wurden 1999 auf einer Konferenz im Hamburger Institut für Sozialforschung zum Thema „Das Foto als historische Quelle“ einige Empfehlungen formuliert, die *mutatis mutandis* auch für historische Film-

3 Buchmann, Wolf, Bilder in Archiven. Empfehlungen für den Umgang mit historischen Fotografien, in: Müller, Gisela (Hg.), Ein Jahrhundert wird besichtigt. Momentaufnahmen aus Deutschland, Koblenz 2004, S. 27-41. <<http://www.barch.bund.de/imperia/md/content/abteilungen/abtb/bildarchiv2/1.pdf>> (13.04.2006).

editionen gelten können:

- Herkunft, Entstehungszusammenhang, Überlieferungsgeschichte, Urheber/innen und Rechtsinhaber/innen sind zu dokumentieren.
- Archivisches Ordnungsprinzip ist die Provenienz; von dieser Ordnung ist die Inhaltserschließung getrennt.
- Wird Bildmaterial aus konservatorischen Gründen umkopiert, so wird die früheste überlieferte Fassung (*idealiter* das Negativ) aufbewahrt.
- Restaurierungsmaßnahmen müssen sorgfältig dokumentiert werden; es muss jeweils erkennbar sein, inwieweit die restaurierte Fassung vom Original abweicht.
- Bei der Veröffentlichung von Bildmaterial ist die früheste Überlieferungsform und Bildbeschriftung zu verwenden – oder zumindest als Verweis zu nennen.

Was ändert sich im Internet-Zeitalter, was bleibt sich gleich?

Etwas, das sich stark ändert, ist der Publizitäts- bzw. Verbreitungsgrad einer historischen Filmedition. Das heißt: Fehler fallen eher auf als früher. Dies ist in zweierlei Hinsicht günstig: Zum einen steigert es die ohnehin hohe Bearbeitungssorgfalt, zum anderen kann mit hilfreichem kritischen Input aus mehr Quellen gerechnet werden. Dies kann durchaus einen kontinuierlichen Optimierungsprozess anstoßen.

Zugleich muss man sehr vorsichtig sein, welches Filmmaterial man ins Internet stellt. Stark tendenziöse Filme, zum Beispiel Hasspropaganda oder gewaltverherrlichende Filme, können entweder gar nicht oder nur unter der Verwendung von mächtigen Authentifizierungsverfahren online gestellt werden. Die Gefahr, dass ein 16mm-Film in die falschen Hände geriete, war zwar vor der Veröffentlichungen der Filmeditionen im Internet nicht an sich geringer, aber die möglichen Folgen waren deutlich weniger dramatisch. Kopien waren nur mit hohem Kostenaufwand herstellbar, für die Vorführung benötigte man Geräte, die wenig verbreitet und auch nicht von jedermann bedienbar waren.

Zumindest in Deutschland muss man stark auf das Recht der im Film gezeigten Personen am eigenen Bild achten. Auch bei älteren Filmen ist nicht ausgeschlossen, dass hier Persönlichkeitsrechte noch lebender Perso-

nen tangiert sein könnten. Erst vor kurzem meldete sich jemand mit der Behauptung, er sei der Junge mit den erhobenen Händen auf dem sehr bekannten und weit verbreiteten Foto aus dem Warschauer Ghetto.⁴ Unabhängig davon, ob diese Behauptung stimmt oder nicht, sie könnte stimmen und wenn sie stimmt, würde das deutsche Recht am eigenen Bild greifen. Selbst wenn man die vielen Fälle außer Betracht lässt, bei denen absolute oder relative Personen der Zeitgeschichte gezeigt werden, oder bei denen im Sinne des Kunsturheberrechtsgesetzes⁵ Personen, die „in der Öffentlichkeit erschienen“ sind entweder als „Beiwerk“⁶ oder als Teil einer Versammlung⁷ eingestuft werden können, verbleiben immer noch viele Möglichkeiten, wie Personen zufällig und ungewollt vor die Kameralinse gelaufen sein könnten. Diese Personen zumindest genießen weiterhin ihr Recht am eigenen Bild.

Caroline-Urteil

Mit ‚Caroline‘ wird kurz der Fall Caroline von Hannover (früher von Monaco) gegen verschiedene deutsche Boulevardzeitungen bezeichnet. Die Klägerin hatte sich gerichtlich gegen die Veröffentlichung von Aufnahmen aus ihrem Alltag gewehrt. Die deutschen Gerichte sahen bei einem Teil der umstrittenen Aufnahmen (Caroline von Hannover beim Pferdesport, im Skiurlaub usw.) die Veröffentlichung durch Paragraph 23 KunstUrhG gedeckt, der die zustimmungsfreie Veröffentlichung bei Bildnissen aus dem Bereich der Zeitgeschichte zulässt. Der Europäische Menschenrechtsgerichtshof hat jedoch auf eine Beschwerde von Caroline von Hannover hin entschieden, dass Artikel 8 der europäischen Konvention zum Schutz der Menschenrechte das Privatleben auch dann schützt, wenn die Wechselbeziehungen in den öffentlichen Raum reichen. Die Veröffentlichung von Bildern ist nur dann zulässig, wenn die Bilder zu einer öffentlichen Dis-

4 U.a. unter <<http://www.auschwitz.dk/holocaust2.htm>> (13.04.2006) zu sehen.
 5 Gesetz betreffend das Urheberrecht an Werken der bildenden Künste und der Photographie (Kunsturheberrechtsgesetz – KunstUrhG) vom 9. Januar 1907 (RGBl. S. 7, BGBl. III 440-3); Zuletzt geändert durch Gesetz vom 16. 02. 2001).
 6 Wie definiert in § 23 Abs. 1 Nr. 2 KunstUrhG.
 7 Wie definiert in § 23 Abs. 1 Nr. 3 KunstUrhG.

kussion über eine Frage allgemeinen Interesses beitragen und Personen des politischen Lebens betreffen. Die Bilder aus dem Privatleben von Caroline von Hannover genügten diesen Anforderungen nicht und durften daher nicht veröffentlicht werden.

Potentiell kann diese Rechtsprechung auch ähnliche Probleme bei zeitgeschichtlichen Filmdokumentationen verursachen. Würde Konrad Adenauer heute noch leben, so könnte er aufgrund des Caroline-Urteils die Verbreitung des berühmten Bildes unterdrücken, auf dem er beim Boccia-Spiel zu sehen ist.⁸



Abbildung 1: Konrad Adenauer beim Boccia-Spiel

Technische Fragen

Sieht man erst einmal von den Entwicklungs- bzw. Erwerbungskosten

8 Quelle: <<http://www.mediaculture-online.de/>> (13.04.2006).

eines entsprechenden Systems ab, ist die Online-Veröffentlichung durchaus billiger als andere Verlagsplattformen. Außerdem kann man davon ausgehen, dass es künftig mehr Dienstleister geben wird, die gegen Entgelt ihre Systeme für Angebote Dritter öffnen werden.

Leider – und auch sehr zur Überraschung mancher technikgläubigen Menschen – hat das digitale Zeitalter das Problem der Filmkonservierung nicht zu lösen vermocht. Hießen die einschlägigen Probleme früher „Sprödigkeit der Farbemulsionsschicht“, „Essig-Syndrom“, „Pilzbefall“ oder „Entmagnetisierung“, so haben wir es inzwischen mit einem rasanten Wandel der technischen Formate zu tun, die gerade kleine, staatlich alimentierte Einrichtungen mit erheblichen Problemen der rechtzeitigen Datenmigration konfrontieren. Noch größer als das Kostenproblem ist hier möglicherweise sogar das Problem der fehlenden Personalkapazitäten. Innerhalb eines Zeitraums, der nach archivischem Verständnis geradezu grotesk kurz war, mussten Filmarchive ihre Bestände zunächst auf Video abtasten lassen. Wer das Pech hatte, auf das falsche Format – etwa U-matic – zu setzen, durfte diesen Vorgang sogar wiederholen. Kurz darauf kam schon der Generationswandel von den analogen zu den digitalen Videoformaten. Wer hier nicht mitmacht, hat kaum Aussichten, in der Online-Welt mitzuhalten. Anschließend müssen die Bänder jedoch in den diversen Datei-Formaten – etwa MPEG1, MPEG2, AVI, WMV, Real-Video, Quicktime usw. – ausgespielt werden. Die einzige belastbare Gewissheit, die wir in Bezug auf diese Datei-Formate haben, ist, dass es sie nicht sehr lange geben wird, und zwar weil sie technisch überaltert sind und von einer neuen Technikgeneration abgelöst werden. Die digitalen Bandformate, die die Basis der Ausspielung sind, werden nur wenig länger existieren.

Beispielhafte Filmeditionen: Stadtgeschichte Hannovers im Film

Die Auseinandersetzung mit den *Hannover-Filmen* und die Bemühungen um ihre Rettung begannen bereits in den 1980er Jahren, denn das hannoversche Filmerbe war gefährdet. Bereits seit ihrer Gründung bemüht sich die Gesellschaft für Filmstudien e.V. (GFS) um die Rettung des hannoverschen Filmerbes.

Zunächst oftmals für rein aktuelle Zwecke produziert, wurden viele

Filmkopien anschließend nicht unter dem Gesichtspunkt der nachhaltigen Quellensicherung behandelt, sondern eher lieblos verstaut. Als Folge der jahrelangen unsachgemäßen Lagerung erlitten sie unterschiedliche Schäden (Schrumpfung, Austrocknung, Farbverlust). Andere waren durch zahlreiche Vorführungen verschlissen. Von manchen Filmen existierte gar nur eine einzige Kopie, bei deren Verlust ein Stück bildhafter Stadtgeschichte unwiederbringlich verloren gegangen wäre. Zudem sind die filmaufbewahrenden Institutionen oft durch Umzüge, Zusammenlegungen, Sparmaßnahmen oder gar Auflösungen in ihrer Arbeit eingeschränkt oder in ihrer Existenz bedroht gewesen.

Oftmals waren es die städtischen Behörden selbst, die zur Image- oder Tourismusförderung Filme in Auftrag gegeben haben und so zur Entstehung eines Bestands von bewegten Bildern beigetragen haben. Diese Filme, aber auch andere, die stadthistorische Themen zum Inhalt haben, vermitteln wertvolle Einblicke in die jüngste Geschichte einer deutschen Stadt.⁹ Es ist daher verdienstvoll, dass das Projekt *Hannover-Filme* unter der wissenschaftlichen Leitung von Dr. Peter Stettner sich die Sicherung, Dokumentation und Präsentation der Hannover-Filme auf die Fahne geschrieben hat. Bei der Sicherung wurden aus dem besten vorhandenen Filmmaterial neue Archiv- und Gebrauchskopien (in unterschiedlichen Formaten: Film zur Vorführung im Kino, Beta-SP zum Grabben von Klammerteilen für Fernsehproduktionen, VHS-Video für Sichtungszwecke) in einem professionellen Filmkopierwerk hergestellt. Dazu mussten vorher die einzelnen Kopien und eventuell vorhandenen sonstigen Materialien detailliert gesichtet, geprüft und verglichen werden. Kriterien waren dabei die Vollständigkeit von Bild und Ton sowie der Zustand des Materials. Sichtung und Prüfung des Materials wurden dokumentiert. Die Dokumentation erfolgte in Form von Einstellungs- und Sequenzprotokollen und

9 Stettner, Peter, Stadtportraits. Die Geschichte Hannovers im Dokumentar- und Kulturfilm, in: Lichtspielräume. Kino in Hannover 1896-1991, Hannover 1991; ders., Auf dem steinigen Weg zum Erfolg. Der Aufbau in den 50er Jahren im Spiegel der zeitgenössischen Informationsfilme der Stadt Hannover, in: von Saldern, Adelheid (Hg.), Bauen und Wohnen in Niedersachsen während der fünfziger Jahre, Hannover 1999.

in einer eigenen Datenbank.

Die Ergebnisse des Projekts *Hannover-Filme*¹⁰ werden regelmäßig der Öffentlichkeit vorgestellt. Eine erste Präsentation fand am 4. November 2004 im Kino im Künstlerhaus in Hannover statt. Weitere sind ihr gefolgt, zuletzt im vorigen November. Begleitend wurde eine Datenbank erstellt, die erstmals einen zentralen Überblick über alle als solche identifizierten *Hannover-Filme* einschließlich der bekannten Kopien ermöglicht.¹¹

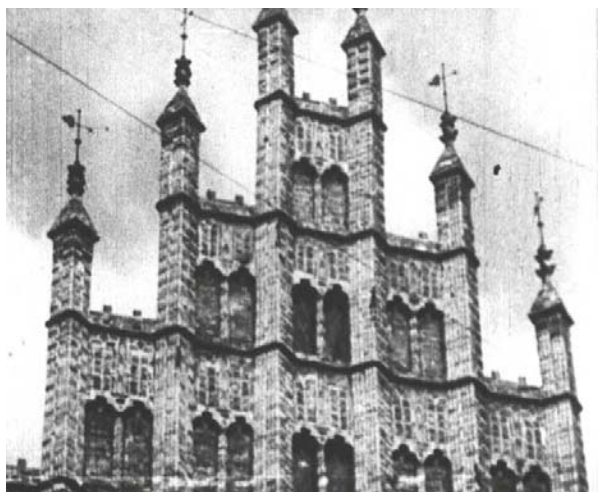


Abbildung 2: Standfoto aus dem Film „Gesicht einer Stadt, Dachgiebel des Alten Rathauses“

10 Eines dieser Ergebnisse ist der Film „Gesicht einer Stadt“ (Produzent: Döring-Film Hannover (1932); Bearbeitung (1981): Horst Latzke; Regie: August Koch; Fotografie: August Lutz; Wissenschaftliche Beratung: Dr. Bojunga, Dr. Langemann, Oberbaurat Meffert, Museumsdirektor Dr. Peßler, Vermessungsdirektor Siedentopf; Länge: 24:33 Minuten), unter: <http://www.hist.uni-hannover.de/kulturarchiv/h_film/filme/gesicht_standfotos.php> (13.04.2006), vgl. auch Abb. 2.

11 <http://www.hist.uni-hannover.de/kulturarchiv/h_film/ausgangssituation.htm> (13.04.2006).

Beispielhafte Filmeditionen: Der ewige Jude

„Der ewige Jude“¹² – ein berüchtigter Propagandafilm aus dem Dritten Reich – wurde von einer Arbeitsgruppe unter der Leitung des dänischen Historikers Stig Hornshoj-Moller aus mehreren konkurrierenden Filmfassungen rekonstruiert und kritisch neu ediert. War der Film selbst relativ rasch rekonstruiert, so dauerte die Erstellung, Verifizierung und Veröffentlichung eines detaillierten Einstellungsprotokolls und einer quellenkritischen Analyse von 1975 bis in die 1990er Jahre hinein. Hauptfilmquelle war eine – allerdings lückenhafte – 16mm-Filmkopie des Bundesarchivs, die auf ein 35mm-Duplikatnegativ der Gaufilmstelle Hannover zurückging. Ergänzt wurde der Film unter anderem mit Material aus dem Filmmuseum Kopenhagen.¹³



Abbildung 3: Standfoto aus dem Film „Der ewige Jude“, aus dem Vorspann

12 Hippler, Fritz; Taubert, Eberhard, „Der ewige Jude“. Film, 16 mm, LT, 756 m; SW, 69 ½ min; Komm. de; Orig. <Signatur: G 171>, 1940 (Erstveröffentlichung); IWF (Göttingen), 1995 (kritische Edition).

13 Hornshoj-Moller, Stig, „Der ewige Jude“. Quellenkritische Analyse eines antisemitischen Propagandafilms, Beiträge zu zeitgeschichtlichen Filmquellen, Band 2, Göttingen 1995.

Die Editionsarbeiten wurden zusätzlich durch die Existenz konkurrierender Kinofassungen erschwert. Am 28. November 1940 wurde „Der ewige Jude“ im Berliner „UFA-Palast“ am Zoo uraufgeführt. Nach einem Vorprogramm, das aus dem Kurzfilm „Ostraum – deutscher Raum“ bestand, wurden Jugendliche und Frauen um 16 Uhr zu einer für „empfindsame Gemüter“ gekürzten Fassung zugelassen, bei der die Schächtungsszenen durch eine Trickfilmsequenz ersetzt wurden. Die Vorführung für männliche Erwachsene fand in der vollen Langfassung mit kompletten Schächtausschnitten um 18.30 Uhr statt. Möglicherweise wurde auch eine andere Fassung in Österreich gezeigt.

Beispielhafte Filmeditionen: Bilder aus Ostpommern – Kreis und Stadt Bütow



Abbildung 4: Standbild: das alte Rathaus von Bütow

Begonnen hat es, wie schon oft in der Filmgeschichte, mit einem Dachbofund ... Genauer gesagt, mit einem Kohlenkellerfund. Aufgenommen

in den 1920er Jahren von dem Bütower Bürger Rudolf Steineck¹⁴, wurde der Film nach dessen Tod 1945 von Steinecks Witwe zur Sicherung an einen Freund der Familie verschenkt. Dieser versteckte den Film vor den sowjetischen Besatzern und von den späteren Machthabern der DDR in seinem Kohlenkeller – und zwar über vierzig Jahre lang. Dort wurde der Film von Herrn Gubes Erben/innen gefunden und sie schenkten ihn wiederum dem Heimatkreis Bütow, der sich der Pflege des Andenkens der ehemaligen Bütower an ihre alte Heimat verschrieben hat. Der Film war teilweise beschädigt, aber – wie durch ein Wunder – war der allgemeine Erhaltungszustand durchaus annehmbar. So war es möglich, den Film auf Digital Beta abzutasten und ihn online zu sequenzieren. Als weitere Arbeitsperspektive ist jetzt geplant, Zeitzeugen/innen für die Kommentierung der Filmszenen zu gewinnen. Der Film als solcher wird zwar voraussichtlich stumm bleiben müssen – im wesentlichen eine Kostenfrage –, aber die Kommentare könnten als Online-Beschreibung hinterlegt und somit durchsuchbar gemacht werden.

Damit wäre zugleich eine große Stärke des Internets als Veröffentlichungsmedium angesprochen: Man kann leichter neue Aggregatzustände herstellen, als es mit linearen Medien möglich wäre. Auch wenn der Film *per se* nicht leichter zu modifizieren ist, die Beigabe von Annotationen und anderen Begleitmaterialien wird leichter, weil man sowohl laufend neue solche hinzukommen lassen kann als auch vorhandene leichter redigieren kann. Müsste man bei einem gedruckten Band als kritischem Anmerkungsapparat die Kosten und Mühen einer Neuauflage auf sich nehmen, genügt es im Online-Betrieb, die entsprechenden Textdateien zu überarbeiten und sie erneut auf den Server hochzuladen.

Beispielhafte Filmeditionen: Panorama-Farb-Monatsschauen

Ein derzeit laufendes Projekt des IWF Wissen und Medien ist die Edition der vier Panorama-Farb-Monatsschauen aus der Zeit von circa November 1944 bis Januar 1945. Geplant sind folgende Arbeiten:

- Vollständige Wiedergabe der vier „Panoramen“ nach dem Ausgangs-

¹⁴ Quellennachweis für Abbildung 4: IWF Wissen und Medien gGmbH.

material im Bundesarchiv

- Einstellungsprotokolle der vier „Panoramen“ nebst historischen und formalen Erläuterungen, gefertigt von Hans-Peter Fuhrmann
- Ein Text, der das Unternehmen „Panorama“ erklärt, eventuell. der Text von Karl Stamm aus dem „Filmblatt“ 5. Jahrgang, Nummer 12, ergänzt durch eine chronologische Liste der aktenmäßigen Erwähnungen von „Panorama“
- Die erhaltenen „Farbbriefe“ (Kopien von K. Stamm)
- Die Filmographie von Jean-Paul Goergen aus dem „Filmblatt“
- Ein Beitrag zur speziellen Farbfilm-Ästhetik der „Panoramen“ (möglicher Autor: Gert Koshofer)

Anhand dieses Editionsprojektes kann man sehr gut die Möglichkeiten einer Online-Datenbank mit der Funktionalität der Sequenzerschließung beleuchten. Besonders wichtig: Jede Sequenz kann individuell beschrieben werden; weitere Informationsressourcen können sequenzgenau zugeordnet werden.

Das folgende Schaubild ist ein Screenshot aus dem Redaktionssystem einer solchen Sequenzdatenbank. Zu sehen sind einzelne Sequenzannotationen aus der Panorama-Ausgabe Nummer 1.¹⁵

Mithilfe von definierten Keyframes werden Anfang und Ende einer Sequenz festgelegt. Sechzehn unterschiedliche Felder wurden für die sequenzweise Erschließung vorgesehen – andere Felder dienen zusätzlich dazu, eine bibliografische Werkstitelaufnahme herzustellen, um das AV-Medium insgesamt zu beschreiben. Die Sequenzerschließungsfelder können unterschiedlich sequenziert werden. Das ist zum Beispiel nützlich, wenn bestimmte Sequenzen zwar unterschiedliche Sachverhalte zeigen, aber dieselbe rechtliche Konstellation – etwa, weil sie in demselben Vertragswerk geregelt wurden – aufweisen. Dann werden die Felder für Inhaltsangaben entsprechend den gezeigten Bildinhalten feiner unterteilt als das Rechte-Feld. Ein anderer Anwendungsfall tritt dann auf, wenn ein Film sowohl gesprochenen Originalton als auch Untertitel hat. Da die Untertitel

¹⁵ Quelle: IWF Wissen und Medien gGmbH.

in aller Regel nur eine verkürzte Version des gesprochenen Textes darstellen, bietet es sich an, hierfür unterschiedliche Felder vorzusehen, die dann auch unterschiedlich nach Sequenzen aufgeteilt werden können.

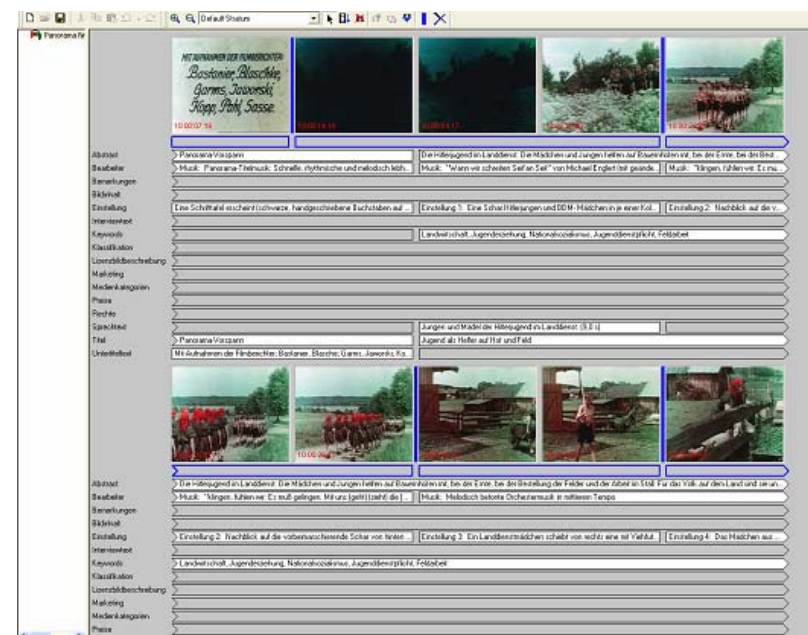


Abbildung 5: Screenshot aus dem Sequenzerschließungsmodul des *media archive-Systems*

Fazit

Auch im Internet-Zeitalter bleibt die historische Filmeditionsarbeit schwierig und zeitaufwändig – schon deshalb, weil man nach wie vor das Problem der Konservierung und der Restaurierung des meist filmischen Originalmaterials hat. Außerdem müssen Zeitgeschichtler/innen immer wieder erleben, dass audiovisuelle Medien mit deutlich mehr Argwohn betrachtet werden als zum Beispiel vergleichbare schriftliche Quellen. Dennoch bieten nicht-lineare Online-Systeme für Annotation, Suche und

Präsentation von AV-Medien bisher ungekannte Möglichkeiten für die Durchführung und Verbreitung von Editionsarbeiten. Insbesondere die damit einhergehende Möglichkeit, Bewegtbilder, Fotos, Ton- und Textmedien in einer einheitlichen Arbeitsumgebung ohne die früher unvermeidlichen Medienbrüche zusammenzuführen, stellt eine dramatische Verbesserung der entsprechenden Rahmenbedingungen für Arbeiten dieser Art dar. Trotz aller fortbestehenden Erschwernisse darf man hoffen, dass die neuen Arbeitsmöglichkeiten dazu beitragen, die historische Filmeditionsarbeit aus ihrer eher randständigen Nische herausholen.

Dr. Christopher N. Carlson ist Wissenschaftlicher Mitarbeiter für Geschichte und Informationswissenschaften bei der IWF Wissen und Medien gGmbH. E-Mail: christopher.carlson@iwf.de

ZEITHISTORISCHE ONLINE-EDITIONEN:
„DIE KABINETTSPROTOKOLLE DER BUNDESREGIERUNG“ UND
RETRODIGITALISIERUNG DER EDITION „AKTEN DER
REICHSKANZLEI – WEIMARER REPUBLIK“

von Jörg Filthaut

Der Beitrag stellt Herausforderungen, Erfahrungen und Leistungen der Internetpräsentation zeitgeschichtlicher Quelleneditionen am Beispiel der Edition „Die Kabinettsprotokolle der Bundesregierung“ und der Edition „Akten der Reichskanzlei – Weimarer Republik“ dar und erläutert insbesondere den Mehrwert der Online-Edition der Kabinettsprotokolle, die durch vielfältige Navigations- und Recherchefunktionen einen multidimensionalen Zugriff auf die präsentierten Informationen ermöglicht. Der Mehrwert an Erschließungsinformationen und ihr möglicher Ausbau wird ebenso thematisiert wie die eingesetzte Software und deren Nachnutzung im Kooperationsprojekt des Bundesarchivs und der Historischen Kommission München mit dem Ziel der Retrodigitalisierung der Edition „Akten der Reichskanzlei – Weimarer Republik“.

