

Workshop der Arbeitsgruppe Elektronisches Publizieren
der Union der deutschen Akademien der Wissenschaften



Düsseldorf, 04. bis 06. Oktober 2010 in der
Nordrhein-Westfälischen Akademie der Wissenschaften und der Künste

in Zusammenarbeit mit:

Nordrhein-Westfälischen Akademie der Wissenschaften und der Künste

Kompetenzzentrum für elektronische Erschließungs-
und Publikationsverfahren in den Geisteswissenschaften

Sonderforschungsbereich 600 „Fremdheit und Armut“

Zentrum für Informations-, Medien- und Kommunikationstechnologie
(ZIMK), Universität Trier

Abstracts

INHALTSVERZEICHNIS

Eröffnungsvortrag	3
Die Wissenschaftskommunikation der Zukunft - Aspekte von Science 2.0.....	3
Beiträge aus den Akademien: Alte Geschichte und Inschriftenforschung	3
Das Hethitologie Portal	3
Vom Zettelarchiv zur Datenbank: Die Inschriften des Tempels von Edfu in Oberägypten ...	4
Die epigraphische Datenbank Heidelberg.....	4
Die deutschen Inschriften online	5
Beiträge aus den Akademien: Repositorien zur Überlieferung heiliger Schriften.....	6
Die Überlieferung des Neuen Testaments	6
Die Überlieferung des Korans (Corpus Coranicum)	7
Buddhistische Steininschriften in China	8
Beiträge aus den Akademien: Wörterbücher und Editionen.....	9
Das Deutsche Rechtswörterbuch - ein Repitorium?	9
Vom Buch zum "Deutschen Text-Archiv"	10
Zwischen Datenbank und Wörterbuch im Dictionnaire étymologique de l'ancien français.....	11
Stand und Folgen der Retrodigitalisierung des Althochdeutschen Wörterbuchs	13
Beiträge aus den Akademien: Digitale Bibliotheken und Informationssysteme	14
Digitale Bibliotheken am Deutschen Referenzzentrum für Ethik in den Biowissenschaften.....	14
Die Neue Deutsche Biographie (NDB) - von der Digitalisierung zur digitalen Dienstleistung.....	15
Akademien in Not? Überlegungen der BAdW und der BSB zur Langzeitarchivierung ihrer Forschungsdaten	16
Konzepte und Entwicklungen I.....	16
Forschungsdaten in interdisziplinären Langfristvorhaben: Bereitstellung und Nachnutzung	16
Perspektiven der Langzeitarchivierung der Universität Trier	17
Perspektiven der Langzeitarchivierung der BSB	18

TextGrid - Virtuelle Forschungsumgebung, Forschungsdaten und Langzeitarchivierung (BMBF)	18
Digitale Quellen sichern und verfügbar halten: Die Lösung des Bundesarchivs	20
Das Personendaten-Repository der BBAW	21
Repositorieninfrastruktur der MPDL.....	22
Öffentlicher Vortrag	22
Repositorien - Digitalmüll der Zukunft oder Zukunft digitaler Bibliotheken?	22
Einzelaspekte und Kooperationen	24
Rechtliche Probleme der digitalen Langzeitarchivierung wissenschaftlicher Primärdaten.....	24
Digitalisierungsstrategien in der SAGW	26
Langzeitarchivierung von Literatur im DLA	27
Repositorien und Langzeitarchivierung aus der Sicht der Bibliotheken	28
Abschlussvortrag	28
Geisteswissenschaftliche Forschungskommunikation der Zukunft im Dreieck Autor – Verlag - Bibliothek.....	28
Leipziger Empfehlungen der Arbeitsgruppe „Elektronisches Publizieren“ der Union der deutschen Akademien (21. September 2009)	29
Autorenverzeichnis	30

ERÖFFNUNGSVORTRAG

DIE WISSENSCHAFTSKOMMUNIKATION DER ZUKUNFT - ASPEKTE VON SCIENCE 2.0

R. Ball

Die Vorstellung von wissenschaftlichen Publikationen und die Formalisierung des Outputs in Form von Zeitschriften, Konferenzbeiträgen und Büchern ist seit mehr als fünfhundert Jahren gängige Praxis. Doch seit rund zehn Jahren existiert durch die digitale Revolution eine neue Möglichkeit, den wissenschaftlichen Output in Form digitaler elektronischer und dynamischer Dokumente zu generieren. Aber diese Form der Wissenschaftskommunikation ist längst noch nicht in den Köpfen und den Schreibstuben der Wissenschaftler präsent. Zu kurz ist die Phase des digitalen Zeitalters und zu lang die Tradition der klassischen, schriftlichen, gedruckten Formen der Wissenschaftskommunikation. Der Vortrag zeigt die Entwicklung der Wissenschaftskommunikation von der Erfindung des Buchdrucks bis heute und macht klar, dass das Ende der gedruckten Wissenschaftskommunikation gekommen ist.

BEITRÄGE AUS DEN AKADEMIEN: ALTE GESCHICHTE UND INSCHRIFTENFORSCHUNG

DAS HETHITOLOGIE PORTAL

G. Wilhelm

Das 2001 eröffnete und seitdem stetig ausgebaut Hethitologie Portal bietet vielfältige Recherche-Möglichkeiten, z.B. durch eine Konkordanz sämtlicher hethitischer Texte mit multiplen Abfragemöglichkeiten und Verknüpfung mit der digitalisierten Photosammlung, die bei vollem Ausbau über 60.000 Photos verfügbar macht. Das "Portal" zielt insbesondere auch auf eine Zusammenführung und Vernetzung der internationalen hethitologischen Forschung - vor allem bei der Quellenedition, der Bibliographie und der Lexikographie - ab. Bei der Textedition ist eine spezifische Darstellungsform entwickelt worden, die den Besonderheiten keilschriftlicher Texte Rechnung trägt.

In einem Teilprojekt geht es um die digitale dreidimensionale Erfassung und Darstellung hethitischer Keilschrifttafeln.

Das "Portal" hat die Forschungsbedingungen in der Hethitologie nachhaltig verändert; in einer Rezension heißt es: "studying Hittite is unthinkable without the Portal." (Bibliotheca Orientalis 64, 2007, 187).

Der Vortrag stellt das Projekt in wesentlichen Zügen vor und erörtert die projektspezifischen Probleme einer Langzeitsicherung.

Link:

www.hethiter.net

VOM ZETTELARCHIV ZUR DATENBANK: DIE INSCRIFTEN DES TEMPELS VON EDFU IN OBERÄGYPTEN

J. Graeff

Seit Mitte der neunziger Jahre nutzt das in Hamburg beheimatete Edfu-Projekt der Akademie der Wissenschaften zu Göttingen bei seinen Übersetzungsarbeiten für die Suche nach internen Parallelen eigene Datenbanken. Im Einzelnen sind dies die Formulardatenbank, die Götterliste, die Ortsnamenliste und die Wörterliste. Jede dieser Datenbanktabellen kann auf ihrem Gebiet wichtige Parallelen liefern, die eine Übersetzung und ein Verständnis beschädigter oder inhaltlich unverständlicher Textstellen oft auch in solchen Fällen möglich macht, wo ein Übersetzer andernfalls an seine Grenzen stößt.

Da das Projekt jedoch über große Materialmengen verfügt, die kaum auf unmittelbarem Wege in eine Datenbank einpflegbar sind, wurde schon früh damit begonnen, die eigentlichen Dateninhalte von der ausführenden Programmlogik zu trennen. So werden die Datenbanken des Projektes heute nicht mehr von einem vorgefertigten System, wie beispielsweise MS Access, angetrieben, sondern durch den eigens entwickelten "Edfu-Explorer". Dieses Datenbankprogramm greift nicht nur auf die eigentlichen Datenbanken zu, sondern vernetzt sie darüber hinaus mit externen Materialien, wie z.B. dem Photo-Archiv, das mit seinen rund 15.000 Bildern einen Datenumfang von mehreren Gigabyte hat. Auch die Scans der hieroglyphischen Texte, publiziert von Émile Chassinat, können auf diese Weise mit den jeweiligen Datensätzen verglichen werden. Zudem sind etwaige Kollatierungsergebnisse einblendbar, um ein dritte Quelle zu liefern.

Ein zusätzlicher Aspekt, der sich auf die Programmgestaltung des Edfu-Explorers auswirkte, war die Notwendigkeit, Textzusammenhänge aufgrund ihrer jeweiligen Position im Tempel zu bestimmen. Da die Verwendung einer Textgattung ebenso wie die darin genutzten Formulierungen Bezug auf das damit verbundene Bauelement des Tempels nimmt, stellt der Edfu-Explorer einen zusätzlichen Zusammenhang zwischen Suchergebnissen und Fundposition einer Datenbankabfrage her.

Die Möglichkeiten dieses modernen Werkzeugs haben in den vergangenen Jahren maßgeblich dazu beigetragen, das Konzept der Anlage des Tempels, seine Dekoration, seine Texte und dadurch nicht zuletzt die gesamte ägyptische Kultur in weiten Teilen besser zu verstehen als dies mittels konventioneller Übersetzungsmethoden bislang möglich war.

Links:

<http://www1.uni-hamburg.de/Edfu-Projekt//Datenbanken.html>

<http://www1.uni-hamburg.de/Edfu-Projekt//Edfu.html>

DIE EPIGRAPHISCHE DATENBANK HEIDELBERG

F. Grieshaber

Das an der Heidelberger Akademie der Wissenschaften beheimatete Forschungsprojekt "Epigraphische Datenbank Heidelberg (EDH)" hat in erster Linie die Aufnahme lateinischer Inschriften des Imperium Romanum in einer umfassenden Datenbank zum Ziel; daneben werden seit 2004 in kleiner Auswahl auch griechische Inschriften desselben Zeitraums aufgenommen. Es besteht aus den drei Teildatenbanken "Epigraphische Text-Datenbank", "Epigraphische Bibliographie" und "Epigraphische Fotothek" und steht auf internationaler Ebene neben anderen Datenbankvorhaben, die als Instrumentarium zur raschen und

einfachen Sammlung, Sichtung, Ergänzung und interdisziplinären Auswertung epigraphischer Zeugnisse dienen.

Die "Epigraphische Text-Datenbank" umfasst derzeit über 60.000 Inschriften und bietet die Inschriften überwiegend in überprüfter und häufig verbesserter Lesung dar. Eine derartige Kontrolle ist vor allem im Fall älterer Inschriftenpublikationen unverzichtbar, die modernen textkritischen Standards vielfach nicht mehr genügen. Zum anderen werden nicht allein die Texte erfasst, sondern diese mit allen verfügbaren Informationen zu Inschrift und Inschriftträger, mit den einschlägigen bibliographischen Angaben und, wo möglich, mit Fotos verbunden. Dank dieser aufwendigen und zeitintensiven Arbeitsweise wird die Datenbank hohen wissenschaftlichen Ansprüchen gerecht.

Die bibliographischen Angaben werden in einer gesonderten Datenbank, der "Epigraphischen Bibliographie", mit aktuell knapp 13.000 Datensätzen gesammelt. Der Rückgriff auf zeichnerische und fotografische Dokumentation, u. a. aus den Beständen der "Epigraphischen Fotothek" (aktuell ca. 22.000 Datensätze), erweitert die Informationsbasis.

In Vorbereitung befindet sich eine weitere Datenbank, die geographische Informationen zu antiken Ortsnamen, deren modernen Entsprechungen sowie geographische Koordinaten für die präzise Anzeige in Google Maps umfassen wird.

Vor dem Hintergrund zunehmender internationaler Zusammenarbeit und Vernetzung mit anderen verwandten Datenbankvorhaben stellt sich immer mehr heraus, dass das bisher in der EDH genutzte, proprietäre Format der Textdaten der Inschriften einen Datenaustausch und die Integration der Epigraphischen Datenbank Heidelberg in ein geplantes internetbasiertes, alle Datenbanken mit lateinischen Inschriften umfassendes Metaportal sehr aufwendig macht. Hierfür müssen für die Epigraphische Datenbank Heidelberg zukünftig andere, neue Lösungen auf Basis des EpiDoc-Standards erarbeitet werden.

Links:

EDH Website: <http://www.epigraphische-datenbank-heidelberg.de/>

EpiDoc: <http://epidoc.sourceforge.net/>

DIE DEUTSCHEN INSCRIFTEN ONLINE

T. Schrade, Aline Deike

Das Projekt "Die Deutschen Inschriften" ist das älteste moderne Unternehmen zur Sammlung und Edition mittelalterlicher und frühneuzeitlicher Inschriften bis zum Jahr 1650. Begründet wurde es vor über 70 Jahren als Gemeinschaftsunternehmen der wissenschaftlichen Akademien in Deutschland und Österreich. Die Ergebnisse der Sammlung werden in der Reihe "Die Deutschen Inschriften" publiziert. Ein Band beinhaltet dabei entweder die Inschriften eines bzw. mehrerer Stadt- oder Landkreise (in Österreich: Politischer Bezirke) oder die Inschriften einzelner Städte.

