

# IIE aktuell

Mitteilungen aus der Abteilung Überregionale Bibliographische Dienste (II E)

## Inhalt

Allheilmittel Digitalisierung? <i>Bernd Augustin</i>	1
100 Jahre Pressedokumentation im HWWA <i>Helmut Leveknecht</i>	3
Inhaltliche und technische Darstellung der retrospektiven Digitalisierung der historischen Pressearchive der ZBW und des HWWA <i>Thomas Sergej Huck / Max-Michael Wannags</i>	7
Alles digitalisiert! Was nun? <i>Dr. Günter Mühlberger</i>	13
Archive vor der digitalen Herausforderung <i>Rainer Hering</i>	15
Digitale Archive <i>Dr. Martin Warnke</i>	20
IIE-Adressen	28

Zu dieser Ausgabe:

Da in der vorliegenden Nummer dieses Newsletters entgegen der sonstigen Gepflogenheit mehrere längere Beiträge enthalten sind, haben wir uns entschlossen, zur besseren Lesbarkeit ein zweispaltiges Layout zu verwenden.

IIE aktuell wird  
im Auftrag  
der Generaldirektorin  
der Staatsbibliothek zu Berlin  
herausgegeben  
von der Abteilung  
Überregionale Bibliographische Dienste (II E)

## Allheilmittel Digitalisierung?

### 9. Tagung der AG Zeitungen zur Digitalisierung und Langzeitarchivierung

Bernd Augustin

Digitalisierung und Langzeitarchivierung, so das offizielle Thema der 9. Tagung der AG Zeitungen im Forum Zeitschriften (GeSIG), dürfen nicht nur unter den Aspekten betrachtet werden, welche Quellen bekomme ich mit welchen Technologien, ohne zu große finanzielle und personelle Aufwendungen, in einem möglichst überschaubaren Zeitraum, aber in erstklassiger Qualität optimal präsentiert, in möglichst unbegrenzter «Haltbarkeit», sprich ohne Datenverluste trotz fortschreitender Technologieentwicklungen.

Auf dieser Tagung, die am 13. und 14. September 2005 im Hamburgischen Welt-Wirtschafts-Archiv (HWWA) stattfand und wie üblich moderiert wurde von Dr. Hartmut Walravens, wurden nicht nur die «klassischen» Digitalisierungsprojekte von Zeitungen behandelt, sondern ebenso Aspekte der Digitalisierung, die auch für nicht vordergründig nur an Zeitungen Interessierte höchst interessant sind. Dies gilt vor allem für die Beiträge von *Dr. Rainer Hering* (Staatsarchiv Hamburg) und von *Dr. Martin Warnke* (Universität Lüneburg).

Hering beschreibt die archivarischen Strategien nach innen und nach außen im Umgang mit der Digitalisierung von Unterlagen, konkret die Handlungsfelder am Beispiel der Einführung der elektronischen Dokumentenverwaltung in der Hamburger Verwaltung mit dem Namen ELDORADO. Die Aufgaben der Archive in den E-Government-Strategien der öffentlichen Verwaltungen werden aufgezeigt und die oft vernachlässigten aber erforderlichen mentalen Veränderungen durch die elektronische Dokumentenverwaltung herausgestellt.



*Frauke Nordmann*  
Foto © HWWA – Olaf Hoppe

Warnke beschäftigt sich mit Überlegungen, was die Folgen wären, wollte man einfach alles digital archivieren, beginnend vom menschlichen Gedächtnis, über die Aufzeichnung aller produzierten Daten in verschiedenen medialen Formen, seien es Printmedien, audiovisuelle Medien und sämtliche digital geführte Kommunikation. Er stellt die Frage, ob die allseits signifikante Ewigkeitsvorstellung des Digitalen tatsächlich der Realität entspricht oder nicht doch eher ein archivarischer Alptraum bleibt.

*Dr. Günter Mühlberger* (Universität Innsbruck) bringt das Problem auf den Punkt: Alles digitalisiert! Was nun? Am Beispiel des Innsbrucker Zeitungsarchivs



*Ein Mitglied des Organisationsteams des  
HWWA*  
Foto © HWWA – Olaf Hoppe

(IZA), das seit Oktober 2000 Zeitungsartikel mit Hilfe des LAURIN-Systems scannt und elektronisch gespeichert wird beschrieben, dass nach erfolgter Digitalisierung dieser Dokumente keineswegs wohlgefällig die Hände in den Schoß gelegt werden können, sondern die durch die Digitalisierung gewonnenen Daten und

Dokumente im steten Wandel der Datenformate einer aufmerksamen Pflege bedürfen. Der Schwerpunkt dieser Pflege muss nicht unbedingt bei den Digitalisierungsdateien liegen, sondern die Datenbanksysteme und ganz besonders die Applikationssoftware werden einem schnellen Wandel unterzogen sein.

Auf dieser vom HWWA gestalteten Tagung wurden auch hausspezifische Inhalte präsentiert. *Helmut Leveknecht* fasst die Geschichte des Hauses zusammen und stellt sie aus aktuellem Anlass unter das Thema «100 Jahre Pressedokumentation im HWWA».

Das Hauptthema des HWWA war die Vorstellung des Projekts der retrospektiven Digitalisierung der historischen Pressearchive des Instituts für Weltwirtschaft in Kiel bzw. der Zentralbibliothek für Wirtschaftswissenschaften (ZBW) und des Hamburgischen Welt-Wirtschafts-Archivs (HWWA). *Dr. Thomas Huck* und



*Hartmut Walravens*  
Foto © HWWA – Olaf Hoppe

*Max-Michael Wannags* beschreiben die in diesem Projekt durchgeführte Digitalisierung der Zeitungsausschnitte, die teils im Original, teils als Sicherheitsverfilmungen vorlagen, und die mit diesem Prozess verbundene, formale und inhaltliche Erschließung des Materials.

Zu danken ist Frau *Frauke Nordmann* vom HWWA und ihrem Team für die Planung und Durchführung dieser gelungenen und gut besuchten Tagung, die die Teilnehmer nicht nur mit Themen rund um die Digitalisierung sondern auch mit der gastgebenden Institution vertraut machte.

Die 10. Tagung der AG Zeitungen im Forum Zeitschriften (GeSIG) wird am 27./28.4.2006 in Leipzig stattfinden.

# 100 Jahre Pressedokumentation im HWWA<sup>1</sup>

Helmut Leveknecht

## Entstehung und Entwicklung des HWWA

Als 1908 das Hamburgische Kolonialinstitut errichtet wurde, begann die Geschichte des HWWA, welches im nächsten Monat sein 97-jähriges Jubiläum begeht.

Das Kolonialinstitut hatte zwei Zwecke:

- Zum einen sollte es eine Hochschulausbildung für Beamte des Reichskolonialamtes und anderer Behörden anbieten, aber auch für andere Personen, wie Kaufleute, Landwirte, Lehrer und Missionare, die beabsichtigten, in die sog. Schutzgebiete zu gehen,
- zum anderen – und hier liegt die Keimzelle des HWWA – stand die Schaffung einer Zentralstelle im Mittelpunkt, in der sich alle wissenschaftlichen kolonialen Bestrebungen konzentrieren konnten – d.h. die Sammlung und Auswertung von Informationsmaterial deutscher und ausländischer Kolonien, welches aus Zeitungen, Zeitschriften und Forschungsberichten entnommen und neben den Lehrenden und Lernenden auch der Hamburger Kaufmannschaft zur Verfügung gestellt wurde. Daraus entstand ebenso die Bibliothek des HWWA.

Das Kolonialinstitut wurde nach dem Ersten Weltkrieg aufgelöst, da das Deutsche Reich seine Kolonien abtreten musste. Aus dem Institut entstand 1919 die neu gegründete Hamburger Universität. Aus der Zentralstelle entstand im selben Jahr das «Hamburgische Welt-Wirtschafts-Archiv». Das HWWA wurde in die Wissenschaftlichen Anstalten Hamburgs eingereiht. Die neue Aufgabe sollte sein, wirtschafts empirische Informationen zu sammeln und der hamburgischen Kaufmannschaft zur Verfügung zu stellen.

Nach dem Zweiten Weltkrieg durfte das HWWA offiziell erst 1948 wieder eröffnet werden. Von diesem Zeitpunkt an liegen die Schwerpunkte in der Erforschung der wirtschaftlichen Entwicklung der Bundesrepublik Deutschland und ihrer weltwirtschaftlichen Verflechtungen sowie weiterhin in den Sammlungen des Archivs und der Bibliothek und deren Nutzung

durch Fachpublikum und Öffentlichkeit. Der immer stärker werdenden wissenschaftlichen Tätigkeit wurde 1970 Rechnung getragen. Der neue Name lautete von hier an: «HWWA – Institut für Wirtschaftsforschung – Hamburg».



*Helmut Leveknecht  
Foto © HWWA – Olaf Hoppe*

Bedingt durch inhaltliche Kritik und der sich daraus ergebenden neuen Schwerpunktsetzung wurde das Institut 1999 wieder in das «Hamburgische Welt-Wirtschafts-Archiv» umbenannt. Die Rechtsform einer öffentlichen Stiftung löste den bisherigen Status einer Dienststelle des Hamburger Hochschulamtes ab. Heute besteht das HWWA aus den Hauptbereichen Bibliothek, Pressedokumentation und Informationsdienstleistungen sowie aus drei wissenschaftlichen Abteilungen. Als Mitglied der Wissenschaftsgemeinschaft Leibniz (WGL) wird das Institut wie bisher von Bund und Ländern finanziert.