Die Edition „Die Kabinettsprotokolle der Bundesregierung“

Mit Beschluss vom 20. Juni 1979 hatte die Bundesregierung das Bundesarchiv beauftragt, die Sitzungsniederschriften des Bundeskabinetts und seiner Ausschüsse nach einer Mindestfrist von 30 Jahren in wissenschaftlicher Form zu edieren.¹ Dieser Auftrag bedeutete für das Bundesarchiv,

¹ Vgl. hierzu das Geleitwort von Bundeskanzler Helmut Schmidt in: Die Kabinettsprotokolle der Bundesregierung. Hg. für das Bundesarchiv von Hans Booms (bis 1989). Band 1: 1949. Bearb. von Ulrich Enders und Konrad Reiser,

unter Wahrung der bestehenden Schutzvorschriften die bis dahin geheim gehaltenen Kabinettsprotokolle nach den allgemein üblichen textkritischen und wissenschaftlichen Kriterien zu bearbeiten und nach den etablierten Standards wissenschaftlicher Quelleneditionen zu veröffentlichen. Unter diesen Vorgaben und Ansprüchen nahm das Bundesarchiv 1980 die Arbeiten an der Editionsreihe der Kabinettsprotokolle der Bundesregierung auf. Einschließlich der Protokolle der Kabinettsausschüsse (für Wirtschaft und für die Sozialreform) umfasst die inzwischen auf 19 Bände angewachsene Reihe den Zeitraum von 1949 bis 1962.² Der Jahresband 1963 wird im Herbst dieses Jahres vorliegen.

Boppard 1982, S. V.

- 2 Vgl. Die Kabinettsprotokolle der Bundesregierung. Hg. für das Bundesarchiv von Hans Booms (bis 1989), Friedrich P. Kahlenberg (bis 1999) und Hartmut Weber (ab 2000).
- Band 1: 1949. Bearb. von Ulrich Enders und Konrad Reiser, Boppard 1982.
 Band 2: 1950. Bearb. von Ulrich Enders und Konrad Reiser, Boppard 1984.
 Band 3: 1950. Wortprotokolle. Bearb. von Ulrich Enders und Konrad Reiser, Boppard 1984.
 Band 4: 1951. Bearb. von Ursula Hüllbüsch, Boppard 1988.
 Band 5: 1952. Bearb. von Kai von Jena, Boppard 1989.
 Band 6: 1953. Bearb. von Ulrich Enders und Konrad Reiser, Boppard 1989.
 Band 7: 1954. Bearb. von Ursula Hüllbüsch und Thomas Trumpp, Boppard 1993.
 Band 8: 1955. Bearb. von Michael Hollmann und Kai von Jena, München 1997.
 Band 9: 1956. Bearb. von Ursula Hüllbüsch, München 1998.
 Band 10: 1957. Bearb. von Ulrich Enders und Josef Henke, München 2000.
 Band 11: 1958. Bearb. von Ulrich Enders und Christoph Schawe, München 2002.
 Band 12: 1959. Bearb. von Josef Henke und Uta Rössel, München 2002.
 Band 13: 1960. Bearb. von Ralf Behrendt und Christoph Seemann, München 2003.
 Band 14: 1961. Bearb. von Ulrich Enders und Jörg Filthaut, München 2004.
 Band 15: 1962. Bearb. von Uta Rössel und Christoph Seemann, München 2005.
 Band 16: 1963. Bearb. von Ulrich Enders und Christoph Seemann, München 2006.
- Kabinettsausschuß für Wirtschaft. Band 1: 1951-1953. Bearb. von Ulrich Enders, München 1999.
 Kabinettsausschuß für Wirtschaft. Band 2: 1954-1955. Bearb. von Michael Hollmann, München 2000.
 Kabinettsausschuß für Wirtschaft. Band 3: 1956-1957. Bearb. von Ralf Behrendt und Uta Rössel, München 2001.
 Ministerausschuß für die Sozialreform 1955-1960. Bearb. von Bettina Martin-Weber, München 1999.

Die Edition präsentiert den Text der Protokollausfertigungen in ungekürzter, aber für das Layout aufbereiteter Darstellung. Die Sitzungsniederschriften des Kabinetts und der Kabinettsausschüsse sind in der Regel in Form von zusammenfassenden Kurzprotokollen überliefert, Wortprotokolle zu einzelnen Sitzungen kommen bisweilen vor. Jede Sitzung ist in der Buchedition so aufbereitet, dass zunächst die Teilnehmer aufgelistet werden, danach Ort und Zeit der Kabinettsitzung. Es folgt die Tagesordnung, wie sie in der Serie der Ausfertigungen überliefert ist, angereichert mit den Angaben zu den Kabinettsvorlagen. Danach werden die Tagesordnungspunkte in der Reihenfolge ihrer Behandlung präsentiert und im Anmerkungsapparat kommentiert. Die Netzeditionsgrundsätze halten dazu fest: „Neben der textkritischen Beschreibung des Textkorpus dienen die Anmerkungen zunächst dazu, den Protokolltext verständlich zu machen und an den Beratungsgegenstand heranzuführen. Nicht allgemein bekannte Ereignisse und Institutionen sowie heute ungebräuchliche Begriffe werden bei ihrer ersten Erwähnung erläutert. Im Protokolltext erwähnte Besprechungen, Schreiben, Vermerke, Gesetzestexte, Drucksachen, Zeitungsartikel usw. werden verifiziert und deren Fundstellen dokumentiert. Insbesondere ist beabsichtigt, die zu den Beratungsgegenständen entstandene und im Bundesarchiv, aber auch in anderen Archiven verwahrte archivarische Überlieferung so weit möglich nachzuweisen. Vor allem gilt das für die den Kabinettsberatungen zu Grunde liegenden Kabinettsvorlagen, die in den Beständen des jeweils federführenden Ressorts im Bundesarchiv oder im Politischen Archiv des Auswärtigen Amtes bzw. auch im Bestand Bundeskanzleramt im Bundesarchiv überliefert sind. Darüber hinaus soll dem Benutzer der Edition durch die Bezeichnung weiterer Unterlagen der zuständigen und qualifiziert beteiligten Ressorts der gezielte Einstieg in die einschlägige Sachaktenüberlieferung ermöglicht werden. Auch auf Unterlagen in den verfügbaren Nachlässen von Kabinettsmitgliedern im Bundesarchiv und in den Parteiarchiven wird gegebenenfalls hingewiesen.“

Die als Hyperlinks erkennbaren Rückverweise, Fortgänge und Querverweise lassen im Kommentar den Verlauf der Beratungen einer Angelegenheit sowohl im Kabinett als gegebenenfalls auch im Kabinettsausschuss

für Wirtschaft bzw. im Ministerausschuss für die Sozialreform nachvollziehbar werden. Gleiches gilt für die Beratungen in den Ausschüssen“³.

Jeder Band besteht des Weiteren aus einer Einleitung, welche in die Grundlinien und Hauptthemen der Politik der Bundesregierung für das betreffende Jahr einführt und die Zusammenhänge zu den entsprechenden Kabinettsitzungen herstellt; ein Abbildungsteil illustriert dies in optischer Hinsicht. Vor dem eigentlichen Protokolltext sind noch die Editionsgrundsätze und eine ausführliche Übersicht der Sitzungen und Tagesordnungspunkte vorgeschaltet. Der Anhang eines jeden Bandes führt die Geschäftsordnung der Bundesregierung auf sowie ein Quellen- und Literaturverzeichnis. Besonders interessant sind die Biogramme der regelmäßigen Sitzungsteilnehmer/innen, der sonstigen Teilnehmer/innen sowie anderer im Protokolltext genannter Personen, die insbesondere den beruflichen Werdegang in einer biografischen Skizze nachzeichnen. Eine Zeittafel ordnet die Kabinettsberatungen in die allgemeinen Ereignisse auf nationaler und internationaler Ebene des betreffenden Jahres ein. Schließlich runden ein Personen- sowie ein Sach- und Ortsindex den Jahresband ab, die das Auffinden bestimmter Inhalte erleichtern. Im Gegensatz zu den Jahresbänden der Kabinettsprotokolle enthalten die Bände der Kabinettsausschüsse zwecks Entlastung des Kommentars zusätzlich ausgewählte Schlüsseldokumente im Anhang.

Das Digitalisierungsprojekt

Die Bedeutung und Aussagekraft der Kabinettsprotokolle als historische Quelle⁴ sowie die weiter zunehmende Informationsfülle legten es nahe, mit Hilfe der modernen Informationstechnik eine Internetversion der Kabinettsprotokolle herzustellen, mit dem Ziel, diese besser und mit einem höheren Nutzungskomfort zugänglich zu machen. In diesem Vorhaben

3 Vgl. Netzeditionsgrundsätze, unter: <<http://www.bundesarchiv.de/kabinettsprotokolle/web/index.jsp>> (31.03.2006) und Editionsgrundsätze, in: Die Kabinettsprotokolle der Bundesregierung, hg. für das Bundesarchiv von Hartmut Weber, Bd. 15: 1962, bearb. von Uta Rössel und Christoph Seemann, München 2005, S. 11 f.

4 Vgl. hierzu etwa Kleßmann, Christoph, Der bleibende Wert von Quelleneditionen, in: Mitteilungen aus dem Bundesarchiv 3 (2003), S. 14-16.

wurde das Bundesarchiv von der Deutschen Forschungsgemeinschaft (DFG) finanziell unterstützt, die im Rahmen ihres Förderschwerpunkts „Retrospektive Digitalisierung von Bibliotheksbeständen“ bereits ähnliche große Projekte zur Digitalisierung forschungsrelevanter Quellen und Literatur gefördert hatte und auch weiterhin fördert.⁵

Im Rahmen des von der DFG geförderten Projektes standen dem Bundesarchiv Personalmittel für die zeitlich befristete Beschäftigung eines wissenschaftlichen Angestellten im Umfang von 24 Monaten und Sachmittel für die Systementwicklung in Höhe von 30.000 Euro zur Verfügung. Daneben hat das Bundesarchiv in erheblichem Umfang Eigenleistungen in das Vorhaben eingebracht. Das Projekt begann im Februar 2002, und bereits Ende Oktober 2003 konnten die ersten vier Bände⁶ im Rahmen eines Pressetermins im Bundeskanzleramt freigeschaltet werden.⁷ Nach der Retrodigitalisierung der nur in Buchform vorhandenen Bände 1949 bis 1956 sowie des Kabinettsausschusses für Wirtschaft 1951 bis 1953 sind seit September 2004 alle verfügbaren Bände eingestellt worden.⁸

5 Vgl. die Übersichten zu den geförderten Projekten: <<http://www.hki.uni-koeln.de/retrodig/>> (31.03.2006) und <<http://www.zvdd.de/sammlungen.html>> (31.03.2006) sowie den Evaluationsbericht zum Förderschwerpunkt der DFG: Manfred Thaller u.a., Retrospektive Digitalisierung von Bibliotheksbeständen. Evaluationsbericht über einen Förderschwerpunkt der DFG, Köln 2005, <http://www.dfg.de/forschungsfoerderung/wissenschaftliche_infrastruktur/lis/download/retro_digitalisierung_eval_050406.pdf> (31.03.2006).

6 Vgl. Die Kabinettsprotokolle der Bundesregierung, hg. für das Bundesarchiv von Friedrich P. Kahlenberg (bis 1999) und Hartmut Weber (ab 2000). Band 10: 1957. Bearb. von Ulrich Enders und Josef Henke, München 2000. Band 11: 1958. Bearb. von Ulrich Enders und Christoph Schawe, München 2002. Kabinettsausschuß für Wirtschaft. Band 3: 1956-1957. Bearb. von Ralf Behrendt und Uta Rössel, München 2001. Ministerausschuß für die Sozialreform 1955-1960. Bearb. von Bettina Martin-Weber, München 1999.

7 Zum Kontext dieser Veranstaltung sowie zu den Redebeiträgen vgl.: Die Kabinettsprotokolle online. Präsentation im Bundeskanzleramt in Berlin, in: Mitteilungen aus dem Bundesarchiv 3 (2003), S. 12-18.

8 Dieser Umstand wurde mit einer eintägigen Tagung im Bundesarchiv am 27.09.2004 zum Thema „Digitale Erschließung und Edition: Archivische Dienstleistungen im Informationszeitalter“ gewürdigt. Die Beiträge können unter <<http://www.bundesarchiv.de/aktuelles/projekte/00023/index.html>> abgerufen werden. Tagungsbericht von Jörg Filthaut, Digitale Erschließung und Edition:

Zu Beginn des Projekts stand die eingehende Analyse der Struktur der Kabinettsprotokolle und die Erarbeitung der Anforderungen an eine am Bedarf der Nutzer/innen orientierten Bereitstellung der Protokolle im Internet. Dazu wurden in- und ausländische Digitalisierungsprojekte ausgewertet und Befragungen bei potentiellen Nutzern/innen durchgeführt. In enger Zusammenarbeit mit dem IT-Dienstleister Startext Unternehmensberatung, Bonn, wurden die Anforderungen in einem Präsentationsmodell für das Internet umgesetzt.

Da es sich bei der Edition „Kabinettsprotokolle der Bundesregierung“ um eine laufende Edition handelt, lagen ab dem Jahresband 1957 Bände in der elektronischen Form vor, wie sie für die Buchausgabe benötigt wurde. Zielstellung war es daher, ein Software-Werkzeug zu erstellen, das es ermöglicht, die Druckausgangsdatei automatisch unter Einsatz herkömmlicher IT-Kenntnisse selbständig durch die Editoren in die Internetpräsentation zu konvertieren und dabei maximal eine Woche Zusatzarbeit zu erfordern.

Soweit von den vorhandenen Bänden keine maschinenlesbaren Texte vorlagen, wurden diese retrodigitalisiert. Dies geschah in Zusammenarbeit mit dem Kompetenzzentrum für elektronische Erschließungs- und Publikationsverfahren in den Geisteswissenschaften an der Universität Trier im sogenannten „double-keying“-Verfahren, bei dem zwei Erfassungsgruppen getrennt voneinander den Text manuell eingeben. Anschließend protokolliert ein Computerprogramm die Abweichungen beider Versionen, die intellektuell korrigiert werden. Diese Methode ist mit 99,9975 Prozent Genauigkeit noch wesentlich genauer als das Scanverfahren mit anschließender OCR-Erkennung. Angesichts des Mehrwerts der Volltextsuche war die höhere Genauigkeit ausschlaggebend.⁹

Archivische Dienstleistungen im Informationszeitalter“, in: H-Soz-u-Kult, <<http://hsozkult.geschichte.hu-berlin.de/tagungsberichte/id=585>> (31.03.2006).

9 Vgl. zum Kompetenzzentrum für elektronische Erschließungs- und Publikationsverfahren in den Geisteswissenschaften an der Universität Trier <<http://germazope.uni-trier.de/Projects/KoZe2>> (31.03.2006).

Vom Mehrwert der Internetpräsentation

Die Bereitstellung der Kabinettsprotokolle im Internet¹⁰ vereinigt mehrere Vorteile in sich: Die Kabinettsprotokolle stehen als zentrale Quelle für die Geschichte der Bundesrepublik Deutschland weltweit einem interessierten Publikum zur Verfügung. Die komfortable Navigation durch die Protokolle und das Angebot mehrdimensionaler Zugriffsmöglichkeiten beschleunigt und vereinfacht den Zugriff auf die Protokolle und die Recherchen erheblich. Dabei werden die Jahresgrenzen der Druckausgabe überwunden: Die Verknüpfung der einzelnen Bände untereinander eröffnet die Möglichkeit auch bandübergreifender Suchabfragen.

Der entscheidende Mehrwert der Internetversion ergibt sich aus der Mehrdimensionalität des Zugriffs. Neben der gezielten Suche und dem raschen Zugriff auch auf detaillierte Informationen ermöglicht die digitale Form der Bereitstellung der Protokolle mittels kombinierter Suchstrategien auch Abfragen komplexerer Sachverhalte sowie neue Methoden und Wege der Textauswertung. Die Volltextsuche erlaubt die Einschränkung der Suchabfrage, die wiederum mit Booleschen Operatoren versehen werden kann, nach Bandarten, Zeitraum und Suchbereichen (Einleitung, Protokolltext, Fußnoten, Tagesordnungspunkte, Teilnehmer). Gleichzeitig erleichtert die Online-Version durch den Navigationsbaum die Orientierung in einem immer umfangreicher werdenden Text. Darüber hinaus ist eine digitale Version besonders dazu geeignet, das im Editions-konzept enthaltene System von Vor- und Rückverweisen zu unterstützen. Eine Verlinkung der einzelnen Tagesordnungspunkte untereinander erlaubt es nunmehr, deren Beratungsverlauf im Kabinett und seinen Ausschüssen auch über die bisherigen Band- und Jahrgangsgrenzen hinaus im gesamten anwachsenden Corpus der Protokolltexte zu verfolgen.

Bestimmte Strukturbestandteile der Druckbände, wie zum Beispiel Editionsgrundsätze, Geleitwort, die Geschäftsordnung der Bundesregierung, werden zusammen mit Hilfefunktionen (Suchtipps, FAQ's) auf der Startseite angeboten. Die Bildleiste am oberen rechten Bildschirmrand ist bandbezogen angelegt und mit einer Vergrößerungsfunktion versehen. Die

10 Vgl. <<http://www.bundesarchiv.de/kabinettsprotokolle>> (31.03.2006).

vergrößerte Version ermöglicht nicht nur die genaue Betrachtung bzw. Lektüre, sondern gibt auch Bildunterschrift mit Signatur und Herkunftsarchiv wieder. Das Abkürzungsverzeichnis der Druckbände wurde ersetzt durch Auflösung der Abkürzungen im Text per Mouse-Over, das heißt bei Berührung mit dem Mauszeiger erfolgt die Auflösung. Da die Volltextsuche auch die Indices ersetzt, konnte auf diese verzichtet werden. Erhalten blieb aus dem Personenindex die Liste der im jeweiligen Band auf die eine oder andere Weise vorkommenden Personen, die mit der Suchfunktion verknüpft ist. Darüber hinaus wurde ein bandübergreifender navigierender Zugriff auf alle Kurzbiografien erstellt, der neben den bandbezogenen Angaben das rasche Auffinden von Teilnehmern, Gästen und sonstigen Personen in den Kabinettsitzungen gewährleistet. Damit liefert die Online-Edition einen weiteren Mehrwert durch konzentrierte Erschließung der Biogramme, zumal diese gerade mit Personen aus dem Bereich der Ministerialbürokratie zu den besonderen Vorzügen der Edition zählen. Optisch ist das Subsystem der Online-Edition über den einheitlichen Styleguide in den Gesamtauftritt des Bundesarchivs integriert.

Arbeitsabläufe

Die Editionsgruppe *Kabinettsprotokolle im Bundesarchiv* erarbeitet jeden Band in einem stark arbeitsteiligen Produktionsprozess, der die stabile jährliche Erscheinungsweise der laufenden Reihe gewährleistet. In diese Arbeitsabläufe galt es, die Internetedition möglichst effektiv und effizient zu integrieren, um Synergieeffekte zu nutzen.

Am Anfang stehen die Erfassung der Protokolltexte sowie die technische Einrichtung des Textes und seine inhaltliche Konstituierung. Nach einem ersten Korrekturgang erfolgt dann arbeitsteilig und selbständig durch jede/n Editor/in die Kommentierung der ihm zugeteilten Tagungsordnungspunkte. In der Regel ist ein/e Editor/in für mehrere Ressorts zuständig. Dadurch ist es möglich, sowohl der Struktur der Protokolltexte als auch der Aktenüberlieferung Rechnung zu tragen und einen bestimmten Sachverhalt über die Jahrgrenzen eines Bandes hinaus zu kommentieren und entsprechende Expertise in diesem Bereich zu sammeln. Die Kommentare werden wieder gebündelt, indem die beiden Hauptbearbeiter/innen

das zu edierende Jahr chronologisch und quantitativ in überschaubare Tranchen unterteilen, die jeweils Gegenstand von Manuskriptbesprechungen aller Editoren/innen ist. Nach Abschluss der Kommentierung und eines zweiten Korrekturdurchlaufs verantworten die beiden Hauptbearbeiter/innen die Einleitung, Bilder, Zeittafel, Kurzbiografien und Register. Es entsteht eine finale Datei, die für den Buchdruck in eine Postscriptdatei überführt wird. Hier setzt das entwickelte Software-Werkzeug an: Bedingungen, welche die MS-Word-Datei ohnehin für den Buchdruck erfüllen muss, wie zum Beispiel Formatvorlagen und Formalia der Zitation, macht sich das Werkzeug zunutze, um in einem ersten Schritt die Word-Datei in eine XML-Datei zu überführen, die dann in jeweils eine XML-Datei pro Tagungsordnungspunkt zergliedert wird, für den jeweils auch eine rtf-Datei generiert wird. In einem dritten Schritt werden alle übrigen Dateien (Bilder, Startseiten, pdf) in eine WAR-Datei komprimiert, wobei auch der Navigationsbaum erzeugt, die Suche indiziert und die Softlinks der Querverweise hergestellt werden. Die WAR-Datei kann dann über den Application-Server Tomcat¹¹ im Internet gestartet werden. Auch die sehr leistungsfähige Suchmaschine Lucene¹² ist ein Open-Source Produkt auf Java-Basis, das für einige spezielle Suchabfragen modifiziert wurde. Mit Lucene fiel die Entscheidung für eine Volltextsuchmaschine, die nicht datenbankgestützt operiert, sondern den Text vollständig indiziert, was schnellere Zugriffszeiten im Internet gewährleistet.

Der effektive Aufwand, einen neuen Band zu konvertieren und in die Webpräsentation einzustellen, beträgt zwei Arbeitstage, womit die Vorgabe erreicht wurde. Die Editionsreihe wird zukünftig parallel in gedruckter und elektronischer Form weitergeführt. Damit bleibt für das sequentielle Lesen der Protokolle und für deren langfristige Sicherung weiterhin die bewährte Form erhalten. Den künftig erscheinenden Bänden wird – nicht zuletzt mit Rücksichtnahme auf Verlagsinteressen – in einem Abstand von eineinhalb Jahren die Bereitstellung des jeweiligen Jahresbandes im Internet folgen.

11 Vgl. <<http://tomcat.apache.org/>> (31.03.2006).

12 Vgl. <<http://lucene.apache.org/java/docs/index.html>> (31.03.2006).

Barrierefreiheit

Der bestehende Internetauftritt der Kabinettsprotokolle ist zunächst für den Internet-Explorer 6 optimiert. Anfang 2007 wird die Kompatibilität mit älteren Versionen und anderen Browsern (Netscape, Opera, Mozilla und Derivate) hergestellt sowie insbesondere die Barrierefreiheit vollständig realisiert sein. Unter Barrierefreiheit versteht man die Zugänglichmachung von Webseiten für behinderte Mitmenschen, meistens mit Sehschwächen oder völligem Verlust des Augenlichts. Bundesbehörden sind nach dem Gesetz der Gleichstellung behinderter Menschen (BGG) vom 27. April 2002¹³ gehalten, ihre Angebote im Internet barrierefrei zu gestalten. Diesem Gesetz folgte 2003 die „Verordnung zur Schaffung barrierefreier Informationstechnik nach dem Behindertengleichstellungsgesetz“ (BITV)¹⁴, die sehr umfangreich und detailliert die Anforderungen an Barrierefreiheit definiert. Der Auftritt wird zur Zeit für Vorleseprogramme, sogenannte Screenreader optimiert. Auch wenn eine völlig deckungsgleiche Umsetzung nicht möglich ist, so wird der neue Auftritt weitestgehend die gleichen Funktionalitäten für beide Nutzergruppen bieten, sich aber dadurch optisch und technisch erheblich vom jetzigen Präsentationsmodell unterscheiden.

Das Folgeprojekt Retrodigitalisierung der Edition „Akten der Reichskanzlei. Weimarer Republik“

Die Edition „Akten der Reichskanzlei. Weimarer Republik“ wurde von 1968 bis 1990 gemeinsam von der Historischen Kommission bei der Bayerischen Akademie der Wissenschaften und dem Bundesarchiv in 23 Bänden herausgegeben.¹⁵ Die Serie setzt mit der Berufung Philipp Schei-

¹³ Vgl. Bundesgesetzblatt I, S. 1467.

¹⁴ Vgl. hierzu die Hinweise auf der einschlägigen Seite für barrierefreies Webdesign <<http://www.einfach-fuer-alle.de/artikel/bitv/>> (31.03.2006).

¹⁵ Vgl. Akten der Reichskanzlei. Hg. Von Karl Dietrich Erdmann und Wolfgang Mommsen (bis 1972) bzw. Hans Booms (ab 1972).
Kabinett Scheidemann. Bearb. von Hagen Schulze, Boppard 1971
Kabinett Bauer. Bearb. von Anton Golecki, Boppard 1980
Kabinett Müller I. Bearb. von Martin Vogt, Boppard 1971
Kabinett Fehrenbach. Bearb. von Peter Wulf, Boppard 1972

demanns zum Reichsministerpräsidenten im Februar 1919 ein und endet mit der Entlassung des Reichskanzlers Kurt von Schleicher und seines Kabinetts am 30. Januar 1933. Bei den Akten der Reichskanzlei handelt es sich um eine Auswahledition aus dem Aktenfonds der Reichskanzlei. Neben den Sitzungsprotokollen des Reichskabinetts, die den roten Faden der Edition bilden, enthält sie weitere Archivalien, insbesondere Besprechungs- und Konferenzprotokolle, ferner Vermerke und Korrespondenzen der Ressorts untereinander, mit den Länderregierungen, Parteien, Gewerkschaften usw. und gibt auf diese Weise Einblick in die zentralen politischen Vorgänge dieser Epoche. Ferner informiert sie über die Wirtschafts-, Sozial- und Verfassungsgeschichte der Weimarer Republik. Diese im Gegensatz zu den Kabinettsprotokollen der Bundesregierung abgeschlossene Edition wird seit Mai 2005 retrodigitalisiert¹⁶. Historische Kommission und Bundesarchiv warben erfolgreich Sach- und Personalmittel von der DFG ein, um in einem auf zwei Jahre terminierten Projekt diese Edition im Volltext ins Internet zu stellen. Dabei soll das für die Kabinettsprotokolle erstellte Software-Werkzeug nachgenutzt werden. Dies stellt einerseits sicher, dass verwandte Benutzer- und Präsentationsoberflächen zu einer gewissen Standardisierung im Bereich der elektronischen Quellenedition führen, andererseits soll das Software-Werkzeug in seiner Funktionalität optimiert werden. So ist daran gedacht, diejenigen Nachweise im Anmerkungsapparat, die bereits digital vorliegen, zu verlinken. Dies gilt zum

Kabinette Wirth I und II. 2 Bde. Bearb. von Ingrid Schulze-Bidlingmaier, Boppard 1973

Kabinett Cuno. Bearb. von Karl-Heinz Harbeck, Boppard 1968

Kabinette Stresemann I und II. 2 Bde. Bearb. von Karl Dietrich Erdmann und Martin Vogt, Boppard 1978

Kabinette Marx I und II. 2 Bde. Bearb. von Günter Abramowski, Boppard 1973

Kabinette Luther I und II. 2 Bde. Bearb. von Karl-Heinz Minuth, Boppard 1977

Kabinette Marx III und IV. 2 Bde. Bearb. von Günter Abramowski, Boppard 1988

Kabinett Müller II. 2 Bde. Bearb. von Martin Vogt, Boppard 1970

Kabinette Brüning I und II. 3 Bde. Bearb. von Tilman Koops, Boppard 1982, 1990

Kabinett von Papen. 2 Bde. Bearb. von Karl Heinz Minuth, Boppard 1989

Kabinett von Schleicher. Bearb. von Anton Golecki, Boppard 1986.