"Deutsche Inschriften Online" ist ein interakademisches Online-Projekt der Inschriften-Arbeitsstellen an den deutschen Akademien der Wissenschaften. Ziel des Projektes ist die Digitalisierung der Inschriftenbände im Rahmen einer umfassenden epigraphischen Forschungsplattform. Die "Digitale Akademie" Mainz und die Inschriften-Arbeitsstelle Greifswald übernehmen hierbei die konzeptionelle und technische Leitung. Neben der Verfügbarmachung der Inschriftenkataloge liegt der Fokus vor allem auf der Schaffung neuer Recherchemöglichkeiten in Ergänzung zur Publikationsreihe.

Bei der Umsetzung des Projektes konnte die "Digitale Akademie" ihr Know-How aus dem Pilotprojekt "Inschriften Mittelrhein-Hunsrück" einbringen, in dessen Rahmen bereits 2008

der Inschriftenband Nr. 60 "Rhein-Hunsrück-Kreis" im Web digital bereitgestellt wurde. Zum Einsatz kam damals wie heute das Content-Management-System TYPO3. Die Greifswalder Arbeitsstelle steuerte umfangreiches XML-Know-How bei. Die Textdaten werden nicht per Hand übertragen, sondern in Open Office mit speziellen Formatvorlagen ausgezeichnet und anschließend via XSLT in XML überführt. Mit Hilfe eines eigens entwickelten TYPO-Moduls erfolgt der Import in die Datenbank. Im TYPO3-Backend können die technischen Redakteure die Datensätze kontrollieren und ggf. verbessern. Somit wird sichergestellt, dass alle Daten werkstreu in die Datenbank gelangen. Bis heute sind 11 Bände der Mainzer und Göttinger Reihe in die Plattform überführt. Bis zum Jahr 2012 sollen weitere Inschriftenbände aus allen Inschriften-Arbeitsstellen folgen.

- Adresse: www.inschriften.net
- Umfang: Bisher 11 digitalisierte Bände (5 davon Live, 6 in Arbeit), ca. 4000 Abbildungen
- Besonderheiten: TYPO3 Inschriften-Extension & XML-Importmodul, Mediendatenbank, Literaturdatenbank, XSL-Transformator, Datierungsapplikation, Applikation zur Registerkumulierung
- Technologien: TYPO3, XHTML/CSS, Mootools, PHP, MySQL, XSLT
- Partner: Institut für Geschichtliche Landeskunde an der Universität Mainz e.V., Kompetenzzentrum Trier, i3 Mainz, ADW Göttingen, ADW Mainz

BEITRÄGE AUS DEN AKADEMIEN: REPOSITORIEN ZUR ÜBERLIEFERUNG HEILIGER SCHRIFTEN

DIE ÜBERLIEFERUNG DES NEUEN TESTAMENTS

H. Strutwolf, V. Krüger

Die Editio Critica Maior dokumentiert die griechische Textgeschichte des ersten Jahrtausends anhand der überlieferungsgeschichtlich wichtigen griechischen Handschriften, alten Übersetzungen und neutestamentlichen Zitate in der antiken christlichen Literatur. Auf der Basis genealogischer Untersuchungen des erstmals mit dieser Vollständigkeit aufbereiteten Materials wird der Ausgangstext des griechischen Neuen Testaments rekonstruiert.

Zu diesem Zweck hat das Institut für Neutestamentliche Textforschung (INTF) zunächst alle bekannten Handschriften erfasst und katalogisiert. Es besitzt von über 90 % des bekannten Materials Mikrofilme oder Fotos.

Das Institut setzt dabei auf allen Ebenen seiner philologischen Arbeit digitale Methoden ein: Handschriften werden in größtmöglicher Qualität als digitale Fotos erfasst. Diese Fotos sind Grundlage für die am Computer zu erstellenden Transkriptionen. Mittels der Software Collate wird dann die Erstellung eines kritischen Apparates vorbereitet.

Neben der in Buchform publizierten Edition wird eine digitale Plattform entwickelt, die den Online-Zugriff auf alle Daten erlaubt, die während der Arbeit an den ECM gesammelt werden. Dazu gehören vor allem die diplomatischen Transkripte aller in der ECM verwendeten Handschriften und die Datenbanken, auf denen die Editio Critica Maior beruht. In Kooperation mit dem Zentrum für Informationsverarbeitung der Universität Münster und der Universitätsbibliothek werden die Ausgangsdaten so gespeichert, dass eine ständige Verfügbarkeit garantiert ist. Um die mittel- und langfristige Sicherung der Daten gewährleisten zu können, ist die Arbeitsstelle Novum Testamentum Graecum - Editio Critica Maior auf die Kooperation mit größeren Institutionen angewiesen. Besonders die ULB

Münster ist hier ein wichtiger Partner, weil sie intensiv am Thema Langzeitarchivierung arbeitet und sich gegenüber dem ECM-Projekt als sehr kooperativ erweist.

Neben der Langzeitarchivierung der Daten, ihrer Auswertungen und der dafür nötigen Programme ist auch das Problem der langfristigen Stabilität der digitalen Veröffentlichungsplattform zu bearbeiten: Zum einen sind zurzeit die Bilder und die dazugehörigen Metadaten noch nicht unter unveränderbaren Adressen zu erreichen. Zum anderen ist es in vielen Fällen nötig, Digitalisate anderer Institutionen in die digitale ECM und den Virtuellen Handschriftenlesesaal einzubinden. Hier muss eine Möglichkeit gefunden werden, die Zugänglichkeit dieser Daten auf Dauer zu gewährleisten, obwohl sich die Internetadressen dieser Materialien immer wieder ändern.

Schließlich stellt sich bei verschiedenen Handschriften auch die Frage, wie der Zugang zu online vorhandenen Materialien auf einen von den besitzenden Institutionen definierten Nutzerkreis eingeschränkt werden kann. Diese Probleme werden in den nächsten Jahren die Arbeit des Projektes begleiten und gelöst werden müssen.

Links:

<http://www.uni-muenster.de/NTTextforschung>

<http://intf.uni-muenster.de/vmr/NTVMR/IndexNTVMR.php>

DIE ÜBERLIEFERUNG DES KORANS (CORPUS CORANICUM)

M. Marx

Das Vorhaben "Corpus Coranicum" wurde 2007 an der BBAW eingerichtet. Es knüpft an die Arbeiten von Theodor Nöldeke (1860-1930), Heinrich Speyer (gest. 1935) und Gotthelf Bergsträßer (1886-1933) an. Mit Bezug zur Erforschung der Textgeschichte des Korans ist besonders die von Gotthelf Bergsträßer 1930 eingerichtete "Korankommission" an der Bayerischen Akademie der Wissenschaften in München zu erwähnen, die nach dem zweiten Weltkrieg nicht weitergeführt wurde.

Im Vortrag werden zunächst die Grundzüge der Textüberlieferung des Korans beschrieben. Nach heutiger Quellenlage wurde der Text sowohl schriftlich (d.h. in Handschriften, Papyri, Inschriften u.a.) als auch mündlich überliefert (Lesartenwerke der islamischen Tradition). Es soll die Frage diskutiert werden, wie beide Überlieferungswege wissenschaftlich beschrieben werden können und welche weiteren Fragestellungen sich bei der Auswertung der Daten ergeben. Dabei sollen die beiden Projektdatenbanken "Manuscripta Coranica" (schriftliche Überlieferung) und "Variae Lectiones Coranae" (mündliche Überlieferung) in ihren Grundzügen vorgestellt werden. Das heute zur Verfügung stehende Quellenmaterial in Form von Handschriften und mündlichen Lesarten wird im Rahmen des Projektes zunächst dokumentiert. Die für die Erforschung der Textgeschichte grundlegende Frage, ob eine kritische Edition des Korans überhaupt möglich ist, kann vor dem Hintergrund der Quellen noch nicht eindeutig beantwortet werden. Aufgrund der spezifischen Situation der Überlieferung können die verschiedenen Zeugnisse zwar erschlossen und dokumentiert werden, die Rekonstruktion eines "Urtextes" scheint allerdings kaum möglich zu sein.

Das Projekt "Corpus Coranicum" umfasst neben der Dokumentation der Textgeschichte, zwei weitere Abteilungen, die sich mit der "Umwelt des Textes" und einer kritisch-literaturwissenschaftlichen Kommentierung beschäftigen. Zum Schluss soll ein kurzer Überblick über die thematische und die datentechnische Struktur des gesamten Vorhabens gegeben werden.

BUDDHISTISCHE STEININSCHRIFTEN IN CHINA

Claudia Wenzel, W. Meier

Die digitale Dokumentation heiliger buddhistischer Schriften in China, die zum größten Teil zwischen dem 6. und 8. Jahrhundert in Stein gemeißelt wurden, erfolgt auf der Grundlage von Photos des behauenen Steins, der auf Papier mit Tusche angefertigten Abreibungen, sowie von Scans der bemeißelten Steinoberflächen. Weitere Primärdatenquellen sind Geodaten und 3D-Modelle der Inschriftenorte und ihrer Topographie.

Die Transkription der so erfassten Texte wird wesentlich erleichtert durch eine gleichzeitige Kollation mit den Schriften des buddhistischen Kanons, der mittlerweile in einer digitalen Ausgabe der Chinese Buddhist Electronic Text Association (CBETA) vorliegt. Bei der Kollation von chinesischen Texten gilt es, den Besonderheiten von Schreibvarianten der chinesischen Schriftzeichen Rechnung zu tragen. Die transkribierten und kollationierten Texte werden gemäß TEI in XML erfasst.

Der Standard des TEI(Text Encoding Initiative)-Konsortiums ist international anerkannt und in den Geisteswissenschaften weit verbreitet. Die bibliographischen Metadaten werden gemäß den Metadaten-Schemata (MODS/MADS) der Library of Congress eingepflegt. TEI-XML erfasst vorrangig die Semantik der Daten gemäß den fachlichen Erfordernissen, nicht aber deren Darstellung. Die unterschiedlichen Publikationsformen, z.B. Web oder Print, sind nur Sichten auf dieselben fachlichen Daten und werden weitgehend automatisch daraus generiert.

Die Beschränkung auf anerkannte Standards ermöglicht eine nachhaltige Nutzung, einen einfachen Austausch mit anderen Projekten sowie die direkte Einbindung externer Datenquellen. Bezüglich der buddhistischen Steinschriften sind hier besonders die bereits erwähnte digitale CBETA Version des buddhistischen Kanons zu nennen, die ebenfalls in TEI-XML vorliegt; außerdem die Einbindung von digital verfügbaren Wörterbüchern, oder spezialisierten Datenbanken mit Geographie-Bezug, die aus der Auswertung von chinesischen Mönchsbiographien im Dharma Drum College auf Taiwan (<http://buddhisticinformatics.ddbc.edu.tw/>) gewonnen wurden.

Alle Daten werden direkt, d.h. ohne weitere Aufbereitung, in einer XML-Datenbank gespeichert und sind über standardisierte Schnittstellen und Abfragesprachen (XQuery, XSLT) ansprechbar, welche sowohl vom Webangebot als auch der Printausgabe genutzt werden. Datenhaltung und Zugriff sind somit grundsätzlich technologieunabhängig, wodurch eine langfristige Verfügbarkeit und nachhaltige Nutzung gewährleistet wird.

Link:

http://www.haw.uni-heidelberg.de/forschung/forschungsstellen/buddhistische_inschriften.de.html

BEITRÄGE AUS DEN AKADEMIEN: WÖRTERBÜCHER UND EDITIONEN

DAS DEUTSCHE RECHTSWÖRTERBUCH - EIN REPOSITORYUM?