## Pressedokumentation und Archive damals

Was braucht ein Pressearchiv?

- Menschen, die Gedrucktes lesen und nehmenswerte Presseartikel mithilfe von Ordnungsprinzipien bzw. Notationen einer Sammlung von Themen-

<sup>1</sup> Überarbeitete Fassung des auf der 9. Tagung der AG Zeitungen im Forum Zeitschriften (GeSIG) am 13. und 14. September 2005 in Hamburg gehaltenen Vortrages. Address all communications to: Hamburgisches Welt-Wirtschafts-Archiv; Retro-Digitalisierung; Neuer Jungfernstieg 21; D-20347 Hamburg; e-Mail: [helmut.leveknecht@hwwa.de](mailto:helmut.leveknecht@hwwa.de)

mappen zuordnen sowie diese den Archivkunden beratend zu vermitteln – das sind die Pressedokumente bzw. das «Lektorat»,

- ebenso Menschen, die nach denselben Ordnungsprinzipien Presseauschnitte ordnen, verwahren und den Nutzern und Nutzerinnen aushändigen und wieder einordnen – die «Archivare»,
- und schließlich Menschen, welche die Presseartikel auf DIN-A-4-Papier umbrechen – die «Kleberei» bzw. technische «Aufbereitung».
- Natürlich benötigt ein Pressearchiv ebenso viel Platz – heute wie damals.

Aus der anfänglichen Beschränkung der Themenausrichtung auf deutsche Kolonien und andere überseeische Länder wurde wegen der starken Nachfrage der hamburgischen Kaufmannschaft sowie der Dozenten und Studenten des Kolonialinstituts das Archiv ausgeweitet auf weltweite Themen aus Wirtschaft, Politik und Gesellschaft. Das Archivmaterial wurde immer umfangreicher und zusehends unübersichtlich, deshalb wurde 1911 von Heinrich Waltz, einem ehemaligen Reichsbankbeamten und von da an neuer Leiter des Archivs, das «Hamburger System» erfunden.

Das Archiv wurde in vier Subarchive nach verschiedenen Ordnungsprinzipien strukturiert:

- Das Sach- bzw. Länderarchiv mit Sachthemen aus allen Ländern, Teilkontinenten, Kontinenten und der gesamten Welt,
- das Warenarchiv mit nationalen sowie internationalen Rohstoffen, Halbfertig- und Fertigerzeugnissen,
- das Firmenarchiv mit Pressemeldungen über in- und ausländische Unternehmen
- und schließlich das Personenarchiv mit Artikeln über weltweite Persönlichkeiten aus Wirtschaft, Politik und Gesellschaft.

Für das Sach- bzw. Länderarchiv und für das Warenarchiv wurden je spezielle Systematiken entwickelt, denen eine gesonderte Ländersystematik zugeordnet wurde. Alle diese Systematiken wurden bis zum Jahre 1998 verwendet und bis dahin immer wieder an die aktuellen Erfordernisse angepasst. Die Pressedokumente des Firmen- und Personenarchivs wurden ausschließlich nach den entsprechenden Namen angeordnet.

Jeder nehmenswerte Presseartikel wurde von der Aufbereitung ausgeschnitten und auf Papierträger geklebt. Danach wurde er vom Lektorat mit einer Signatur versehen. Das Schreibwerkzeug war ein Bleistift. Das Lektorat bestand in der Anfangszeit aus – wie es damals hieß – «Damen, die das Lehrerinnenexamen für fremde Sprachen abgelegt hatten» zur Kennzeichnung und Signierung der auszuschneidenden Artikel. Die Signaturenordnung bildete zugleich die Aufstellungsordnung der Themenmappen im Archiv ab. Die Themenmappen sind in sich nach wie vor chronologisch geordnet. Für die stabile Verwahrung bzw. Aufstellung der Mappen werden bis heute Pappkästen verwendet.

1914 waren bereits 36 Mitarbeiter bei der Zentralstelle beschäftigt – wissenschaftliche Beamte, Bibliothekarinnen, Lektorinnen, Archivare und technische Arbeitskräfte in der «Kleberei». Zu diesem Zeitpunkt verfügte das Archiv bereits über ca. 350.000 Presseauschnitte.

Franz Stuhlmann, der Leiter der Zentralstelle und später erster Direktor des HWWA, formulierte 1918 den damaligen Zweck so: «Den Kern unserer Tätigkeit bilden die Archivsammlungen, deren hoher Wert für Wissenschaft und Praxis heute kaum mehr bezweifelt wird, und die jetzt in Deutschland an den verschiedensten Plätzen nachgeahmt werden – aus der Erkenntnis heraus, daß sie dringend notwendig sind. Ein ungeheures Gegenwartsmaterial wird darin angehäuft über politische und wirtschaftliche Vorgänge in allen Ländern der Erde; aber größtenteils liegt es ungenutzt. [...] Nur die emsige, ruhige Forschungstätigkeit fleißiger Studenten und Gelehrter vermag aus den Sammlungen die wissenschaftlichen Werte herauszuholen, die darin enthalten sind. Ich denke dabei hauptsächlich an Historiker, Nationalökonomien und Wirtschaftsgeographen; aber auch für viele Juristen, Botaniker, Geologen, Sprachforscher, Ethnologen werden die Archive eine reiche Fundgrube bilden.<sup>2</sup>»

Für die alten Bestände gilt natürlich immer noch Stuhlmanns Aufzählung des Nutzerkreises. Doch im Laufe der Zeit hat dieser sich auf die Leser und Leserinnen aus Wirtschaftspraxis und Wirtschaftstheorie beschränkt. Seit der Evaluation der Pressedokumentation durch den Wissenschaftsrat 1996 soll der Schwerpunkt in der Informationsversorgung der überregionalen Wirtschafts- und Sozialwissenschaften liegen.

Kurz vor seinem Tode im Jahre 1928 ging Stuhlmann auf den Wert der Presse und die Informationsveredelung der Presseartikel vom HWWA ein: «Bereits viel früher, schon Jahre vor dem Kriege hat das Hamburgische Welt-Wirtschafts-Archiv, die Notwendigkeit er-

---

2 Stuhlmann, Franz: «Zur Hochschulfrage», in: *Neue Hamburger Zeitung* Nr. 210 v. 25.4.1918.

kannt, die Presse als eine der wichtigsten Informationsquellen zum Gegenstand systematischer Beobachtung und wissenschaftlicher Verarbeitung zu machen. Die Aufgabe des Welt-Wirtschafts-Archivs, das Zeitgeschehen auf wirtschaftlichem und politischem Gebiet in aller Welt aus möglichst vielseitigen Informationsquellen zu erfassen, hätte sich gar nicht erfüllen lassen, ohne die unendliche Fülle informatorischer Arbeit, wie sie von der Tagespresse fortlaufend geleistet wird, in möglichst weitgehendem Maße auszuwerten. So bilden die Zeitungsausschnittarchive, die aus mehr als 50 Zeitungen und über 100 Spezialzeitschriften einen Tageszuwachs von etwa 1200 Ausschnitten in ihren nach feinstem System ausgearbeiteten, nach Hunderttausenden zählenden Fächern aufsaugen, gleichsam ein dauerndes Denkmal der unendlichen Arbeit, die von der Presse tagaus, tagein und für den Tag geleistet wird. Aber es ist mehr als ein Denkmal, es ist der Versuch, die unendliche Fülle des wertvollen Informationsstoffes über den Tag hinaus dem künftigen Bedarf dienstbar zu machen, den Wert der einzelnen Informationen durch spätere oder aus anderer Quelle stammende Informationen zu erhöhen und so für alle Gebiete der schaffenden Arbeit einen leicht zugänglichen, nie versagenden Quell lebendigen Wissens zu schaffen.<sup>3</sup>

Stuhlmanns Beschreibung gilt für die inhaltliche Erschließung im Prinzip weiterhin. Heute ist die Pressedokumentation in allen Fragen über das nationale und internationale Wirtschaftsgeschehen und deren Einschätzungen ein «nie versagender Quell». Alle anderen Themen über Politik, Gesellschaft und Kultur sind nur noch am Rande vertreten.

Von 1933 bis 1936 wurden die Bestände auf ihre Übereinstimmung mit der nationalsozialistischen Ideologie eingeschränkt, so dass viele wichtige Informationen aus dem Ausland fehlten. Von 1936 an wurden zwar alle weltweiten Themen wieder gesammelt, aber nur einem beschränkten Personenkreis zugänglich gemacht, der sich aus den gleichgeschalteten leitenden Unternehmensvertretern und staatlichen Dienststellen zusammensetzte.

### **Pressedokumentation und Archive heute**

Heute umfasst das Archiv fast 20 Millionen Presseartikel. Pro Jahr kamen zuletzt ca. 120.000 neue Ausschnitte hinzu. Es ist das größte öffentlich zugängliche Pressearchiv Deutschlands, wahrscheinlich auch Europas.

Um die Themen der letzten 97 Jahre finden zu können, greifen die Lektorinnen und Lektoren bei der Leserberatung auf ihr eigenes Wissen zurück oder benutzen

Zettelkataloge, die bis 1998 fortgeführt wurden. Seit 1999 wird eine interne elektronische Datenbank hierbei verwendet. Die Stichwörter von besonderen Ereignissen bzw. Sachverhalten werden in regelmäßig stattfindenden Sitzungen des Lektorats abgestimmt und somit einheitlich zugeordnet.