¹⁶ Projektverantwortlicher Bearbeiter ist Matthias Reinert, Kontakt: m.reinert@bundesarchiv.de.

Beispiel für Referenzen auf die Stenografischen Berichte des Reichstages, die von der Bayerischen Staatsbibliothek digitalisiert wurden bzw. werden.¹⁷ Zudem soll eine editionenübergreifende Suche den Mehrwert beider digitaler Quellengebote beträchtlich erhöhen.

Rückwirkungen auf die historische Forschung

Welche Auswirkungen haben diese Online-Angebote für die geschichtswissenschaftliche Forschung? Oberflächlich betrachtet zunächst keine, denn beide Editionen sind in jeder wissenschaftlichen Bibliothek für den/die Forscher/in greifbar. Und doch ändern die ubiquitäre Verfügbarkeit und komfortable Recherche die Möglichkeiten, Forschungsfragen zu beantworten. Die Herangehensweise ändert sich und damit auch die Fragen, die gestellt werden können. Leichtere Verfahren erleichtern das Stellen von Fragen, die früher mühsamer zu beantworten gewesen wären, zum Beispiel quantifizierende Fragestellungen, aber auch Fragen entlang chronologischer Längsschnitte.

Es wird der Forschung in zunehmendem Maße möglich sein, große Mengen an Quellen auszuwerten. Zum einen, weil diese digital präsentiert und recherchierbar sind, zum anderen, weil der/die Forscher/in selbst seine Exzerpte elektronisch – etwa in Form einer Datenbank – organisiert. Dadurch können Aussagen auf tragfähigere Grundlagen gestellt, quantitative und qualitative Fragestellungen einfacher verfolgt werden. Durch die Möglichkeit, größere Materialmengen handhabbar organisieren zu können, werden die Ansprüche an das mengenmäßige Quellengerüst von wissenschaftlichen Arbeiten wachsen.

Dass das Online-Angebot der Kabinettsprotokolle angenommen wird, belegen die Nutzerzahlen: Durchschnittlich 500 verschiedene Besucher/innen, 1.000 Besuche und knapp 30.000 Zugriffe pro Monat verdeutlichen die hohe Akzeptanz bei den Nutzern/innen. Damit liegen die Kabinettsprotokolle als eine Form der Erschließung der Quellenbestände

17 Vgl. „Die Reichstagsprotokolle (1919-1939)“, unter: <<http://mdz1.bib-bvb.de/cocoon/reichsblatt/start.html>> (31.03.2006) sowie auch das Vorgängerprojekt „Protokolle des Reichstages (1867-1895)“, unter: <<http://mdz.bib-bvb.de/digbib/reichstag>> (31.03.2006).

des Bundesarchivs nach den Sachthemen der Kabinettsitzungen auf der Rangliste der Webangebote des Bundesarchivs auf dem dritten Platz, da sie sich mit der Bundesrepublik auf nur eine Epoche der deutschen Geschichte beziehen. Die Spitzenreiter Beständeübersicht¹⁸ mit Online-Findbüchern¹⁹ sowie die zentrale Datenbank Nachlässe²⁰ umfassen alle Epochen der deutschen Geschichte, die sich in der Überlieferung des Bundesarchivs sowie im Falle der Nachlassdatenbank auch aller teilnehmenden Archive widerspiegeln.

Der Webauftritt der Kabinettsprotokolle nahm zweimal erfolgreich an E-Government-Wettbewerben teil²¹ und wurde mittlerweile auch positiv rezensiert.²²

Neben den beschriebenen Effekten nach außen auf die Benutzer/innen gibt es zudem eine Rückwirkung nach innen; denn die Online-Edition unterstützt auch die archivfachliche Arbeit der Archivarinnen und Archivare in der Beantwortung von Benutzeranfragen, deren sachlicher Kern im Kabinett behandelt wurde.

Perspektiven

Was die Weiterentwicklung der Online-Edition nach Abschluss der Barrierefreiheit und Herstellung der Browserkompatibilität angeht, so ist mittelfristig an die weitere Vernetzung mit internen Webangeboten des Bundesarchivs zu denken. Dies betrifft insbesondere die Verlinkung der

18 Vgl. <http://www.bundesarchiv.de/bestaende_findmittel/bestaendeuebersicht/index_frameset.html> (31.03.2006).

19 Vgl. <http://www.bundesarchiv.de/bestaende_findmittel/findmittel_online/index.html> (31.03.2006).

20 Vgl. <<http://www.nachlassdatenbank.de/>> (31.03.2006).

21 Vgl. zum E-Government-Wettbewerb <http://www.egovernment-wettbewerb.de/site/front_content.php> (31.03.2006) und <http://www.bearingpoint.de/content/news/6_3410.htm> (31.03.2006) sowie zum Wettbewerb BundOnlineStar <http://www.wms.bundonline.bund.de/nn_1304/Content/99_shareddocs/Publikationen/Oeffentlichkeitsarbeit/BundOnlineStar/g2c_kabinettsprotokolle_online,templated=raw,property=publicationFile.pdf/g2c_kabnettsprotokolle_online.pdf> (31.03.2006).

22 Vgl. die Rezension von Stephan Scheiper, WWW: Kabinettsprotokolle, in: H-Soz-u-Kult, <<http://hsozkult.geschichte.hu-berlin.de/rezensionen/id=73&type=rezwww>> (31.03.2006).

Fundstellen des Kommentars mit dem entsprechenden Eintrag des Online-Findmittels. Langfristig ist die Verlinkung mit externen Angeboten in Erwägung zu ziehen, falls diese digital verfügbar sein werden. In Frage kommen etwa das Bundesgesetzblatt, die Drucksachen von Bundestag und Bundesrat oder auch die Urteilssammlungen des Bundesverfassungsgerichts. Desgleichen könnten Links zu den Fundstellen anderer Quelleneditionen hergestellt werden, deren digitale Ausgabe zum Teil realisiert ist, so etwa die Dokumente zur Deutschlandpolitik²³ oder die Foreign Relations of the United States.²⁴ Konsequentermaßen multimedial zu Ende gedacht, könnten Belegstellen aus Reden mit Audiodateien oder Filmdateien verknüpft werden; die Anbindung von Bilddateien wäre analog zu sehen.²⁵

Hinsichtlich internationaler Standards für die Beschreibung und Strukturierung elektronischer Texte ist eine Anlehnung an das Modell der Text Encoding Initiative (TEI) erstrebenswert, wobei im Falle der Quellengattung Protokoll eine spezifische Weiterentwicklung vorstellbar ist, wie sie etwa für den Bereich der Edition der Quellengattung Urkunde schon vorangetrieben wird.²⁶

Schließlich sollte auf einer Metaebene überlegt werden, wie die im Internet bereits verfügbaren digitalen Angebote besser nachgewiesen werden, da sie meistens nur von Spezialisten gefunden werden. Es wäre wünschenswert, wenn Findmittel, wie zum Beispiel der Karlsruher Virtuelle Katalog (KVK)²⁷, beim Suchergebnis als erste Treffer die Verweise zu den digitalen Angeboten liefern, dann zu den Nachweisen der Buchform. Hierzu ist es erforderlich, sich auf Beschreibungsstandards zu verständi-

23 Vgl. die Bahr-Kohl-Gespräche der Dokumente zur Deutschlandpolitik in: <http://www.bundesarchiv.de/bestaende_findmittel/editionen/dzd/01077/index.html> (31.03.2006).

24 Vgl. Foreign Relations of the United States online, unter: <http://www.state.gov/www/about_state/history/frusonline.html> (31.03.2006).

25 Der Ansatz der Kabinettsprotokolle Nordrhein-Westfalen verfolgt diesen Weg, vgl. dazu Andreas Pilger, Stand und Perspektiven einer digitalen Edition der Kabinettsprotokolle der Landesregierung von Nordrhein-Westfalen (Legislaturperiode 1966-1970), unter: <<http://www.bundesarchiv.de/imperia/md/content/abteilungen/abt/48.pdf>> (31.03.2006).

26 Vgl. dazu die Charters Encoding Initiative (CEI), unter: <<http://www.cei.lmu.de/>> (31.03.2006).

27 Vgl. <<http://www.ubka.uni-karlsruhe.de/kvk.html>> (31.03.2006).

gen, damit solche seriösen Nachweise funktionieren. Eine kommerzielle Suchmaschine kann hinsichtlich Authentizität und Vertrauenswürdigkeit den Ansprüchen der wissenschaftlichen Welt nicht genügen.

Abschließende Betrachtung

Editionen haben für die historische Forschung Rückgratfunktion. Im Unterschied zu Darstellungen veralteter Quelleneditionen kaum, bieten vielmehr die Grundlage, neue Fragestellungen der Forschung anzuregen und zu überprüfen. Für die Zeitgeschichte bilden die „Akten der Reichskanzlei – Weimarer Republik“ und die „Kabinettsprotokolle der Bundesregierung“ solche Rückgratserien. Ihre Digitalisierung bietet in der Kombination von vielfältigen Navigations- und Recherchefunktionen in bandübergreifendem Rahmen einen erheblichen Mehrwert neben der zeitlich und räumlich unbegrenzten Verfügbarkeit.

Die Kabinettsprotokolle der Bundesregierung sind meines Wissens auch die erste umfassende Edition der Sitzungsniederschriften des höchsten exekutiven Gremiums eines kurrenten politisch-gesellschaftlichen Systems.²⁸ Damit wird auch die retrospektive Inspektion der Entscheidungsabläufe dieses Gremiums ermöglicht, „[...] den Bürgern bisher verschlossene Einblicke in die Arbeit der Bundesregierung [...]“ eröffnet und „[...] das Verständnis für die Arbeitsweise einer Regierung in der parlamentarischen Demokratie [...]“²⁹ vertieft. Die Online-Edition erhöht dieses Gebot der Transparenz noch einmal in besonderem Maße und ist gleichfalls im Bereich der Zeitgeschichte international Vorreiter.

Schließlich sind die „Kabinettsprotokolle der Bundesregierung“ ein

28 Vgl. auch den Überblick bei Karl-Ulrich Gelberg, Die Protokolle des Bayerischen Ministerrates 1945-1954. Ein Editionsprojekt der Historischen Kommission bei der Bayerischen Akademie der Wissenschaften und der Generaldirektion der Staatlichen Archive Bayerns, in: Arbeitsgemeinschaft außeruniversitärer historischer Forschungseinrichtungen in der Bundesrepublik Deutschland (Hg.), Jahrbuch der historischen Forschung in der Bundesrepublik Deutschland. Berichtsjahr 2003, München 2004, S. 75 f.

29 Geleitwort von Bundeskanzler Helmut Schmidt in: Die Kabinettsprotokolle der Bundesregierung, hg. für das Bundesarchiv von Hans Booms (bis 1989), Band 1: 1949, bearb. von Ulrich Enders und Konrad Reiser, Boppard 1982, S. V.

Beitrag des Bundesarchivs zur E-Government-Initiative der Bundesregierung „BundOnline 2005“, um eine moderne und dienstleistungsorientierte Verwaltung zu schaffen.³⁰ Zudem ist die Online-Edition auch Bestandteil der Strategie des Bundesarchivs, als Dienstleister für die Forschung, die Verwaltung und den Bürger den Zugang zu seinen Beständen offen, effektiv und komfortabel zu gestalten.

Jörg Filthaut arbeitet als Wissenschaftlicher Archivar am Bundesarchiv, Dienort Koblenz, in der Abteilung „Bundesrepublik Deutschland“ und ist neben Fragen der Schriftgutverwaltung und Bewertung vor allem für die archivische Bearbeitung des Bestandes „Bundeskanzleramt“ zuständig. Zugleich ist er Projektverantwortlicher der Internetpräsentation der Edition „Die Kabinettsprotokolle der Bundesregierung“. E-Mail: j.filthaut@bundesarchiv.de

30 Vgl. <<http://www.wms.bundonline.bund.de/>> (31.03.2006).

von Thomas Burch und Andrea Rapp

Nachschlagewerke aller Art verzeichnen, ordnen und bewahren das ständig wachsende Weltwissen. Sie ermöglichen die Orientierung in verschiedenen Fachgebieten und können so die Wissensbildung bei einer in ihrer Ganzheit oft nicht mehr überschaubaren Informationsfülle entscheidend befördern. Digitale Nachschlagewerke sind auf vielfältige Art aufeinander bezogen und damit in gewisser Weise implizit „vernetzt“. Nachschlagewerke, die durch inhaltlich-strukturelles Markup in standardisierte und damit vergleichbar gemachte Informationseinheiten gegliedert und durch Metadaten angereichert sind, machen die impliziten Vernetzungen explizit. Dadurch kann eine neue Qualität der Informationsgewinnung erreicht werden und die Lücke zwischen der schwerfälligen Benutzbarkeit und eingeschränkten Verfügbarkeit der Buchversionen einerseits und der fehlenden Systematik und Beliebigkeit der Information im Internet andererseits geschlossen werden. Der Aufsatz beschreibt Methoden und Verfahren zum Aufbau eines intelligenten Wörterbuchnetzes für deutschsprachige Wörterbücher. Von einem „Meta-Wörterbuch“ der deutschen Sprache, dem ambitionierten Fernziel dieses Vorhabens, könnten alle an Sprachinformationen interessierten Disziplinen profitieren.

Ausgangslage: Kompetenzzentrum für elektronische Erschließungs- und Publikationsverfahren

Das seit 1998 an der Universität Trier bestehende „Kompetenzzentrum für elektronische Erschließungs- und Publikationsverfahren in den Geisteswis-

senschaften“ hat mehrere Arbeitsschwerpunkte. Dazu gehören nicht nur eigene (Retro-)Digitalisierungsprojekte, sondern auch Serviceleistungen für Projekte und Verlage sowie Programmierarbeiten für interdisziplinäre Projektverbünde.¹ Im Folgenden soll einer der zentralen Arbeitsschwerpunkte des Zentrums ausführlicher vorgestellt werden, nämlich die Retrodigitalisierung und weitere Aufbereitung bzw. Erschließung und Vernetzung von Wörterbüchern.

Wörterbücher als vernetzte Strukturen

Bereits in den ersten Vorschlägen der „Facharbeitsgruppe Inhalt“ zur Vorbereitung des Programms „Retrospektive Digitalisierung von Bibliotheksbeständen“ der Deutschen Forschungsgemeinschaft (DFG) haben Wörterbücher eine dominierende Rolle gespielt, nicht nur weil sie besonders intensiv benutzt werden, sondern auch weil elektronische Nachschlagewerke Auswertungsmöglichkeiten bieten, die in ihrer gedruckten Form kaum jemals ausgeschöpft werden können.²

Nachschlagewerke aller Art verzeichnen, ordnen und bewahren das ständig wachsende Weltwissen. Sie ermöglichen die Orientierung in verschiedenen Fachgebieten und können so die Wissensbildung bei einer in ihrer Ganzheit oft nicht mehr überschaubaren Informationsfülle entscheidend befördern. Schwierigkeiten bereitet jedoch die zunehmende Diversifizierung der Disziplinen und die daraus resultierende Isoliertheit der Einzelinformationen.

1 Die Tätigkeitsfelder reichen von der Organisation der Erfassung bereits vorhandener Grundlagenwerke, von der Anreicherung der Daten mit metasprachlichen Informationen über die Beratung bei der Konzeption von elektronischen Publikationen neuer Werke bis zur Entwicklung und Implementierung netzbasierter Arbeitsumgebungen in dezentralen Arbeitsstellen. Die Zusammenarbeit von Informatik, Informationswissenschaft, philologischer Datenverarbeitung sowie verschiedenen Fachwissenschaften und die daraus resultierende Methoden- und Perspektivenvielfalt kann als Markenzeichen des Trierer Kompetenzzentrums bezeichnet werden: <<http://www.kompetenzzentrum.uni-trier.de>>. Alle im Folgenden angeführten Websites zuletzt eingesehen am 21.2.2006.

2 <http://www.sub.uni-goettingen.de/index_menu.html?http://www.sub.uni-goettingen.de/ebene2/vdf/empfehl.htm>.

Zugleich entspricht ihre Anlage als alphabetisch oder systematisch aufbereitete Informationsquelle einer verbreiteten, von ihnen selbst begründeten Kulturtechnik der Informationssuche, ganz unabhängig von den Inhalten, die durch ihre jeweilige Struktur erschlossen werden. Das macht sie zu zentralen Hilfsmitteln von Forschung und Lehre, denn die von Internet-Suchmaschinen gesammelten Daten können weder die Verlässlichkeit noch den Grad an Systematisierung und Tiefenerschließung der in Nachschlagewerken aufbereiteten Informationen erreichen und bleiben derzeit weit hinter deren Präzision und Umfang zurück.

Digitale Nachschlagewerke sind, wie auch ihre gedruckten Entsprechungen, auf vielfältige Art aufeinander bezogen und damit in gewisser Weise implizit „vernetzt“; eine übergreifende, integrierte Recherche ist jedoch wegen der Unterschiede in der Anlage, Anordnung und Struktur der einzelnen Werke nicht ohne Weiteres möglich.³ Nachschlagewerke, die durch inhaltlich-strukturelles Markup in standardisierte und damit vergleichbar gemachte Informationseinheiten gegliedert und durch Metadaten angereichert sind, machen die impliziten Vernetzungen explizit. Dadurch kann eine neue Qualität der Informationsgewinnung erreicht werden und die Lücke zwischen der schwerfälligen Benutzbarkeit und eingeschränkter Verfügbarkeit der Buchversionen einerseits und der fehlenden Systematik und Beliebigkeit der Information im Internet andererseits geschlossen werden.

Konzepte, die solche Strukturen abbilden, sollten auf der Grundlage einer repräsentativen Auswahl aus den verschiedenen Kategorien von Nachschlagewerken bzw. Wörterbuchtypen entwickelt werden, da diese Typen unterschiedliche und je spezifische wissensorganisierende Strukturen aufweisen. Der Trierer Arbeitsgruppe stehen unterschiedliche Typen tief annotierter Wörterbücher in digitaler Form zur Verfügung, die bereits in unterschiedlicher Dichte untereinander vernetzt, durch Symptomwertangaben und Metadaten intellektuell erschlossen und recherchierbar oder

3 Diese unterschiedlichen Strukturen erklären sich zum einen wissenschaftshistorisch aus der Entwicklung der lexikografisch-lexikologischen Methoden, zum anderen aus den spezifischen Zielsetzungen der verschiedenen Wörterbuchtypen.

auch mit Primärquellen vernetzt sind:

- Typ Sprachstadienwörterbuch: Mittelhochdeutsches Wörterbuch, Mittelhochdeutsches Handwörterbuch (einschlägige Nachträge), Findebuch zum mittelhochdeutschen Wortschatz (als Verbund mit multidirektionalen Verweisen)⁴, Neues Mittelhochdeutsches Wörterbuch einschließlich digitalem Quellen- und Belegarchiv⁵
- Typ sprachstadienübergreifendes Wörterbuch: Deutsches Wörterbuch von Jacob und Wilhelm Grimm⁶ (DWB)
- Typ Dialektwörterbuch: Pfälzisches Wörterbuch, Rheinisches Wörterbuch, Wörterbuch der elsässischen Mundarten, Wörterbuch der

4 Lexner, Matthias, Mittelhochdeutsches Handwörterbuch. 3 Bde. Nachdruck der Ausgabe Leipzig 1872-1878. Mit einer Einleitung von Kurt Gärtner, Stuttgart 1992; Mittelhochdeutsches Wörterbuch. Mit Benutzung des Nachlasses von Georg Friedrich Benecke ausgearbeitet von Wilhelm Müller und Friedrich Zarncke. 4 Bde. Nachdruck der Ausgabe Leipzig 1854-1866. Mit einem Vorwort und einem zusammengefaßten Quellenverzeichnis von Eberhard Nellmann, Stuttgart 1990 (=BMZ); Gärtner, Kurt u.a., Datenverarbeitung Hanrieder, Gerhard, Findebuch zum mittelhochdeutschen Wortschatz. Mit einem rückläufigen Index, Stuttgart 1992; Burch, Thomas u.a. (Hgg.), Mittelhochdeutsche Wörterbücher im Verbund. CD-ROM und Begleitbuch, Stuttgart 2002. CD-ROM-Version auf dem CD-ROM-Server der UB Trier; Internetversion unter <<http://www.mwv.uni-trier.de>>.

5 Gärtner, Kurt u.a. (Hgg.), Mittelhochdeutsches Wörterbuch. Erster Band, Doppellieferung 1/2, Lieferung 1: a-amurschaft bearbeitet in der Arbeitsstelle der Akademie der Wissenschaften und der Literatur Mainz an der Universität Trier von Ralf Plate und Jingning Tao, Lieferung 2: an-balsieren bearbeitet von der Arbeitsstelle der Akademie der Wissenschaften zu Göttingen von Susanne Baumgarte, Gerhard Diehl und Bernhard Schnell. Mit einer CD-ROM, Stuttgart 2006 (vgl. <<http://www.mhdwb.uni-trier.de>>); Digitales Mittelhochdeutsches Textarchiv: <<http://www.mhgt.uni-trier.de>>.

6 Deutsches Wörterbuch von Jacob und Wilhelm Grimm. Nachdr. der Erstbearbeitung München 1984; Der Digitale Grimm. Deutsches Wörterbuch von Jacob und Wilhelm Grimm. Elektronische Ausgabe der Erstbearbeitung, bearbeitet von Hans-Werner Bartz, Thomas Burch, Ruth Christmann, Kurt Gärtner, Vera Hildenbrandt, Thomas Schares, Klaudia Wegge. Herausgegeben vom Kompetenzzentrum für elektronische Erschließungs- und Publikationsverfahren in den Geisteswissenschaften an der Universität Trier in Verbindung mit der Berlin-Brandenburgischen Akademie der Wissenschaften. 2 CD-ROMs, Benutzerhandbuch, Begleitbuch. 1. Aufl., Frankfurt a.M. 2004; vgl. <<http://www.dwb.uni-trier.de>>.

- deutsch-lothringischen Mundarten, Luxemburgische Wörterbücher⁷
- Typ Autorenwörterbuch: Goethe-Wörterbuch⁸ (GWB)
- Typ Sachwörterbuch: Ökonomische Enzyklopädie von J.G. Krünitz (Projekt der UB Trier)⁹

Ziel der Überlegungen ist, anhand der verschiedenen Wörterbücher zum ‚Deutschen‘ einen standardisierten Einstieg zu schaffen, von dem aus sich die nicht standardisierten, heterogen organisierten und verstreuten Informationen verzweigen – heterogen im Hinblick auf synchrone wie diachrone Varianz. Dieses Konzept lässt sich nicht nur auf Wörterbücher anwenden, sondern auch auf die Erschließung großer Korpora übertragen. Das skizzierte Ziel soll an einigen Beispielen verdeutlicht werden.

7 Fournier, Johannes, Vorüberlegungen zum Aufbau eines Verbundes von Dialektwörterbüchern, in: Zeitschrift für Dialektologie und Linguistik 70 (2003), S. 155-176; Wörterbuch der elsässischen Mundarten. Bearbeitet von Ernst Martin und Hans Lienhart. 2 Bände, Strassburg 1899-1907. [Nachdruck Berlin/New York 1974]; Wörterbuch der deutsch-lothringischen Mundarten. Bearbeitet von Ferdinand Follmann, Leipzig 1909. [Nachdruck Hildesheim/New York 1971]; Rheinisches Wörterbuch. Im Auftrag der Preußischen Akademie der Wissenschaften, der Gesellschaft für Rheinische Geschichtskunde und des Provinzialverbandes der Rheinprovinz auf Grund der von Johannes Franck begonnenen, von allen Kreisen des Rheinischen Volkes unterstützten Sammlung bearbeitet und herausgegeben von Josef Müller, Heinrich Dittmaier, Rudolf Schützeichel und Mattias Zender. 9 Bände, Bonn/Berlin 1928-1971; Pfälzisches Wörterbuch. Begründet von Ernst Christmann, fortgeführt von Julius Krämer, bearbeitet von Rudolf Post unter Mitarbeit von Josef Schwing und Sigrid Bingenheimer. 6 Bände, Wiesbaden/Stuttgart 1965-1997; Gangler, J. F., Lexicon der Luxemburger Umgangssprache, Luxemburg 1847. [Unveränderter Nachdruck, Vaduz 2002]; Wörterbuch der luxemburgischen Mundart, Luxemburg 1906; Luxemburger Wörterbuch. Bd. 1-4 und Nachtragsband, Luxemburg 1950 – 1977; <<http://www.dvw.uni-trier.de>>.