A. Deutsch

Seiner ursprünglichen Bedeutung nach meint "Repositoryum" schlicht ein Lager. In diesem Sinne ist bereits das ab 1897 erstellte "Archiv" des Deutschen Rechtswörterbuchs (DRW), in dem rund 2,5 Millionen Belege zu Wörtern der älteren deutschen Rechtsprache alphabetisch sortiert in mehreren hundert Zettelkästen gesammelt liegen, ein Repositoryum. Zur engeren Bedeutung eines Repositoryums als "Verwahrungsort von geordneten Dokumenten, die öffentlich und allgemein zugänglich sind", wird aber wohl eher das fertiggestellte Wörterbuch gerecht. Heute sind vom DRW bereits mehr als elf Bände mit insgesamt knapp 90.000 Artikeln und etwa 450.000 Belegen (Belegzitate und Fundstellenangaben) zu Wörtern mit den Anfangsbuchstaben A-S gedruckt - und so wissenschaftlich aufbereitet für jedermann verfügbar. Jedes Jahr kommen 320 Druckspalten mit etwa 1000 neuen Artikeln hinzu. Dokumentiert werden im DRW nicht nur Rechtswörter im engeren Sinne, sondern auch Wörter der Alltagssprache, soweit sie in einem rechtlichen Kontext besondere Bedeutung erhalten. So finden sich beispielsweise neben "Litiskontestation" und "Reichsabschied" auch "Kuss" und "Saubär" im DRW erläutert. In Band 11, Sp. 917, kann ferner nachgelesen werden, dass "Repositur" (zu lat. "Repositoryum", das in älteren Rechtstexten nicht gebraucht wurde) ein "Aufbewahrungsort für Schriftgut in einem Archiv oder einer Bibliothek" ist; dazu findet sich unter anderem ein Beleg von 1584, wonach den Registratoren angeordnet wurde, sie sollten alle Schriften "in der Repositur verwahrlich behalten". Mehr zum Projekt unter: www.deutsches-rechtsworтерbuch.de/ueber.htm.

Der moderne Wortgebrauch schränkt die Bedeutung eines Repositoryums allerdings noch einmal ein: Gemeint ist eine Art digitale Bibliothek zur Veröffentlichung und Archivierung von elektronischen Publikationen. Ein wichtiger Aspekt ist dabei die Persistenz der Daten. Inwieweit das breitgefächerte Online-Angebot des DRW (vgl. www.deutsches-rechtsworтерbuch.de) die Kriterien eines solchen Repositoryums erfüllt, wird zu diskutieren sein. Neben der elektronischen Vollversion des Wörterbuchs mit zahlreichen Suchfunktionen bietet es u.a. ein elektronisches Textarchiv mit mehreren Millionen Wörtern und Faksimiles zu zahlreichen der vom DRW zitierten rund 8400 Textquellen. Hinzu kommt neuerdings das aus der Arbeit am Deutschen Rechtswörterbuch heraus entstandene, dessen Angebot ergänzende Projekt "DRQEdit" (kurz für: "Deutschsprachige Rechtsquellen in digitaler Edition", vgl. www.drqed.it). Dieses von der DFG finanzierte Projekt der Heidelberger Akademie der Wissenschaften, des Max-Planck-Instituts für europäische Rechtsgeschichte und der Professur für Historisch-Kulturwissenschaftliche Informationsverarbeitung der Universität zu Köln hat zum Ziel, die deutschsprachige juristische Literatur des 15. und 16. Jahrhunderts als Corpus im WWW verfügbar und nach verschiedenen Metadaten recherchierbar zu machen. Kürzlich konnte in einem ersten Schritt bereits ein großer Teil der etwa 400 zum Corpus gehörenden Werke als elektronische Faksimile online gestellt werden. Vielfach steht zusätzlich ein (voll recherchefähiger) elektronischer Volltext zur Verfügung; auf Wunsch können beide Ansichten parallel aufgerufen werden. Da die in DRQEdit publizierten Werke zum größten Teil dem Corpus des Rechtswörterbuchs angehören, stellt das Projekt auch einen erheblichen Zugewinn für die Nutzer des DRW dar. Selbstredend

werden die entsprechenden Fundstellen im DRW auf DRQEdit verlinkt. In einem weiteren Schritt soll eine "Nachschlagfunktion" in DRQEdit eingebaut werden: Ein Doppelklick auf ein beliebiges Wort in einem DRQEdit-Quellentext wird dann genügen, um das Wort im DRW nachzuschlagen.

Links:

<http://drw-www.adw.uni-heidelberg.de/drw/>

<http://drw-www.adw.uni-heidelberg.de/drqedit/>

VOM BUCH ZUM "DEUTSCHEN TEXT-ARCHIV"

M. Schulz

In dem von der Deutschen Forschungsgemeinschaft (DFG) geförderten Projekt Deutsches Textarchiv (DTA), das im Juli 2007 an der Berlin-Brandenburgischen Akademie der Wissenschaften (BBAW) seine Arbeit aufgenommen hat, wird ein linguistisch annotiertes Volltextkorpus frei zugänglich im Internet zur Verfügung gestellt.

Grundlage der elektronischen Volltexte sind Erstausgaben von Druckwerken der Jahre 1650 bis 1900. Derzeit bereitet das Deutsche Textarchiv die Veröffentlichung von 650 Texten aus den Zeitepochen von 1780-1900 vor. Daneben beginnt die Digitalisierung, Volltexterfassung und Annotation von weiteren ca. 700 Drucken aus der Zeit des Barock und der Aufklärung, die schrittweise bis 2013 veröffentlicht werden sollen. Ziel der Projektarbeit ist es einen Umfang von 1300 Texten zu erreichen.

Das sind im Verhältnis zum Digitalisierungsvorhaben der Bayrischen Staatsbibliothek in Verbindung mit Google Books 1,3% Bücher. Zum ersten Mal entsteht hier jedoch ein frei zugängliches Volltextkorpus der deutschen Sprache in einem flexiblen maschinenlesbaren Format (TEI-P5) mit linguistischen Annotationen. Dieses Korpus kann als Beginn für das Anlegen von weiteren speziellen Korpora dienen, für die tiefere textkritische Arbeit verwendet werden und als Baustein für ein zukünftiges lexikographisches System des Deutschen angesehen werden.

Die elektronischen Volltexte sind einzeln oder korpusübergreifend durchsuchbar und enthalten neben dem "reinen" Text des jeweiligen Werks biographische und bibliographische Metadaten sowie reichhaltige Informationen zur Mikro- und Makrostruktur der Vorlage.

Gegenwärtig können sich die Nutzer des DTA die digitalen Faksimiles der Werke vollständig oder in Ausschnitten anzeigen lassen sowie die aufbereiteten Volltexte sowohl als benutzerfreundliche Leseausgabe als auch in Form der XML-Quelltexte ansehen. Darüber hinaus können die Texte im HTML- oder XML-Format heruntergeladen werden. Die Transkriptionen sind entsprechend der aktuellen Richtlinie der Text Encoding Initiative (TEI) in der Version P5 codiert, was den Austausch mit anderen Projekten ermöglicht.

Über den primär gesetzten linguistischen Fokus hinaus ist das DTA aufgrund der thematisch-inhaltlichen Vielfalt der ausgewählten Texte eine ideale Arbeitsumgebung für die angrenzenden Geisteswissenschaften, und auch Naturwissenschaftler, Juristen und Ökonomen finden wertvolle Quellen aus der Geschichte ihres jeweiligen Fachs.

Das DTA soll zu einem Aktiven Archiv weiterentwickelt werden, das den Benutzern verschiedene Möglichkeiten eröffnet, online mit den Texten zu arbeiten und sie zu bearbeiten. Dazu gehören Funktionen wie die Zusammenstellung einer eigenen Textauswahl

(beispielsweise für Seminarapparate), das Setzen persistenter Lesezeichen auf Textpassagen und das Hinzufügen von Annotationen.

In unserem Vortrag wollen wir einen Einblick in die Arbeitsweise des DTA bieten sowie die Anforderungen eines "Aktiven Archivs" und weitere Möglichkeiten der Vernetzung des Projekts innerhalb der angedeuteten Wissenschaftsbezüge mit Ihnen diskutieren.

Link:

<http://www.bbaw.de/bbaw/Forschung/Forschungsprojekte/dta/de/Startseite>

ZWISCHEN DATENBANK UND WÖRTERBUCH IM DICTIONNAIRE ÉTYMOLOGIQUE DE L'ANCIEN FRANÇAIS

S. Dörr

Der Dictionnaire étymologique de l'ancien français (DEAF) ist ein Grundlagenwörterbuch des Altfranzösischen, das den Zeitraum von 842 bis in die Mitte des 14. Jahrhunderts umfasst. Als historisch wichtigste Vernakularsprache des europäischen Mittelalters ist das Altfranzösische in vielen Tausend Texten überliefert, die nicht nur aus Frankreich, sondern beispielsweise auch aus England, Italien, Zypern und dem Heiligen Land stammen. Seine Wirkungsmacht und sein Prestige lassen sich darüber hinaus auch am Einfluss der altfranzösischen Literatur auf alloglotte Literaturtraditionen - darunter etwa die mittelhochdeutsche Klassik - ablesen.

Der DEAF wurde in den sechziger Jahren des letzten Jahrhunderts von Georges Straka, Kurt Baldinger und Jean-Denis Gendron in Québec initiiert und erscheint seit Beginn der neunziger Jahre in einem Publikationsrhythmus von etwa einem Faszikel mit 192 Spalten pro Jahr. 25 Faszikel, in denen die Buchstaben G, H, I, J und K behandelt werden, sind bisher erschienen. Alle verfügbaren Materialien - Wörterbücher, Glossare, literarische und nichtliterarische Texte, Sekundärliteratur - werden im DEAF kritisch ausgewertet; derzeit umfasst der DEAF-Fichier etwa 1,5 Millionen Zettel, die auf ca. 12 Millionen Belege verweisen.

Seit 2008 hat die DEAF-Redaktion in Zusammenarbeit mit dem Karlsruher Institut für Programmstrukturen und Datenorganisation ein elektronisches Redaktions- und Publikationssystem (DEAFél) entwickelt, das am 2. Juli 2010 der Öffentlichkeit vorgestellt wurde. Im Hinblick auf die verbleibende Projektlaufzeit wurde hierbei ein zweistufiger Ansatz verfolgt: die Alphabetstrecke D-F aus wissenschaftshistorischen Gründen das "schwarze Loch" der galloromanischen Lexikographie - wird weitestgehend in gewohnter Form lexikographisch erschlossen; die Ergebnisse werden sowohl in gedruckter Form als auch im Internet publiziert (DEAFplus). Parallel dazu werden sämtliche Materialien zu den Buchstaben A-C und L-Z lemmatisiert und mit einer groben semantischen Struktur versehen, um auch die für diese Alphabetbereiche bereits erarbeiteten Daten langfristig zu konservieren und für eine weitergehende lexikographische Erschließung zugänglich zu machen (DEAFprés). Kernstück des elektronischen Redaktionssystems ist eine MySQL-Datenbank; weiterhin werden die mannigfaltigen in den Artikeln vorliegenden Informationen (etwa Belege, Datierungen, grammatikalische Kategorien, Definitionen, Sprach- und Dialektangaben) mit einem komplexen XML-Markup semantisch ausgezeichnet. Neben Zeitersparnissen insbesondere bei der Klassifizierung und Strukturierung der Materialien und komfortablen Suchfunktionen, die weitaus umfassendere Abfragemöglichkeiten als ein herkömmlicher Index erlauben, bietet die computergestützte

Redaktion und Publikation erstmals die Möglichkeit, die vom DEAF erarbeiteten wissenschaftlichen Ressourcen in einem für die langfristige Archivierung und Bereitstellung geeigneten digitalen Format zu speichern. Im Einzelnen erscheint für die folgenden Materialien eine digitale Archivierung wünschenswert:

1. Primärdaten: Das Ausgangsmaterial für die Redaktion der Wörterbuchartikel liegt zunächst in handgeschriebenen - im Hinblick auf ihren Informationsgehalt und ihre Lesbarkeit ausgesprochen heterogenen - Zetteln vor, die von Hilfskräften mittels einer Eingabemaske in einer XML-Struktur erfasst werden. Da aufgrund der hohen anfallenden Datenmengen (mehrere Tausend Zettel pro Monat) keine durchgängige manuelle Fehlerkorrektur möglich ist, wurde für zentrale Elemente - insbesondere für die Sigel in den Literaturangaben - ein erweiterbares Regelsystem geschaffen, das derzeit etwa 10.400 Ersetzungsregeln umfasst und die automatische Korrektur von 83 bis 97% der jeweils in einer Buchstabenstrecke auftretenden sigelspezifischen Eingabefehler erlaubt.

2. Bibliographische Daten: Das Complément bibliographique des DEAF bietet eine umfassende kritische Bibliographie zum Altfranzösischen und wird ständig erweitert und aktualisiert. Derzeit beinhaltet es 6.410 Literaturverweise mit Angaben zu 953 mittelalterlichen Autoren, 2.663 Datierungen und 1.672 Lokalisierungen. Die XML-Struktur der Bibliographie erlaubt es, die elektronisch erfassten Zettel beim Import in die Datenbank mit aktuellen bibliographischen Angaben zu versehen und lässt sich darüber hinaus problemlos für die Online-Publikation in ein HTML-Format überführen. Ausgehend von den in XML vorliegenden bibliographischen Daten wurde das Complément bibliographique darüber hinaus zuletzt 2007 in gedruckter Form veröffentlicht.