90 Jahre lang wurden die Presseausschnitte von der «Kleberei» bzw. technischen «Aufbereitung» mit der Schere ausgeschnitten und mit Kleister auf vorgedruckte Papierträger geklebt. Das Kleben wurde mit Maschinen vorgenommen – am Anfang mit Handkurbel, später mit Elektromotor. Seit 1999 werden die vom Lektorat gekennzeichneten Zeitungsseiten zunächst gescannt und darauf die vom Lektorat gekennzeichneten Artikel elektronisch «ausgeschnitten» und «geklebt» (Cut & Paste).

Das Pressearchiv ist weiterhin wegen des Urheberrechts ein Papierarchiv. Alle Presseartikel werden ausgedruckt und dann elektronisch gelöscht. Zur Zeit sind noch Presseausschnitte seit 1991 auf Papierträgern vorzufinden. Bis 1990 sind alle Ausschnitte auf Mikrofilm bzw. Mikrofilmen verfilmt. Dieses geschah bisher alle zehn Jahre, da das Archiv sonst aus allen Nähten platzen würde.

Von 1996 bis 1999 hat die Pressedokumentation des HWWA in Kooperation mit dem Wirtschaftsarchiv in Kiel eine neue Sachklassifikation als neues Erschließungsinstrument entwickelt. Die Klassenbenennungen stimmen mit dem im HWWA entwickelten Standardthesaurus Wirtschaft (STW) weitgehend überein. Beide Institutionen haben ihre jeweiligen Quellenprofile aufeinander abgestimmt und eine gemeinsame überregional recherchierbare Referenzdatenbank – EconPress – für die zu erschließenden Presseartikel aufgebaut, welche seit Januar 2001 im Internet zu nutzen ist. Mit EconPress können nicht nur Presseartikel über Titel bzw. Titelstichwörter, Klassenbenennungen, Inhaltsstichwörter und anderen inhaltlichen Angaben recherchiert, sondern auch bestellt werden. D.h. einzelne Presseartikel können mit einem relativ günstigen Gebührensatz zugesandt oder im HWWA in chronologisch geordneten Themenmappen gelesen bzw. kopiert werden.

Das Pressearchiv besteht seit 1999 aus vier anders gewichteten Subarchiven:

- Das größte ist das «Länderarchiv» bzw. «Sacharchiv». Heute bilden inländische und internationale Wirtschaftsthemen (Betriebswirtschaft, Volkswirtschaft, Wirtschaftspolitik, Märkte, Branchen, Innovationen) den Fokus der Sammlung sowie am

3 100 Jahre Hamburger Fremdenblatt 1828-1928. [Jubiläumsschrift]. Hg. Alfred Herrmann. [Hamburg 1928]: S. 41.

Rande Politik, Gesellschaft und Umwelt. Die aktuelle Sachklassifikation weist 726 Sachthemen auf, die weltweit mit allen Ländern und Kontinenten verknüpft werden können. Das bis 1998 bestandene «Warenarchiv», in welchem Produkt- bzw. Marktmeldungen gesammelt wurden, ist 1999 in das Sacharchiv integriert worden.

- Das «Firmenarchiv» mit heute 70.000 in- und ausländischen Firmen seit Beginn (neben den Presseauschnitten hat es einen sehr großen Bestand von Geschäftsberichten).
- Das «Personenarchiv» ebenso seit 1908, worin Meldungen über Persönlichkeiten aus Wirtschaft, Politik, Wissenschaft und Gesellschaft gesammelt werden.
- Im «Körperschaftenarchiv» werden neben den erwähnten Firmenmeldungen seit 1999 speziell Berichte über ca. 100 Institutionen und internationale Organisationen gesammelt.

Um die ausgewählten Presseartikel ohne Rücksichtnahme auf Vorder- und Rückseite ausschneiden zu können, war es vom Anfang des Archivs bis 1998 nötig, pro Printmedium zwei Exemplare zu bearbeiten. Zunächst wurde die jeweils rechte Seite, darauf die linke Seite lektoriert und danach entsprechend zerschnitten und aufgeklebt. Dieses Prinzip erübrigte sich bei der Einführung des elektronischen Cut & Paste. Das Quellenprogramm des HWWA und seines Kooperationspartners in Kiel umfasst heute ca. 100 Tages- und Wochenzeitungen, Zeitschriften und Informationsdienste in zehn Sprachen aus ca. 25 Ländern.

Für die gesuchten Presseartikel wird bis heute eine Beratung angeboten, die zwar wegen der Möglichkeit von Selbstrecherchen in der o.g. Referenzdatenbank EconPress weniger in Anspruch genommen wird, aber für das alte Material unerlässlich ist.

Zur Zeit wird eine Retrodigitalisierung des alten Materials der Jahre 1908 bis 1930 vorbereitet. Die alten Mikrofilme (Rollfilme) werden digitalisiert und können später als elektronische Volltexte recherchiert werden. Das Projekt wird durch die Deutsche Forschungsgemeinschaft gefördert.



*Thomas Huck und Bernd Augustin  
Foto © HWWA – Olaf Hoppe*

### **Das Ende der Pressedokumentation und Archive**

Wir schreiben das Jahr 2005. Vor genau 400 Jahren kam die erste Zeitung – eine Wochenzeitung – in Straßburg auf die Welt, vor 355 Jahren die erste Tageszeitung in Leipzig.

Am Ende dieses Jahres wird die aktuelle Pressedokumentation des HWWA – ebenso die des Wirtschaftsarchivs in Kiel – eingestellt. Das Material in den Pressearchiven des HWWA wird nach 97 Jahren eingefroren und nur noch als historischer Bestand nutzbar sein. Die Referenzdatenbank EconPress wird für die Jahre 2001 bis 2005 als Recherche- und Bestellinstrument bestehen bleiben. Mit unserer kompetenten Beratung werden wir weiterhin für Fachpublikum und Öffentlichkeit da sein. Anlass für die Entscheidung, mit der aktuellen Auswertung aufzuhören, ist die Evaluation der Wissenschaftsgemeinschaft Leibniz (WGL) im Jahre 2003.

Es stirbt die größte öffentlich zugängliche Pressedokumentation Deutschlands. Was bleibt, ist eine vielfältige Informationsquelle von einzigartiger Bedeutung in einer Kontinuität von fast 100 Jahren.

# Inhaltliche und technische Darstellung der retrospektiven Digitalisierung der historischen Pressearchive der ZBW und des HWWA

Thomas Sergej Huck und Max-Michael Wannags<sup>1</sup>

Die Archive der Pressedokumentationen des Instituts für Weltwirtschaft in Kiel bzw. der Zentralbibliothek für Wirtschaftswissenschaften (künftig: ZBW) und des Hamburgischen Welt-Wirtschafts-Archivs (künftig: HWWA) bieten seit Gründungszeiten bis heute ein außerordentlich breites Spektrum an Themen aus Wirtschaft, Politik, Gesellschaft und Kultur aus – zu verschiedenen Zeiten – bis zu 1.150 zeitgenössischen Quellen der deutschen und internationalen Zeitungs- und Zeitschriftenlandschaft in chronologischer und sachlicher Ordnung zur Nutzung an<sup>2</sup>.

## Das Projekt

Diese Materialien liegen in den Archiven beider Institute entweder als originale Presseauschnitte vor, aufgeklebt auf Trägerpapier, oder in Form von Rollfilmen und Mikrofiches, sofern es sich um Sicherheitsverfilmungen handelt, die in etwa zehnjährigen Abständen seit den 1960er Jahren durchgeführt werden.

Besonders das noch in Originalform erhaltene, papierne Material ist vom Zerfall bedroht, aber auch die Sicherheitsverfilmungen in Form von Rollfilmen und Mikrofiches zeigen bereits Verfallserscheinungen. Daher sollen die Materialien durch Digitalisierung gesichert werden. Dies, die Bestandssicherung, ist einer der beiden Projektschwerpunkte.

Ein zweiter Schwerpunkt ist die mit diesem Prozeß verbundene, formale und inhaltliche Erschließung des Materials. Sie verfolgt das Ziel, die Altbestände öffentlich über das Internet zugänglich zu machen und recherchefähig zu präsentieren. Zunächst wird nur das Material bis in die 1930er Jahre digitalisiert. Hier ist eine kritische Grenze gegeben, einerseits aus urheberrechtlichen Gründen bezüglich der elektronischen Präsentation der Dokumente, andererseits auch im Hinblick auf die Bestandssicherung.

Sowohl die originalen als auch die bereits als Sicherheitsverfilmungen vorliegenden Archivalien des Bearbeitungszeitraumes werden für die Digitalisierung durch einen externen Dienstleister vorbereitet. Die zu scannenden Dokumente werden auf ihre urheberrechtlich unbedenkliche Präsentierbarkeit geprüft und ggf. freigeschaltet bzw. bis zum Zeitpunkt einer möglichen Freigabe gesperrt.

## Bearbeitung des papiernen Materials

Für den Rücklauf des Materials in Form digitaler Abbildungen der Originale wird für die bibliographisch-formale Erfassung und sachlich-inhaltliche Erschließung ein von der Firma NEWBASE entwickeltes Presse-Clippingsystem benutzt. Diese Anwendung wird bereits seit längerem auch für die aktuelle Presseauswertung eingesetzt.

Die besonderen Erschließungsanforderungen, die sich aus der Verarbeitung des Altmaterials ergeben, machen gewisse Modifikationen an diesem Programm notwendig. Erhalten bleibt insgesamt aber eine den Mitarbeitern beider Häuser weitgehend vertraute Arbeitsumgebung, so daß die Fortführung der Erschließungsarbeit auch über die Projektlaufzeit hinaus gewährleistet ist.