8 Goethe-Wörterbuch. Hg. v. der Berlin-Brandenburgischen Akademie der Wissenschaften [bis Bd. 1, 6. Lfg.: Deutsche Akademie der Wissenschaften zu Berlin; bis Bd. 3, 4. Lfg.: Akademie der Wissenschaften der DDR], der Akademie der Wissenschaften in Göttingen und der Heidelberger Akademie der Wissenschaften. Bd. 1 (A-azurn), 1978; Bd. 2 (B-einweisen), 1989; Bd. 3 (einwenden-Gesäusel), 1989; Bd. 4, Lieferung IV/1-10 (Geschäft-hinzutreten); <<http://www.gwb.uni-trier.de>>.

9 Oeconomische Encyclopädie oder allgemeines System der Land-, Haus- und Staats-Wirtschaft : in alphabetischer Ordnung. Bd. 1 - 242. Berlin : Pauli, 1773-1858; <<http://www.kruenitz.uni-trier.de>>.

Konkrete Vernetzungen

Im Gegensatz zu einer bloßen, unverbundenen Bereitstellung verschiedener Nachschlagewerke bieten Wörterbuchverbände mit multidirektionalen Verlinkungen komplexe und gezielt spezifizierbare Zugänge zum Material. Diese Verlinkungen gehen weit über eine rein ausdrucksseitige Verknüpfung hinaus, indem sie philologische und informationswissenschaftliche Methoden verbinden. Die Entwicklung von Verfahren, Methoden und Technologien, die eine intelligente, dynamische Vernetzung von Nachschlagewerken unterschiedlichen Typs einschließlich ihrer Quellenverzeichnisse und Primärquellen ermöglichen und dadurch eine neue Qualität des Informations- und Wissensmanagements erreichen, erfordert neben informationstechnologischen Konzepten auch eine stärkere Einbeziehung philologischer und lexikografischer Verfahrensweisen.

Neben der Digitalisierung der Basiskomponenten wird ein Mehrwert durch zusätzliche Inhalte in Form von Meta-Daten geschaffen: Einfachstes Beispiel ist die Abbildung explizit genannter Verweise in andere Wörterbücher auf zunächst monodirektionale Hyperlinks, die jedoch mit Hilfe von automatischen Verfahren auch multidirektional erzeugt werden können. Eine weitere Möglichkeit ist beispielsweise die Auswertung von Sprachangaben im Innern des Artikels, zum Beispiel im Pfälzischen Wörterbuch die Abbildung der Angaben ‚*mhd. Lemma*‘ auf die mittelhochdeutschen Wörterbücher.

Die Stichwörter verschiedener Nachschlagewerke lassen sich beispielsweise über eine ‚Hyper-Lemmaliste‘ zusammenführen, um einen integrierten Zugang zu den unterschiedlichen Lemma-Ansätzen (zum Beispiel dialektal versus hochsprachlich, sprachstufenbezogen versus neuhochdeutsch) zu schaffen. Die Kategorisierungen von Quellen nach so genannten Symptomwerten bieten weitere Möglichkeiten (Raum, Zeit, Textsorte, Stil etc.). Auch die inhaltliche Klassifizierung von Nachschlagewerken und Enzyklopädien (zum Beispiel durch die DDC oder ähnliches)¹⁰ ist ein solcher Mehrwert, der neue Erschließungs- und damit Suchmöglichkeiten schafft.

¹⁰ DDC = Dewey Decimal Classification <<http://www.oclc.org/dewey/>>.

Im Falle der ‚Mittelhochdeutschen Wörterbücher im Verbund‘ beispielsweise wurde die multidirektionale Vernetzung anhand einer standardisierten ‚Hyper-Lemmaliste‘ aller Stichwörter aus den vier eingebundenen Werken realisiert. Hierzu mussten etwa die Stichwortansätze eines der Wörterbücher, das für Verben jeweils die erste Person Singular Indikativ Präsens ansetzt, mit den infinitivischen Lemmaansätzen der übrigen mittelhochdeutschen Wörterbücher verbunden werden.¹¹ Eine rein ausdrucksseitige, alphabetische Kumulierung der jeweiligen Lemmata hätte im Falle der Mittelhochdeutschen Wörterbücher also nicht zu einem brauchbaren Ergebnis geführt. Ferner galt es, Homographen voneinander zu unterscheiden, um auf das semantisch entsprechende Lemma verweisen zu können. Dies wurde anhand halbautomatischer Verfahren realisiert, indem Konkordanzen mit automatisch erzeugten Zuordnungsvorschlägen von philologisch geschulten Bearbeitern/innen geprüft wurden. Bei den automatischen Zuordnungen können zum Beispiel auch grammatische Angaben aus den Wörterbuchartikeln mit herangezogen werden, was die Trefferquote erhöht.¹²

Die Quellenverzeichnisse der Mittelhochdeutschen Wörterbücher wurden darüber hinaus mit Metadaten angereichert, die die Primärquellen kategorisieren, so dass eine spezifizierte Abfrage nach den Kategorien Raum, Zeit und Textsorte möglich ist. Zusammen mit der differenzierten inhaltlich-strukturellen Annotierung der Artikel können im Mittelhoch-

¹¹ Die Stichwortansetzung bei Verben nach der Form der 1. Person Singular Präsens Indikativ ist auch für lateinische Wörterbücher üblich. Die strikt alphabetische Anordnung aller Stichwörter gilt im ‚Mittelhochdeutschen Wörterbuch‘ (BMZ) nur für diejenigen Wörter, die die Wortstämme repräsentieren, das sind in der Regel Simplizia. Unter dem Stichwort für den Wortstamm sind dann alle Glieder der Wortfamilie systematisch (nicht alphabetisch) aufgeführt. Die über eine Hyper-Lemmaliste erfolgte Zuordnung aller Lemmata zu der strikt alphabetischen Anordnung der anderen Komponenten des Verbundes (Lexer, Findebuch) war eine erhebliche Herausforderung, die im Projekt zu bewältigen war.

¹² Thomas Burch, Johannes Fournier, Lexikografische Information per Mausclick: Die wichtigsten Wörterbücher zum Mittelhochdeutschen auf einer CD-ROM, in: ZfdA 130 (2001), S. 306-318; dies., Zur Anwendung der TEI-Richtlinien auf die Retrodigitalisierung Mittelhochdeutscher Wörterbücher, in: Lemberg, Ingrid u.a. (Hgg.), Probleme und Perspektiven computergestützter Lexikografie, Tübingen 2001, S. 133-153.

deutschen Wörterbuchverbund komplexe Recherchen wie beispielsweise die nach allen Ableitungen mit dem Suffix *-unge* in religiösen Texten des 13. Jahrhunderts aus dem Bairischen Raum durchgeführt werden.

Mit der digitalen Erstbearbeitung des „Deutschen Wörterbuchs von Jacob und Wilhelm Grimm“ (DWB) steht ein weiteres, sprachstadienübergreifendes Wörterbuch zur Verfügung, das die zu schaffende Hyperlemmaliste des Deutschen um das Frühneuhochdeutsche und Neuhochdeutsche erweitern wird. Mit seinem reichen Sprachschatz kann das DWB ein zentraler Bestandteil eines Informationsnetzes zur deutschen Sprache sein, von dem aus zahlreiche weitere Verzweigungen, vor allem auch zu den in den Belegzitate repräsentierten Primärquellen, realisierbar sind.

Eine Vernetzung mit den für die deutsche Sprache wichtigen Autorenwörterbüchern liegt also auf der Hand. Zum Individualwortschatz Goethes, des im DWB am häufigsten zitierten Autors, wird seit 1946 ein vollständiges Wörterbuch erstellt, dessen elektronische Aufbereitung derzeit mit DFG-Förderung in Trier erfolgt und das mit dem DWB verknüpft wird. Als Typus des ‚Autorenwörterbuchs‘ wird es eine Verbundkomponente bilden, die beispielsweise die Möglichkeit bietet, über die Belegzitate eine Verfeinerung der Auszeichnung des DWB vorzunehmen und damit weitere Recherchen zu spezifizieren. Auch das Goethewörterbuch selbst wird durch das Einbringen von Metadaten weiter erschlossen: Im GWB werden Belege nach Band und Seite der Sophienausgabe zitiert, nicht nach Werken Goethes. Ideal wäre sicherlich eine direkte und multidirektionale Verknüpfung mit dieser Ausgabe selbst, die jedoch derzeit nicht in einer frei zugänglichen digitalen Version zur Verfügung steht. Als Behelf bietet sich folgendes Verfahren an: Verknüpft man eine Auswertung der Sophienausgabe, die die Seitenangaben mit Werktiteln hinterlegt, mit dem Wörterbuch, lassen sich werkbezogene Recherchen im Goethewörterbuch durchführen. Auch eine Klassifizierung dieser Auswertung nach Textsorten bzw. Gattungen ist geplant, so dass zum Beispiel Recherchen nach Adjektivbildungen in der Lyrik Goethes möglich werden.

Im Falle des derzeit im Aufbau befindlichen „Digitalen Verbundes von Dialektwörterbüchern“ muss neben der genauen Feststellung der Etymologie und Bedeutung der Stichwörter auch eine Auswertung des

Inhalts der einzelnen Artikel erfolgen, um am hochdeutschen Standard orientierte („Pfälzisches Wörterbuch“, „Rheinisches Wörterbuch“) und dialektale Lemmaansätze („Wörterbuch der elsässischen Mundarten“, „Wörterbuch der deutsch-lothringischen Mundarten“, „Luxemburgische Wörterbücher“) korrekt aufeinander beziehen zu können. Überdies müssen die Dialektwörterbücher auch über ein Sachstichwort aufeinander beziehbar gemacht werden, um die reiche Synonymie und Heteronymie der Dialekte recherchierbar zu machen. Damit werden Vorarbeiten nicht nur für eine semasiologisch, sondern auch für eine onomasiologisch orientierte Verbindung von einzelnen Informationspositionen geleistet, die auch die automatische Erstellung von Umkehrwörterbüchern und eine entsprechende Nutzung der angewandten Methoden vorantreiben können.¹³

Bei den Dialektwörterbüchern ließe sich darüber hinaus anhand der Belegorte eine punktgenaue Verortung recherchieren, aber auch kartografisch visualisieren bzw. eine Vernetzung mit Sprachatlanten erreichen¹⁴, so dass die großräumigeren Angaben der Sprachstadienwörterbücher durch kleinräumige Kontrastierungen ergänzt werden könnten.

Eine andere Art der semantischen Erschließung wird von der Universitätsbibliothek Trier bei der Digitalisierung der 242 Bände umfassenden „Ökonomischen Enzyklopädie“ von Johann Georg Krünitz angewendet. Die Klassifizierung der Artikel mit Hilfe der Dewey Decimal Classification (DDC) ermöglicht auch international standardisierbare inhaltliche Recherchen.¹⁵ Als Perspektive ist eine Verbindung der onomasiologischen

13 Eine besondere Schwierigkeit bei der semantischen Auswertung stellen Phraseologismen dar, die sowohl für historische Sprachstufen als auch für Dialekte nur unzureichend erforscht sind. In Trier werden derzeit zwei Datenbanken zur historischen (Dr. Natalia Filatkina, Prof. Dr. Claudine Moulin) und zur dialektalen Phraseologie (Kerstin Knop, M.A.) erstellt und ausgewertet. Die Ergebnisse dieser Arbeiten fließen in die Erschließung der Wörterbuchbibliothek ein.

14 Die Erprobung einer solchen Vernetzung anhand des Verbunds von Dialektwörterbüchern und dem Digitalen Wenkeratlas ist in Arbeit <<http://www.diwa.info>>; vgl. dazu auch Fournier (wie Anm. 7), S. 174.

15 Darüber hinaus stellt dieses Sachwörterbuch eine hervorragende Ausgangsbasis für eine Einbeziehung weiterer Enzyklopädien und Nachschlagewerke dar, indem sie etwa das „Grammatisch-kritische Wörterbuch“ von Johann Christoph Adelung (<<http://mdz.bib-bvb.de/digbib/lexika/adelung>>) oder die „Systema Naturae“ Carl

Erschließung mit der Sacherschließung gut vorstellbar, denn gerade solche Arten der Vernetzung ermöglichen die Überwindung von Disziplingrenzen.

Das Prinzip der Interoperabilität ermöglicht außerdem die Verknüpfung mit externen Ressourcen. Neben der Vernetzung der durch das Kompetenzzentrum bereitgestellten Wörterbuchressourcen bietet das Wörterbuch-Netz auch die Möglichkeit, aus externen elektronischen Dokumenten angesprochen zu werden. Über die vor den Wortartikeln eingerichteten Icons kann zum einen die persistente Adresse des betreffenden Artikels abgefragt werden, zum anderen stellt das Kompetenzzentrum auf Anfrage auch die vollständigen Lemmalisten mit den zugehörigen Adressen zur Verfügung, um auf diese Weise eine umfassende Vernetzung zwischen einer externen Wörterbuchinstallation und den Ressourcen im Wörterbuch-Netz realisieren zu können. So existieren bereits Verknüpfungen aus dem *Deutschen Rechtswörterbuch*¹⁶ und dem *Wörterbuch der Deutschen Winzersprache*¹⁷ in das Wörterbuchnetz, die demonstrieren, wie Vernetzungen dezentral realisiert werden können.

Schließlich ergibt sich im Kontext einer vernetzten ‚Wörterbuchbibliothek‘ die Möglichkeit der weiterführenden Vernetzung der Wörterbuchbelege mit Primärquellen. Eine solche Verknüpfung von großen Textkorpora und Wörterbüchern bietet eine neue Qualität wissenschaftlichen Arbeitens, da die Vernetzung in beide Richtungen erfolgen kann, von den Wörterbüchern zu den Quellen und umgekehrt. Elektronische Texteditionen, aber auch alle anderen Formen von Primärquellen (Audio, Video, Karten, Musiknoten, Diagramme etc.) können damit auf sehr einfache Weise mit Lemmata oder einzelnen Textstrecken verknüpft werden. Dies macht je nach Maßgabe des Auszeichnungsgrades kritische Apparate unmittelbar transparent und überprüfbar. Modellhaft wurde dies bereits

von Linnés (<http://www-gdz.sub.uni-goettingen.de/cgi-bin/digbib.cgi?PPN371257700>) zitiert, so dass die digitalen Versionen dieser beiden und weiterer Quellen mit der Krünitzschen Enzyklopädie verlinkt werden können.

16 <http://www.deutsches-rechtswörterbuch.de>; Akademie der Wissenschaften Heidelberg.

17 <http://www.winzersprache.de/onlinewb>; Akademie der Wissenschaften und der Literatur Mainz.

durch die Verknüpfung des ‚Findebuchs zum mittelhochdeutschen Wortschatz‘ und des ‚Mittelhochdeutschen Textarchivs‘ realisiert.

Auswertungsbeispiel

Ein konkretes Beispiel soll im Folgenden das Vernetzungs- und Auswertungspotenzial der entsprechend annotierten Nachschlagewerke und Quellen veranschaulichen.¹⁸

Bezeichnungen für Realien, insbesondere für Früchte oder landwirtschaftliche Erzeugnisse, sind zum einen für die Sprachwissenschaft, zum anderen aber auch für kulturwissenschaftliche wie naturwissenschaftliche Disziplinen besonders interessant, da diese Bereiche zumeist eine reiche und häufig hochspezialisierte regional gegliederte Synonymik aufweisen. Die ‚Brombeere‘ (*Rubus fruticosus*) als eher selten domestizierte typische Wildfrucht bietet ein vergleichsweise gut überschaubares Beispiel. Neben dem durchsichtigen Grundwort *-beere* geht das heute nicht mehr verständliche Bestimmungswort *Brom-* vermutlich auf die indogermanische Wurzel **bher* für ‚etwas Spitzes, Kantiges, Eckiges‘ zurück.

Eine semasiologische Gliederung und Auswertung anhand der vorhandenen Wörterbücher zeigt, dass ‚Brombeere‘ regional gegliedert durchaus verschiedene Realien bezeichnet (Brombeere, Brombeerstrauch, dornige Ranke, Heidelbeere, Preiselbeere etc.) und auch im übertragenen Sinn gebraucht werden kann (Gewitterwolke, schwarzhaariges Mädchen etc.). Eine onomasiologische Gliederung des Materials fördert ferner die reiche Synonymik des Begriffs mit unterschiedlichen semantischen Konnotationen zu Tage (Schmerzbeere, Heckenbeere, Bachbeere, Schwarzbeere, Pferdsbeere, Hundsbeere, More etc.). Als Motivation für die Benennung der Brombeere beschreiben die Wörterbücher verschiedene Aspekte, zu denen Eigenschaften, Aussehen und Verwendung der Pflanze gehören können.

Eine sachbezogene Erschließung macht deutlich, dass die volksspra-

18 Die Auswertungsmöglichkeiten können hier nur kurz angerissen werden, vgl. daher ausführlicher unter <http://www.dvw.uni-trier.de/onomasiologie/start.html>.

chigen Bezeichnungen für die Beerenfrüchte biologisch nicht eindeutig sind, das heißt, dass Bezeichnungen wie ‚Heidelbeere, Himbeere, Preiselbeere, Brombeere‘ nach ihrer biologisch korrekten Bezeichnung disambiguiert werden müssen. Die Klassifikation der Artikel des Krünitz gemäß DDC ermöglicht weitergehende Fragestellungen und Vergleichsmöglichkeiten, zum Beispiel ob für Pflanzen der Klassifikation ‚Rosales‘ (im Krünitz unter anderem die Lemmata *Eberäsche*, *Erd=Beere*, *Geiß=Bart*, *Odermennige*) bei ähnlichem Aussehen und Eigenschaften ähnliche Benennungen verwendet werden. Abgesehen von dieser eng an die im Wörterbuch angelegte Thematik eröffnet die DDC-Klassifikation vielfältige Anknüpfungspunkte; für ‚Brombeere‘ sind dies neben den schon genannten ‚Rosales‘ beispielsweise die Themenkreise „Ökologie“, „Schäden, Krankheiten, Schädlinge an Pflanzen“, „Kochen“, „Wein“, „Fütterung (Viehzucht)“, „Pharmakologie“ etc. Nicht nur die Sprachwissenschaft, sondern auch die Wissenschaftsgeschichte bis hin zur modernen Botanik oder Biodiversitätsforschung kann also von einer solch differenzierten Aufschlüsselung entsprechender Wortfelder profitieren.

Die Vorgehensweise für die Realisierung solcher Konzepte kann folgendermaßen aussehen: Die Mikrostruktur der Wörterbuchartikel wird ausgezeichnet, zum Beispiel nach den Informationseinheiten ‚Stichwortansatz (abgebildet auf ein Hyper-Lemma)‘, ‚Etymologie‘, ‚Bedeutungsangabe‘ und ‚Ortsangabe‘, zudem erfolgt eine Annotierung über eine standardisierte Klassifizierung (DDC, Ontologien).

Die Kombination der zusätzlich eingebrachten Metadaten ermöglicht dann spezifizierte Recherchen: Eine semasiologische Recherche funktioniert ausdrucksseitig und greift auf den Wortlaut der Stichwörter zurück. Durch die Abbildung auf das Hyper-Lemma führt die Recherche auch bei grafischer bzw. diachroner Varianz zu etymologisch gleichen Stichwörtern (zum Beispiel mhd. *brâm-ber*, nhd. *Brombeere*). Dabei werden alle Artikel zur Zeichenfolge ‚Brombeere‘ gefunden, auch wenn zum Beispiel semantisch die ‚Maulbeere‘ dahinter steht. Eine onomasiologische Recherche geht semantisch-begrifflich vor, es wird der Begriff, die „Sache“ Brombeere gefunden, auch wenn sie mit *Schwolzel*, *Mouze* oder *Schmierbat* bezeichnet wird. Eine sachbezogene Recherche findet auch die Artikel, in

denen es im weiteren bzw. impliziten Sinne um die ‚Brombeere‘ geht, auch dann, wenn in dem entsprechenden Artikel die Brombeere gar nicht selbst namentlich, sondern nur implizit – zum Beispiel durch die Nennung der biologischen Klasse der ‚Rosales‘ – genannt ist.

Methoden und Verfahren zur rechnerunterstützten Vernetzung

Aus der Sicht der Informatik stellt sich die Frage, mit welchen Methoden sich der Aufbau eines Wörterbuchnetzes algorithmisch unterstützen lässt. Eine erste Ansatzmöglichkeit, die im Folgenden näher vorgestellt werden soll, bietet hier der Bereich des Information Retrieval. In einem ersten Schritt wird von der Realität abstrahiert und ein geeignetes Modell für ein Wörterbuchnetz erstellt, welches dann als Ausgangspunkt zur Implementierung der Algorithmen dient. Ganz allgemein entspricht einem Wörterbuchnetz dabei die informationstheoretische Datenstruktur eines aus Knoten und Kanten bestehenden Graphen.¹⁹ Die Knoten des Graphen repräsentieren zunächst die Wörterbücher und die Kanten beschreiben die Verbindungen, das heißt die Verweise zwischen den Wörterbüchern. Diese grobe Struktur lässt sich weiter verfeinern und genauer modellieren, indem nicht das gesamte Wörterbuch als Knoten aufgefasst wird, sondern die einzelnen Wortartikel. Die Kanten verlaufen dann zwischen den Artikeln verschiedener Wörterbücher bzw. auch innerhalb eines einzelnen Wörterbuchs. Die Kanten sind gerichtet, das heißt Ausgangs- und Zielpunkt sind eindeutig bestimmt.

Ausgehend von den gedruckten Wörterbüchern lassen sich die Verweise zunächst in drei Klassen einteilen: 1. Verweise, die explizit im Wörterbuch genannt sind, beispielsweise durch Angabe einer Seiten-, Spalten-, Zeilen- oder Artikelreferenz; 2. Verweise, die sich aufgrund von statistischen Berechnungen der Wörter in den Artikeln identifizieren lassen; 3. Verweise, die sich durch Graphalgorithmen aus den Klassen 1 und 2 berechnen lassen, beispielsweise durch transitiven Abschluss. Mit

¹⁹ Vgl. Mehlhorn, Kurt, *Data Structures and Algorithms 2: Graph-Algorithms and NP-Completeness*, in: *EATCS Monographs on Theoretical Computer Science*, Berlin 1984.

den im Folgenden vorgestellten Verfahren sollen zunächst Verweise der Klasse 2 bestimmt werden.

Ausgangsbasis zur Implementierung der Algorithmen bildet ein Datenbanksystem (siehe Abbildung 1), in welchem die Wörterbücher in strukturierter Form abgespeichert sind. Die Architektur des Gesamtsystems basiert auf einer standardisierten SGML/XML-Kodierung²⁰ der Wörterbuchdaten, die sich nach unterschiedlichen Document-Type-Definitionen (DTDs) richten kann. Zu jeder der vorgegebenen Kodierungen existiert ein zugehöriger Importfilter, über den die Wörterbuchdaten in ein Datenbankmanagementsystem (DBMS)²¹ übernommen werden können. In umgekehrter Richtung wird ein Exportfilter entwickelt, über den die durch die Bearbeitung innerhalb des Systems neu eingefügten Informationen in Form von zusätzlichen Metadaten zu den grundlegenden Wörterbuchdaten ebenfalls in standardisiertem XML-Format aus dem Datenbanksystem herausgezogen werden können. Auf diese Weise stehen sowohl die Wörterbuchdaten als auch die Beziehungen zwischen den Wörterbüchern in einer plattformunabhängigen und damit langfristig nutzbaren Form zur Verfügung.

Kernstück des gesamten Systems bildet ein relationales Datenbankmanagementsystem, in dem jeweils ein Wörterbuch durch eine zugehörige Datenbank verwaltet wird. Der Aufbau der einzelnen Datenbanken richtet sich dabei nach der Granularität der Auszeichnung. Als kleinste gemeinsame Schnittmenge wird nur festgelegt, dass separate Tabellen zur Abfrage der Stichwörter existieren. Weitere Konstituenten der Wörterbuchartikel wie beispielsweise Angaben zur Wortart, Bedeutungserläuterungen, Belege, Angaben zur Etymologie usw. werden dann in eigenen Datentabellen verwaltet, wenn sie durch die XML-Auszeichnung innerhalb des Gesamtartikels markiert wurden. Je feiner die Auszeichnung ist, desto differenzierter sind die Suchmöglichkeiten und desto genauer die Suchergebnisse.²²

20 Vgl. Goldfarb, Charles, *The SGML Handbook*, Oxford 1990.

21 Vgl. Rob, Peter; Coronel, Carlos, *Database Systems. Design, Implementation and Management*, in: *The Wadsworth Series in Management Information Systems*, Belmont, California 1993.

22 Bei Wörterbüchern, bei denen die Artikelmikrostrukturen kaum differenziert

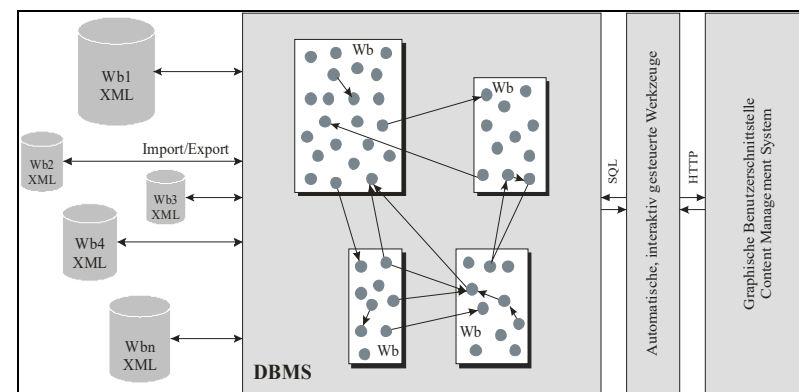


Abbildung 1: Architektur des Datenbanksystems zur Speicherung der Wörterbuchdaten

Ausgehend von den in den Datenbanken abgespeicherten Wörterbuchartikeln können nun automatische Verfahren entwickelt werden, die zunächst zur Aufgabe haben, sämtliche Wortformen der Wörterbücher zu gewichten. Diese Gewichtung erfolgt durch quantitative Angaben über die Häufigkeit einer Wortform (=Term) innerhalb eines Wortartikels und innerhalb des gesamten Wörterbuchs. Berechnet werden die so genannte relative Häufigkeit und die inverse Dokumentfrequenz für jeden Term. Aus beiden Werten ergibt sich dann durch Multiplikation das so genannte Termgewicht. Ein Beispiel für diese Berechnungsvorlage zeigt Abbildung 2.

werden, bliebe im ungünstigsten Fall nur der Zugriff über eine herkömmliche Volltextrecherche.