3. Wörterbuchartikel (DEAF und DEAFplus): Die mithilfe des elektronischen Redaktionssystem erarbeiteten Artikel zu den Buchstaben D, E und F (DEAFplus), die zunächst in einer MySQL-Datenbank gespeichert werden, können in verschiedene Formate - Latex für den Druck, HTML für die Online-Publikation, XML für die Archivierung - exportiert werden. Für die bereits publizierten Artikel der Alphabetstrecke G bis K, die lediglich als Image-Dateien vorliegen, ist eine Retrodigitalisierung wünschenswert.

4. Strukturierte Materialien (DEAFpré): Die lemmatisierten und mit einer groben semantischen Struktur versehenen Materialien zu den Alphabetstrecken A-C und L-Z, die im Rahmen der verbleibenden Laufzeit nicht vollständig bearbeitet werden können, können analog zu den Artikeln des DEAFplus aus der Datenbank in HTML und in XML exportiert werden.

Mit der konsequenten Speicherung aller im Redaktionsprozess erarbeiteten Daten - vom Zettel über die Bibliographie bis zum fertigen Wörterbuchartikel - in XML erfüllt der DEAF die in den Empfehlungen der AG Elektronisches Publizieren umrissenen Anforderungen an die Standardisierung der Daten und ist für die Langzeitarchivierung seiner Materialien in einem digitalen Repositorium bestens gerüstet. Als Desiderat wäre in erster Linie die Bereitstellung eines zentralen Servers, auf dem die anfallenden Daten langfristig sicher abgelegt werden können, zu nennen.

Link:

<http://www.deaf-page.de/>

STAND UND FOLGEN DER RETRODIGITALISIERUNG DES ALTHOCHDEUTSCHEN WÖRTERBUCHS

Brigitte Bulitta

Das seit 1948 an der Sächsischen Akademie der Wissenschaften zu Leipzig beheimatete "Althochdeutsche Wörterbuch" nutzt und verwaltet zwar große Datenbestände - sei es das ca. 750.000 Zettel umfassende Belegarchiv, das von der Haager Konvention als schützenswertes nationales Kulturgut eingestuft wurde, sei es die umfangreiche Spezialbibliothek - , doch geschieht dies auch nach Abschluss des fünften Bandes (von ca. zehn Bänden) noch immer weitestgehend auf traditionell-mechanische Weise. Nach wie vor werden auch die Arbeitsergebnisse ausschließlich in gedruckter und nicht in digitaler Form an die Benutzer weitergereicht. Ein "Repositorium" im neueren Sinne des Wortes ist das Vorhaben angesichts dieser Umstände und angesichts der - verglichen mit Forschungseinrichtungen anderer Zielstellungen wie Bibliotheken - wiederum geringen Datenmengen also nicht. Überhaupt werden die Ressourcen, die große Wörterbuchunternehmen zu einer speziellen Forschungsaufgabe über lange Zeit zusammengetragen haben, normalerweise nicht gespeichert oder zugänglich gehalten: sie gelten nach Abschluss des Vorhabens als ausgewertet und werden, nicht zuletzt aus Platzgründen, aufgelöst.

Eine Digitalisierung des Zettelarchivs des Althochdeutschen Wörterbuchs wird aus verschiedenen Gründen vorerst nicht in Betracht gezogen. Vorrangig angestrebt wird eine digitale Nutzbarkeit der publizierten Bände des Wörterbuchs. Während 1999 ein erster Versuch in diese Richtung scheiterte, konnte in diesem Jahr aus Mitteln der SAW eine Retrodigitalisierung als Volltexterkennung per Double-Keying über das Kompetenzzentrum Trier in Angriff genommen werden. Die mittels TUSTEP verarbeiteten Daten sind ausschließlich zur arbeitsstelleninternen Nutzung bestimmt. Sie sind mit einer flachen XML-Auszeichnung versehen und stehen dem Vorhaben im rtf-Format zur Verfügung, so dass bereits mit der vorhandenen Computerausstattung ohne zusätzlichen informationstechnischen Aufwand auf die Daten zugegriffen werden kann.

Die Retrodigitalisierung bleibt nicht ohne Folgen für die künftige Arbeitsroutine. Zum einen muss die Codierung der vorhandenen projekteigenen Schriftsätze mit Sonderzeichen an UNICODE angepasst werden. Mit dem Übergang auf unicodebasierte Sonderzeichen muss auch ein unicodefähiges Textverarbeitungsprogramm (MS-Office 2010, Open-Office o. a.) eingeführt werden. Überhaupt muss die Auszeichnung der künftig zu erarbeitenden Daten mit der der Retrodigitalisierung übereinstimmen. Mit dem Satzunternehmen sind entsprechende Abstimmungen vorzunehmen. Für die Einführung eines Artikelredaktionssystems sind die grundsätzlichen Voraussetzungen (seine Entwicklung, das Vorhandensein einer kritischen Masse digital aufbereiteter Daten und Texte, kontinuierliche informationstechnische Betreuung) erst noch zu schaffen. Bis das Althochdeutsche Wörterbuch als Datenbank im Internet recherchierbar sein wird, ist noch ein weiter Weg zurückzulegen.

Link:

<http://www.saw-leipzig.de/forschung/projekte/althochdeutsches-woerterbuch>

BEITRÄGE AUS DEN AKADEMIEN: DIGITALE BIBLIOTHEKEN UND INFORMATIONSSYSTEME

DIGITALE BIBLIOTHEKEN AM DEUTSCHEN REFERENZZENTRUM FÜR ETHIK IN DEN BIOWISSENSCHAFTEN

U. Marder, Claudia Leuker

Das Deutsche Referenzzentrum für Ethik in den Biowissenschaften (DRZE, <http://www.drze.de>) ist Arbeitsstelle der Nordrhein-Westfälischen Akademie der Wissenschaften und der Künste zum Projekt "Grundlagen, Normen und Kriterien der ethischen Urteilsfindung in den Biowissenschaften - Referenzzentrum". Als Dokumentations- und Informationszentrum für den gesamten Bereich der Ethik in den biomedizinischen Wissenschaften hat es u.a. die Aufgabe, die Informationen aus den unterschiedlichsten Bereichen zur aktuellen bioethischen Diskussion in Wissenschaft, Gesellschaft und Politik aufzubereiten und zugänglich zu machen. Eines der zentralen Informationsangebote des DRZE ist die integrative Bioethik-Literaturdatenbank BELIT (<http://www.drze.de/BELIT>), die Zugriff auf rund 420.000 bibliographische Datensätze von fünf Bibliotheken bzw. Dokumentationszentren bietet.

Publikationen zu bioethischen Themen erscheinen zunehmend in elektronischer Form. Aus diesem Anlass hat das DRZE die Digitale Bibliothek entwickelt, um direkten Zugang zu den Volltexten bieten zu können. Da die Verwaltung von Links zu anderen Servern einen zu großen Arbeitsaufwand darstellt, werden unter Wahrung des Urheberrechts nach schriftlicher Erlaubnis die elektronischen Dokumente auf dem DRZE-Server abgespeichert und mit bibliographischen Metadaten versehen. Dabei handelt es sich in der Regel um Graue Literatur, die nicht kommerziell verlegt wird. Derzeit kann über BELIT auf die Volltexte zugegriffen werden.

Obwohl der Anstoß zur Entwicklung der Digitalen Bibliothek aus dem BELIT-Projekt kam, wurde sie von Beginn an als unabhängiger Dienst entwickelt. Dadurch ist es leicht möglich, die Digitale Bibliothek in andere Anwendungen einzubauen. So arbeiten wir derzeit an der Einbindung der elektronischen Dokumente in die Bioethik-Links (<http://www.drze.de/bioethik-links>), wo sie sowohl nach geographischen als auch thematischen Aspekten strukturiert dargestellt werden sollen. Im EU-Projekt Ethicsweb (<http://www.ethicsweb.org>), an dem das DRZE federführend beteiligt ist, kommt die Digitale Bibliothek ebenfalls zum Einsatz. In Ethicsweb wird derzeit ein Web-Portal zu europäischen Informationsquellen zur Bioethik entwickelt. In diesem Rahmen wird die Digitale Bibliothek um eine Volltext-Retrieval-Komponente erweitert.

Die gesamte Entwicklung der Software wie auch das Hosting der Dienste erfolgt im Hause durch die Digitale Abteilung des DRZE. Software-technisch basiert die Digitale Bibliothek überwiegend auf frei verfügbarer Standardsoftware wie dem relationalen Datenbanksystem MySQL 5, dem Information Retrieval System Apache Lucene 3 und dem Web Application Server Zope 2. Die darauf entwickelten Anwendungen der Digitalen Bibliothek wurden in Python und SQL geschrieben und sind somit weitgehend unabhängig von der Betriebssystemplattform (aktuelles OS ist Windows Server 2003 R2). Für das Hosting stehen moderne 19-Zoll-Serversysteme sowie redundante Server-RAIDs zur Verfügung. Die Datensicherung erfolgt täglich auf Ultrium-2 Magnetbänder. Zur Absicherung gegen Totalverluste (Diebstahl, Feuer etc.) steht ein Banksafe zur Verfügung, in dem Sicherungsbänder regelmäßig eingelagert werden.

DIE NEUE DEUTSCHE BIOGRAPHIE (NDB) - VON DER DIGITALISIERUNG ZUR DIGITALEN DIENSTLEISTUNG

M. Reinert

Im Februar 2010 ist die digitale Volltextfassung der "Neuen Deutschen Biographie" (NDB) und der "Allgemeinen Deutschen Biographie" (ADB) im Internet freigeschaltet worden.

Die Redaktion der NDB hat gemeinsam mit der Bayerischen Staatsbibliothek (BSB), gefördert durch die DFG, die Bände 1-22 der NDB und alle 55 Bände der ADB in XML gemäß TEI-P5 in der Grundstruktur aufbereitet (vgl. Bernhard Ebner (BAW) auf dem Workshop in Leipzig 2009, online:

<http://www.akademienunion.de/gremien/elektronischespublizieren/Personendateien.pdf?page=82>).

Nach Erscheinen des aktuellen Bandes (Bd. 24 im Februar 2010) wurde der vorherige aufbereitet und online gestellt.

Damit bietet die Website rund 47.000 Biographien und ein integriertes Register mit rund 93.000 Einträgen (seit 2001 online). In einem laufenden Folgeprojekt (Digitalisierung der Hauptkartei) werden die Informationen von rund 150.000 Karteikarten zu Personen und Familien erfasst und nach Abgleich hinzugefügt.

In beiden Projekten steht die Identifizierung der Personen anhand des Nachweissystems für Personen bei den Bibliotheken (Personennamendatei, PND) im Vordergrund. Der Abgleich mit diesen Normdaten ist der Schlüssel für die Vernetzung der biographischen Vorhaben im deutschsprachigen Raum und die Vernetzung von bibliographischen, fachwissenschaftlichen und lexikalischen Quellen und Angeboten im Web. Über eine Schnittstelle für Metadaten (PND-BEACON) wird die gegenseitige Verlinkung ermöglicht.

Neben komfortabler Suche und leichter Pflege von Korrekturen im Volltext ist die XML-Fassung Ausgangspunkt für die weitere, modulare Kodierung von interessanten Informationen im Text - Personen, Orte, Sachbezüge.

Die Bezüge zwischen den Artikeln lassen sich über die Nennung im Text (Co-occurrence), die gemeinsam zitierte Literatur (Co-citation) und gemeinsame Auszeichnungen erschließen und dann visuell präsentieren, z.B. als "Wissenschaftler-Stammbäume". Mittels automatisierter Georeferenzierung werden Artikel ortsbezogen auffindbar.

Eine automatisierte Herausarbeitung der Sachbezüge in der ADB kann das alte (Artikel)Register erweitern und das von Ludwig Quidde 1920 formulierte Desiderat "[e]in systematisches Sachregister zur Allgemeinen Deutschen Biographie" versuchen zu erfüllen (vgl. Hinweis auf das Archiv der Historischen Kommission, Band 298 in: Gohle, Peter: Archiv der Historischen Kommission 1858-1961. Repertorium. [München 2005], online: <http://www.historischekommission-muenchen.de/seiten/archivfindbuch.pdf?page=96>).

Die NDB erstellt nunmehr nicht nur Lexika, sondern erbringt eine digitale Dienstleistung - sie liefert "zertifiziertes Wissen" (Hans Günter Hockerts) zu Personen - bisher durch weitgehend automatisierte Konversion des Gedruckten für das Internet. Dieses Aufgabengebiet wächst und soll zu einem webbasierten Redaktionssystem führen, das die Artikel zuerst im Medium Internet aufbereitet und für den Druck automatisiert konvertiert.