Mit dem Clipping-Modul (Abb. 1, s. nächste Seite) werden der Bestand Personenarchiv sowie zwei Sondersammlungsbestände bearbeitet. Die einzelnen Dokumente werden (wenn nötig) geschnitten und mit formalen und inhaltlichen Daten angereichert. Technisch gesehen werden diese Metadaten einerseits vollständig in einer XML-Verwaltungsdatei und andererseits in Schlüsseln kodiert im Dateinamen abgelegt.

Langfristig ergibt sich so erstmalig, wenn die noch bestehende Lücke in der Digitalisierung der Altbestände von der Mitte der 1930er Jahre an bis zur gegenwärtigen

1 Überarbeitete Fassung der auf der 9. Tagung der AG Zeitungen im Forum Zeitschriften (GeSIG) am 13. und 14. September 2005 in Hamburg gehaltenen Vorträge. Besonderer Dank gebührt für kritische Anmerkungen und Durchsicht des Textes Johanna Becker und Kirsten Jeude. Adress all communications to: Hamburgisches Welt-Wirtschafts-Archiv; Retro-Digitalisierung; Neuer Jungfernstieg 21; D-20347 Hamburg; e-Mail: retrodigitalisierung@hwwa.de

2 Aus wirtschaftlichen Gründen wird die Aktualisierung der Archive zum Jahresende 2005 eingestellt. Vgl. auch den Beitrag von H. Leveknecht in dieser Ausgabe.





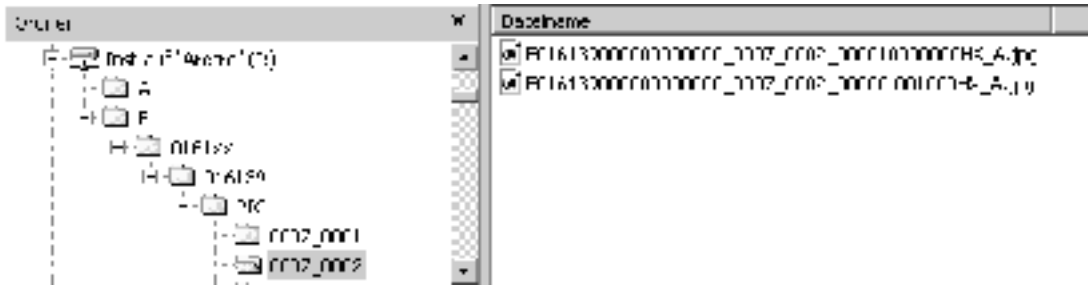


Abb. 3: Kodierung der Metadaten aus der Verwaltungsdatei im Dateinamen sowie in der Server-Architektur  
 Abb. © HWWA – Olaf Hoppe

Zusätzlich werden während dieser Inventarisierungsphase auf dem Projektserver die Meta-Daten zu dem «obersten» Sortierkriterium abgebildet. Im Falle einer Person werden also etwa Lebensdaten, Beruf, usw. oder im Falle einer Institution die dieser entsprechenden Daten sowie weitere Verknüpfungen (z. B. EconPress, Internet-Links) angezeigt (Abb. 4 und 5).

Rallin, Albert 1857-1912	
Beitrag	Beitrag: HAPAG: Vorstand (1888), Generaldirektor (1896)
Lebensdaten	Lebensdaten: Hechtling, Joseph/Karl/Julius Packard&Co. (Hamburg)
Verknüpfungsbereich	W11: Deutschland: Wirtschaft
Nationalität	DE: Deutsche
Pressemappe	1811
Homepage	de.wikipedia.org/wiki/Albert_Rallin
Identifikationsnummern	189694
Über die Person	1 Text
über die Person	183 Grafische Textdateien (Anzahl: 183)

Abb. 4: Stammdatensatz, hier zu einer Person  
 Abb. © HWWA – Olaf Hoppe

Körperschaftsname	Atlantic Hotel Kempinski (Hamburg)
Sitz der Institution	2DE Deutschland
Web-Adresse	<a href="http://www.kempinski.atlantic.de">www.kempinski.atlantic.de</a>
benutze für	Hotel Atlantic (Hamburg), Atlantic-Hotel AG (Hamburg)
Klassifikation	W11-000 Gastgewerbe
Firmensignatur	2DEhh-A.0003
Identifikationsnummern	18037
Pressemappe	Pressemeldung1931-1942,1954-1957,L,1959-1960,L,1963,L,1968-1969,L,1972,L, EconPress
über die Institution	29 Grafische Textdateien

Abb. 5: Stammdatensatz, hier zu einer Institution  
 Abb. © HWWA – Olaf Hoppe

Zu diesem «primären» Sortierkriterium werden in Kurzform die zugehörigen Digitalisate gelistet.

DOC-Nr. 0002 (21 Seiten)	<a href="#">JPEG</a>
DOC-Nr. 0003 (2 Seiten)	<a href="#">JPEG</a>
DOC-Nr. 0004 (3 Seiten)	<a href="#">JPEG</a>
DOC-Nr. 0005 (2 Seiten)	<a href="#">JPEG</a>
DOC-Nr. 0006	<a href="#">JPEG</a>
DOC-Nr. 0007	<a href="#">JPEG</a>
DOC-Nr. 0008 (3 Seiten)	<a href="#">JPEG</a>
DOC-Nr. 0009	<a href="#">JPEG</a>
DOC-Nr. 0010	<a href="#">JPEG</a>

Abb. 6: Reihung der Digitalisate zu einem Stammdatensatz  
 Abb. © HWWA – Olaf Hoppe

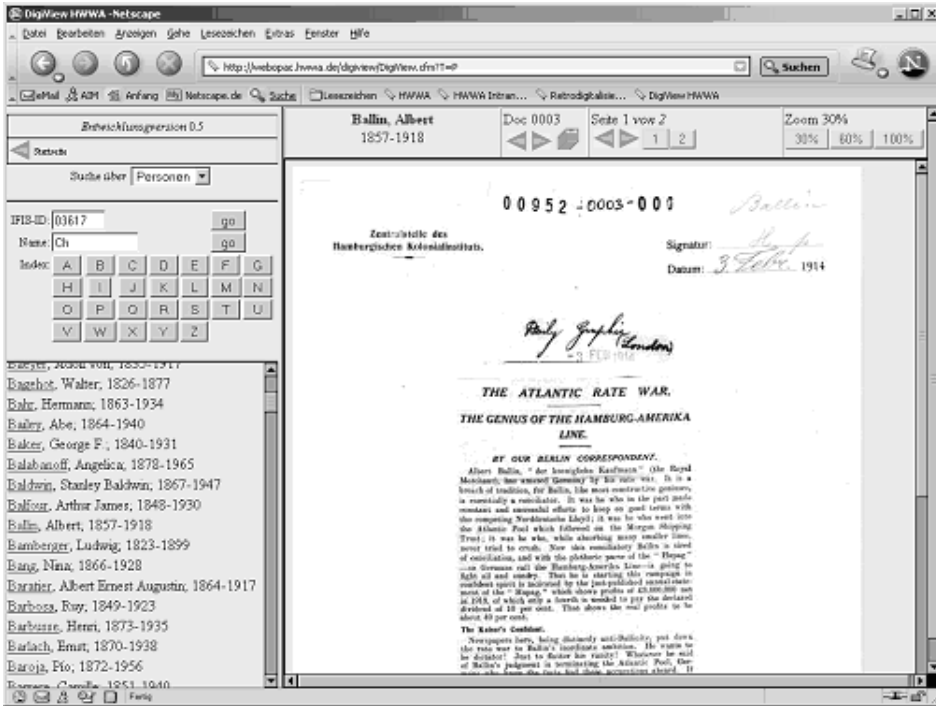


Abb. 7: Ein geöffnetes Dokument zu einem Stammdatensatz mit dokumentübergreifenden Vor- und Rück-Navigationspfeilen sowie Blätterpfeilen innerhalb eines Dokuments und den Zoom-Stufen  
Abb. © HWWA – Olaf Hoppe

Nach der Katalogisierung werden Einträge der Art «DOC-Nr. 0002 (21 Seiten)» durch dokumentbezogenen Einzellerschließungsdaten wie Autor des Artikels, Titel des Artikels, Quelle, Erscheinungsdatum usw. angereichert. Diese Daten müssen in dem NEWBASE-Modul einzeln erzeugt werden, während die automatische Zuordnung der Artikel zu den jeweiligen Personen aufgrund der Codierung im Dateinamen erfolgt. Mit Klick auf «JPEG» wird dann das erste Bild

angezeigt. Von dort kann man dem Menü folgend den Material-Bestand zu dieser Person durcharbeiten. Jeder gescannte Presseartikel kann in der Größe 30, 60 und 100 Prozent der Originalgröße angezeigt werden.

## Bearbeitung des Rollfilmmaterials

Im Gegensatz zu den auf Papier vorliegenden Beständen muß das Arbeitsinstrument zur Bearbeitung von Digitalisaten, die auf der Basis von Rollfilmen oder Mikrofiches erzeugt werden, völlig andersartig konzipiert werden. Ursache dafür ist die Art des weitgehend automatischen Scanprozesses, bei dem den erzeugten Bildern, anders als bei den Bildern auf Basis originaler Vorlagen aus Pa-

pier, im Dateinamen überhaupt keine Merkmale der Formalerschließung beigegeben werden können. Ihr Dateiname enthält lediglich die Filmnummer sowie eine fortlaufende Hochzählung der Bilder eines Rollfilmes oder eines Fiches.