TextID	Wort	Normierte Form	ArtikelID	Häufigkeit der Wortform	Wortanzahl im Artikel	Dokumenten-häufigkeit	Inverse Dokumentenfrequenz	Gewicht
1507749	butterblume	butterblume	GB13619	1	31	21	14.791	-58.486
1507750	f	f	GB13619	1	31	93291	2.674	-10.572
1507751	nnl	nnl	GB13619	1	31	3629	7.280	-26.788
1507752	boterbloem,	boterbloem	GB13619	1	31	1	19.183	-75.854
1507753	gilt	gilt	GB13619	1	31	2996	7.683	-30.361
1507754	für	fuer	GB13619	4	31	40359	3.882	-10.220
1507755	mehreere	mehere	GB13619	1	31	1408	8.724	-34.495
1507756	kräuter,	kraeuter	GB13619	1	31	467	10.316	-40.791
1507757	für	fuer	GB13619	4	31	40359	3.882	-10.220
1507758	caltha	caltha	GB13619	1	31	31	14.229	-56.264
1507759	palustris,	palustris	GB13619	1	31	163	11.834	-46.795
1507760	döterblume,	döterblume	GB13619	1	31	10	15.861	-62.718
1507761	für	fuer	GB13619	4	31	40359	3.882	-10.220
1507762	leontodon	leontodon	GB13619	1	31	42	13.791	-54.531
1507763	taraxacum,	taraxacum	GB13619	1	31	53	13.455	-53.204
1507764	für	fuer	GB13619	4	31	40359	3.882	-10.220
1507765	ranunculus	ranunculus	GB13619	1	31	126	12.183	-48.174
1507766	aunicomus	aunicomus	GB13619	1	31	8	16.183	-63.991
1507767	das	das	GB13619	1	31	94491	2.817	-11.137
1507768	volk	volk	GB13619	1	31	4258	7.127	-28.182
1507769	wähnt,	waeht	GB13619	1	31	74	12.974	-51.300
1507770	davon,	davon	GB13619	1	31	5680	8.711	-26.538
1507771	dasz	dasz	GB13619	1	31	35412	4.071	-16.096
1507772	sie	sie	GB13619	1	31	45488	3.710	-14.870
1507773	die	die	GB13619	2	31	123292	2.271	-7.653
1507774	kühe	kuehe	GB13619	1	31	374	10.636	-42.058
1507775	fressen,	fressen	GB13619	1	31	967	9.266	-36.636

Abbildung 2: Berechnung der inversen Dokumentfrequenz für einen Wortartikel aus dem Deutschen Wörterbuch

Die Tabelle zeigt die durch den Algorithmus berechneten Werte für den Artikel *Butterblume* aus dem Deutschen Wörterbuch von Jacob und Wilhelm Grimm. Entscheidend zur Berechnung des Gewichtes (letzte Spalte) sind die Angaben über die Häufigkeit der Wortform, die Länge des Artikels (gemessen in Wortformen) sowie die Dokumentenhäufigkeit, die die Anzahl der Dokumente (=Wortartikel) repräsentiert, in denen der betreffende Term auftritt. Jedes Dokument wird dabei nur einmal gezählt, auch wenn der Term in ihm mehrfach vorkommt. Aus diesen Angaben lässt sich das Gewicht eines Terms bestimmen: Der Bezug zwischen Häufigkeit der Wortform und Artikellänge ergibt seine relative Häufigkeit im betrachteten Wortartikel; der Bezug zwischen Dokumentenhäufigkeit und der Gesamtanzahl an Dokumenten²³ liefert die inverse Dokumentenfrequenz; das Produkt aus beiden Werten bildet das Gewicht des Terms im vorliegenden

23 Beim Deutschen Wörterbuch sind dies 297.613 Dokumente.

Artikel. Sortiert man nun die Terme eines jeden Artikels nach den berechneten Gewichten, so erhält man einen sogenannten Relevanzvektor der Terme. Für den betrachteten Beispielartikel ist diese Sortierung in Abbildung 3 gezeigt.

TextID	Wort	Normierte Form	ArtikelID	Häufigkeit der Wortform	Wortanzahl im Artikel	Dokumenten-häufigkeit	Inverse Dokumentenfrequenz	Gewicht
1507752	boterbloem,	boterbloem	GB13619	1	31	1	19.183	-75.854
1507766	aunicomus,	aunicomus	GB13619	1	31	8	16.183	-63.991
1507760	döterblume,	döterblume	GB13619	1	31	10	15.861	-62.718
1507749	butterblume	butterblume	GB13619	1	31	21	14.791	-58.486
1507758	caltha	caltha	GB13619	1	31	31	14.229	-56.264
1507762	leontodon	leontodon	GB13619	1	31	42	13.791	-54.531
1507763	taraxacum,	taraxacum	GB13619	1	31	53	13.455	-53.204
1507764	für	fuer	GB13619	4	31	40359	3.882	-10.220
1507765	ranunculus	ranunculus	GB13619	1	31	126	12.183	-48.174
1507766	aunicomus	aunicomus	GB13619	1	31	8	16.183	-63.991
1507767	das	das	GB13619	1	31	94491	2.817	-11.137
1507768	volk	volk	GB13619	1	31	4258	7.127	-28.182
1507774	kühe	kuehe	GB13619	1	31	374	10.636	-42.058
1507756	kräuter,	kraeuter	GB13619	1	31	467	10.316	-40.791
1507779	geb,	geb	GB13619	1	31	522	10.155	-40.156
1507776	butter	butter	GB13619	1	31	666	9.804	-38.766
1507775	fressen,	fressen	GB13619	1	31	967	9.266	-36.636
1507755	mehreere	mehere	GB13619	1	31	1408	8.724	-34.495
1507753	gilt	gilt	GB13619	1	31	2996	7.683	-30.361
1507751	nnl	nnl	GB13619	1	31	3629	7.280	-26.788
1507768	volk	volk	GB13619	1	31	4258	7.127	-28.182
1507776	verde	verde	GB13619	1	31	5146	6.854	-27.101
1507770	davon,	davon	GB13619	1	31	5680	8.711	-26.538
1507771	dasz	dasz	GB13619	1	31	35412	4.071	-16.096
1507772	sie	sie	GB13619	1	31	45488	3.710	-14.870
1507767	das	das	GB13619	1	31	94491	2.817	-11.137
1507750	f	f	GB13619	1	31	93291	2.674	-10.572
1507754	für	fuer	GB13619	4	31	40359	3.882	-10.220
1507757	für	fuer	GB13619	4	31	40359	3.882	-10.220

Abbildung 3: Sortierung der Wortformen nach den berechneten Termgewichten (=Relevanzvektor)

Für jeden Wortartikel eines Wörterbuchs wird der Relevanzvektor nach einer festen Maximalanzahl von Elementen abgebrochen.²⁴ Auf diesem Weg entsteht also für jedes in das Netz zu integrierende Wörterbuch eine Liste von Relevanzvektoren.

Im darauffolgenden Schritt werden nun je zwei dieser Vektorlisten elementweise miteinander verglichen, das heißt pro Vektorvergleich werden die in ihnen enthaltenen Wortformen gegeneinander geprüft und die An-

24 Auf diese Weise werden die für die Vergleiche irrelevanten hochfrequenten Wortformen ausgeschlossen, da sie sich durch sehr niedrige Gewichtung auszeichnen.

zahl der Übereinstimmungen ermittelt. Hier wird zunächst auf exakte Gleichheit getestet, womit eine Basis für die möglichen Verweise zwischen den Wortartikeln geschaffen wird. In weiteren Verfeinerungen dieses Algorithmusschrittes könnten auch Informationen über die Wortformen berücksichtigt werden (zum Beispiel durch Lemmatisierung, Stemming, approximative Vergleiche etc.), um zusätzliche Übereinstimmungen zu ermitteln, die dann auch entsprechend bewertet werden können. Einen Vergleichsvektor zu obigem Beispiel zeigt Abbildung 4. Hierbei handelt es sich um den Artikel *Gackel(s)-blume* aus dem Pfälzischen Wörterbuch.

TextID	Wort	Normierte Form	ArtikelID	Häufigkeit der Wortform	Wortanzahl im Artikel	Dokumenten-häufigkeit	Inverse Dokumentfrequenz	Gewicht
1117999	Gackelblumm	gackelblumm	PG00045	1	43	1	17.251	-76.359
1117918	Gackelblümm	gackelblümm	PG00045	1	43	1	17.251	-76.359
1117927	Gackelblümm	gackelblümm	PG00045	1	43	1	17.251	-76.359
1117930	Kuhschelle	kuhschelle	PG00045	1	43	3	15.666	-69.343
1117936	Gackel(s)-e	gackel(e)s	PG00045	1	43	3	15.666	-69.343
1117993	Gackel(s)-blume	gackel(s)blume	PG00045	1	43	4	15.251	-67.506
1117903	Gackel1	gackel1	PG00045	1	43	4	15.251	-67.506
1117901	Godramst	godramst	PG00045	1	43	5	14.929	-66.081
1117911	[Ranunculus]	ranunculus	PG00045	1	43	7	14.444	-63.933
1117897	[Caltha]	caltha	PG00045	1	43	8	14.251	-63.080
1117910	Hahnenfuß	hahnenfuss	PG00045	1	43	9	14.081	-62.328
1117928	putzattay	putzattia	PG00045	1	43	9	14.081	-62.328
1117924	Küchenschelle	kuechenschelle	PG00045	1	43	10	13.929	-61.859
1117914	Butterblume	butterblume	PG00045	1	43	11	13.792	-61.048
1117996	palustris1	palustris	PG00045	1	43	12	13.666	-60.491
1117917	Akelei	akelei	PG00045	1	43	16	13.251	-58.654
1117922	Glockenblume	glockenblume	PG00045	1	43	25	12.807	-55.904
1117925	[Anemone]	anemone	PG00045	1	43	27	12.486	-55.312
1117934	1049	1049	PG00045	1	43	31	12.297	-54.430
1117896	Dotterblume	dotterblume	PG00045	2	43	10	13.929	-61.859
1117906	Dotterblume	dotterblume	PG00045	2	43	10	13.929	-61.859
1117928	[LA-Rohrb]	larohrb	PG00045	1	43	80	11.344	-50.213
1117900	[LA-Wey]	lawey	PG00045	1	43	170	9.842	-43.563
1117912	[Gemm]	gemmh	PG00045	1	43	334	8.868	-39.250
1117919	[GH-Neubg]	ghneubg	PG00045	1	43	564	8.112	-35.905
1117905	c	c	PG00045	1	43	986	7.298	-32.280
1117916	2	2	PG00045	1	43	9211	4.248	-18.902

Abbildung 4: Relevanzvektor zum Artikel *Gackel(s)-blume* aus dem Pfälzischen Wörterbuch

Vergleicht man die normierten Wortformen der Vektoren aus Abbildung 3 und Abbildung 4 miteinander, so findet man drei exakte Übereinstimmungen in den Formen *caltha*, *ranunculus* und *butterblume* innerhalb der ersten 20 gewichteten Terme. Über einen bestimmten Schwellenwert wird für den Gesamtalgorithmus vorgegeben, wann ein Vergleich positiv zu

bewerten ist, das heißt wann ein Verweis zwischen den verglichenen Wortartikeln etabliert werden soll. In obigem Beispiel bedeutet die Anzahl von drei Übereinstimmungen, dass eine Kante im Wortartikelgraphen vom Artikel *BUTTERBLUME* (Deutsches Wörterbuch) zum Artikel *Gackel(s)blume* (Pfälzisches Wörterbuch) eingerichtet werden soll.

Jeder derart ermittelte Verweis wird in einer zusätzlichen Datenbank als Metainformationen über den eigentlichen Wörterbuchdaten abgespeichert, die Basisdaten bleiben also unverändert. Beim Zugriff auf ein Wörterbuch kann aus dieser Datenbank die Vernetzungsinformation ausgelesen und ebenfalls in der grafischen Benutzeroberfläche visualisiert werden (siehe Abbildung 5).

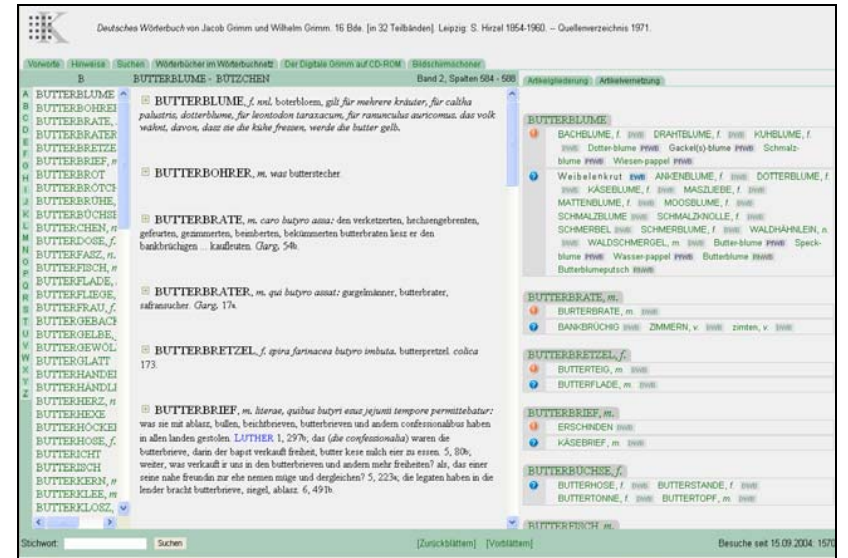


Abbildung 5: Visualisierung der Vernetzungsinformation zur Artikelstrecke *BUTTERBLUME* im Deutschen Wörterbuch von Jacob und Wilhelm Grimm

Abbildung 5 zeigt die Vernetzung des Artikels *BUTTERBLUME* im Deutschen Wörterbuch mit weiteren Artikeln des DWB (zum Beispiel *BACH-*

BLUME, DRAHTBLUME, KUHBLUME, KÄSEBLUME etc.), aber auch mit Artikeln aus anderen Wörterbüchern (zum Beispiel *Dotterblume, Schmalzblume* im Pfälzischen Wörterbuch, *Weibelenkrut* im Elsässischen Wörterbuch oder *Butterblume* im Rheinischen Wörterbuch). Die Symbole vor den Verweisgruppen in Abbildung 5 kennzeichnen die Zahl der Übereinstimmungen während der Vektorenvergleiche, höhere Anzahl von Übereinstimmungen führt zu einer Höhergruppierung des Verweises und bedeutet in der Regel eine höhere Zuverlässigkeit der Verknüpfung.

Perspektiven

Das hier skizzierte ambitionierte „Ideal-Ziel“, ein vielfältig verknüpftes, intelligentes ‚Meta-Wörterbuch‘ der ‚deutschen Sprache‘, steht noch in seinen Anfängen und ist gewiss nur in einer gemeinsamen, „koordinierten“ Forschungs- und Umsetzungs-Anstrengung realisierbar. Die dafür notwendigen Konzepte, Methoden und Verfahren müssen in einem interdisziplinären Methodenbündel aus den folgenden Bereichen zusammenwirken:

- Informatik (Ontologien, Graphalgorithmen, Information Retrieval etc.),
- Computerphilologie/Computerlinguistik (automatisches Markup, Pattern Matching, Konkordanzen, automatische Lemmatisierung etc.),
- Systematische Linguistik (Abbildung von Lautgesetzen synchron und diachron, Erarbeitung von Hyper-Lemmalisten, Phraseologie etc.),
- Lexikografie/Lexikologie (Semasiologie, Onomasiologie, Umkehrlexikografie).

Von einem solchen Meta-Zugriff können alle an Sprachinformationen interessierten Disziplinen in höchstem Maße profitieren, darüber hinaus sind weitreichende Impulse für neue Fragestellungen und Lösungsansätze auf verschiedensten Forschungsfeldern zu erwarten.

Dr. Thomas Burch ist seit 1998 Geschäftsführer des Kompetenzzentrums für elektronische Erschließungs- und Publikationsverfahren in den Geisteswissenschaften an der Universität Trier. Die Konzeption und Organisation von Projekten im Bereich des elektronischen Publizierens nach neuesten informationstechnologischen Standards bildet den Schwerpunkt seiner Tätigkeiten. E-Mail: burch@uni-trier.de

Dr. Andrea Rapp war bis September 2004 Leiterin des Göttinger Digitalisierungszentrums an der Niedersächsischen Staats- und Universitätsbibliothek Göttingen; sie ist derzeit Akademische Rätin im Fach Germanistik/Ältere deutsche Philologie und Geschäftsführerin des Kompetenzzentrums für elektronische Erschließungs- und Publikationsverfahren in den Geisteswissenschaften an der Universität Trier. E-Mail: rappand@uni-trier.de

VOM NACHSCHLAGEWERK ZUM VIRTUELLEN
INFORMATIONSSYSTEM –
PERSPEKTIVEN HISTORISCHER LEXIKOGRAFIE IM INTERNET

von Robert Charlier

Der Beitrag widmet sich den Perspektiven historischer Wörterbücher im weltweiten Wissensnetz aus der Sicht des Wörterbuchmachers. Nach einem Überblick über die traditionellen Techniken enzyklopädischer und lexikografischer Wissensordnung und einem Strukturvergleich zwischen historischer und virtueller „Wörterbuchwelt“ rückt das Goethe-Wörterbuch als wichtigstes Klassikerwörterbuch deutscher Sprache in den Mittelpunkt. Der Einblick in die logisch-semantische Datenstrukturierung dieses paradigmatischen historischen Autorenwörterbuchs ermöglicht einen kurzen Ausblick auf aktuelle Entwicklungstendenzen lexikografischer Infrastrukturen im Internet. Ein vorläufiges Fazit resümiert Vernetzung, Visualisierung und Miniaturisierung als wichtigste Tendenzen im Prozess der virtuellen Optimierung von Wörterbüchern.

Von der Systematik zur Alphabetik

Als bedeutendster Autor und repräsentativer Sprecher seines Zeitalters verwendete Johann Wolfgang von Goethe (1749-1832) den Begriff „Diktionär“ als Eindeutschung des französischen Maskulinums „dictionnaire“. Das eingedeutschte Neutrum gebrauchte er zugleich im Sinne von Enzyklopädie, (Real-)Lexikon und von Wörterbuch (Dictionarium, Thesaurus).¹ Erscheint das Wort zunächst unscheinbar, so markiert es eine historische

Wende in der Evolution der abendländischen Wissensordnung. Im Titel von Diderot und d’Alemberts Jahrhundertwerk *Encyclopédie ou Dictionnaire raisonné des sciences, des arts et des métiers* (1751-80) kündigt die Vorstellung eines umfassend durchdachten, komplexen und universalen ‚Wörterbuchs‘ von der wohl einschneidendsten Zäsur in der Geschichte der Sammlung von Menschheitswissen seit der monumentalen *Naturalis Historia* des Plinius Secundus des Älteren (23-79 n. Chr.). Es handelt sich dabei um die Ablösung des Prinzips der *hierarchischen* Ordnung von Inhalten nach systematischen Kriterien, die von der Antike bis zur Frühen Neuzeit galten. Zwar ist die alphabetische Anordnung bereits Bestandteil des kulturellen Vermächnisses der Antike, und auch das Mittelalter hat eine Fülle alphabetischer Texte (wie zum Beispiel Glossare) hervorgebracht. Doch die Enzyklopädisten und ihre Epigonen ersetzten die thematisch sortierte, jeweils mythisch, theologisch, politisch oder naturkundlich auf- oder absteigende Ordnung der Wörter und Dinge in nie zuvor dagewesener Konsequenz und Universalität durch eine Lemmatisierung der Schlag- oder Stichwörter von A bis Z.² Es handelte sich dabei vermutlich um einen Akt der Übertragung sprachlich arbiträrer Ordnungskriterien, wie sie sich in der Abfolge der Buchstaben des lateinischen Alphabets manifestiert, auf die neue Erfahrungswelt, die sich um die Mitte des 18. Jahrhunderts rasant ausdifferenzierte. Die Rede ist von einer exponentiell wachsenden Welt der exotischen Waren, modernisierten Transportmittel, technischen Erfindungen und neuen Entdeckungen, vor allem im transatlantischen und pazifischen Raum. Die Modernisierungsschübe in Wissenschaft, Wirtschaft und Technik brachten nicht nur eine Fülle von Waren und Dingen, zum Beispiel aus der außereuropäischen Welt, sondern auch eine Flut von Wörtern und Begriffen aus der Abstraktionshöhe neuer Theorien und Terminologien. Dieser Wandel der Wissensstrukturierung hatte wesentlichen Anteil an der Wirkmächtigkeit der *Encyclopédie* als

¹ Vgl. Goethe-Wörterbuch, hg. von der Berlin-Brandenburgischen Akademie der Wissenschaften, der Akademie der Wissenschaften in Göttingen und der Heidelberger Akademie der Wissenschaften. Bd. 2, Stuttgart 1989, Sp. 1205, 24-42.

² Die formalen und sachlichen Ordnungskriterien haben sich innerhalb der Realenzyklopädie bzw. der (Fremd-)Sprachenlexikografie je unterschiedlich entwickelt (vgl. Hass-Zumkehr, Ulrike, Die Sachlexikografie und ihr Verhältnis zur Sprachlexikografie, in: Dies., Deutsche Wörterbücher – Brennpunkt von Sprach- und Kulturgeschichte. Berlin u.a.O. 2001, S. 291-332, bes. S. 297-327).

Instrument der europäischen Aufklärung, die durch eine singuläre Säkularisierung, Rationalisierung und Relativierung des Weltwissens und seiner Repräsentation in die Geschichte eingehen sollte. Was zunächst aus der Notwendigkeit der Bewältigung neuer, ungeahnter Komplexität geboren war – die Instrumentalisierung des unter sachlichen Aspekten weitgehend willkürlichen Alphabets als Ordnungskriterium für die Abfolge innerhalb eines universalen Lexikons – gewann mit der produktiven Rezeption der *Encyclopédie* an Eigendynamik. So relativierte die enthierarchisierte, allein von der alphabetischen Buchstabenabfolge bestimmte Zusammenschau der Inhalte im historisch wohl bedeutendsten Universallexikon der Frühmoderne bestehende Hierarchien in nie dagewesener Weise. So gelangten beispielsweise religionskritische Artikel in die unmittelbare Nachbarschaft zu zentralen Leit- und Wesensbegriffen des christlichen Abendlandes oder der katholischen Kirche.

Historische und virtuelle Wörterbuchwelt

Substituiert man für die ältere Wissensordnung die Kategorie der semantischen Systematik (zum Beispiel die Darstellung von Sachgebieten als Aufstieg von der unbelebten über die organische Natur und die Tiere bis zu den Menschen)³ und für die Lemmatisierung den Aspekt der logischen Serialität (durch alphabetische oder numerische Darbietung, etwa in einer Folge von Datensätzen), so finden sich beide Wissensordnungen in der Welt der Datenbanken und Online-Nachschlagewerke verschränkt. Der Bereich von Daten-Ablage/Speicher (memory) ist dabei logisch (alphabetisch, numerisch und ähnliches) aufgebaut, wogegen die Ebene der Abfrage/Suche (retrieval) nach semantischer bzw. mehr oder weniger

³ Eines der ältesten deutschen Wörterbücher ist der *Vocabularius Sancti Galli* in der Stiftsbibliothek des Klosters St. Gallen. Es handelt sich dabei um eine gut 200 Seiten starke Miniaturhandschrift eines lateinisch-althochdeutschen Wörterverzeichnisses. Das Wörterbüchlein ist nicht alphabetisch organisiert, sondern nach sinnvollen Sachgruppen sortiert. Beginnend mit der Welt der Pflanzen und Gewässer behandelt es den Menschen und dessen Kulturbereiche. Darauf folgen Wortfelder aus den Themenbereichen Tierwelt, Ackerbau, ländliche und dörfliche Kultivation usw. Den Abschluss bildet die Benennung des Himmels, der Wetterphänomene und der Jahreszeiten (siehe Abbildung 1).

systematisierter Vernetzung verlangt. Logisch-semantische Abfrage- und Suchmöglichkeiten zählen gegenwärtig zu den unverzichtbaren Errungenschaften eines ‚durchdigitalisierten‘ Alltags.