Link:

<http://www.ndb.badw-muenchen.de>

AKADEMIEN IN NOT? ÜBERLEGUNGEN DER BADW UND DER BSB ZUR LANGZEITARCHIVIERUNG IHRER FORSCHUNGSDATEN

M. Wesche

Die Bayerische Akademie der Wissenschaften führte von Dez. 2009 bis Jan. 2010 eine Umfrage durch, wie es die einzelnen Unternehmungen mit den "offenen Standards" beim elektronischen Publizieren und bei der Archivierung ihrer Daten hielten. Das Ergebnis war beängstigend: Nur diejenigen Projekte, die von Beginn digital arbeiten oder im Prozess der Digitalisierung stehen, beachten die in den "Leipziger Empfehlungen" für die Akademienprojekte empfohlenen Standards. In der Frage der langfristigen Archivierung ihrer digitalen Daten insgesamt war - und ist weiterhin - die Akademie in Not!

Dieser Beitrag zum Workshop skizziert mit mehreren Beispielen die schwierige Ausgangslage bei den Projekten und beschreibt die Vorstellungen, die in einer vom Vorstand unserer Akademie eingesetzten Arbeitsgruppe zur Steuerung dieser Mängel entwickelt wurden. Hierbei setzte sich sehr bald die Einsicht durch, dass der Langfrist-Archivierung bei der Einführung der "offenen Standards" eine Hebelwirkung zuzuschreiben sei, sie im Angelpunkt langfristiger digitaler Arbeit stehe. Um dies jedoch umzusetzen, sind völlig neue Personalstrukturen für den weiteren Prozess der Digitalisierung nötig.

Link zu den "Leipziger Empfehlungen":

<http://www.akademienunion.de/gremien/elektronischespublizieren/>

KONZEPTE UND ENTWICKLUNGEN I

FORSCHUNGSDATEN IN INTERDISZIPLINÄREN LANGFRISTVORHABEN: BEREITSTELLUNG UND NACHNUTZUNG

T. Burch, Marina Müller

In Projektverbänden ist der schnelle und einfache Zugriff auf gemeinsame Datensammlungen eine wichtige Voraussetzung für die Forschungsarbeit. Für diesen Zweck entwickelt der Sonderforschungsbereich 600 "Fremdheit und Armut" in Kooperation mit dem "Kompetenzzentrum für elektronische Erschließungs- und Publikationsverfahren in den Geisteswissenschaften" das "Forschungsnetzwerk und Datenbanksystem" (FuD). Die virtuelle Forschungsumgebung wird im SFB von derzeit 70 Wissenschaftlerinnen und Wissenschaftlern in über 20 Teilprojekten aus unterschiedlichen geistes- und sozialwissenschaftlichen Fächern eingesetzt. Darüber hinaus verwenden weitere Einrichtungen FuD, vor allem die Mainzer Akademieprojekte "Corpus der Quellen zur mittelalterlichen Geschichte der Juden im Reichsgebiet" und das Projekt "Bildatenbank zur antiken Sklaverei (BizaS)".

Die Arbeitsplattform unterstützt die dezentrale Forschungsarbeit, indem die Forschungsdaten in einer zentralen Datenbank vorgehalten werden. Sie ermöglicht über ein differenziertes Rechte- und Nutzermanagement die strukturierte Erfassung der Daten, ihre Analyse und redaktionellen Bearbeitung bis hin zur Publikation als digitales oder Print-Medium. Zurzeit wird eine Modellösung für die langfristige Archivierung und Nachnutzung

der Forschungsdaten mit dem Zentrum für Informations-, Medien- und Kommunikationstechnologien der Universität Trier (ZIMK) erarbeitet.

Im Vortrag werden die Anforderungen an ein Forschungsdatenrepositorium skizziert. Der Lösungsansatz für die Datenhaltung im FuD-System wird im zweiten Teil des Vortrages vorgestellt. Dabei sind im Umgang mit Primärdaten zwei Zeitphasen zu unterscheiden: 1. während der Projektlaufzeit; 2. nach Projektabschluss. Zunächst wird erläutert, welche Möglichkeiten die Arbeitsplattform FuD während der laufenden Forschungsarbeiten bietet, um den Datenaustausch und die gemeinsame Datenbearbeitung zu organisieren. Anschließend wird gezeigt, welche Arbeitsschritte zur Vorbereitung der Langzeitarchivierung der Primärdaten im Verbund geplant sind. Zu klären sind dabei u.a. folgende Fragen: Wann werden die Daten in das Langzeitarchiv überführt? Welche Daten werden archiviert? Welche Metadaten müssen ergänzt werden, um z.B. die Suche im Langzeitarchiv optimal gewährleisten zu können? Welche Rechte sind bei der Übergabe in ein Langzeitarchiv zu berücksichtigen? Im abschließenden dritten Teil wird dargestellt, welche Funktionen in die Forschungsumgebung für die Vorbereitung der Langzeitarchivierung der Daten implementiert werden müssen. Außerdem wird mit dem ZIMK eine Lösung zur Archivierung des Software-Systems erarbeitet, um zum einen den bei Projektende erreichten Entwicklungszustand zu sichern und zum anderen den langfristigen Zugang zu den Daten auch über das FuD-System zu gewährleisten.

Links:

<http://www.sfb600.uni-trier.de/>

<http://urts55.uni-trier.de:8080/Projekte/KoZe2/>

<http://fud.uni-trier.de/>

PERSPEKTIVEN DER LANGZEITARCHIVIERUNG DER UNIVERSITÄT TRIER

P. Leinen

Wie andere Forschungseinrichtung, so steht auch die Universität Trier vor der Herausforderung, Forschungsergebnisse und auch Primärdaten zugänglich zu machen und den Zugriff auf diese Daten dauerhaft zu gewährleisten. Zusammen mit den Forschenden stellt sich die Universität Trier dieser Aufgabe.

Das "Zentrum für Informations-, Medien- und Kommunikationstechnologie" (ZIMK) hat ein Konzept für den Aufbau und Betrieb eines universitätsweiten Archivs erarbeitet, das auch die spezifischen Anforderungen des Sonderforschungsbereichs 600 "Fremdheit und Armut" berücksichtigt.

Zunächst soll im Rahmen eines Projektes ein Archiv für die Ergebnisse und Daten des Sonderforschungsbereiches aufgebaut und die Grundlage für eine universitätsweite Infrastruktur geschaffen werden. Eine Besonderheit wird dabei die Archivierung von FuD (Forschungsnetzwerk und Datenbanksystem), einem Web-basierten, Datenbank-gestützten Werkzeug zur Erfassung und Vernetzung von Informationen sein.

Im Vortrag werden Aspekte der Archivierung von FuD sowie Rahmenbedingungen und das Arbeitsprogramm zum Aufbau eines prototypischen Langzeitarchivs an der Universität Trier vorgestellt.

Links:

<http://www.uni-trier.de/index.php?id=34>

<http://fud.uni-trier.de/>

PERSPEKTIVEN DER LANGZEITARCHIVIERUNG DER BSB

G. Horstkemper

Eine Institution, die über große Altbestände verfügt, als Forschungsbibliothek agiert und zugleich landesbibliothekarische Funktionen ausübt, wird auf mehreren Wegen mit großen Mengen an digitalen Publikationen konfrontiert:

- Handschriften und Drucke werden in umfangreichen Digitalisierungsprojekten zugänglich gemacht

- E-Journals, E-Books und Fachdatenbanken werden für die lokale oder überregionale Bereitstellung lizenziert

- Im Rahmen der eigenen Publikationsinfrastruktur werden digitale Forschungsveröffentlichungen hergestellt

- Amtliche Schriften werden als elektronische Veröffentlichungen eingeliefert

Ähnlich wie Bibliotheken für den Erhalt ihres hergebrachten Bestands an Druckpublikationen sorgen, wird von ihnen erwartet, dass sie auch die Verantwortung für die langfristige Archivierung der elektronischen Publikationen übernehmen.

Auch wenn moderne Bibliotheken ab einer bestimmten Größenordnung über eine leistungsfähige Netz- und Speicherinfrastruktur verfügen, empfiehlt sich für die Langzeitarchivierung großer Mengen an E-Medien die Zusammenarbeit mit technischen Partnern. Am Beispiel des BABS-Projektes (Bibliothekarisches Archivierungs- und Bereitstellungssystem) wird exemplarisch die Zusammenarbeit zwischen der Bayerischen Staatsbibliothek und dem Leibniz-Rechenzentrum vorgestellt. Kooperation ist jedoch auch im nationalen und internationalen Maßstab notwendig, z. B. im Bereich der Standardisierung und des Erfahrungsaustauschs.

Abschließend werden Zukunftsperspektiven für die Langzeitarchivierungs-Aktivitäten der Bayerischen Staatsbibliothek aufgezeigt. Mit der Einführung des LZA-Systems Rosetta werden die bereits existierenden Strukturen um weitere Funktionalitäten erweitert und leistungsfähiger gemacht.

Links:

<http://www.babs-muenchen.de/>

TEXTGRID - VIRTUELLE FORSCHUNGSUMGEBUNG, FORSCHUNGSDATEN UND LANGZEITARCHIVIERUNG (BMBF)

Heike Neuroth

The TextGrid research group, a consortium of 10 research institutions in Germany, is developing a Virtual Research Environment (VRE) for researchers in the arts and humanities that provides services and tools for the analysis of text data and supports the curation of research data by means of grid technology. Libraries and data centres as well as universities and research institutions are collaborating in a community-driven process that is funded by the German Ministry for Education and Research (BMBF). Initially consisting of two academic communities in 2006, textual philology and linguistics, the TextGrid project was joined by art history, classical philology, and musicology in 2009. As part of the German grid initiative DGrid, TextGrid maintains a common grid resource centre in Göttingen together with grid projects in physics, medicine, astronomy, and climate research. The TextGrid VRE consists of two main components: the TextGrid Lab(oratory), which serves as the entry point to the virtual research environment, and the TextGrid Rep(ository), which is a long-term

humanities data archive ensuring sustainability, interoperability and long-term access to research data.

To support all stages of the research lifecycle, preserve and maintain research data and ensure its long-term usefulness, existing research practices must be supported.

Therefore the TextGridLab provides common functionalities in a sustainable environment to intensify re-use of data, tools and services and the TextGridRep enables researchers to publish and share their data in a way that supports long-term availability and reusability.

TextGridRep is ensuring sustainability, interoperability and long-term access to research data. Researchers can decide how and with whom their data will be shared by using the detailed rights management (based on RBAC). They can publish their findings and research data from the TextGridLab in the repository, and archives and other institutions can ingest enormous amounts of data into the repository via a special interface that uses koLibRI, which supports for example automatic metadata validation. On a basic level TextGrid will offer bitstream preservation with redundant storage and tape backup for 10 years (as recommended in the guidelines of the German Research Foundation). Long-term bitstream preservation and higher security levels such as further distributed storage on multiple sites will be available at greater cost.

All data will be addressable via persistent identifiers that TextGrid will allocate by using a reliable handle service that is provided by the local data centre, which is a main developing partner in EPIC. A portal solution will enable rapid searching across public research data via a second instance of the search utility without connection to the rights management. An open REST interface for individual portal solutions will be provided.

Higher-value long-term preservation services will be provided in 2011 by making use of developments within the WissGrid project, which is also part of the German Grid Initiative and consists of five academic communities from the natural sciences and TextGrid from the humanities. The project is developing a service framework that fulfils more sophisticated long-term preservation needs like a provenance service, metadata extraction, format validation and conversion. Guidelines will be adapted to the specific needs of the humanities and be incorporated in the virtual research environment. The grid storage and all connected resources will be maintained together with those from the other academic disciplines at the common grid resource centre in Göttingen (e.g., 275 terabytes for the humanities). There are plans for the migration of the current repository infrastructure to Fedora and iRODS or dCache/SRM that will probably be implemented in 2011.

Links:

TextGrid (<http://www.textgrid.de/en.html>)

The D-Grid Initiative (<http://www.d-grid.de/>)

GoeGrid (<http://www.goegrid.de/>)

kopal Library for Retrieval and Ingest (http://kopal.langzeitarchivierung.de/index_koLibRI.php.de)

Proposals for Safeguarding Good Scientific Practice:

(http://www.dfg.de/download/pdf/foerderung/rechtliche_rahmenbedingungen/gute_wissenschaftliche_praxis/self_regulation_98.pdf)

The European Persistent Identifier Consortium (<http://www.pidconsortium.eu>)

WissGrid: Grid for Science (http://www.wissgrid.de/index_en.html)

DIGITALE QUELLEN SICHERN UND VERFÜGBAR HALTEN: DIE LÖSUNG DES BUNDESARCHIVS

M. Hollmann

Eine zentrale Aufgabe des Bundesarchivs ist es, Unterlagen von Bundesbehörden, die dort nicht mehr dauerhaft benötigt werden, zu übernehmen, zu bewerten und - soweit sie von dauerhaftem Wert sind - zu erschließen, damit sie anschließend der Allgemeinheit als Archivgut des Bundes und so als Quellen für die Erforschung der deutschen Geschichte zur Verfügung stehen. Die Pflicht der Behörden zur Anbietung ihrer archivreifen Unterlagen gilt ohne Einschränkung für alle Überlieferungsformen und Informationsträger, also auch für genuin elektronische Unterlagen, d.h. für elektronische Akten und sonstige Dokumente behördlichen Ursprungs, für elektronische Fachanwendungen sowie für digitale Bilder, Filme und Töne.