Um dennoch ähnlich wie am Beispiel des Bestandes Personenarchiv vorgeführt, eine Sortierung nach einem «obersten» Sortierkriterium zu ermöglichen, also der

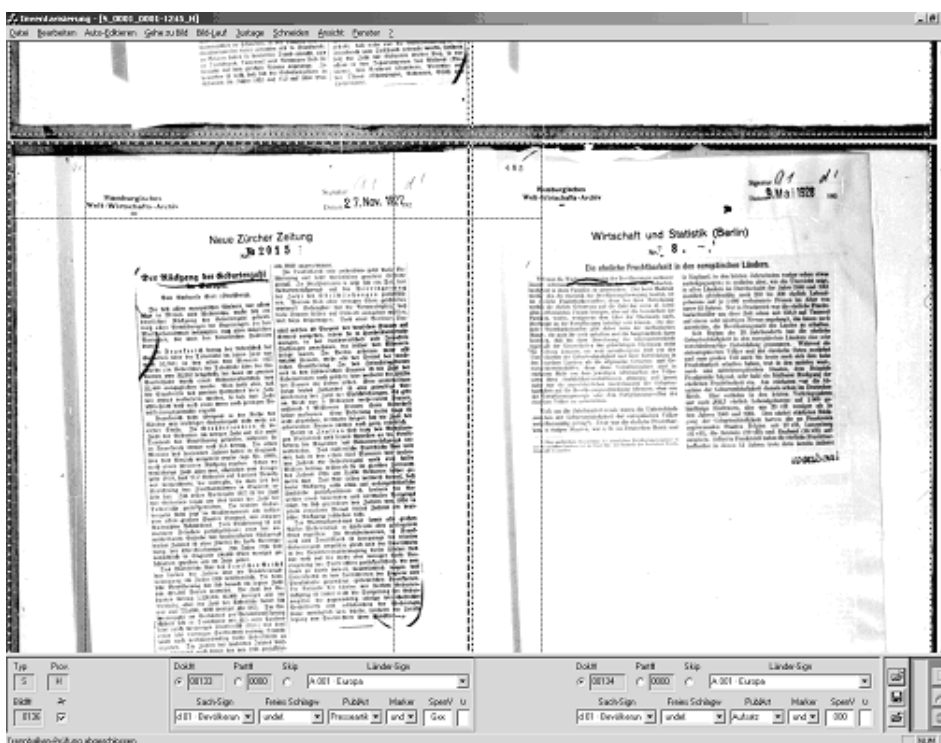


Abb. 8: Doppelbild von einem Rollfilm  
Abb. © HWWA – Olaf Hoppe

konvolutartigen Präsentation, müssen die Themencluster auf den Rollfilmen und Mikrofiches rekonstruiert und die zugehörigen Daten analog zum Bestand Personenarchiv verwaltet werden.

Zu diesem Zwecke wurde ein elektronisches Rollfilm-sichtgerät, ein «Slider», mit dem aber ebensogut auch Mikrofiches dargestellt und bearbeitet werden könnten, entwickelt und mit Eingabefeldern zur formalen Dokumenterschließung versehen. Diesen formalen Daten müssen nur einmal eingetragen werden und werden anschließend durch «Weiterkurbeln» des «Films» (bestehend aus den zusammengeführten Doppelbildern des Scan-Ergebnisses) solange auf alle folgenden Dokumente übertragen, bis der Bearbeiter andere Merkmale vergibt, die dann wiederum übernommen werden. Zwischenzeitlich muß lediglich auf die Einstellung «Dok» und «Part» geachtet werden, womit einerseits die Zählung der Dokumente und andererseits die Zählung der Dokumentseiten organisiert wird.

Auf diese Weise wird in vergleichsweise kurzer Zeit ein Formalerschließungszustand rekonstruiert, wie er vor der Verfilmung der Originale in Form von Presse-mappen einst vorlag. Damit wird es möglich, das natürlich noch nicht als Einzeldokument erschlossene digitale Bild in gleicher Weise zu präsentieren wie im Fall des Personenarchivs. Für den zweiten Schritt, also die Einzeldokumenterschließung, wird ebenfalls der «Slider» benutzt, der nun aber ergänzende Eingabe-felder aufweist, die auch bei der NEWBASE-Software benutzt werden.

Auf diese Weise werden in beiden – den jeweiligen Besonderheiten adaptierten – Modulen grundsätzlich gleichrangige formale und inhaltliche Daten produziert und diese mehrfach archiviert. Die Archivierungsstruktur läßt sich ohne zusätzliche Hilfsmittel mit jedem Tool zum Dateimanagement einsehen und ordnen. Gleichwohl existiert – über den Projekt-Server organisiert – ein Datenbanksystem, das die Meta-Daten verwaltet und für Recherchen nutzt.

Zur Langzeitsicherung jedoch wird dieses Datenbanksystem nicht benötigt. Es kann zwar in einer lauffähigen Version mitgesichert werden, aber für einen dauerhaften Datenzugriff ist das nicht erforderlich. Zwingend gesichert werden müssen lediglich die digitalen Bilder im Rahmen der Server-Architektur, welche wiederum von der Dateinamenstruktur vorgegeben wird, mitsamt zugehörigen Verwaltungsdateien, welche die ermittelten Metadaten nochmals im Klartext enthalten. Nur diese Dateien werden regelmäßigen Sicherungsroutinen unterzogen, wozu in Abständen auch das erzeugen von 1:1-Kopien auf magneto-optischen Speichermedien gehört.

## Wahl des Dateiformates

Abschließend einige Bemerkungen zur Frage der Formate der Digitalisate. Üblicherweise werden für Zwecke wie in diesem Projekt Rastergrafikformate benutzt, da deren Pixel-Matrix selbst keine explizite logische Struktur besitzt, welche die Semantik der vom Scanner gelieferten Inhalte in einer für den Computer verstehbaren Form beschreiben müßte und insofern verkomplizieren würde.

Das «Tagged Image File»-Format hat in der Vergangenheit vermutlich aufgrund seiner Flexibilität einen



*Thomas Sergej Huck  
Foto © HWWA – Olaf Hoppe*

hohen Verbreitungsgrad erreicht und ist zum Quasi-standard avanciert. Es ist weder Druckersprache noch Seitenbeschreibungssprache, sondern reines Raster-Modell. Die originäre Spezifikation dieses Formats stammt von der Firma Aldus, die es 1986 offenlegte und die nun zu Adobe gehört. Mehrfache Weiterentwicklungen wurden vorgenommen. Heute ist man bei Version 6 oder ähnlich, aber alle Versionen sind abwärtskompatibel. Der Algorithmus ist zwar geschützt, aber dennoch frei verwendbar. Häufigste Parameter lauten: 300 x 300 dots per inch, 256 Graustufen, keine Komprimierung. Bei den Papiervorlagen ergibt sich so ein Speicherbedarf in der Größenordnung von 20 GB, bei den Rollfilmobjekten eine Größenordnung von 40 GB.

Bei Größenordnungen von mehreren Millionen Dokumentseiten ergeben sich kaum zu finanzierende Speicherkosten. Daher wurde in dem Projekt eine Ausweichstrategie auf das JPEG-Format verfolgt mit dem

Qualitäts-Parameter «lostless». Damit hat man die Gewährleistung dauerhaft druckbarer Qualität. Dies ergab eine Serie von Proofs mit unterschiedlichen JPEG-Parametern. Aber auch in diesem Fall sind Festplatten im Umfang von 40 bis 60 TeraB erforderlich (wegen der erwähnten Redundanz durch Mehrfachbevorratung). Grundsätzlich ist noch zu erwähnen, daß die technische Durchführung des Langzeitsicherungskonzeptes nicht projekthärent organisiert, sondern im Rahmen der HWWA-eigenen EDV- bzw. IT-Infrastruktur bewältigt wird.

Der Projektkonzeption allerdings entspricht durchaus die logische Datenorganisation, besonders eben auch die erwähnte Tatsache der Mehrfachverwaltung der Erschließungsdaten in codierter Form in den Dateinamen der Bilddateien ebenso wie in den Dateinamen der Verwaltungsdateien, die ihrerseits die Erschließungsdaten ein drittes Mal enthalten, und zwar im Klartext. So konnte darauf verzichtet werden, in den Headern der Grafiken Informationen zu verzeichnen, deren Lesbarkeit unter Umständen nicht auf Dauer gewährleistet ist.