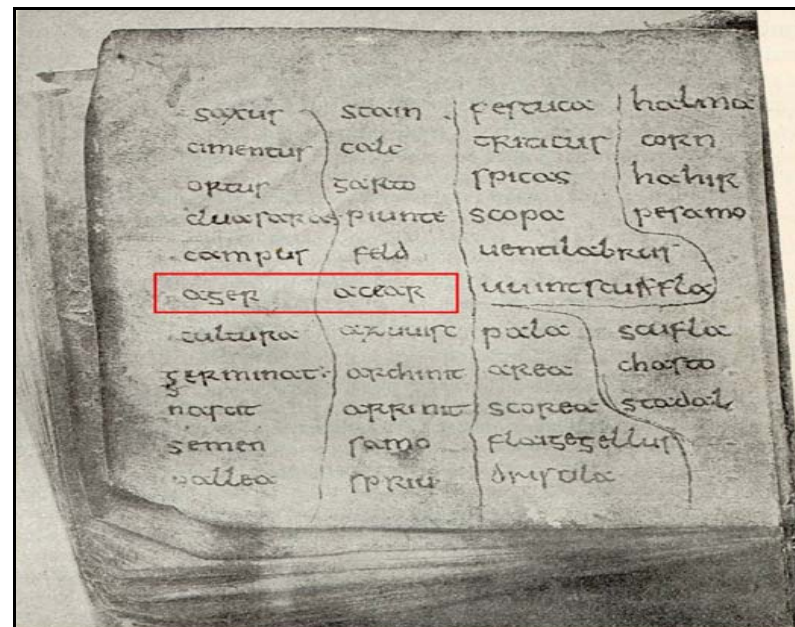


Abbildung 1: *Vocabularius Sancti Galli*, um 790 (althochdt. Systematik lat. Wörter, Beispielwort: *ager*)

Lokale Online-Hilfen für unterschiedlichste Standard-Software, die mehr oder weniger streng nach stichwortbezogenen Kriterien wörterbuchartig bzw. quasi enzyklopädisch aufgebaut sind, bilden dafür nur die naheliegendsten Beispiele. Diese in fast jedem Anwender- und Nutzerbereich der digitalen Medien usuell und ubiquitär gewordenen lexikografischen Infrastrukturen haben den mit der Schriftkultur altüberkommenen Akt der aktiven Wissensaneignung, das „Nachschlagen“, im Prozess des „Durchsuchens“ kraft einer virtuellen Datenabfrage technisch rationalisiert und – semantisch betrachtet – säkularisiert. Ein Informationsvorgang im Internet

oder Digitalmedium, der auf menschliche Empfänger/innen abzielt, basiert auf dem Zusammenspiel eines technisch-physikalischen, logischen Datentransfers mit einer semantischen Schnittstelle. Damit ein Datentransport auch in eine Extraktion von Wissen mündet, also zu einer Information von oder für jemanden werden kann, benötigt diese Schnittstelle einen natürlichsprachlichen Ausgang. Diese anthropomorphe Seite der Schnittstelle zwischen den objektiven Daten und der subjektiven Information aktualisiert zwangsläufig die Ordnungsmuster der althergebrachten (Wörter-)Buch- und Wissenswelt. Damit ist der herkömmliche epistemische Strukturbereich gemeint, der bildhaft auch als „Gutenberg-Galaxis“ firmiert.⁴ So spiegeln die Treffereinträge aus einer Datenabfrage nicht nur Zahl und Inhalt der Felder der zugrunde liegenden Datenbank, sondern auch die semantischen Module eines virtuellen Wörterbucheintrages (Stich-, Schlag- oder Suchwort = Lemma; Fließ- oder Ergebnistext = Belegzitat usw.). Die Akkumulation von Daten im weltweiten Wissensnetz folgt damit mehr oder weniger konsequent den historischen Wissenssystemen, wie sie seit Jahrhunderten durch Institutionen und Techniken wie Archiv, Bibliothek, Index, Kartei, Katalog, Register oder Thesaurus vorgegeben sind.⁵

Die natürlichsprachliche Seite des wichtigsten semantischen Kommunikationsvorgangs in der technischen Medienwelt zum Zwecke der informationellen Mehrwertschöpfung – die Information eines Menschen als Netz- oder Mediennutzer/in – gewinnt ständig an Bedeutung. Im Rahmen des vorliegenden Beitrages kann unmöglich die Problematik semantischer Schnittstellen für (retro)digitale Wörterbücher in ihrer gesamten Breite und

- 4 Vgl. McLuhan, Marshall, *The Gutenberg Galaxy*, Toronto 1962 [Die Gutenberg-Galaxis. Das Ende des Buchzeitalters, Düsseldorf u.a.O. 1968]. Die Thesen McLuhans über das Ende des Buchzeitalters und die elektronische Vernetzung der Weltgesellschaft zu einem „globalen Dorf“ haben zwar viele Aspekte des Internetzeitalters vorausgesehen. Der sich abzeichnenden technischen und semantischen Synthese der Buch- und Internetwelt wurden sie aber nicht gerecht.
- 5 Vgl. Charlier, Robert, Synergie und Konvergenz. Tradition und Zukunft historischer Semantik am Beispiel des Goethe-Wörterbuchs, in: Müller, Ernst (Hg.), *Begriffsgeschichte im Umbruch?* (Archiv für Begriffsgeschichte, Sonderheft 2004), Hamburg 2005, S. 167-183, bes. S. 175-177.

Tiefe ausgeleuchtet werden. Ich beschränke mich daher im Folgenden auf einige wenige exemplarische Aspekte.

Logisch-semantische Datenstrukturierung

Die elektronische Redaktion des *Goethe-Wörterbuchs* basiert auf der Standard-Software Microsoft (MS) Word. Die SGML/XML-kompatible Datenauszeichnung erfolgt unmittelbar bei der Texteingabe für die Artikel-erstellung. Dabei dienen die standardmäßig vorgegebenen Zeichen- und Absatz-Formatvorlagen als Parameter bzw. Attributräger. Dieses kombinierte Mark-up folgt etwas vereinfacht im Wesentlichen den Kriterien einer Document Type Definition (DTD) des Typs ‚Wörterbuch‘ (*dictionary.dtd*) nach den Richtlinien der Text Encoding Initiative (TEI) und garantiert die logische und semantische Strukturierung⁶ der Text- und Sinnelemente eines Wörterbuchartikels (Textdaten). Die Standard-Software MS Word fungiert durch makrobasierte Modifikationen der Entwicklungsumgebung von MS Office, Visual Basic for Applications, quasi als *Textdatenverarbeitung*. Zwar bleibt die Funktionsmächtigkeit eines klassischen Spezialprogramms der geisteswissenschaftlichen Informationsverarbeitung wie TUSTEP damit unerreicht, dennoch lassen sich für die Erfordernisse der elektronischen Redaktion eines historischen Wörterbuchs erstaunliche Rationalisierungen erzielen.⁷

Die logisch-semantische Datenauszeichnung beschreibt genotypisch, was als Phänotyp auf der sichtbaren Oberfläche der Software-Anwendung vom virtuellen Wörterbuch erscheint. Die Tiefe der Auszeichnung entscheidet über die Breite der Information, die man über ein Wort, einen Begriff oder eine Sache abfragen kann. Die grafische Erscheinungsform rekuriert dabei auf den zugrundeliegenden Datenschatz, eine Textdaten-

- 6 Vgl. dazu etwa Münz, Stefan; Nefzger, Wolfgang, Logische und physische Auszeichnung, in: Dies., *HTML Referenz: Grundlagen, Alternativen und Erweiterungen*. Feldkirchen 1997, S. 81 f. und 18* f. (Systematische Kurzreferenz) sowie im Folgenden öfter ebd.
- 7 Vgl. Charlier, Robert, *Goethe-Lexikographie im Internet – Aspekte prospektiver Digitalisierung*, in: Reeg, Gottfried; Schubert, Martin J. (Hgg.), *Edieren in der elektronischen Ära (Informationstechnologie und Geisteswissenschaften)*; 1), Berlin 2004, S. 17-37, bes. S. 22-35.

struktur, die wie ein *Palimpsest* aufgebaut ist. Je nach Auszeichnungstiefe enthält ein strukturierter Text mehrere Textdatenschichten, so etwa die Nutzeransicht im Dokumentenformat (*what you see is what you get*), die versionsunabhängige Fassung im Rich Text Format (RTF) oder die ansichtsunabhängige, gleichsam ‚grammatische‘ SGML- bzw. XML-Fassung im reinen ASCII-Format (*what you mark is what you get*). Die Berührung von Informationstechnologie und Philologie führt allenthalben zu bildhaften Übertragungen, das belegen zum Beispiel die Metaphern von *Buch*, *Buchrolle* („Scrolling“) oder *Kartei/Karte* im Hypertextbereich. Durch die digitale Revolution sind traditionelle Konzepte des kulturellen Gedächtnisses aber auch auf ihre ursprüngliche Bedeutung zurückgeführt worden. So erweisen sich moderne Bibliotheken aufgrund ihrer Anstrengungen um die elektronische Erschließung von Büchern und Texten in gewisser Weise als digitale *Skriptorien*. Dies ganz im Sinne einer älteren Bedeutung von (Kloster-)Bibliothek, die sich vor allem auch als Ort der Herstellung und Vervielfältigung von (Hand-)Schriften verstand.⁸

Grafische Oberflächen

Die visuelle Anmutung und Struktur der Benutzeroberfläche eines Online-Wörterbuchs ist zwar datenlogisch eher der Peripherie zuzuordnen. Im Zeitalter des Bildes und des Primats der visuellen Kommunikation ist sie jedoch von kaum zu unterschätzender Bedeutung. Daher sei im Folgenden eine Typologie der Erscheinungsbilder virtueller Wörterbücher entworfen. Zu unterscheiden wäre zunächst der Wörterbuchtypus mit windows-kompatibler Vollanzeige, basierend auf einer komplexen Schachtelung von Fenstern und Spalten, wie sie unter anderen das Trierer Kompetenzzentrum für elektronische Erschließungs- und Publikationsverfahren in den Geisteswissenschaften optimiert hat (1). Ein weiterer grafischer Oberflächentyp stellt die integrierte Teil- oder Kompaktanzeige einer (externen) Wörterbuchabfrage als Toolbar oder Add-In dar (2). Neben der Lösung

8 Vgl. Charlier, Robert; Habermehl, Peter, Die Bibliothek von Babel. Impressionen einer Kölner Tagung, in: Circular der Berlin-Brandenburgischen Akademie der Wissenschaften 8/29 (2004), S. 36f. sowie <<http://www.goethe-wortschatz.de/index.php?c=Tagungen#ad4>> (06.09.2006).

von Google-Toolbar bietet das Leipziger *Wortschatz-Lexikon* dafür einen technisch erwähnenswerten Ansatz. Schließlich sind die virtuellen Kleinst-Wörterbücher und miniaturisierten Indizes zu nennen, die eine kontextsensitive Minimalanzeige lexikografischer Angaben in Form von Bubble-Helps oder Pop-Ups gewährleisten (3). Exemplarisch sind hierfür die Thesaurus-, Synonym- und Übersetzungsfunktionen in den neueren Versionen von MS Word zu nennen.

(1) Buchform und virtuelle Vollanzeige: Fenster und Spalten

Das Kompetenzzentrum Trier hat mit einschlägigen Referenzprojekten im Bereich der geisteswissenschaftlichen (Retro-)Digitalisierung zweifellos auch im Bereich der Formensprache digitaler historischer Lexikografie Maßstäbe gesetzt. Das gilt für den Bereich der retrospektiven bzw. prospektiven Datenstrukturierung ebenso wie für die Gestaltung der Benutzeroberflächen historischer Standardwerke, wie sie das *Deutsche Wörterbuch von Jacob und Wilhelm Grimm*⁹ oder das *Goethe-Wörterbuch* darstellen. Die in Trier digitalisierten Wörterbücher dokumentieren, wie weitgehend die Doppelstruktur von „Fenstern/Eingabezeilen“ einerseits und „Karten und Spalten“ andererseits zum grafischen Quasi-Standard in der Web- und Windows-Welt geworden sind. Das virtuelle Formschema, das sich mit der Formel ‚Fenster und Spalten‘ charakterisieren läßt, vereint synoptisch-hypertextuelle und traditionelle hierarchische Elemente (zum Beispiel Sucheingabezeile versus Lemmalaufleiste; Gliederungsansicht versus Artikelfließtext – siehe Abbildung 2).

9 Der Digitale Grimm. Deutsches Wörterbuch. Elektronische Ausgabe der Erstbearbeitung von Jacob Grimm und Wilhelm Grimm, hg. vom Kompetenzzentrum für elektronische Erschließungs- und Publikationsverfahren in den Geisteswissenschaften an der Universität Trier in Verbindung mit der Berlin-Brandenburgischen Akademie der Wissenschaften. 2 CD-ROMs sowie Begleit- und Benutzerhandbuch, Frankfurt am Main 2004, vgl. <<http://www.Zweitausendeins.de>>.

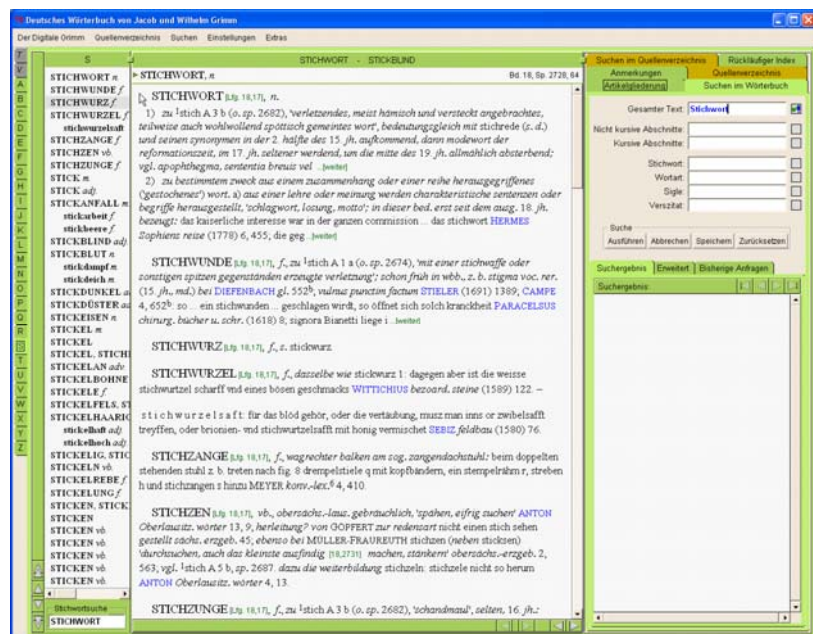


Abbildung 2: Der Digitale Grimm, lokal installierte Vollversion (Stand der Edition von 2004)

Ohne der Feinstruktur dieser virtuellen Bildschirm-Repräsentation eines monumentalen historischen Wörterbuchs bis in alle Verästelungen nachgehen zu können, sei in diesem Zusammenhang auf die einheitliche visuelle Morphologie verwiesen, wie sie sich in der Selbstdarstellung des Trierer Zentrums offenbart.¹⁰ Kulturell aus der Wörterbuch- und Wissenschaftswelt übernommene Formvorgaben wie Buch- oder Seitenmetapher, Karteikarte und Kartenreiter, stehen rein technisch bedingten Modulen gegenüber, wie sie die Eingabezeile für die Stichwortsuche oder das An-

¹⁰ Vgl. Mittelhochdeutsche Wörterbücher auf CD-ROM und im Internet (vgl. <<http://www.mwv.uni-trier.de>>) sowie die Informationsbroschüre: Elektronische Erschließungs- und Publikationsverfahren: Volltextdigitalisierung – Textstrukturierung – SGML/XML-Markup – Graphische Oberflächen – Typographie. Trier 2003, S. 3; 7-11 (vgl. <<http://www.kompetenzzentrum.uni-trier.de>>).

zeige-feld für die erweiterten Suchfunktionen rechts im Bildschirm repräsentieren. Morphologisch und semantisch lassen sich auf diese Weise die Archiv-, Buch- und Drucksatz-Elemente (Seite, Karte, Reiter, Kolumne) von den computerspezifischen Technischelementen separieren (Anzeigefeld, Hilfefenster, Pfeilnavigation). Da es sich bei allen Elementen lediglich um eine virtuelle Repräsentation handelt, kann man bestimmte Strukturen auch dichotomisch deuten. So entspricht zum Beispiel die Anzeige einer Wörterbuchseite einem (oder einer bestimmten Anzahl von) Datenbankfeld(ern) für den Artikelfließtext oder die Synopse mehrer Buchseiten der Schachtelung einer bestimmten Anzahl von Frames. Statische Lesetextzeilen stehen dynamischen Zeilenformen gegenüber, die durch einen Hyperlink oder als Eingabezeile für die Stichwortabfrage über die Funktionalität einer gleichsam starr gedruckten Zeile hinausgehen.

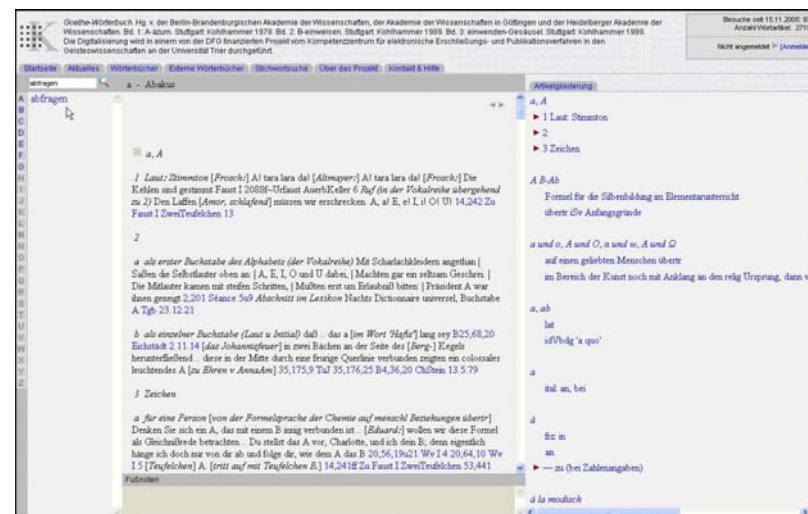


Abbildung 3: Retrodigitalisierung des Goethe-Wörterbuchs, Band I-III (Stand bei Drucklegung)

Ein Blick auf den Trierer Prototyp *Das Goethe-Wörterbuch im Internet* veranschaulicht, inwieweit sich eine derartig gestaltete Oberfläche als

gelungene Synthese traditioneller und technischer Strukturelemente für die Wiedergabe eines historischen Autorenwörterbuches erweist (siehe Abbildung 3).¹¹

Die Beispiele veranschaulichen, in welchem hohem Maße Traditionen aus dem Buchdruck und der Typografie die Gestaltung digitaler Medien geprägt haben. So belegt das Beispielprojekt ‚Goethezeitportal‘ anschaulich, wie der Kolumnensatz des klassischen Zeitungs- und Zeitschriften-drucks das visuelle Informationsmanagement im Rahmen eines modernen Content-Management-Systems (CMS) dominieren kann.¹²

(2) Integrierte Kompaktanzeige externer Wörterbücher

In der expandierenden Welt der lexikografischen Informationssysteme haben sich neue Wörterbuchtypen etabliert. Bis auf Ausnahmen sind diese Online-Ressourcen auf korpuslinguistische Wissensextraktion spezialisiert. Sie lassen daher Kulturwissenschaftler/innen, Historiker/innen oder Philologen/innen oft eher ratlos zurück. Zudem erscheint ihr lexikografischer und wissenschaftlicher Anspruch im Einzelfall einer populären Breitenwirksamkeit untergeordnet.¹³ Mit Blick auf ihre technische Interaktivität und offene Nutzbarkeit bieten derartige Projekte allerdings durchaus interessante Anregungen. So können die Nutzer/innen auf der Website des Leipziger *Wortschatz-Lexikons* durch Kopieren des bereitgestellten HTML-Codes die externe Abfrage der zugrunde liegenden Korpusdatenbank in ihre lokale oder netzgestützte Anwendung integrieren. Dies geschieht über ein HTML-Formular (Applet), das auf ein entsprechend eingerichtetes Common Gateway Interface (CGI) des Leipziger Wortschatz-Servers zugreift. So ergibt sich per Skript-Übernahme nach dem Open-Source-Prinzip die Möglichkeit einer Wörterbuchabfrage, die voll-

11 Vgl. <<http://germazope.uni-trier.de/Projects/GWB>> (17.01.2006).

12 Vgl. <<http://www.goethezeitportal.de>> (Stand bei Drucklegung).

13 Vgl. den Pfad: Wortschatz-Downloads-Suchabfrage-Homepage unter <<http://wortschatz.uni-leipzig.de/>> (Stand bei Drucklegung). Eine Fülle sprechender Beispiele für die fachwissenschaftliche Nützlichkeit solcher Wörterbuchschnittstellen präsentiert die in der Germanistik einschlägige „Erlanger Liste“, vgl. <<http://www.erlangerliste.de/ressourc/liste.html>> (Stand bei Drucklegung).

ständig in das Erscheinungsbild der jeweiligen Benutzeroberfläche eingehen kann (siehe Abbildung 4).

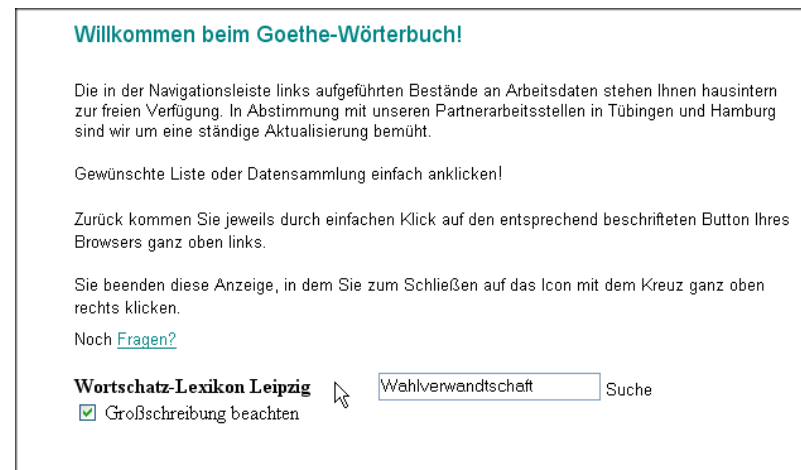


Abbildung 4: Abfrage Wortschatz-Lexikon Leipzig in Datensammlung *GWb intern*

Erst mit der Ergebnisanzeige wird die externe Wortschatz-Datenbank sichtbar (siehe Abbildung 5).

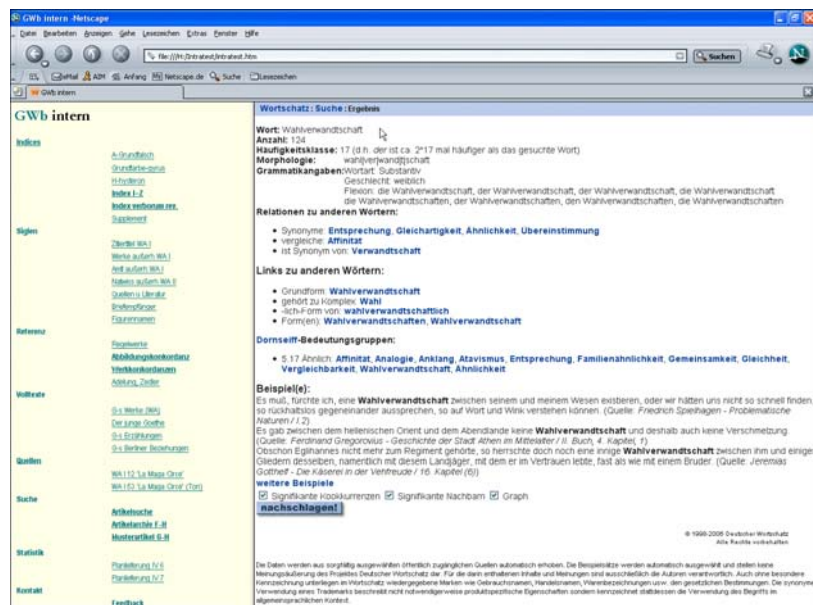


Abbildung 5: Ergebnis-Anzeige Wortschatz-Lexikon, Stichwort *Wahlverwandtschaft*

(3) Kontextsensitive Minimalanzeige lexikografischer Angaben

Die technischen Entwicklungen seit der Jahrtausendwende haben ihre immensen Beschleunigungstendenzen rigoros vom Hardware- auf den Software-Bereich übertragen. So verlangt das Wireless Access Protocol (WAP) der schnurlosen Telekommunikation nicht nur eine noch ökonomischere Rationalisierung der Datenpakete im Internet, sondern auch eine grafische und typografische Verkleinerung der semantischen Schnittstellen. Ein Index oder Thesaurus, der per Mobiltelefon abgefragt werden soll, stellt ganz neue Anforderungen an die Wörterbuchmacher/innen. Auch im zumeist auf größere Displays optimierten Bereich der Standard-Software für Personal Computer spiegeln integrierte Kleinst- oder Minimalwörterbücher diesen Trend. So sind die Wörterbuchfunktionen im Falle von MS Word in der Version 2003 zu kontextsensitiven Hilfenfenstern ausge-

baut. Mehrsprachige Thesauri, Synonymenverweise oder Übersetzungsvorschläge können so per Markierung oder rechter Maustaste eingeblendet werden (siehe Abbildung 6). Neuere Versionen des Internet-Explorers bieten Wörterbucheinträge sogar automatisch per Mouse-Over als Pop-Up bzw. Bubble-Help.

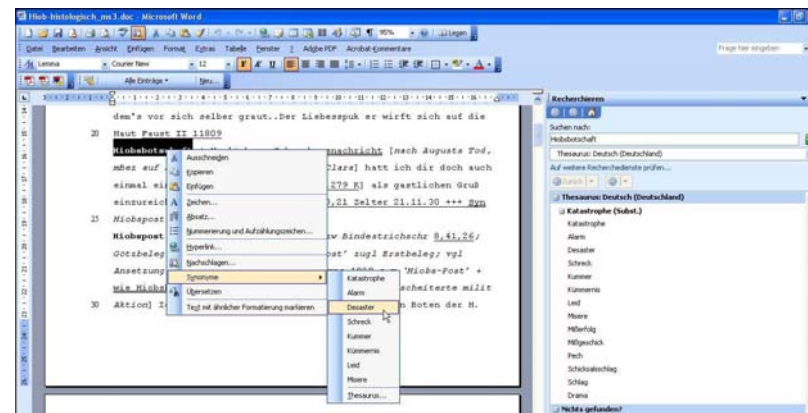


Abbildung 6: Wörterbuch-Hilfen in MS Word 2003, Beispielwort *Hiobsbotenschaft*

Das Wörterbuch der Zukunft: Visualisierung und Miniaturisierung

Folgt man der geschichtlichen Dynamik von *Lexikon* und *Wörterbuch* als Medium der Repräsentation und Anordnung von ‚Geschichte im Netz‘, so scheinen drei Tendenzen die Zukunft der Lexikografie im Internet und in den neuen Medien zu bestimmen: die Tendenzen zu *Vernetzung*, *Visualisierung* und *Miniaturisierung* lexikografischer Informationsstrukturen. Für die Wörterbuch*vernetzung*, das heißt die integrative Entwicklung mächtiger virtueller Wörterbuchplattformen, die ganze Bündel oder sogar Verbünde von historischen Fachwörterbüchern lemmabezogen miteinander verbinden, haben das Berlin-Heidelberger Wörterbuchportal¹⁴ mit dem retrodigitalisierten *Wörterbuch der Deutschen Gegenwartssprache* bzw.