Die Anfänge der Archivierung genuin elektronischer Unterlagen gehen in die frühen 1990er Jahre zurück, als das Bundesarchiv umfangreiche Datenbankanwendungen aus den zentralen Rechenzentren der ehemaligen DDR zu übernehmen und für die Wahrung berechtigter Belange von Bürgern zugänglich machen musste, insbesondere für die Feststellung von Versorgungsansprüchen.

An der Diskussion um die Einführung der "elektronischen Akte" hat das Bundesarchiv sich frühzeitig beteiligt. Ein besonderer Schwerpunkt seiner Tätigkeit lag und liegt auf der Mitarbeit an der Normierung und Standardisierung auf europäischer Ebene (Moreq), auf deutscher Ebene (DIN-NABD 15) und im Bereich der Bundesverwaltung (DOMEA).

Beim Aufbau eines eigenen, den Prinzipien des OAIS folgenden Archivierungssystems hat das Bundesarchiv sich zunächst auf den Bereich elektronischer Akten und Dokumente aus Vorgangsbearbeitungs- (VBS) und Dokumentenmanagementsystemen (DMS) konzentriert. Dabei musste das Bundesarchiv davon ausgehen, dass ihm e-Akten und Dokumente uneinheitlicher Formate und Struktur angeboten werden. Die erste Ausbaustufe des Digitalen Archivs, die 2008 in den Wirkbetrieb übernommen wurde, ist daher in der Lage, mit Hilfe des Standardarchivierungsmoduls (SAM) elektronische Unterlagen aus heterogenen Systemen zu übernehmen und im Rahmen des Ingest-Prozesses hinsichtlich der Datenstrukturen, der Datenformate und der Metadaten zu vereinheitlichen. Für die Konvertierung und Migrierung der Dateiformate wurde eine spezielle Konvertierungsplattform und für die Vereinheitlichung der Datenstrukturen ein "Strukturierungsektor" entwickelt. Ein hierarchisches Speichersystem, das redundant an den Dienstorten Koblenz und Berlin gehalten wird, übernimmt die physikalische Sicherung der Daten.

Die Konvertierungsplattform stellt auch das Herzstück der Langzeiterhaltungsstrategie dar. Sie wird jedoch ergänzt um ein Werkzeug zur langfristigen Beobachtung der Entwicklung von Dateiformaten und zur Steuerung von Konversions- und Migrationsprozessen, das derzeit entwickelt wird.

Die archivische Erschließung und die Bereitstellung der Erschließungsdaten für Benutzer im Bundesarchiv und über das Internet erfolgen unter Einbindung des Archivverwaltungssystems des Bundesarchivs BASYS. So wird - ungeachtet der verschiedenen Typen und Formen - eine integrierte Präsentation des gesamten Archivguts aus dem Bereich der Bundesverwaltung sichergestellt. Für die Bereitstellung elektronischen Archivguts stehen spezielle Viewer zum Teil bereits zur Verfügung, zum Teil befinden sie sich in Entwicklung.

Spezielle elektronische Fachanwendungen einzelner Bundesbehörden werden, sofern sie als archivwürdig bewertet werden, ebenfalls beim Ingest in ein standardisiertes Format überführt, damit sie später unabhängig vom Entstehungssystem über ein Datenbanksystem für die Benutzung bereitgestellt werden können. Die entsprechende Systemkomponente des Digitalen Archivs befindet sich gleichfalls in Entwicklung.

Geplant ist der Ausbau des Digitalen Archivs des Bundesarchivs zu einem zentralen Dienstleistungszentrum zunächst für die Bundesministerien und in einem weiteren Schritt für die gesamte Bundesverwaltung, das den Bundesbehörden die mittelfristige Sorge um die Verfügbarkeit elektronischer Dokumente abnimmt. Wie schon seit mehr als 40 Jahren im Bereich klassischer Akten soll ein Digitales Zwischenarchiv elektronische Akten und Dokumente unmittelbar dann übernehmen, wenn sie für die aktuelle Aufgabenwahrnehmung nicht mehr benötigt werden. Im Digitalen Zwischenarchiv sollen sie dann solange vorgehalten werden, bis eine weitere Speicherung aus Sicht der Behörde nicht mehr geboten ist. Soweit die elektronischen Unterlagen dauerhaften archivistischen Wert haben, sollen sie in das Digitale Archiv überführt werden, falls das nicht der Fall ist, werden sie endgültig gelöscht. Ein derartiges Dienstleistungszentrum bietet langfristig ein erhebliches Kostenvermeidungspotential und entlastet die beteiligten Bundesbehörden weitgehend von der Entwicklung und dauerhaften Unterhaltung eigener Systeme zur mittelfristigen Datenspeicherung.

Einen Sonderfall stellt das Digitale Bildarchiv dar. Es wurde nicht primär mit dem Ziel der langfristigen Sicherung genuin digitaler Bilder aufgebaut, sondern als ein modernes Medium der Präsentation und Bereitstellung von überwiegend analogem Bildmaterial für wissenschaftliche, journalistische und andere Publikationszwecke.

Link:

www.bundesarchiv.de

DAS PERSONENDATEN-REPOSITORY DER BBAW

G. Neumann, T. Roeder

Die Berlin-Brandenburgische Akademie der Wissenschaften baut, unterstützt durch die DFG, ein Personendaten-Repository auf, in dem sie biographische Daten, die in verschiedenen Forschungsvorhaben und unter unterschiedlichen wissenschaftlichen Perspektiven entstanden sind, sammelt und für eine wissenschaftliche Auseinandersetzung aufbereitet.

Ein zentraler Bestandteil der Arbeiten des ersten Jahres war eine ausführliche Evaluation von Repository-Software, die für den Aufbau eines Personendaten-Repositorys geeignet ist. Ein weiterer Bestandteil war die Erarbeitung eines Inhaltsmodells für biographische Forschungsdaten, die aus unterschiedlichen Kontexten stammen. Das Ziel war die konsequente strukturelle Isolierung der biografischen Daten von der Person, die stringente Atomisierung der Informationen, eine semantische Erschließungsmöglichkeit der Daten und die Möglichkeit, Beziehungen zwischen Daten aus verschiedenen Quellen transparent abbilden zu können. Der dritte, aktuelle Bestandteil ist die Integration größerer Datenmengen in das Repository, anhand des vordefinierten Inhaltsmodells.

Der Vortrag gibt einen kurzen Überblick über das Projekt. Es werden das Inhaltsmodell und die sich daraus ergebenden Konsequenzen für die Datenaufbereitung beschrieben. Anhand der aktuell verfügbaren Daten werden praktische Anwendungsfälle, die sich aus dem derzeitigen Entwicklungsstand des Personendaten-Repositorys ergeben, illustriert.

Link:
<http://pdr.bbaw.de/>

REPOSITORIENINFRASTRUKTUR DER MPDL

M. Dreyer

eSciDoc wurde als eResearch-Plattform speziell für die Benutzung in Forschung und Wissenschaft entwickelt, um die geografisch verteilte und disziplinübergreifende Zusammenarbeit in virtuellen Arbeitsgruppen zu unterstützen. Es umfasst Kernfunktionen inklusive eines zentralen Repositories ("eSciDoc Infrastructure"), diversen ergänzenden Diensten ("eSciDoc Services") und Applikationen, die auf der Infrastruktur aufsetzen und nach Bedarf ergänzende Dienste einbinden ("eSciDoc Solutions"). Diese Applikationen stellen bereits einige interessante Lösungen für Forscher bereit, aber der tiefergehende Nutzen von eSciDoc liegt in der Bereitstellung eines "Baukastens" im Sinne einer serviceorientierten Architektur (SOA), über die sich einfach und schnell weitere e-Research-Szenarien (zum Beispiel sogenannte "Virtual Research Environments") realisieren lassen. Basierend auf der eSciDoc Infrastruktur entwickelt die MPDL ein Set an Services und Anwendungen, die als Repository für unterschiedliche Forschungsdaten, u.a. Publikationsdaten, in der MPG eingesetzt werden. Der Beitrag skizziert ein möglichst umfassendes Bild der Ausgangslage, der fachlichen und technischen Anforderungen sowie des gewählten Entwicklungsansatzes und der entwickelten Dienste.

Link:
<http://www.mpd1.mpg.de>

ÖFFENTLICHER VORTRAG

REPOSITORIEN - DIGITALMÜLL DER ZUKUNFT ODER ZUKUNFT DIGITALER BIBLIOTHEKEN?

G. Schneider

Die Vorteile digitaler Dokumente liegen auf der Hand: man kann die Daten, die auf einem Server liegen, quasi mit Lichtgeschwindigkeit über das Internet an den eigenen Arbeitsplatz kopieren. Die Präzision dieser umständlich klingenden Aussage ist beabsichtigt, insbesondere im Licht der aktuellen Tagespolitik mit ihren Patentrezepten für das Internet. Da in der Vergangenheit Dokumente auf schwer beweglichem Papier erschienen sind, wird in vielen Projekten die Digitalisierung der vorhandenen Schätze - und da gibt es viele - vorangetrieben. Ein bekanntes und erfolgreiches Vorhaben ist das von der DFG geförderte VD16/VD17. Ein Nutzer steht aber viel zu oft vor dem Problem, herausfinden zu müssen, ob ein gesuchtes Dokument bereits digitalisiert ist und wo es gefunden werden kann. Der Ruf nach der zentral ordnenden Hand erschallt zwar immer wieder - auch in Form einer Projektförderung -, ist jedoch aufgrund der Natur des Internet und der menschlichen Sozialisierung (insbesondere in einem föderalen System) zum Scheitern verurteilt. Immerhin sind Repositorien entstanden, um dem lokalen Vergessen entgegenzuwirken und die Auffindbarkeit zu steigern. Repositorien sind lediglich geordnete Datensammlungen mit einer mehr oder weniger definierten Verantwortlichkeit. Eine Datensammlung, die von einer Bibliothek betrieben wird, genießt aufgrund der Vorgeschichte und des Ansehens dieser

Einrichtung einen Vertrauensvorschuss in Bezug auf die langfristige Verfügbarkeit, der im Einzelfall aber nicht begründet sein muss und den Streichungen der Finanzmittel durch die Geldgeber hilflos ausgesetzt sein wird.

Das Urheberrecht ist überraschenderweise eine große Bedrohung für Repositorien. Viele Datensammlungen sind heute unvollständig, weil die Rechte an jüngeren Digitalisaten nicht geklärt werden können - oder schlichtweg nicht übertragen werden. Öffentliche Institutionen verfügen nicht über die Möglichkeit großer Firmen wie Google, sich im Streitfall das Recht einfach zu kaufen - bzw. den begangenen Rechtsbruch ohne Folgen zu heilen. Ein jüngst im Spiegel (<http://www.spiegel.de/spiegel/0,1518,709761,00.html>) erschienener Artikel beschreibt überraschend negative Folgen des Urheberrechts in der Geschichte.

Die Zentralisierung auf wenige Betreiber von Repositorien - also Besitzern von Daten und damit Information - ist per se gefährlich. Nicht nur, weil im Orwellschen Sinne störende Dokumente einfach ohne Spuren ausgetauscht werden können - wie oft werden Versuche der Geschichtsfälschungen durch Funde vergessener / nicht vernichteter Dokumente erkannt? Diese Wenigen könnten auf die Idee verfallen, den Zugriff auf ihre Datensammlungen zu reglementieren und einzuschränken. Der Ausschluss irakischer Wissenschaftler von der Nutzung US-amerikanischer Wissensdatenbanken ist ein Beispiel hierfür, wenngleich die Ursache in diesem Fall für uns verständlich erscheint. Ohne belastbare langfristige Verträge, sowohl national wie international, und entsprechender gesetzlicher Grundlagen sind von privater Hand oder ausländischen Institutionen betriebene Repositorien nicht ohne Gefahr für die Gesellschaft.