Was den Aufbau der Verwaltungsdatei im XML-Format betrifft, ist zu erwähnen, daß dieser Aufbau inhaltlich, soweit das den Bedingungen des Projektes angemessen erscheint, den 15 Elementen des Dublin-Core-Standards entspricht, wobei klar ist, daß nicht alle Elemente zum Einsatz kommen. Im Gegensatz dazu beschreibt der Aufbau unserer dokumentbezogenen bibliographischen Angaben das jeweilige Dokument präziser als dies beispielsweise mit der DC-Auszeichnung möglich ist. Blickt man mit Abstand auf die projektbezogene Langzeitsicherungsstrategie, fällt auf, daß das Design der zu bearbeitenden Daten dem Referenzmodell des Open Archive Information System in den Grundzügen entspricht:

Abbildung 9: Grundgerüst OAIS

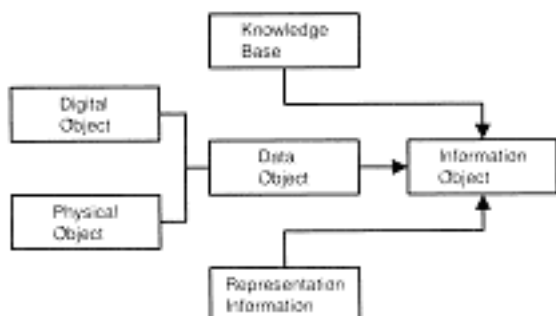


Abb. © HWWA – Olaf Hoppe

Dennoch: Das in diesem Projekt gewählte Prinzip ist von keinem Datenbankmodell und auch nicht von spezifischen Dokumentstrukturen abhängig. Die Verkettung erfolgt nur auf dem Wege gleicher Namen für



Max-Michael Wannags  
Foto © HWWA – Olaf Hoppe

Grafik und XML – dies gewährleistet einfachstes hardwareseitiges Handling., das wie folgt aussieht:

1. In einer ersten Stufe werden während des Zuwachses der Bilder und der zugehörigen Verwaltungsdateien aus Text und XML die Daten inkrementell im Rahmen der täglichen Sicherungsroutinen des HWWA mitgesichert. Dies betrifft zugleich die Verwaltung der Bild- und Bildbeschreibungdateien im Rahmen der Erschließungsmodule und deren anschließender Datenbank-Verwahrung.
2. Sind später alle Bilder und Beschreibungsdateien auf dem Server vorhanden, werden Kopien auf magneto-optischen Datenträgern für externe Verwahrung z. B. in einem Tresor gezogen. Parallel fallen die zwischenzeitlich als Sicherungsmedium dienenden DVDs der Dienstleister, welche die Bild-Scans geliefert haben, in den Rang eines Benutzermediums für die Bibliothek zurück.

**Projekthomepage** (momentan in der Entwicklungsversion 0.5): <http://webopac.hwwa.de/digiview/>

# Alles digitalisiert! Was nun?<sup>1</sup>

Dr. Günter Mühlberger

(Universität Innsbruck, Abt. für Digitalisierung und elektronische Archivierung)

Das IZA (Innsbrucker Zeitungsarchiv)<sup>2</sup> verfügt über eine der größten Sammlungen von Zeitungsausschnitten zur deutsch- und fremdsprachigen Literatur. Seit Oktober 2000 werden alle Artikel mit Hilfe des LAURIN Systems gescannt und elektronisch gespeichert, derzeit befinden sich bereits mehr als 140.000 Artikel im System. Der Altbestand, immerhin rd. 800.000 Seiten aus den Jahren 1960 bis 2000, wird nun ebenfalls rückwirkend digitalisiert und in einem eigens dafür entwickelten System gespeichert und verfügbar gemacht. Im Sommer 2006 wird das IZA daher eines der ersten öffentlichen Zeitungsausschnittsarchive sein, das alle seine Artikel in elektronischer Form vorliegen hat.



*Dr. Günter Mühlberger*  
Foto © HWWA – Olaf Hoppe

Zwei Aspekte sind zu berücksichtigen: zum einen die technische Umsetzung des Retrodigitalisierungsprojekts und zum anderen der Aspekt der Langzeitarchivierung. Die technische Umsetzung wird in einem Artikel von Albert Greinöcker, der 2006 im Tagungsband

der ODOK '05 (11. Österreichisches Online-Informationstreffen ; 12. Österreichischer Dokumentartag. Bozen, 13.-16.9.2005)<sup>3</sup> erscheinen wird und als Preprint auf den Webseiten der Abteilung für Digitalisierung und elektronische Archivierung bereits verfügbar ist, im Einzelnen beschrieben.<sup>4</sup>

Die wichtigsten Überlegungen des Vortrags in Bezug auf die Langzeitarchivierung sollen hier in Kürze erörtert werden:

1. Digitale Archivierung muss auf die komplexen Datenstrukturen von digitalen Daten Rücksicht nehmen.
2. Digitale Archivierung muss die unterschiedlichen Geschwindigkeiten im Alterungsprozess bereits von Anfang in der Softwarearchitektur berücksichtigen.
3. Digitale Archivierung ist nicht nur eine technische, sondern mindestens ebenso eine organisatorisch-finanzielle Frage.

Ad 1:

Im Digitalisierungsprozess für das IZA sind folgende Daten angefallen

- Imagedateien, z.B. JPEG und TIFF Dateien
- Metadaten bezogen auf einen Artikel, z.B. Titel oder Zeitung
- Metadaten bezogen auf eine Sammlung (Dossier) von Artikeln, z.B. alle Interviews von Günter Grass
- Volltexte, z.B. die automatisch erkannten Texte der gescannten Artikel
- Prozessdaten, z.B. Logfiles, die bei der automatischen Verarbeitung produziert werden, um die Qualitätskontrolle zu ermöglichen

Während sich die Diskussion oftmals nur auf Imagedaten oder Metadaten konzentriert, ist zu betonen, dass alle Daten, die für das umfassende Verständnis des Archivs notwendig sind, von der digitalen Archivierung erfasst werden müssen. Die Daten müssen dabei in einer Art und Weise vorliegen, dass auch die Beziehungen zwischen den einzelnen Komplexen eindeutig re-

1 Zusammenfassung eines Vortrags, gehalten auf der 9. Tagung der AG Zeitungen im Forum Zeitschriften (GeSIG). Hamburg, 13/14.9.2005

2 <http://iza.uibk.ac.at/>

3 <http://www2.uibk.ac.at/voeb/odok2005/>

4 [http://www2.uibk.ac.at/ub/dea/projekte/iza\\_beschreibung.html](http://www2.uibk.ac.at/ub/dea/projekte/iza_beschreibung.html)

konstruierbar sind. Mit anderen Worten: Scheinbare Selbstverständlichkeiten sollten ebenfalls in der Datenstruktur abgebildet werden.



*Frauke Nordmann*  
Foto © HWWA – Olaf Hoppe

Ad 2:

Während Datenformate wie JPEG, TIFF oder ASCII sich nur sehr langsam ändern und auch langfristig Programme verfügbar sein werden, die eine Konvertierung in andere Formate erlaubt, sind Datenbanksysteme und ganz besonders Applikationssoftware einem wesentlich rascheren Wandel unterworfen. Mit anderen Worten: Die Anwendungssoftware, die für die Benutzung des digitalen Archivs des IZA notwendig ist, wird bereits in 3-5 Jahren als «veraltet» angesehen werden und in weniger als 10 Jahren wird ein dringender Bedarf nach einer Neulösung entstanden sein. Dieser Generationswechsel muss bereits jetzt in der Architektur berücksichtigt werden und alle oben erwähnten Daten müssen daher so vorliegen, dass sie erstens ohne «Insiderwissen» verstanden und zweitens ohne größeren technischen Aufwand in ein neues System eingelesen werden können. Die Verwendung von XML ist dabei ein notwendiger erster Schritt, noch

empfehlenswerter ist allerdings die Verwendung allgemeiner Standards wie METS, die die eigentliche Ausformung der XML Dateien definieren.

Ad 3

Auch wenn die oben genannten Voraussetzungen gegeben sind, so ist doch heute nicht absehbar, ob in 10 oder 15 Jahren eine bestimmte Institution, in unserem Fall also die Universität Innsbruck, das Geld ausgeben möchte, um die alten Daten in ein neues System zu transferieren. Umgekehrt ist klar, dass – auch wenn in heutigen Systemen nicht alle Überlegungen zur digitalen Archivierung voll umgesetzt sind, aber gesichert ist, dass es sich um für die Institution so wichtige Daten handelt, dass sich «immer jemand darum kümmern wird», die Chancen für eine erfolgreiche digitale Archivierung wesentlich erhöht werden. Oder anders ausgedrückt: Laufende Systeme sind grundsätzlich leichter am Leben zu erhalten, als Systeme, die einmal «ausgeschieden» wurden und später «wiederbelebt» werden müssen.



*Hartmut Walravens*  
Foto © HWWA – Olaf Hoppe

# Archive vor der digitalen Herausforderung<sup>1</sup>

Rainer Hering

## 1. Einleitung

Digitalisierung wird besonders unter technischen Gesichtspunkten betrachtet. Das ist insbesondere für die Printmedien auch zweifelsohne eine sinnvolle Herangehensweise. Dass der Komplex der Digitalisierung für Archive in erster Linie ganz andere Komponenten mit sich bringt, soll in diesem Beitrag dargestellt werden. Aus in der Sache liegenden Gründen ist die Zusammenarbeit von Archiven, Bibliotheken und Dokumentationen wichtig, und dazu gehört auch die gegenseitige Information über die eigene Vorgehensweise. Diese weist bei aller Nähe doch manche Unterschiede auf, was für das gegenseitige Verständnis von Bedeutung ist. Sicherlich haben auch Archive technische Fragen im Verfahren der Digitalisierung zu klären, die denen anderer Bereiche vergleichbar sind. So ist z.B. in den Archiven die Langzeitarchivierung digitaler Daten über Jahrhunderte nicht wirklich gelöst.<sup>2</sup> Im Folgenden sollen jedoch die archivischen Strategien nach innen wie nach außen im Umgang mit der Digitalisierung von Unterlagen erläutert werden. Konkret sollen die Handlungsfelder am Beispiel der Einführung der elektronischen Dokumentenverwaltung in der Hamburger Verwaltung mit dem Namen ELDORADO aufgezeigt werden. Die Freie und Hansestadt Hamburg hat für die Begleitung dieses Prozesses einen Beirat eingesetzt, dessen Vorsitz beim Staatsarchiv Hamburg liegt.<sup>3</sup> Einleitend werden kurz die Aufgaben von Archiven zusammengefasst, anschließend die Rolle von Archiven in den E-Government-Strategien der öffentlichen Verwaltungen aufgezeigt, die erforderlichen mentalen Veränderungen durch die elektronische Dokumentenverwaltung herausgestellt und schließlich die konkrete Arbeit des Staatsarchivs Hamburg bei der Einführung von ELDORADO aufgezeigt.