14 Vgl. <<http://www.woerterbuch-portal.de>> (Stand bei Drucklegung).

dem digitalen *Wörterbuch der deutschen Sprache des 20. Jahrhunderts* (DWDS) sowie das Trierer Wörterbuchnetz¹⁵ mit dem *Deutschen* und dem *Goethe-Wörterbuch* wissenschaftliche und technische Maßstäbe gesetzt. Für Detailinformationen zu diesem Aspekt sei auf die entsprechenden Beiträge im Rahmen dieses Tagungsbandes verwiesen.

Abschließend bleibt zu resümieren, dass sich die Entwicklungen der Visualisierung und Miniaturisierung wiederum in der Geschichte bedeutender lexikografischer Werke im Geiste der europäischen Aufklärung antizipiert finden. So war es nicht zuletzt die Diderot/d'Alembertsche *Encyclopédie*, die mit ihren monumentalen, separat edierten Tafelbänden die Fülle der schönen neuen Alltags- und Technikwelt veranschaulichte und damit das abendländische Bildgedächtnis revolutionierte. Und hat Voltaire mit seinem *Dictionnaire philosophique, portatif* (1764) nicht in gewisser Weise die usuelle und ubiquitäre Mobilisierung des Wissens vorweggenommen, die mit der heutigen technischen Allverfügbarkeit von Informationen Realität geworden ist? Als virtuelle Informationssysteme – werk- und korpusübergreifend vernetzt, miniaturisiert und multimedial abrufbar, werden wissenschaftliche Wörterbücher in ihrer digitalen Derivation einmal mehr an ihre Ursprünge zurückkehren: als digitales *Encheiridion* (Handbüchlein) oder elektronisches „Allbuch“, als allgegenwärtiges und unverzichtbares informationelles Gebrauchsmedium der Zukunft.

Abbildungsnachweise

1. Bildzitat: Vogt, Friedrich; Koch, Max, *Geschichte der deutschen Literatur von den ältesten Zeiten bis zur Gegenwart*, 2. Auflage, Bd. 1 (Leipzig 1904), S. 29 – Wiedergabe einer Seite aus dem *Vocabularius Sancti Galli*. Nach dem Original in der Stiftsbibliothek zu St. Gallen; Hervorhebung R. C.
2. Bildzitat [Screenshot]: *Der Digitale Grimm. Deutsches Wörterbuch. Elektronische Ausgabe der Erstbearbeitung von Jacob Grimm und Wilhelm Grimm*, hg. vom Kompetenzzentrum für elektronische

15 Vgl. <<http://germazope.uni-trier.de/Projects/WBB/woerterbuecher/>> (Stand bei Drucklegung).

- Erschließungs- und Publikationsverfahren in den Geisteswissenschaften an der Universität Trier in Verbindung mit der Berlin-Brandenburgischen Akademie der Wissenschaften. 2 CD-ROMs sowie Begleit- und Benutzerhandbuch, Frankfurt am Main 2004, Ergebnisanzeige der Abfrage nach ‚Stichwort‘.
3. Bildzitat [Screenshot]: <<http://germazope.uni-trier.de/Projects/GWB/>> [Stand bei Drucklegung, September 2006].
 4. Bildzitat [Screenshot]: BBAW intern, lokale Datensammlung; vgl. dazu Demo-Ansicht unter <<http://www.goethe-wortschatz.de/>> | Link ‚Intern‘ (Passwort) [17.01.2006].
 5. Bildzitat [Screenshot]: Lokale Datensammlung „GWb intern“ [siehe zu Abbildung 4].
 6. Bildzitat [Screenshot]: Arbeitsbildschirm Microsoft Word für Windows XP Professional, Version 2003 [19.01.2006].

Dr. Robert Charlier ist Wissenschaftlicher Mitarbeiter der Berlin-Brandenburgischen Akademie der Wissenschaften (BBAW), Akademienvorhaben Goethe-Wörterbuch, und lehrt an der Technischen Universität Berlin sowie an der Universität Potsdam. E-Mail: charlier@bbaw.de

WÖRTERBÜCHER ALS DIGITALE RESSOURCEN FÜR MENSCH UND
MASCHINE – DIE WÖRTERBUCHPROJEKTE DER BERLIN-
BRANDENBURGISCHEN AKADEMIE DER WISSENSCHAFTEN

von Gerald Neumann

Der Vortrag berichtet über die Wörterbuchprojekte der Berlin-Brandenburgischen Akademie der Wissenschaften. Die Bearbeitung der akademischen Wörterbücher im Rahmen der Telota-Initiative wird beschrieben. Erläutert wird ein neuer, dynamischer Ansatz in der Aufbereitung von digitalen Wörterbuchressourcen. Dieser Ansatz wird an vier bearbeiteten Projekten vorgestellt.

Einführung

Der überwiegende Teil der hier vorgestellten digitalen Wörterbuchressourcen wurde im Rahmen der Telota¹-Initiative der Berlin-Brandenburgischen Akademie der Wissenschaften (BBAW) erstellt. Im Jahre 2002 ins Leben gerufen, soll Telota Instrumente entwickeln, mit denen die BBAW ihre Forschungsergebnisse digital erarbeitet, dokumentiert und präsentiert. Sie unterstützt und befördert die Standards der digitalen geisteswissenschaftlichen Arbeit an der Akademie.

Die aktuellen Aufgaben von Telota sind die Evaluation des digitalen Zustands der Forschungsvorhaben der BBAW, die Entwicklung von Konzepten und das Setzen von Zielen der Entwicklung digitaler Arbeitsmethoden sowie die Initiierung von Projekten zur Förderung der digitalen Forschungsarbeit. Das aktuelle Projekt (2005-2006) ist das „Telota-Projekt

des Monats“ (TPM). Es präsentiert jeden Monat ein neues, digitales Projekt aus den Akademienvorhaben. Es soll das Informationsangebot für die geisteswissenschaftliche Forschung bereichern und einer interessierten Öffentlichkeit Einblicke in die aktuelle Forschung geben. Das TPM arbeitet seit Februar 2005.

Die Schwerpunkte der Projekte liegen, induziert durch die an der Akademie angesiedelten Vorhaben, im Bereich der digitalen Editionen, Wörterbücher und Prosopografien. Darüber hinaus gibt es verschiedene andere Vorhaben, für die im Rahmen des TPM die unterschiedlichsten Einzelfalllösungen erarbeitet wurden, wie zum Beispiel einen georeferenzierten Zugang zu einer kunsthistorischen Datenbank per GoogleMaps (Census), die wechselseitige Verknüpfung von Datenbank und Scan-Galerie (Ritter-Katalog der Leibnizforschung) oder die Entwicklung des Archiv-Editors für das Vorhaben *Preußen als Kulturstaat*, der aber allen „Archivgängern“ nützlich sein wird.² Die bisher vorliegenden Ergebnisse vom TPM lassen sich nur unter den zeitlichen und personellen Bedingungen angemessen einschätzen, unter denen sie entstanden sind. Der Koordinator, die beiden Mitarbeiter und die zwei Studenten sind jeweils mit einer halben Stelle für die Aufgaben vom TPM angestellt. Das Projekt trägt den Namen „Projekt des Monats“ mit voller Berechtigung, da es de facto nur jeweils einen Monat für die Arbeit in einem Vorhaben zur Verfügung hatte. Da die Monate zu Beginn der Arbeit von TPM den Vorhaben zugeordnet wurden, gab es auch keine Möglichkeit, an diesem Zustand etwas zu ändern. Zu Beginn jedes neuen Monats stand das aktuelle Vorhaben mit seinen Forderungen und Aufgaben schon bereit. Um nicht eine Vielzahl von zusammenhangslosen Einzellösungen zu produzieren, wurde von Beginn an versucht, eine einheitliche technische Grundlage zu schaffen und weiterzuentwickeln, die mit möglichst geringem technischen Aufwand die spezifischen Anpassungsaufgaben, die sich aus den Einzelprojekten ergaben, ermöglichen sollte. Wie diese technische Grundlage für die

¹ Telota = The electronic life of the academy.

² Eine vollständige Liste aller Projekte findet sich unter <<http://www.bbaw.de/pom/pom.html>> (09.11.2006). Zum Projekt *Preußen als Kulturstaat* vgl. Czmiel, Alexander; Holtz, Bärbel, Quellenarbeit im Projekt „Preußen als Kulturstaat“, in diesem Band.

Wörterbuchprojekte aussieht und welche Anwendungen sie ermöglicht, soll im Folgenden vorgestellt werden.

Überblick

Akademische Wörterbücher in eine angemessene digitale Form zu verwandeln, zählt zu den derzeit größten Herausforderungen der elektronischen Textaufbereitung. Neben historisch-kritischen Editionen sind Wörterbücher aus elektronischer Sicht diejenige Textsorte mit der höchsten Informationsdichte. Die adäquate Darstellung der über Generationen entstandenen Informationskodierung in einem Wörterbuchartikel und die gleichzeitige Umsetzung der neuen digitalen Möglichkeiten verlangen eine sorgfältige Analyse des gedruckt Bestehenden und eine umfassende Kenntnis des digital Möglichen. Dieser Aufgabe hat sich das TPM in den vergangenen Monaten verstärkt gewidmet. Nachdem schon im April 2005 das *Wörterbuch der deutschen Gegenwartssprache* (WDG) bearbeitet wurde, waren im Juni und August 2005 das *Deutsche Wörterbuch* (DWB) und das *Goethe-Wörterbuch* (GWB) an der Reihe. Von den beiden Letzteren wurde jeweils die aktuelle Lieferung digital aufbereitet und im Internet präsentiert. Die Einbeziehung dieser drei Projekte folgt dem Plan, schrittweise ein experimentelles, digitales System für die akademischen Wörterbücher der BBAW aufzubauen. Dieses System soll nicht nur dem unterschiedlichen Aufbau des jeweiligen Wörterbuches gerecht werden, sondern auch konsequent die Vorteile einer strukturierten digitalen Aufarbeitung deutlich machen.

Die Aufnahme der drei Wörterbücher in das Wörterbuch-Portal³ der Akademie verdeutlicht darüber hinaus die Vernetzungsmöglichkeiten, die sich aus einer strukturierten Aufarbeitung der Wörterbücher ergeben.

Alle vorgestellten digitalen Wörterbuch-Projekte basieren auf der gleichen Technologie. Wie inzwischen üblich, lagen die Wörterbücher als XML-Dokumente vor, bzw. wurden in eine XML-Struktur überführt. Die Vorteile einer solchen strukturierten Erfassung von lexikografischen Daten

³ Das Wörterbuch-Portal ist unter <http://www.woerterbuch-portal.de> (22.11.2006) zu erreichen.

hinsichtlich Langzeitarchivierbarkeit und Softwareunabhängigkeit liegen auf der Hand. Das vom TPM entwickelte System geht aber noch einen Schritt weiter. Auf der Grundlage der nativen XML-Datenbank eXist⁴ wurde die Möglichkeit geschaffen, auf die zugrunde liegenden XML-Strukturen während des Recherche- und Anzeigeprozesses zuzugreifen. Dies ist dadurch möglich, dass das für die Anzeige benötigte Format, im Normalfall sind dies HTML-Dokumente, erst auf eine Nutzeranfrage hin direkt aus den XML-Dokumenten generiert wird. Die XML-Dokumente werden unter Benutzung des W3C-Standards XQuery⁵ abgefragt. Das Ergebnis dieser Abfrage, das ein XML-Fragment ist, wird dann unter Benutzung des W3C-Standards XSLT⁶ in ein HTML-Dokument transformiert und an den anfragenden Web-Browser ausgeliefert. Dieser bringt es dann den Benutzern/innen zur Ansicht.

Hier kann sich natürlich die Frage stellen, warum so kompliziert, wenn es auch einfacher geht? Auf diese Frage will der Vortrag zwei Antworten geben. Erstens, das beschriebene Verfahren ist nicht so kompliziert, wie es auf den ersten Blick scheint. Zweitens, das gewählte Verfahren eröffnet Möglichkeiten, die anders nur schwer und mit erheblichem Mehraufwand oder aber gar nicht zu erreichen sind.

Die Aufbereitung von XML-Daten zur Anzeige für menschliche Benutzer/innen ist ein Vorgang, der nicht nur dann anfällt, wenn die Anzeige, wie in unserm Fall, direkt aus den XML-Dokumenten generiert wird. Das heißt, die Konvertierung in ein lesbares Format, wie zum Beispiel HTML oder PDF findet bei jedem elektronisch verfügbaren Wörterbuch statt, sei es nun in einer relationalen Datenbank gespeichert oder als statisches HTML angelegt. Die Generierung des Anzeigeformats zur Laufzeit erzeugt also keinen zusätzlichen Arbeitsaufwand, sondern hat allenfalls positive Auswirkungen auf die direkten Korrekturmöglichkeiten der Quellen und

⁴ Der Terminus „native XML-Datenbank“ bezeichnet die besondere Eigenschaft einer Datenbank, spezielle technische Voraussetzung für die genuine Verarbeitung von XML-Dokumenten zu besitzen. Vgl. <http://www.exist-db.org> (22.11.2006).

⁵ Vgl. <http://www.w3.org/XML/Query/> (22.11.2006).

⁶ Extensible Style Language Transformation, vgl. <http://www.w3.org/TR/xslt20/> (22.11.2006).

eventuell negative Auswirkungen auf die Performance der Anzeige. Die Trennung von Daten-, Anfrage- und Anzeigeschicht bringt zwar einen erhöhten Implementierungsaufwand mit sich, lässt dafür aber im Gegenzug auch den gezielten Eingriff in nur einen Bestandteil des Gesamtsystems zu. So wird es möglich, die XSL-Transformation der HTML-Seiten zu ersetzen durch eine XSL-FO-Transformation, die aus den Suchergebnissen ein PDF-Dokument erzeugt, ohne dass an der Daten- oder Anfrageschicht etwas geändert werden muss. Dadurch wird es für Nutzer/innen möglich, Wörterbuchauszüge bzw. ganze Wörterbücher auf der Grundlage von speziellen Nutzerinteressen als PDF-Dokumente zu generieren. Ein auf der .hist 2006 vorgestelltes Beispiel ist die Generierung eines suffix-geleiteten Teilwörterbuchs auf der Basis des GWB für die Suffixe *-heit* und *-keit*.

Auch an den anderen Schichten lässt sich ansetzen. So kann der Output der Anfrageschicht, ein XML-Fragment, auch direkt an einen Nutzer/eine Nutzerin weitergereicht werden, ohne vorherige Transformation in ein Anzeigeformat. Dadurch entsteht eine Schnittstelle für eine mögliche maschinelle Nutzung der Wörterbuchressource. Die Anfrageschicht lässt über XQuery zudem die Formatierung des ausgegebenen XML-Fragmentes zu. Das heißt, die weiterverarbeitende Maschine kann ihr XML-Fragment in genau der Form bekommen, in der sie es braucht. Die dafür notwendigen Arbeitsschritte können und sollen von den Nachnutzern/innen des Wörterbuchs ausgeführt werden und nicht von denjenigen, die es bereitstellen.

Aber das gewählte technische Verfahren eröffnet nicht nur interessante technische Möglichkeiten. Auch die inhaltliche Arbeit und Nutzung des Wörterbuchs gewinnt durch die Teilung des Verfahrens in die unterschiedlichen Schichten. So ermöglicht der Nutzerzugriff auf die Anfrageschicht das Erstellen von Anfragen, die nur durch die vorhandenen Strukturen der XML-Dokumente beschränkt sind und nicht durch eine vorgegebene Suchmaske mit zwangsläufig vorgegebenen Recherchemöglichkeiten. Als ein sprechendes Beispiel sei hier die Neubearbeitung des *Deutschen Wörterbuchs* von Jacob Grimm und Wilhelm Grimm beschrieben. Die Tatsache, dass in diesem Wörterbuch jeder Beleg eine strukturierte zeitliche Einordnung besitzt und dass die Belege innerhalb der Artikelstruktur

eindeutig der Bedeutungsvariante eines Stichworts zugeordnet sind, lässt eine Wörterbuchbenutzung zu, wie sie bisher nicht realisiert wurde. So ist es möglich, durch eine XQuery-Anfrage alle Stichwörter des Wörterbuchs zu bekommen, die über eine Bedeutungsvariante verfügen, deren erster Beleg nicht vor dem Jahre 1832 liegt. Ebenso ist es möglich, alle die Stichwörter auszugeben, die nach 1832 nicht mehr belegt sind.

Wozu ist das gut? Mit diesen beiden Anfragen lässt sich eine Liste von Stichwörtern zusammenstellen, die die Lektüre historischer Texte, in diesem Fall von Goethetexten, unterstützt. Die zweite Anfrage generiert offensichtlich eine Liste von Wörtern, die heute nicht mehr benutzt werden. Diese Wörter würden natürlich bei der Lektüre eines Textes ohnehin auffallen, so dass mit der vorangehenden Textaufbereitung nur der Nachschlageprozess überflüssig wird. Die erste Anfrage hingegen listet Wörter auf, die aus der heutigen Perspektive eine andere, zusätzliche Bedeutungsvariante haben als zur Goethezeit. Das heißt, die Möglichkeit einer Fehlinterpretation der Wortbedeutung ist gegeben. Wollte man die Anforderungen an eine digitale Wörterbuchressource auf die Spitze treiben, so könnte man von ihr verlangen, dass sich das Wörterbuch von selbst zu Wort meldet, wenn eine solche Möglichkeit der Fehlinterpretation beim Lesen historischer (elektronischer) Texte besteht.

Die Trennung der drei Aufbereitungsschichten ermöglicht also den getrennten, dynamischen Zugriff und auch eine schichtenabhängige Nutzerinteraktion mit dem System. Dies erweitert die Möglichkeiten der menschlichen und der maschinellen Benutzung der gesamten Wörterbuchressource beträchtlich. Außerdem enthebt es den Betreiber eines elektronisch verfügbaren Wörterbuchs von der unlösbaren Aufgabe, allen Nutzerinteressen zu genügen, bzw. selbst festlegen zu müssen, welche Nutzungsmöglichkeiten die ‚richtigen‘ sind. Was nicht heißen soll, dass er für die Nutzbarkeit seiner Ressource keine Verantwortung trägt. Im Gegenteil, mehr als bisher sollten die verschiedenen Nutzungsmöglichkeiten einer digitalen Wörterbuchressource Kriterium für die qualitative Bewertung ihrer elektronischen Präsentation sein.

Die auf dem Vortrag vorgestellten digitalen Wörterbuchressourcen sind folgende:

Wörterbuch der deutschen Gegenwartssprache (WDG)

Beim ersten Wörterbuchprojekt handelte es sich um die Erweiterung der Abfragemöglichkeiten des *Wörterbuchs der deutschen Gegenwartssprache* (WDG)⁷, dessen Online-Ausgabe seit März 2003 vom Projekt „Digitales Wörterbuch der deutschen Sprache“ bereitgestellt wird. Die Aufgabe von TPM bestand in der so genannten ‚Auszeichnung‘ von besonderen inhaltlichen Strukturen in den Wörterbuchartikeln. Es wurden sämtliche Belegstellen und Siglen ausgezeichnet und mit dem speziell dazu digitalisierten Quellenverzeichnis von ungefähr 1.000 Werken verknüpft. Dadurch ergaben sich völlig neue Funktionalitäten in der Wörterbuchabfrage. So ist es jetzt nicht nur möglich von den einzelnen Siglen an die zutreffende Stelle im Quellenverzeichnis zu springen, sondern auch vom Quellenverzeichnis zu allen Stichwörtern, die sich aus einer bestimmten Quelle speisen.

Deutsches Wörterbuch von Jacob Grimm und Wilhelm Grimm (DWB)

Beim *Deutschen Wörterbuch*⁸ wurde erstmals eine Schnittstelle für den freien Zugriff auf die Daten nach einem vom W3C verabschiedeten Standard XQuery eingefügt, womit das Wörterbuchsystem über die Begriffsabfrage hinaus erweitert wird. Die Nutzer/innen können jede mögliche Anfrage stellen, die sich aus der Strukturierung der Daten ergibt, niemand gibt ihnen mehr vor, welche Fragen sie haben dürfen. Ein Beispiel für eine solche freie Anfrage ist etwa: Welches sind die am häufigsten zitierten Autoren in der Lieferung des DWB? Als Antwort erscheinen die ersten sechs Autoren des Rankings: Luther (53), Schiller (31), Franck (30), Goethe (29), Lessing (28), Herder (27), wobei die Zahlen in den Klammern die Häufigkeit angeben. Diese zusätzlichen Funktionen machen das Wörterbuchsystem zu einer echten digitalen Ressource, die in ihrer Komplexität nicht nur für die Mitarbeiter/innen bereitsteht, sondern auch von jedem/jeder anderen genutzt werden kann.

Darüber hinaus wurde beim DWB ein halb-automatisches Verfahren entwickelt, das die originalen Tustep-Daten in XML-Daten umwandelt.

7 Vgl. <<http://www.dwds.de>> (22.11.2006).

8 Vgl. <<http://pom.bbaw.de/dwb>> (22.11.2006).

Goethe-Wörterbuch (GWB)

Auch die Aufarbeitung der Lieferung des *Goethe-Wörterbuchs*⁹ passt sich nahtlos in das Wörterbuchsystem ein. Die strukturierten Daten werden direkt aus dem lexikografischen Workflow gewonnen. Auch die XQuery-Schnittstelle des Wörterbuchsystems kann genutzt werden.

In der Digitalversion des GWB neu hinzugekommen ist die Möglichkeit der Online-Bearbeitung von XML-Dokumenten. Damit ist der Anfang eines Online-Artikel-Redaktionssystems gemacht. Bisher gestattet das Redaktionssystem das Anlegen neuer, strukturierter Artikel und das Editieren bestehender Artikel auf der Basis und in den Grenzen der vorgegebenen XML-Strukturen.

Als eine weitere Erweiterung gibt es beim GWB die Option der PDF-Ausgabe nach selbst bestimmbareren inhaltlichen und gestalterischen Merkmalen. Dabei werden die PDF-Dokumente unter Benutzung des XML-FO-Standards direkt aus den XML-Dokumenten generiert.

Zusatz: Wörterbuch der manichäischen Texte (WMT)

Eine weitere Möglichkeit des Wörterbuchsystems wurde durch das „Telota – Projekt des Monats Januar 2006“ eindrucksvoll demonstriert. Durch die strukturierte Aufbereitung und Bereitstellung des *Wörterbuchs der manichäischen Texte*¹⁰ wurde eine automatische Verknüpfung von digitaler Edition und digitalem Wörterbuch möglich. Für jede Wortform aus den digital vorliegenden mittelpersischen Texten konnte so eine Verknüpfung zum entsprechenden Lemma im Wörterbuch ohne großen Aufwand realisiert werden. Am digitalen WMT zeigen sich die bisher weitgehend ungenutzten Perspektiven des gleichermaßen maschinen- und menschenlesbaren Formats XML/TEI.

Schlussfolgerungen und Konsequenzen

Die vier oben beschriebenen, konkreten Projekte haben einiges über die praktische Seite der Implementierung eines digitalen Wörterbuchsystems gezeigt. Die Erfahrungen seien hier kurz zusammengefasst.

9 Vgl. <<http://pom.bbaw.de/gwb>> (22.11.2006).

10 Vgl. <<http://pom.bbaw.de/turfan>> (22.11.2006).

Die verfügbaren Datenbanken, Webserver, andere Software und Hilfsmittel aus dem Open-Source-Bereich erfüllen die notwendigen Voraussetzungen an Funktionalität und Performance überraschend gut. Die Standards für Datenauszeichnung (XML/TEI), Datenkonvertierung (XSL, FO), und Datenabfrage (XQuery) sind für grundlegende Anliegen hinreichend weit entwickelt. Die Unicode-Unterstützung in den verwendeten Applikationen deckt fast alle Bedürfnisse der Zeichendarstellung und Suchfunktionalität ab. Die Datenmodellierung der XML-Dokumente und die konsequente Nutzung von Open Source-Technologien sichert ein hohes Maß an Unabhängigkeit und Langzeitverfügbarkeit der erschlossenen Quellen.

Die XML-Strukturierung der Daten ist nur der erste wichtige Schritt. Die XML-Strukturierungen müssen auch für Recherche und Weiterverarbeitung zur Verfügung stehen. Elektronische Ressourcen von Wörterbüchern und anderen komplexen Texten, wie zum Beispiel Digitale Editionen, müssen maschinenlesbar und –auswertbar sein. Die Anwendung von internationalen Standards ist Voraussetzung für eine Vernetzung von Informations- und Wissensbeständen.

Gerald Neumann ist Mitarbeiter beim Akademievorhaben „Digitales Wörterbuch der Deutschen Sprache“ und Koordinator des „Telota – Projekt des Monats“ an der Berlin-Brandenburgischen Akademie der Wissenschaften. Wichtigste Arbeitsgebiete: Digitale Wörterbücher, Strukturierte Erfassung der Ergebnisse geisteswissenschaftlicher Grundlagenforschung. E-Mail: gneumann@bbaw.de

BAYERISCHES MUSIKER-LEXIKON ONLINE

von Josef Focht

Das Projekt der Deutschen Forschungsgemeinschaft (DFG) Bayerisches Musiker-Lexikon Online (BMLO) wird als virtuelles biografisches Lexikon zur bayerischen Musikgeschichte erarbeitet. Es soll eine auf Biografien gestützte grundlegende Dokumentation der bayerischen Musikgeschichte in zeitgemäßer und allgemein zugänglicher Form bereitstellen. Derzeit (Februar 2006) informiert es über seine Strukturen und den Fortgang seiner Vorbereitung und Umsetzung. Zukünftig wird es als Lexikon Benutzern/innen aller Fächer und Bildungshorizonte unbeschränkt und kostenfrei zugänglich sein. Der Editionsbeginn des BMLO ist für das zweite Quartal 2006 vorgesehen; bis 2008 soll es bis zur vollständigen Editionsreife geführt werden, sodass dann Datensätze von deutlich über 20.000 Personen medial angereichert abgerufen werden können. Der folgende Aufsatz gibt einen Überblick über die fächerübergreifende Struktur des Projekts und beschreibt Verfahren zur Umsetzung des Lexikons sowie Prioritäten in der Auswahl der Quellen.