Dagegen sehen die technischen Herausforderungen nahezu beherrschbar aus. Fast allen Nutzern ist das technische Problem der Trägermedien vertraut: Daten, die einst mit Liebe auf Disketten geschrieben wurden, sind heute nicht mehr lesbar - weil es keine Laufwerke mehr gibt. Die CD-ROM, einst als dauerhaftes Speichermedium gepriesen, weist eine beschränkte technische Lebensdauer von 1-50 Jahren auf und wird dabei trotzdem vermutlich die Verfügbarkeitsdauer der Lesegeräte übertreffen. Das Problem der veraltenden Trägermedien ist im Prinzip geklärt: man kopiert die Daten automatisch (mittels Robotern) auf die jeweils aktuelle Trägergeneration. Die logistischen Herausforderungen sind dennoch nicht unbeträchtlich. Wohl organisierte Repositorien finden ihre Daten dann auch am neuen Speicherort wieder; für die Nutzer sind die Vorgänge vollkommen transparent. Kopiergeschützte CD-ROMs hingegen fallen bei solchen Prozessen unter den Tisch.

Die viel größere Herausforderung stellt die Interpretierbarkeit der Dokumente bzw. ihre Darstellbarkeit dar. So kann selbst Word2010 nicht mehr eine Publikation einlesen, die mit einem frühen Vorläufer von Word2010 erstellt wurde. Man könnte glauben, dass solche Schwierigkeiten durch permanente aktive Dokumentenmigration in die jeweils aktuelle Darstellungsform gelöst werden könnten - verbunden mit der Hoffnung, dass das Ergebnis 100% mit dem Original übereinstimmt. Dieser Ansatz versagt jedoch völlig, wenn die gesamte Laufzeitumgebung notwendig ist, um ein Ergebnis darstellen zu können. Hier hilft nur die Emulation, also die softwaremäßige Aufbereitung der Vergangenheit auf aktuellen Rechnern. Gerade zu diesem Thema laufen in Freiburg, im Rahmen des EU-Projekts PLANETS, aktuelle Forschungsarbeiten mit sehr erfolgversprechenden Ergebnissen.

Es zeigt sich, dass auf die Betreiber von Repositorien wesentlich höhere Anforderungen als zu Beginn erwartet zukommen, wenn sie ihre Daten auch für die Zukunft nutzbar halten wollen. Diese Anforderungen gehen mit Betriebskosten und Betriebskompetenz einher. Werden die oben aufgezeigten Transportketten in der Zukunft nur einmal unterbrochen,

dann sind die gespeicherten Daten nur noch Datenmüll, weil niemand sie mehr zu interpretieren vermag.

Wenn Repositorien aber diese Herausforderung meistern, so werden sie unverzichtbare Grundpfeiler der Informationsvermittlung, die das Wissen in die fernsten Ecken der vernetzten Welt liefern können - ähnlich wie der Buchdruck einst zur Wissensverbreitung beitrug.

EINZELASPEKTE UND KOOPERATIONEN

RECHTLICHE PROBLEME DER DIGITALEN LANGZEITARCHIVIERUNG WISSENSCHAFTLICHER PRIMÄRDATEN

T. Hillegeist

Im Rahmen der elektronischen Langzeitarchivierung wissenschaftlicher Primärdaten können vor allem urheberrechtliche und datenschutzrechtliche Probleme auftreten.

Dabei lässt sich zunächst feststellen, dass wissenschaftliche Primärdaten grundsätzlich nicht dem Schutz des Urheberrechtsgesetzes unterliegen. Dies gilt in der Regel auch dann, wenn die Daten in Tabellen oder auf andere Art zusammengefasst werden. Eine solche Zusammenstellung könnte zwar ein Datenbankwerk nach § 4 Abs. 2 UrhG oder eine Datenbank gem. § 87a UrhG darstellen. Zum einen wird es dafür jedoch regelmäßig an der erforderlichen geistigen Schöpfungshöhe fehlen, die bei Datenbankwerken nach § 4 Abs. 2 UrhG in der individuellen Auswahl oder Anordnung der Daten bestehen muss. Eine solche Individualität wird in den vorliegend relevanten Fällen grundsätzlich nicht vorliegen, da die Anordnung nach logischen Gesichtspunkten erfolgen wird.

Zum anderen wird der sui generis-Schutz der Datensammlung nach § 87a UrhG in den meisten Fällen daran scheitern, dass für die vorliegend in Betracht kommenden Datensammlungen in der Regel keine wesentliche Investition im Sinne der Vorschrift erforderlich ist. Investitionen werden vielmehr bei der Datenerhebung getätigt werden, deren Kosten jedoch im Rahmen des § 87a UrhG gerade nicht zu berücksichtigen sind. Trotzdem sollte im Einzelfall stets genau geprüft werden, ob nicht ausnahmsweise doch eine Datenbank i.S.d. § 87a UrhG vorliegt.

Sofern dies der Fall sein sollte, müsste die archivierende Einrichtung prüfen, ob sie Datenbankherstellerin und damit Inhaberin der erforderlichen Nutzungsrechte ist, da bezüglich der Archivierungsmaßnahmen und einer eventuellen öffentlichen Zugänglichmachung der archivierten Daten grundsätzlich keine urheberrechtlichen Schranken zugunsten der archivierenden Einrichtung greifen.

Im Rahmen der elektronischen Archivierung werden dabei vor allem das Recht der Vervielfältigung und das Recht der öffentlichen Zugänglichmachung im Internet betroffen. Insbesondere in den Fällen von Drittmittelprojekten sowie den Fällen der Auftragsforschung, in denen ein Dritter das Investitionsrisiko trägt, ist daher zu prüfen, ob die Rechte des Datenbankherstellers nicht bei der finanzierenden Einrichtung liegen.

Sofern die Nutzungsrechte ursprünglich bei einem Angestellten der archivierenden Einrichtung entstanden sind, könnte diese sie bereits aufgrund des Arbeitsverhältnisses erlangt haben, da insoweit der Grundsatz gilt, dass die Nutzungsrechte, die ein Arbeitnehmer im Rahmen seiner Angestelltentätigkeit erlangt, dem Arbeitgeber zustehen. Zu beachten ist aber, dass dieser Grundsatz aufgrund der durch Art. 5 Abs. 3 GG garantierten Wissenschaftsfreiheit nicht für Hochschulprofessoren gilt, so dass hier nur dann von einem

Übergang der Rechte ausgegangen werden kann, wenn der jeweilige Arbeitsvertrag eine ausdrückliche Rechteübertragungsklausel enthält.

Hat die archivierende Einrichtung die erforderlichen Nutzungsrechte nicht bereits aufgrund eines Arbeitsvertrages oder der Vorschrift des § 137 I UrhG erlangt, bleibt ihr nur die Möglichkeit, sich die benötigten Nutzungsrechte von dem Rechteinhaber rechtsgeschäftlich übertragen zu lassen. Relevant ist dies vor allem bei Forschungsprojekten, bei denen auch ehrenamtliche Beiträger mitgewirkt haben und demzufolge in keinem Angestelltenverhältnis zu der archivierenden Einrichtung standen.

Aus dem oben Gesagten folgt, dass Forschungseinrichtungen besonderes Augenmerk auf die Ausgestaltung ihrer Arbeitsverträge legen sollten. Insbesondere sollten sie detaillierte Rechteübertragungsklauseln in ihre Arbeitsverträge einfügen, um spätere Rechtsunsicherheiten hinsichtlich der Inhaberschaft von Nutzungsrechten von vornherein zu vermeiden. Aufgrund des im Urheberrecht geltenden Bestimmtheitsgrundsatzes genügt hierfür die pauschale Vereinbarung, dass "... alle bestehenden Rechte übertragen werden ..." hingegen nicht.

Rechtliche Probleme können sich bei der Archivierung wissenschaftlicher Primärdaten ferner aus dem Bundesdatenschutzgesetz, den einzelnen Landesdatenschutzgesetzen sowie dem Sozialgesetzbuch X ergeben, sofern die zu archivierenden Daten einen Personenbezug aufweisen. Dies ist vor allem bei medizinischen Forschungsdaten denkbar.

Aus diesem Grund müssen archivierende Einrichtungen vor der Archivierung personenbezogener Daten eine formgerechte Einwilligung des jeweiligen Betroffenen einholen. Sollte dies nicht möglich sein, könnte eine Archivierung der Daten aufgrund der Erlaubnisnorm des § 10 NDSG oder einer entsprechenden Vorschrift der übrigen Landesdatenschutzgesetze zulässig sein. Dies gilt jedoch nur dann, wenn die betreffenden Daten weder zu einem anderen Zweck als dem der wissenschaftlichen Forschung erhoben worden sind noch im Rahmen eines anderen Forschungsvorhabens gewonnen wurden.

Sofern eine wirksame datenschutzrechtliche Einwilligung des Probanden vorliegt, ist darin außerdem gleichzeitig eine (zumindest konkludent erteilte) Entbindung des behandelnden Arztes von seiner ärztlichen Schweigepflicht zu sehen, was vor allem dann relevant ist, wenn medizinische Untersuchungsdaten archiviert werden sollen.

Die Entbindung von der Schweigepflicht entspricht dabei konsequenterweise in ihrer Reichweite dem Umfang, in welchem der Proband auch in die datenschutzrechtlich relevante Nutzung seiner Daten eingewilligt hat.

Der Betroffene kann seine Einwilligung allerdings jederzeit widerrufen, womit diese mit Wirkung für die Zukunft entfällt. Des Weiteren ist aus datenschutzrechtlicher Sicht zu beachten, dass die Daten sowohl nach den Vorschriften des BDSG als auch nach den Landesdatenschutzgesetzen zu anonymisieren sind, sobald der Forschungszweck dies zulässt.

Sofern dies nicht möglich sein sollte, hat die archivierende Einrichtung die Merkmale, mit denen ein Personenbezug hergestellt werden kann, zumindest gesondert zu speichern.

Bevor personenbezogene Daten archiviert werden, ist also stets zu prüfen, ob die Merkmale, mit denen ein Personenbezug hergestellt werden kann, wirklich noch erforderlich sind. Einer solchen Prüfung sind darüber hinaus regelmäßig die bereits archivierten Daten zu unterziehen. Da die elektronische Archivierung personenbezogener Daten regelmäßig eine automatisierte Datenverarbeitung darstellt, hat die archivierende Einrichtung die jeweiligen datenschutzrechtlichen Vorschriften hinsichtlich der zu ergreifenden technischen und organisatorischen Maßnahmen zu beachten, namentlich des § 9 S. 1 i.V.m. der Anlage zu § 9

S. 1 BDSG sowie des § 7 Abs. 2 NDSG beziehungsweise der entsprechenden Vorschriften der übrigen Landesdatenschutzgesetze.

Sofern die archivierende Einrichtung gegen die datenschutzrechtlichen Vorgaben verstößt, kann dies zum einen eine Straftat oder Ordnungswidrigkeit darstellen. Darüber hinaus stehen dem Betroffenen unter Umständen auch Schadensersatzansprüche zu, wobei sowohl das BDSG als auch die Landesdatenschutzgesetze in den Fällen einer automatisierten Datenverarbeitung eine verschuldensunabhängige Haftung vorsehen.

In den Fällen einer nicht automatisierten Datenverarbeitung kann sich die archivierende Einrichtung zwar exkulpieren, trägt allerdings die Beweislast für das fehlende Verschulden. Besondere Bedeutung erlangt die verschuldensunabhängige Haftung im Rahmen der Auftragsdatenverarbeitung, da der Anspruchsgegner in diesen Fällen der Auftraggeber, im Rahmen der Langzeitarchivierung also die archivierende Einrichtung bleibt. Diese sollte aus diesem Grund in die Verträge mit der beauftragten Einrichtung für den Fall eines Verstoßes gegen datenschutzrechtliche Vorschriften und eines daraus entstehenden Schadensersatzanspruches des Betroffenen eine Freistellungsklausel in den Vertrag aufnehmen.

Abschließend lässt sich feststellen, dass der rechtliche Problemschwerpunkt der elektronischen Langzeitarchivierung wissenschaftlicher Primärdaten nicht, wie man zunächst annehmen könnte, urheberrechtlicher, sondern datenschutzrechtlicher Natur ist. Diese Probleme lassen sich aus Sicht der archivierenden Einrichtung insoweit relativ einfach vermeiden, indem nur Daten ohne Personenbezug archiviert werden. Nichtsdestotrotz sollte vor jeder Digitalisierung unbedingt genau geprüft werden, ob nicht aus den oben beschriebenen Gründen ausnahmsweise doch ein urheberrechtlicher Schutz an der zu archivierenden Datenbank besteht. Unabhängig davon sollte bei dem Abschluss neuer Arbeitsverträge ein besonderes Augenmerk auf eine klar und detailliert formulierte Rechteübertragungsklausel gelegt werden.