## 2. Archivische Aufgaben

Welche Aufgaben nehmen Archive wahr? Archive sichern und bewahren Unterlagen, in erster Linie Schriftgut, aber auch andere Informationsträger wie Karteien, Dateien, Karten, Pläne, Bild-, Film-, Ton- und maschinenlesbare Datenträger sowie Siegelstempel. Der Zugang für die Benutzer und die Pflicht der Abgabe an das Archiv sind in den Archivgesetzen geregelt: Entsprechend dem Zuständigkeitsbereich des Archivs – für das Staatsarchiv



*Rainer Hering*  
Foto © HWWA – Olaf Hoppe

Hamburg sind das die Verfassungsorgane, Gerichte, Behörden und Dienststellen der Freien und Hansestadt Hamburg und der ihrer Aufsicht unterstehenden juristischen Personen des öffentlichen Rechts – müssen die betroffenen Stellen die nicht mehr zur Erfüllung ihrer Aufgaben benötigten Unterlagen dem Archiv anbieten. Dieses entscheidet dann über deren Archivwürdigkeit, d.h. die dauernde Aufbewahrung.

- 1 Vortrag, gehalten am 13. September 2005 auf der 9. Tagung der Arbeitsgemeinschaft Zeitungen im Forum Zeitschriften (GeSIG) «Digitalisierung und Langzeitarchivierung» im Hamburgischen Welt-Wirtschafts-Archiv. Für die Druckfassung wurden Anmerkungen mit weiterführenden Hinweisen und Belegen ergänzt.
- 2 Christian Keitel: Erste Erfahrungen mit der Langzeitarchivierung von Datenbanken. Ein Werkstattbericht. In: *Digitales Verwalten – Digitales Archivieren*. 8. Tagung des Arbeitskreises «Archivierung von Unterlagen aus digitalen Systemen» in Hamburg. Hrsg. von Rainer Hering; Udo Schäfer Hamburg 2004, S. 71-81 (Veröffentlichungen aus dem Staatsarchiv der Freien und Hansestadt Hamburg 19).
- 3 Vgl. Rainer Hering: Die archivische Begleitung der Einführung digitaler Systeme in der Verwaltung. In: *2. Norddeutscher Archivtag 23. bis 24. Juni 2003 in Schwerin*. Hrsg. von Rainer Hering; Michael Mahn. Nordhausen 2003, S. 423-432.

Archive sind nach dem Provenienzprinzip geordnet, also nach der Herkunft der Unterlagen, die in dem Zusammenhang archiviert werden, in welchem sie entstanden sind. Jede abliefernde Stelle bildet einen eigenen Bestand, innerhalb dessen die vorgefundene Ordnung, in der Regel der Aktenplan, beibehalten wird. Dadurch bleiben Zusammenhänge sowie Entscheidungsabläufe erhalten, und die Tätigkeit der jeweiligen Akten führenden Stelle ist nachvollziehbar. Zu bedenken ist: Archivalien sind in der Regel Unikate und liegen nur an einem Ort vor.

Ursprünglich hatten Archive ausschließlich rechtliche Funktionen, indem sie der Aufbewahrung juristisch wichtiger Urkunden, Verträge etc. dienten. Erst seit dem 19. Jahrhundert konzentrierte sich die Forschung immer mehr auf Archive als Ort historisch wertvoller und aussagekräftiger Quellen. Dennoch spielen die rechtlichen Aufgaben nach wie vor eine große Rolle, so werden z.B. in Hamburg die Originale der Bebauungspläne im Staatsarchiv verwahrt und jedem Interessierten – unter Glas und unter Aufsicht – vorgelegt. Darüber hinaus gibt es weitere Aufgaben von Archiven – in Hamburg z.B. die Zuständigkeit für die Verkehrsflächenbenennung.<sup>4</sup>

### 3. Archive und E-Government-Strategien

Eine klare strategische Ausrichtung der staatlichen wie der kommunalen Electronic-Government-Konzepte ist erforderlich, wenn sie für die Verwaltungen ertragreich sein sollen. Diese Konzepte dienen einer effizienteren und rationelleren Verwaltung, einem erhöhten internen Informationsaustausch und größerer Transparenz des Verwaltungshandelns; sie sollen also eine höhere Qualität der Dienstleistung herbeiführen. In diesem Zusammenhang dürfen veraltete und ineffiziente Verwaltungsprozesse nicht einfach *elektrifiziert* werden, vielmehr müssen diese zu Beginn der Implementierung von Dokumentenmanagementsystemen analysiert und optimiert werden, damit sie in der Verwaltung eine hohe Akzeptanz finden und erfolgreich eingesetzt werden können.

Und bereits an diesem Punkt sind die Archive gefragt und gefordert. Die Verwaltung, so könnte man einwenden, könnte diese Aufgabe auch erledigen, wenn es kein Archiv gäbe. Doch nur wenn die Schriftgutverwaltung *hinter den Kulissen* effektiv, medienbruchfrei und den rechtlichen Vorgaben entsprechend funktioniert, kann auch der Verkehr nach außen, die Kommunikation mit den Bürgerinnen und Bürgern, op-

timal gestaltet werden. Schriftgutverwaltung ist daher ein zentrales Element für alle E-Government-Verfahren.

Die Ausgangslage in den Verwaltungen ist allerdings weitgehend desolat: Schriftgutverwaltung hat in der Regel keinen angemessenen Stellenwert. Nicht selten sind ganze Akten in den bürokratischen Bermuda-Dreiecken zumindest mittelfristig, manchmal bis zur Pensionierung der Verantwortlichen oder noch darüber hinaus verschollen. Es hängt ganz von der jeweiligen Dienststellenkultur ab, wie genau eine ordnungsgemäße Schriftgutverwaltung praktiziert wird. Die Folgen für die Verwaltung sind offensichtlich: Ein nicht geringer Teil der Arbeitszeit – in manchen Behörden geht man von bis zu zehn Prozent aus – wird für das Suchen von Vorgängen benötigt, oftmals müssen Vermerke, Berichte und Schreiben mehrfach erstellt werden, weil sie oder das für ihre Ausarbeitung notwendige Material nicht greifbar waren. Von daher sollte es schon im Interesse jeder Dienststelle sein, durch die ordnungsgemäße Aktenführung deutlich effektiver arbeiten zu können. Faktisch jedoch scheint man eher einen enormen unnötigen Aufwand zu tolerieren, anstatt sich grundsätzlich dieser Problematik zu stellen und die vorhandenen Vorschriften wirklich in der Praxis durchzusetzen – im Regelfall lassen Führungskräfte suchen und müssen sich nicht selbst darum kümmern. Die so entstehenden Kosten sind enorm.

Es ist offensichtlich: Die archivischen Kompetenzen im Bereich der Schriftgutverwaltung können nachhaltig dazu beitragen, die beschriebenen Missstände zu beseitigen und somit die erfolgreiche Umsetzung der E-Government-Strategien nachhaltig zu unterstützen. Auch im digitalen Zeitalter muss gewährleistet sein, dass es keine Lücken in der Überlieferungsbildung gibt und die Nachvollziehbarkeit des Verwaltungshandelns gewährleistet ist.<sup>5</sup>

### 4. Elektronische Dokumentenverwaltung und mentale Veränderungen

Die Einführung neuer Technologien kann effektiv nicht ohne eine mentale Begleitung erfolgen. In den Verwaltungen arbeiten Menschen, und Menschen müssen mit der Technik umgehen – daher bestimmen die vorherrschenden Mentalitäten auch, ob der Einsatz neuer Technologien erfolgreich oder zum Scheitern verurteilt ist. Dies gilt in einem ganz wesentlichen Maße auch für die Speicherung von Unterlagen in digitalen

4 Hierzu knapp mit weiterführenden Hinweisen: Rainer Hering: Archive in Hamburg. In: *Auskunft* 18 (1998), S. 314-323.

5 Rainer Hering: Change Management und Archive. Archivische Aufgaben im Rahmen der Implementierung von Dokumentenmanagementsystemen in der Verwaltung – Das Beispiel ELDORADO in Hamburg. In: *Digitales Verwalten – Digitales Archivieren*. 8. Tagung des Arbeitskreises «Archivierung von Unterlagen aus digitalen Systemen» in Hamburg. Hrsg. von Rainer Hering; Udo Schäfer Hamburg 2004, S. 167-182, bes. S. 170-172 (Veröffentlichungen aus dem Staatsarchiv der Freien und Hansestadt Hamburg 19).