Einleitung

Der Bedarf an wissenschaftlich qualifizierten und kontrollierten Informationsangeboten im Internet ist in den vergangenen Jahren für alle Fächer sprunghaft angestiegen. Deshalb wurden in jüngster Zeit in verschiedenen Wissenschaftsbereichen digitale Lexika konzipiert und realisiert. Das Bayerische Musiker-Lexikon Online (BMLO)¹ will für die historische Musikwissenschaft ein vergleichbares Pilotprojekt entwickeln. Es soll eine auf Biografien gestützte grundlegende Dokumentation der bayerischen

1 Vgl. <<http://www.bml0.uni-muenchen.de>>.

Musikgeschichte in zeitgemäßer und allgemein zugänglicher Form bereitstellen, die modular aufgebaut, vertiefungs- und erweiterungsfähig ist.

Das BMLO wird derzeit (März 2006) als virtuelles Lexikon vorbereitet; es informiert über seine Strukturen und den Fortgang seiner Bearbeitung. Zukünftig wird es Benutzern/innen als biografisches Lexikon zugänglich sein.

Trägerschaft, Finanzierung und Förderung

Das BMLO ist ein Projekt zur Erfassung, Digitalisierung, Erschließung und Bereitstellung biografischer Daten und personenbezogener Medien. Es wird seit Anfang 2004 vom Institut für Musikwissenschaft an der Universität München in Kooperation mit der Bayerischen Staatsbibliothek (BSB) und der Gesellschaft für Bayerische Musikgeschichte durchgeführt. Als Partner im wissenschaftlichen Austausch fungieren in- und ausländische Universitäten, Akademien, Forschungseinrichtungen und Lexikonprojekte sowie einzelne Autoren/innen. Bei der Durchführung des Projekts stehen neben den Eigenmitteln der Träger – also der Universität München, der BSB und der Gesellschaft für Bayerische Musikgeschichte – auch Drittmittel der Edith-Haberland-Wagner-Stiftung und der DFG zur Verfügung.

Die BMLO-Struktur im Überblick

Das BMLO ist als Online-Lexikon zu Personen der bayerischen Musikgeschichte (Komponisten/innen und Dirigenten/innen, Sänger/innen und Instrumentalisten/innen, Musikverleger/innen, Instrumentenbauern/innen und Musikforscher/innen aus allen Epochen) angelegt. Seine Daten- und Editionsstruktur bietet eine recherchefähige Datenbank, darüber hinaus Lexikonartikel und multimediale Ergänzungen.

Infolge seiner interdisziplinären Verortung – zwischen der Musikwissenschaft und der Landesgeschichte und mit wesentlichen Beiträgen des Bibliothekswesens sowie der Informatik – erfordert das BMLO ebenso dringend eine bedarfsgerechte und aufwändige Methoden-Entwicklung wie infolge seines Pilotprojekt-Charakters und seines großen Volumens.

Das BMLO führt die biografische Forschungsarbeiten aus der musikwissenschaftlichen und landeskundlichen Literatur zusammen, retrokon-

vertiert lexikalische Quellen, erschließt diese sowie weitere Bibliotheks- und Archivbestände; schließlich vernetzt es hochwertige biografische Internet-Ressourcen. Die Datenfülle des BMLO zielt mit über 20.000 Lemmata auf eine Dokumentationsbreite ab, die den Themenkomplex weiträumig abdeckt; ferner bemüht es sich um den Nachweis relevanter Quellen in möglichst großer Erschließungstiefe.

Warum gerade Bayern?

Aus methodischen und pragmatischen Gründen ist das BMLO zunächst auf Bayern beschränkt:

Die Lexikografie zur bayerischen Musikgeschichte setzte schon 1811 mit dem „Baierischen Musiklexikon“ von Felix Joseph von Lipowsky ein. Dies war gleichzeitig die erste auf einen „Nationalstaat“ konzentrierte Sammlung von Musikerbiografien, die im 19. und 20. Jahrhundert dann lebhaft nachgeahmt wurde.

Im 20. Jahrhundert fanden viele Biografien bayerischer Musiker/innen Eingang in allgemeine musikalische Enzyklopädien², ihr überwiegender Anteil wurde jedoch in Monografien und Einzelstudien (etwa über einzelne Städte, Klöster, Hofkapellen oder den Instrumentenbau) bzw. in Periodika veröffentlicht – sie sind also weit verstreut und oft schwer auffindbar. Die relevanten Periodika sind auf die Fächer Musikwissenschaft und Landesgeschichte verteilt sowie in dem bibliografisch nur unvollständig erfassten Bereich der Grauen Literatur zu suchen.³

Bayern ist wie kaum ein anderes Bundesland seit zwei Jahrhunderten in seinen Grenzen stabil und klar definiert, was günstige Konsequenzen namentlich für das Bibliotheks- und Archivwesen hat. Es ist groß genug und soziologisch wie konfessionell differenziert genug, um innerhalb seiner Grenzen eine große Vielfalt musikhistorischer Phänomene exemplarisch abzubilden: katholisch-klösterliche ebenso wie protestantisch-reichsstädtische Musikkultur, höfisches Musikleben an kleinen, weltlichen

2 Unter anderem Die Musik in Geschichte und Gegenwart, (New) Grove Dictionary of Music and Musicians, Riemann Musiklexikon usw.

3 Wie z. B. Programmhefte, Chroniken und Festschriften von Verbänden, Vereinen, Theatern, Orchestern, Instrumentenbau-Firmen.

und geistlichen Residenzen wie an großen von europäischem Rang, ländliche ebenso wie großstädtische Popularkultur.

Bayern ist eine europäische Kernregion, die nach allen Seiten hin offen und durchlässig für Einflüsse aus allen Nachbarregionen ist. Es ist ein paradigmatisches Durchgangsland, in dem sich die wichtigsten europäischen Migrationsrouten kreuzen, und ringförmig umgeben von Musikmetropolen wie Salzburg, Wien, Prag, Dresden, Mannheim, Paris, Mailand und Venedig – was viele Musiker/innen durch Bayern führen musste. Bayern stand stets in intensiver Wechselwirkung mit musikhistorisch führenden oder besonders innovativen Nachbarregionen wie Italien, Österreich, Thüringen/Sachsen oder Böhmen.

Die BSB stellt nicht nur wegen ihres Sondersammelgebiets Musikwissenschaft einen unschätzbaren Standortvorteil dar. Hinzu kommt in Bayern eine beispiellose Dichte von historisch oder musikwissenschaftlich ausgerichteten Institutionen und Vereinigungen, vom engmaschigen Netz der Universitäten bis zu den circa 200 historischen Vereinen, den Komponisten-Instituten, Museen, Archiven und Bibliotheken. Im kultur- und musikhistorischen Bereich verfügt Bayern für ein Projekt wie das BMLO wegen der bereits seit zwei Jahrhunderten betriebenen landeskundlichen Forschung also über günstige Ausgangsbedingungen.

Auch wenn das BMLO aus pragmatischen Gründen nur „bayerische“ Musiker/innen berücksichtigt, also diejenigen, die historisch nennenswert in Bayern gewirkt haben und/oder genealogisch in Bayern verwurzelt sind, deckt es damit doch – wie oben erwähnt – einen Personenbestand ab, der vielfach auch international tätig war. Stichprobenartige Vergleiche mit Nachbarregionen (wie zum Beispiel Tschechien oder Baden-Württemberg) haben bereits Überschneidungen der relevanten Personenkreise von etwa 40 Prozent ergeben. Die enge Kooperation mit Wissenschaftlern/innen und Projekten aus allen Nachbarländern wird diesen bereits in der Sache selbst angelegten überregionalen und internationalen Charakter deutlich herausarbeiten.

Nach Abschluss der (ersten) „bavaristischen“ Phase ist das BMLO strukturell offen für Ausweitungen in alle Richtungen und Kooperationen jeder Art. Es versteht sich dann als Grundstock eines überregionalen

europäischen Musiker-Lexikons.

Interdisziplinäre Struktur

Gemäß der bereits beschriebenen fächerübergreifenden Struktur wird das BMLO in mehreren spezialisierten und differenziert profilierten Arbeitsgruppen vorbereitet: in der BMLO-Redaktion am Institut für Musikwissenschaft der Universität München (fachwissenschaftliche Redaktion), in der IT-Gruppe Geisteswissenschaften der Universität München (EDV- und Datenbankpflege), in der Digitalen Bibliothek der BSB (Retrodigitalisierung, Abgleich mit bibliothekarischen Normdaten, Langzeitarchivierung) sowie von Kooperationspartnern (musikwissenschaftliche Forschung), die auf relevante Themenbereiche spezialisiert sind.

Prioritäten in der Bearbeitung

Die Lexikondaten des BMLO müssen aus Gründen der Quantität und der Methoden in mehreren Phasen bearbeitet und ediert werden. Die höchste Priorität genießt die Entwicklung der Datenbank mit ihrer differenzierten Recherche-Funktionalität. In der nachfolgenden Ausbaustufe werden verschiedenartige Lexikonartikel hinzugefügt; sie entstehen wahlweise durch Erschließung bereits publizierter Beiträge und/oder durch Neuformulierung. In der abschließenden Ausbaustufe werden Medien hinzugefügt: Bilder, Noten, Audios, Videos.

Bearbeitungsphasen im Detail

In der Vorbereitung des BMLO sind mehrere Arbeitsphasen in der folgenden Reihenfolge nötig, die sich jedoch in beachtlichem Maß überlappen:

- Definition von Strukturen und Normen (für die Dateneingabe und -ausgabe, Datenmigration und -konversion, Redaktion und Erschließung, die Langzeitarchivierung und Verwaltung)
- Entwurf einer Namensliste, Verdichtung zur Lemmaliste, die als Grundlage personenbezogener Datensätze dient
- Quellenerschließung und Datensammlung

- Kumulation⁴ bzw. Individualisierung⁵ der Datensätze, das heißt: Zusammenführung verschiedener Namen zu einer Person bzw. bedarfsgerechte Trennung eines Namens in mehrere Personen
- Entwicklung und Optimierung von Methoden vom Versuchsstadium bis zur Serienreife
- Abgleich mit bibliothekarischen und amtlichen Normdateien (Personennamendatei und Normansetzung von Namen, Schlagwortnormdatei, Ortskennzahlen usw.)⁶
- redaktionelle Vervollständigung der Datensätze
- Transformation der plattformabhängigen Eingabedaten und Edition in spezifischen Ausgabe-Oberflächen⁷, die dem virtuellen Medium angemessen sind und den divergierenden Anforderungen des breiten Nutzerkreises gerecht werden
- Erschließung (Retrokonversion)⁸ und Bearbeitung von Lexikonartikeln

4 Ein Beispiel: Der Oberpfälzer Komponist Johann Georg Holler (* 1751 – † nach 1823) trat 1775 als Benediktiner in das Kloster Attl am Inn ein, wo er später Prior wurde. Dort führte er den Namen Pater Aegidius. Ab 1787 hielt er sich längere Zeit in Siena auf, wo seine Werke mit der Autorenangabe Diego Orelli überliefert sind. Die Zusammenführung der Namensformen „Johann Georg Holler“, „Pater Aegidius“ und „Diego Orelli“ in einen Datensatz ist als Kumulation zu verstehen.

5 Ein Beispiel: Gegenwärtig enthält das BMLO etwa 70 Datensätze zu Personen mit dem Nachnamen „Fischer“. Bei einigen liegen sehr wenige biografische Daten vor, gleichzeitig erstrecken sich ihre Wirkungsdaten auf so große Zeiträume, dass sie sich wohl auf verschiedene gleichnamige Personen – etwa Vater und Sohn – beziehen. Diese Datensätze bedürfen dann einer Individualisierung.

6 Die bibliothekarischen Normdateien werden von der Deutschen Bibliothek Frankfurt am Main, vgl. <<http://www.ddb.de>>, als zentrale bibliografische Dienstleistungen betreut. Die Ortsdaten werden von den Statistischen Ämtern der Länder, des Bundes und der Europäischen Gemeinschaft geführt, vgl. <<http://www.statistikportal.de>>.

7 Aus methodischen Gründen hat sich die Trennung von Dateneingabe, -transformation und -ausgabe als zweckmäßig erwiesen. Die Pflege des sehr großen BMLO-Datenbestandes erfordert damit bei geringer Fehlerdichte eine vergleichsweise kurze Bearbeitungszeit. Voraussetzung für diese Trennung der Datenebenen ist eine höchstmögliche Konsistenz in den Ansetzungsformen.

8 Zwei Bestandsgruppen der gedruckten biografischen Literatur kommen für die Retrokonversion in Betracht: erstens überregionale Lexika mit einem nennenswerten Anteil für das BMLO relevanter Personen, zweitens Monografien zu Themen der bayerischen Musikgeschichte mit einem hohen biografischen Informationsgehalt.

- Erschließung weiterer Medien durch Digitalisierung
- fortlaufende fachliche, formale und EDV-technische Kontrolle.

Die Qualität lexikalischer Daten hängt wesentlich vom Grad der Kumulation bzw. Individualisierung der Datensätze ab. Im Gegensatz zu einer Vielzahl von anderen gegenwärtigen Retrokonversionsprojekten leistet die BMLO-Redaktion die unverzichtbare fachwissenschaftliche Beurteilung und Bewertung der höchst inhomogenen Quelldaten, womit die Qualität der BMLO-Daten wesentlich steigt und ihre Benutzer/innen von deren fachlicher Kontrolle entlastet werden, die sie in der Regel weniger leicht leisten können als die BMLO-Redaktion.⁹

Auswahl und Priorität der Quellen

Um gleichzeitig die große Anzahl der Lemmata möglichst effektiv bearbeiten und trotzdem dem Einzelfall gerecht werden zu können, verfolgt das BMLO zwei parallele Strategien bei der Auswahl und Erschließung von Quellen.

- Primäre Strategie der seriellen Erschließung: In der BMLO-Redaktion werden serielle Quellen erschlossen, das heißt Lexika, Monografien, Archivbestände, Internet-Ressourcen vor allem der Musikforschung und der Landesgeschichte.
- In der Digitalen Bibliothek der BSB wird die Retrokonversion historischer musikwissenschaftlicher Lexika und bavaristischer Standardliteratur vorgenommen.
- Sekundäre Strategie der individuellen Vertiefung: Der Ertrag aus seriellen Quellen wird durch individuelle Vertiefung ergänzt, die jedoch von der BMLO-Redaktion allein aus Gründen der Ökonomie nicht zu leisten ist. Sie geschieht in der Regel in Kooperation mit Partner-Institutionen und spezialisierten Autoren/innen.

9 Die Datenkontrolle innerhalb der BMLO-Redaktion geht weit über die Beurteilung musikhistorischer Sachverhalte und die Diskussion von Forschungskontroversen hinaus; sie korreliert etwa auch historische Ortsnamen oder ergänzt Daten, die nicht expressis verbis zu lesen, aber doch aus dem Kontext zu erschließen sind.

BMLO-Datenmodell

Die BMLO-Daten sind plattformunabhängig strukturiert, womit ihre Langlebigkeit besonders wahrscheinlich wird. Die Datensätze sind beliebig aktualisierbar, erweiterungs- und ergänzungsfähig.

Das BMLO-Datenmodell weist in der Eingabemaske insgesamt 48 Datenfelder auf, die qualitativ zu unterscheiden sind. Sie sind in folgende Gruppen gebündelt:

- Gerüstdaten: Namen und Varianten (2 Felder), Lebensdaten (6), Zugehörigkeiten (6), Wirkungsdaten (6)
- Nachweise (6)
- Werke und Verzeichnisse (5)
- BMLO-Lexikonartikel (2)
- Bild- und Audiovisuelle Medien (7)
- Redaktionelle Vermerke (8).

Interdisziplinäre Edition und breite Rezeption

Das BMLO richtet sich an Interessierte verschiedener Fachrichtungen und wird deshalb an verschiedenen Stellen der Öffentlichkeit kostenfrei zugänglich gemacht: als selbständiges Lexikon der Universität München¹⁰ sowie durch die Einbindung in zwei Fachportale der Bayerischen Staatsbibliothek zur Landesgeschichte und zur Musikwissenschaft, nämlich die Bayerische Landesbibliothek Online¹¹ und die Virtuelle Fachbibliothek Musik der BSB.¹² Um das BMLO bekannt zu machen, wird ein umfangreiches Programm zur Öffentlichkeitsarbeit vor allem bei Multiplikatoren/innen, in Mailinglisten und in Internet-Fachportalen durchgeführt.

Ausblick

Beim Projektstart des BMLO im Januar 2004 sah die Planung eine Bearbeitungszeit von fünf Jahren vor. Um die langfristige Weiterentwicklung und nachhaltige Betreuung des BMLO sicherzustellen, ist für die Folgezeit

eine kooperative Fortführung auf drei Schultern vereinbart: Die Gesellschaft für Bayerische Musikgeschichte übernimmt die BMLO-Redaktion und leistet damit die redaktionelle Betreuung, die IT-Gruppe Geisteswissenschaften der Universität München die Pflege der Hardware und der Datenbank sowie die BSB die Langzeitarchivierung der Daten. Die gegenwärtig noch unvollständigen Datensätze von 15.600 Personen sollen bis 2008 bis zur Editionsreife geführt, medial angereichert und auf deutlich über 20.000 vermehrt werden.

Dr. Josef Focht leitet seit Anfang 2004 das DFG-Projekt „Bayerisches Musiker-Lexikon Online“ am Institut für Musikwissenschaft der Universität München. Arbeitsschwerpunkte: Themen der Instrumentenkunde, der historischen Aufführungspraxis, der Regionalmusikgeschichte und der Biographieforschung. E-Mail: focht@lrz.uni-muenchen.de

¹⁰ Vgl. <<http://www.bmlo.uni-muenchen.de>>.

¹¹ Vgl. <<http://www.bayerische-landesbibliothek-online.de>>.

¹² Vgl. <<http://www.vifamusik.de>>.

PERSONENREGISTER FÜR BAND I UND II

A

Axer, Christine · Bd. 2: 5, 518,
532

B

Barteleit, Sebastian · Bd. 1: 4, 73,
393, 410
Berman, Merrick Lex · Bd. 2: 3,
302, 313
Beyer-Thoma, Hermann · Bd. 1:
18, 20; Bd. 2: 1, 71, 84
Blin, Frédéric · Bd. 1: 3, 72, 227,
243
Bodenhamer, David J. · Bd. 1: 10;
Bd. 2: 3, 281, 288
Borgmann, Karsten · Bd. 1: 2, 20,
108, 158, 171
Bröckelmann, Jan · Bd. 2: 5, 518,
532
Burch, Thomas · Bd. 1: 6, 12, 17,
607, 627
Burckhardt, Daniel · Bd. 1: 1, 2,
7, 10, 97, 103; Bd. 2: 3, 257,
269, 271, 278
Burtscheidt, Andreas · Bd. 1: 21;
Bd. 2: 4, 331, 343

C

Carlson, Christopher N. · Bd. 1: 5,
473, 573, 588
Charlier, Robert · Bd. 1: 6, 12,
628, 643

Cornelißen, Christoph · Bd. 1: 55,
56, 60, 62, 65
Cramme, Stefan · Bd. 1: 3, 15, 72,
275, 293
Czmiel, Alexander · Bd. 1: 5, 18,
455, 465

E

Enderle, Wilfried · Bd. 1: 16, 91;
Bd. 2: 1, 18, 35

F

Filthaut, Jörg · Bd. 1: 5, 472, 589,
604
Focht, Josef · Bd. 1: 6, 653, 661
Fournier, Johannes · Bd. 1: 2, 9,
17, 108, 141, 157

G

Geldsetzer, Sabine · Bd. 2: 5, 483,
502
Gersmann, Gudrun · Bd. 1: 2, 9,
16, 105, 110, 125
Gervers, Michae · Bd. 1: 5, 9, 78,
471, 482, 500
Giel, Robert · Bd. 1: 4, 83, 411,
421
Goebel, Ralf · Bd. 1: 9, 20, 89,
90; Bd. 2: 1, 7, 17
Gorißen, Stefan · Bd. 1: 1, 10, 69,
75
Görz, Günther · Bd. 1: 5, 80, 471,
539, 572

Griese, Olivia · Bd. 1: 92; Bd. 2:
1, 61, 70
Gugerli, David · Bd. 1: 20; Bd. 2:
2, 216, 226

H

Haber, Peter · Bd. 1: 12, 97; Bd.
2: 4, 315, 318
Hagedorn-Saupe, Monika · Bd. 1:
3, 70, 216, 226, 333
Hänger, Andrea · Bd. 1: 3, 72,
258, 274
Hein, Dörte · Bd. 2: 4, 344, 363
Hein-Kircher, Heidi · Bd. 1: 16,
18, 95; Bd. 2: 2, 150, 169
Heinz, Karl · Bd. 1: 5, 77, 471,
501, 513
Herren, Madeleine · Bd. 1: 100,
102; Bd. 2: 4, 435, 445, 517
Heuberger, Rachel · Bd. 1: 16, 93;
Bd. 2: 2, 111, 126
Hodel, Jan · Bd. 1: 100, 145; Bd.
2: 4, 407, 432
Hohls, Rüdiger · Bd. 1: 1, 7, 95;
Bd. 2: 2, 170, 186, 190
Holtz, Bärbel · Bd. 1: 5, 18, 455,
465
Horstkemper, Gregor · Bd. 1: 91;
Bd. 2: 1, 36, 60
Huck, Thomas Sergej · Bd. 1: 4,
430, 445
Hummel, Karl-Joseph · Bd. 1: 99;
Bd. 2: 4, 331, 343
Huth, Karsten · Bd. 1: 3, 15, 72,
258, 274

I

Isacson, Kristina · Bd. 2: 2, 216,
226

K

Kaiser, Michael · Bd. 1: 2, 109,
111, 124, 126
Kamke, Hans-Ulrich · Bd. 2: 5,
446, 459
Kellner, Stephan · Bd. 2: 2, 191,
200
Kirsch, Jan-Holger · Bd. 1: 2,
109, 125, 140
Kittler, Friedrich · Bd. 1: 56, 58,
65
Kolbe, Annette · Bd. 2: 2, 99, 110
Krameritsch, Jakob · Bd. 1: 98;
Bd. 2: 4, 364, 386
Kröll, Michael · Bd. 1: 95; Bd. 2:
2, 127, 142
Kuczera, Andreas · Bd. 1: 4, 83,
422, 429
Kunz, Andreas · Bd. 2: 3, 289,
301

L

Löbnitz, Anke · Bd. 1: 4, 73, 368,
392
Lux, Claudia · Bd. 1: 56, 59, 60,
61, 62, 63, 64, 65

M

Margolin, Michael · Bd. 1: 5, 9,
482, 500
Marra, Stephanie · Bd. 1: 2, 10,
87, 96; Bd. 2: 128
Meyer, Thomas · Bd. 2: 3, 244,
256
Minn, Gisela · Bd. 1: 99; Bd. 2: 4,
387, 406
Müller, Harald · Bd. 1: 2, 76, 86

N

Néouze, Valérie · Bd. 1: 3, 9, 227,
243
Nerlich, Daniel · Bd. 1: 5, 446,
454
Neumann, Gerald · Bd. 1: 6, 644,
652
Nippel, Wilfried · Bd. 1: 1, 25, 28

P

Peifer, Karl-Nikolaus · Bd. 1: 2,
16, 109, 172, 190
Perstling, Matthias P. · Bd. 1: 5,
78, 471, 514, 538
Pfanzer Sausgruber, Eva · Bd.
2: 5, 461, 482
Prinz, Claudia · Bd. 1: 1, 7
Przigoda, Stefan · Bd. 1: 4, 350,
367

R

Rahmsdorf, Sabine · Bd. 1: 4,
308, 321
Rapp, Andrea · Bd. 1: 6, 607, 627
Rettinger, Elmar · Bd. 1: 16; Bd.
2: 2, 201, 202, 214

S

Sahle, Patrick · Bd. 1: 5, 13, 76,
467, 481, 522
Schlögl, Daniel · Bd. 1: 1, 4, 56,
322, 329
Schmale, Wolfgang · Bd. 1: 1, 11,
13, 17, 26, 27, 28, 33, 52, 100,
137; Bd. 2: 371, 378
Schrade, Torsten · Bd. 2: 2, 201,
215
Schultes, Kilian · Bd. 2: 5, 504,
517
Schulze, Winfried · Bd. 1: 11, 17,
56, 57, 66, 129
Sorbello Staub, Alessandra · Bd.
1: 92, 93; Bd. 2: 1, 85, 98
Stazic-Wendt, Tamara · Bd. 2: 4,
387, 406
Steyer, Timo · Bd. 1: 3, 294, 307
Stock, Günter · Bd. 1: 1, 53, 55
Strothmann, Meret · Bd. 2: 5, 483,
502, 503

T

Thaller, Manfred · Bd. 1: 9, 14,
17, 56, 58, 61, 63, 64, 66, 476,
523, 538, 543

V

Vogeler, Georg · Bd. 1: 5, 13, 18,
76, 467, 481
Vögler, Max · Bd. 1: 3, 9, 16, 71,
207, 215
von Hagel, Frank · Bd. 1: 4, 74,
331, 349
von Seggern, Jessica · Bd. 1: 4,
368, 392
Voß, Jakob · Bd. 1: 98; Bd. 2: 4,
319, 330

W

Wannags, Max-Michael · Bd. 1:
4, 430, 445
Weber, Cornelia · Bd. 2: 3, 228,
243

Weber, Hartmut · Bd. 1: 1, 26, 29,
32
Winsmann, Anke · Bd. 2: 3, 257,
270
Wirtz, Gudrun · Bd. 1: 16, 92; Bd.
2: 1, 61, 70
Wolf-Klostermann, Thomas · Bd.
1: 2, 109, 191, 205
Wollschläger, Thomas · Bd. 1: 3,
15, 71, 244, 257

Z

Zepf, Robert · Bd. 1: 4, 73, 393,
410
Zündorf, Irmgard · Bd. 2: 2, 145,
149