Hinweis:

Zum Problem vgl. den Beitrag Spindler/Hillegeist, Langzeitarchivierung wissenschaftlicher Primärdaten, in: nestor Handbuch - Eine kleine Enzyklopädie der digitalen Langzeitarchivierung, Version 2.3, Kap. 16:14, abrufbar unter http://nestor.sub.uni-goettingen.de/handbuch/nestor-handbuch_23.pdf.

DIGITALISIERUNGSSTRATEGIEN IN DER SAGW

S. Zala

Die Schweizerische Akademie der Geistes- und Sozialwissenschaften (SAGW) hat im letzten Jahr ein in sich kohärentes strategisches Maßnahmenpaket zur Informationsversorgung insbesondere für die Geisteswissenschaften entworfen, das sich mit der Formel "Bündeln - Digitalisieren - Vernetzen - Sichern" umschreiben lässt. Die Fokussierung auf die Geisteswissenschaften ist Ausdruck eines Nachholbedarfs, den die Akademie gegenüber den Sozialwissenschaften konstatiert.

Bündelung der Ressourcen: Die SAGW ist neben dem Schweizerischen Nationalfonds (SNF) eine der bedeutendsten Trägerinnen geisteswissenschaftlicher Forschungsinfrastrukturen, darunter die Diplomatischen Dokumente der Schweiz (DDS), infoclio.ch oder das Historische Lexikon der Schweiz (HLS). Die Akademie ist bestrebt, diese Position in den kommenden Jahren auszubauen.

Digitalisierung und Open Access: Die SAGW verfolgt eine Vorwärtsstrategie und unterstützt deshalb Digitalisierungsvorhaben in verschiedenen Bereichen. Das jüngste Projekt ist die Online-Schaltung des Wörterbuchs der Schweizerdeutschen Sprache. Ferner sind annähernd die Hälfte der 47 durch die SAGW unterstützten Periodika mittlerweile online zugänglich, die meisten über die Retrodigitalisierungsplattform retro.seals.ch, einem Projekt der Elektronischen Bibliothek Schweiz e-lib.ch. Eng mit der Retrodigitalisierung verbunden, ist auch das Engagement der SAGW für Open Access.

Vernetzungsinitiative: In der Überzeugung, dass die bereit gestellten digitalen Informationen einen erheblichen Mehrwert erfahren, wenn sie miteinander in Beziehung gesetzt werden, unterstützt die SAGW eine Vernetzungsinitiative verschiedener Ressourcen. Im Zentrum steht dabei der von der Forschungsstelle der Diplomatischen Dokumente der Schweiz (DDS/dodis.ch) entworfene Webservice metagrid.ch, der in einer ersten Phase die beiden Ressourcen HLS und DDS miteinander verbindet. Kernstück des Webservices ist eine Datenbankapplikation, die eine frei zu definierende Anzahl an Konkordanzen zwischen den partizipierenden Informationsquellen enthält, die den dynamischen Austausch von Informationen zu Personen, Organisationen, Sachbegriffen, geographischen Orten etc. erlaubt. Für eine Teilnahme an diesem Webservice konnten auch bereits das Projekt LONSEA (League of Nations Search Engine) des Exzellenzclusters der Universität Heidelberg sowie die Schweizerische Nationalbibliothek gewonnen werden. Mittelfristig plant die Akademie, über dieses System der Webservices in verschiedenen Epochenschwerpunkten ein "Historisches Informationssystem Schweiz" zu errichten.

Nachhaltige Sicherung digitaler Forschungsdaten: Die Frage der Langzeitarchivierung von Forschungsdaten wird zurzeit durch das Schweizerische Bundesarchiv geprüft. Für die Forschung geht aber die SAGW von einem Primat des Dienstleistungsangebots vor der Langzeitarchivierung aus. In diesem Zusammenhang unterstützt sie die Entwicklung zweier Forschungswerkzeuge: Dodis 2.0 und SALSAH, ein Editionstool für Quellen und ein Forschungs- und Annotationstool für die Bearbeitung von Digitalisaten.

Informationen finden Sie unter:

<http://www.sagw.ch/>

<http://www.sagw.ch/de/sagw/die---akademie/downloads/mjp2012---2016.html>

<http://www.sagw.ch/de/sagw/die-akademie/unternehmen.html>

<http://www.sagw.ch/open---access>

<http://www.dodis.ch>

<http://www.infoclio.ch>

<http://www.hls.ch>

LANGZEITARCHIVIERUNG VON LITERATUR IM DLA

R. Kamzelak

Jenny Erpenbeck ist aus Sicht der Archive eine ganz gewöhnliche Schriftstellerin. Das Deutsche Literaturarchiv Marbach beobachtet ihre Entwicklung, kauft ihre Bücher, lädt Sie zu Lesungen ein, vielleicht sogar zu einem Workshop mit Schülern im Rahmen seines pädagogischen Programms. Ihr neuester Roman erzählt die Geschichte eines Hauses durch die Jahrhunderte hindurch. Und mit der Geschichte dieses Hauses die Geschichte des geteilten und wiedervereinigten Deutschlands. Es ist ein gutes Buch, auch ein wichtiges Buch, es bewahrt ein Stück kulturelles Erbe und will als kulturelles Erbe verwahrt werden.

Was ist ein Literaturarchiv? Es ist ein verlässlicher Speicherort für literarische Nachlässe, und deshalb oder auch ein Ort der Forschung an diesen oder über diese Nachlässe. Verlässlich in

der Archivwelt bedeutet eine Spanne von 500 Jahren. Es besteht aus Papier, Tinte ... und einer Menge Problemen, denn Nachlässe sind NACH-lässe, weil wir sie "erben", also erst spät zu uns bekommen.

Jedes Literaturarchiv ist auch bereits längst elektrifiziert und besteht zu einem großen Teil aus digitalen Erst- und Zweitformen, auch wenn der allgemeine Zugang zu diesem digitalen Material wegen bestehender Rechte noch eingeschränkt ist.

Die Frage ist, wie können wir ein solches verlässliches Archiv sein? Wie ein verlässlicher Gedächtnisort?

Link:

www.dla-marbach.de

REPOSITORIEN UND LANGZEITARCHIVIERUNG AUS DER SICHT DER BIBLIOTHEKEN

Natascha Schumann, S. Strathmann

Der Beitrag gibt zunächst einen kurzen Überblick über nestor, das Kompetenznetzwerk digitale Langzeitarchivierung. Nestor ist ein Kooperationsverbund, dem Archive, Bibliotheken, Museen und Forschungseinrichtungen angehören. Ein Schwerpunkt liegt in der Standardisierung/Zertifizierung im Bereich vertrauenswürdiger Archive und Repositorien. Darüber hinaus werden Projekte und Erfahrungen der SUB Göttingen mit der Langzeitarchivierung von Forschungsdaten an einem Wissenschaftsstandort vorgestellt.

Link zum Nestor-Handbuch:

<http://nestor.sub.uni-goettingen.de/handbuch/>

ABSCHLUSSVORTRAG

GEISTESWISSENSCHAFTLICHE FORSCHUNGSKOMMUNIKATION DER ZUKUNFT IM DREIECK AUTOR – VERLAG - BIBLIOTHEK

E. Mittler

Die Tagung Repositorien zeigt, wie intensiv in geisteswissenschaftlichen Projekten daran gearbeitet wird, digitale Texte und Editionen sowie Datenbanken zu erstellen, die durch das Internet im Prinzip weltweit zugänglich sind. Konsequente Folge dieser Entwicklung ist, virtuelle Arbeitsumgebungen zu schaffen, die den Zugriff auf diese und andere digitale Dokumente und deren Integration in den Forschungsprozess zu optimieren suchen. Dieser Entwicklung entsprechen die bewährten Formen der gedruckten Forschungskommunikation vom Autor über den Verlag zur Nutzung (mit weitgehender Unterstützung durch die Bibliotheken) nur noch eingeschränkt. Sind sie - und mit ihnen insbesondere die wissenschaftliche Monographie - Auslaufmodelle oder weiterhin der Königsweg wissenschaftlichen Publizierens in den Geisteswissenschaften?

Auf diesem Hintergrund werden die Open-Access-Entwicklung in ihren unterschiedlichen Ausformungen sowie die Entwicklung des elektronischen Publizierens in Repositorien, bei kommerziellen und bei universitären Verlagen und die sich verändernde Rolle der Wissenschaftler, der Verlage und der Bibliotheken diskutiert.

LEIPZIGER EMPFEHLUNGEN DER ARBEITSGRUPPE „ELEKTRONISCHES PUBLIZIEREN“ DER UNION DER DEUTSCHEN AKADEMIEN (21. SEPTEMBER 2009)

Diese Empfehlungen konkretisieren in ihrem letzten Punkt die Münchener Empfehlungen vom 21. Juli 2006.

Die Vorhaben der Akademien erarbeiten in großer Sorgfalt geisteswissenschaftliche Grundlagenwerke, vor allem Wörterbücher und Editionen. Sie erschließen, dokumentieren und bewahren kulturelles Erbe und machen es allgemein auf Dauer zugänglich. Alle großen Forschungseinrichtungen Deutschlands wie auch die großen Bibliotheken, Archive und Museen beteiligen sich gegenwärtig an der Entwicklung von Plänen und Projekten, wie der wesentliche Teil der von ihnen verwalteten Bestände angesichts der gravierenden Umstrukturierung des Informationssektors durch das Internet der Öffentlichkeit frei zugänglich gemacht und miteinander vernetzt werden kann.

Die Arbeitsgruppe „Elektronisches Publizieren“ der Union der deutschen Akademien empfiehlt daher der Union wie auch den Einzelakademien, diese Entwicklung weiter mitzugestalten und voranzutreiben und die Ergebnisse der Akademieforschung unter weitgehender Nutzung des Vernetzungspotentials der elektronischen Medien ebenfalls elektronisch verfügbar zu machen.

Die Arbeitsgruppe empfiehlt, dass die Primärpublikation der Grundlagenwerke weiterhin in der Buchform, die als feste Referenzgröße und als wichtigstes Langzeitarchivierungsmedium erhalten bleiben muss, erfolgen soll und dass zugleich die Daten für den Buchsatz in medienneutraler Form, die prinzipiell Parallelpublikationen im Druck und Internet erlauben, von den Akademien oder in ihrem Auftrag sorgfältig archiviert werden.

Die Arbeitsgruppe empfiehlt weiter, diejenigen Akademievorhaben, die den Weg zur elektronischen Publikation bereits beschritten haben, nachhaltig zu unterstützen und diejenigen Vorhaben, die den neuen digitalen Medien noch eher skeptisch gegenüberstehen, zum elektronischen Publizieren zu ermutigen.

Die Arbeitsgemeinschaft empfiehlt schließlich, alle Neuvorhaben und Förderungsanträge zur Retro-Digitalisierung der Publikationen laufender Vorhaben nach den folgenden Essentials auszurichten:

1. Die Auszeichnung aller Texte soll auf internationalen Standards beruhen; hierfür maßgebende Auszeichnungssprache ist XML. Die XML-Daten bilden die Basis für die Ausgabe in den unterschiedlichen Publikationsmedien.
2. Auf XML aufsetzend soll die Erstellung der Metadaten und die Spezifizierung der Primärdaten nach den Richtlinien der TEI erfolgen.
3. Auf Zeichenebene ist der UNICODE-Standard einzusetzen.
4. Bilddaten sind verlustfrei (komprimiert) im TIF bzw. PNG-Format abzuspeichern.
5. Die Erfassung von Textdaten bei der Retro-Digitalisierung laufender Vorhaben soll zeichenbasiert erfolgen, dabei ist eine hundertprozentige Erfassungsqualität anzustreben.
6. Für alle Phasen der Arbeit an und mit digitalen Ressourcen soll möglichst Open Source-Software benutzt werden, deren Quelltext öffentlich zugänglich ist.

AUTORENVERZEICHNIS

B

Ball, R. **2**
Bulitta, B. **10**
Burch, T. **13**

D

Deike, A. **4**
Deutsch, A. **7**
Dörr, S. **8**
Dreyer, M. **17**

G

Graeff, J. **2**
Grieshaber, F. **3**

H

Hillegeist, T. **19**
Hollmann, M. **16**
Horstkemper, G. **14**

K

Kamzelak, R. **22**
Krüger, V. **5**

L

Leinen, P. **13**
Leuker, C. **11**

M

Marder, U. **11**
Marx, M. **5**
Meier, W. **6**
Mittler, E. **23**
Müller, M. **13**

N

Neumann, G. **17**
Neuroth, H. **15**

R

Raible, W. **23**
Reinert, M. **12**
Roeder, T. **17**

S

Schneider, G. **18**
Schrade, T. **4**
Schulz, M. **8**
Schumann, N. **23**
Strathmann, S. **23**
Strutwolf, H. **5**

W

Wenzel, C. **6**
Wesche, M. **12**
Wilhelm, G. **2**

Z

Zala, S. **21**