Systemen. Daher plädiere ich für eine Mentalitätsoffensive in Richtung auf die Archive und auf die Verwaltungen, um die Einsicht in die Notwendigkeit einer geordneten Schriftgutverwaltung zu befördern.<sup>6</sup> Auch in den Archiven bestehen oft Vorbehalte gegenüber Veränderungen und neuen Techniken. Die Archivarinnen und Archivare müssen daher auf die veränderte Situation rechtzeitig und umfassend vorbereitet werden:

- Wie die Verwaltungsbediensteten müssen auch die im Archiv Tätigen über die Veränderungsprozesse, ihre Auswirkungen auf die Archive und die dahinter stehenden E-Government-Konzepte informiert werden.
- Regelmäßige interne und/oder externe Fortbildungsangebote vertiefen und aktualisieren dieses Wissen. Die rechtzeitige Beratung der Behörden im Vorfeld, das Formulieren der archivspezifischen Anforderungen – z.B. die Definition von Schnittstellen – an die Verwaltungen und ihre Durchsetzung müssen im Vordergrund stehen. Von daher sind auch kommunikative Fähigkeiten im Bereich der Präsentations- und Moderationstechnik, der Rhetorik und des Konfliktmanagements unerlässlich.
- Die archivischen Kontakte zu den Verwaltungen im Sprengel müssen ausgebaut und intensiviert werden, damit die Archivarinnen und Archivare rechtzeitig von der Entwicklung entsprechender Verfahren Kenntnis erlangen und an wichtiger Stelle – z.B. in begleitenden Gremien – beteiligt werden können.
- Die Qualität der Schriftgutverwaltung in den Behörden muss geprüft werden, um diese adäquat beraten zu können. Entsprechende Kurse für Registraturbedienstete – gerade zum Thema Aktenverwaltung – können die Implementierungsprozesse in den Verwaltungen begleiten.

Diese Punkte sind für die Archive von zentraler Bedeutung, da sich bei ausschließlich in digitalen Systemen gespeicherten Unterlagen die Bewertungsentscheidung vom Schluss der Akte auf ihr Anliegen verlagert. Sobald ein Aktentitel gebildet worden ist, muss eine Entscheidung über die Archivwürdigkeit getroffen werden. Daher ist die richtige Zuordnung und Erfassung der einzelnen Schriftstücke unbedingt erforderlich. Die archivische Erschließung wird sich in der Regel auf die Übernahme der Metadaten der elektronischen Unterlagen konzentrieren. Dadurch entfällt ein großer Teil der bisherigen archivischen Tätigkeit, doch setzt eine solche Umstellung intensive Vorfelddarbeit voraus, damit die dann als archi-

vwürdig übernommenen Unterlagen auch benutzt werden können.

Zusammenfassend lässt sich also festhalten, dass sich das Schwergewicht des archivischen Handelns von der retrospektiven auf die prospektive Arbeit verlagert. Auf diese einschneidenden Veränderungen müssen auch die Archivarinnen und Archivare vorbereitet werden.<sup>7</sup>

Im Blick auf die von den Archiven zu betreuenden Verwaltungen müssen Archive auf mehreren Ebenen agieren, um das Bewusstsein für die grundlegende Bedeutung einer ordnungsgemäßen Schriftgutverwaltung sicherzustellen:

- Die Grundlagen der Schriftgutverwaltung müssen in den Ausbildungen aller Laufbahnen sowie in den Kursen für Quereinsteigende einen angemessenen und verbindlichen Stellenwert einnehmen.
- Gerade Vorgesetzte auf allen Ebenen müssen immer wieder, vor allem durch ihr eigenes Verhalten, die Bedeutung der Schriftgutverwaltung herausstellen und bei Problemfällen ihre Führungsverantwortung demonstrieren. Die Behördenkultur prägt auch im Umgang mit Unterlagen das konkrete Verhalten der Mitarbeiterinnen und Mitarbeiter – vorgegeben wird sie von oben.
- Die entsprechenden Regelungen müssen auf den aktuellen Stand gebracht und regelmäßig ins kollektive Behördengedächtnis gerufen werden.
- Darüber hinaus müssen Veränderungsprozesse in Bezug auf den Umgang mit Schriftgut rechtzeitig zuvor durch Veranstaltungen und Hinweise (Flyer, Infozeitungen etc.) unterstützt werden. Genaue Informationen über die Veränderungen und ihre Auswirkungen sowie über die Gesamtplanungen und Hintergründe sind für die Akzeptanz von Umstellungen bei den Beschäftigten eine wichtige Voraussetzung. Dies darf sich nicht nur auf den technischen Bereich beschränken, sondern muss auch die Schriftgutverwaltung als Ganzes im Blick haben.
- Gerade die Einführungsprozesse digitaler Systeme, insbesondere in den Pilotbereichen, müssen mental begleitet werden, um die Akzeptanz für die neuen Verfahren zu erhöhen. Die Eindrücke, die die *Piloten* gewinnen, werden schnell in den Behörden weiter getragen und können so starken Einfluss auf die flächendeckende Einführung nehmen und das Rahmenklima – positiv wie negativ – prägen. Grundsätzlich werden Veränderungen in lange bestehenden Systemen eher mit großen Vorbehalten bzw. sogar einer generellen Ablehnung gesehen, da sie als Bedrohung des Status quo und der eigenen Rolle in der Hierarchie empfunden werden.

6 Hierzu und zum Folgenden: Rainer Hering: Die Mentalitätsoffensive als zentrales Instrument der Archivierung von Unterlagen aus digitalen Systemen. In: *Scrinium*. Zeitschrift des Verbandes Österreichischer Archivarinnen und Archivare 58 (2004), S. 80-87.

7 Hering, Change Management (wie Anm. 5), S. 172-177.

- Generell sollten bei derartigen Veränderungen die Personalräte und die Gewerkschaften eng in den Prozess der Implementierung neuer Technologien einbezogen werden.
- Wichtig ist ebenfalls der Kontakt zu den IuK-Abteilungen und Softwarefirmen, die Hintergrundwissen über Verwaltung und Archivierungsprozesse benötigen, um sie adäquat technisch umsetzen zu können.<sup>8</sup>

Automation, aber auch an der Erarbeitung der dazugehörigen Regelungen und Handreichungen sichergestellt werden. Dieses Gremium trägt dazu bei, die in den letzten Jahrzehnten vernachlässigte Harmonisierung im Umgang mit Schriftgut zu fördern und zwischen allen, die in diesem Kontext arbeiten, eine Vernetzung von archivischen Facherfordernissen und technischen Entwicklungen zu ermöglichen. Aus informellen Strukturen erwachsen, bietet der Beirat hierfür einen Rahmen und stellt eine Schnittstelle für den



*Blick in den Hörsaal*  
Foto © HWWA – Olaf Hoppe

## 5. Die Einführung von ELDORADO in der Hamburger Verwaltung

Am Beispiel des Staatsarchivs Hamburg soll im Folgenden dargestellt werden, welche Schritte ein Archiv unternommen hat, um die Einführung digitaler Systeme in der Verwaltung zu begleiten. Seit Mitte der neunziger Jahre gibt es in verschiedenen Hamburger Behörden Projekte mit dem Ziel, die papiergestützte Aktenhaltung durch die elektronische Dokumentenverwaltung abzulösen. Diese sind unter dem Namen DOKUMENTA bekannt geworden – es ist ein Projekt, kein Produktname, der im letzten Jahr aus rechtlichen Gründen in ELDORADO (= **E**lektronische **D**okumentenverwaltung für **R**egistraturen mit **teraDOC**) geändert wurde. Um diese Aktivitäten zu koordinieren wurde im Februar 2002 der DOKUMENTA-Beirat – jetzt ELDORADO-Beirat – gegründet. Mit dem Beirat, dessen Vorsitz das Staatsarchiv übernommen hat, sollen insbesondere der behördenübergreifende Informationsaustausch sowie eine Mitwirkung aller Beteiligten an der Weiterentwicklung der

Transfer zwischen Praxis und administrativer Ebene dar.

An der Einführung dieser elektronischen Dokumentenverwaltung sind Organisatoren und Projektleitungen, Registratoren, Anwenderinnen und Anwender sowie IuK-Bereiche beteiligt. Allen Beteiligten soll der Beirat ein Forum für den Informationsaustausch über die hierbei zu lösenden Problemstellungen bieten. Konsequenterweise steht er daher allen Behörden und Ämtern offen. Für die eigentliche Arbeit hat der Beirat drei Unterarbeitsgruppen, in denen die Behörden und Ämter vertreten sind, die entsprechende Projekte eingesetzt haben. So wurden in der Unterarbeitsgruppe *Registratur* Empfehlungen für die fachliche sowie organisatorische Einrichtung des Verfahrens und laufend Anforderungen an die Software-Weiterentwicklung des Registraturmoduls verfasst; erarbeitet wurden zuletzt Verfahrensvorschläge zur systemgestützten Anbietung und Abgabe von Papieraktenbänden an das Staatsarchiv sowie Anregungen zu aussagekräftigen und eindeutigen Absenderangaben in E-Mails. Die Unterarbeitsgruppe *Projektleitungen/Organisatoren* dient dem

<sup>8</sup> Hering, Begleitung (wie Anm. 3), S. 427-429.

Austausch der Projektverantwortlichen und der Entwicklung von Verbesserungsvorschlägen, zudem setzt sie sich mit den Auswirkungen dieses Systems auf die Bewertung der künftigen Registraturaufgaben auseinander und hat Richtlinien für die Projektorganisation erstellt. Die Einsetzung einer Unterarbeitsgruppe *Technik*, die den Austausch der technisch Verantwortlichen dient, ist im letzten Jahr erfolgt.

Die nicht mehr bestehende Unterarbeitsgruppe *Regelungen* hatte Vorschläge für die notwendigen Vorschriften im Zusammenhang mit der Einführung von ELDORADO zusammengestellt. Diese sind nunmehr von der Finanzbehörde aufgegriffen worden, die in Hamburg für die Umsetzung der E-Government-Strategie zuständig ist. Für die Zeit vom 1. November 2003 bis zum 30. September 2005 ist ein behördenübergreifendes Projekt ELDORADO eingesetzt worden, das die Behörden und Ämter bei der Umsetzung ihrer Projekte zur Einführung elektronischer Dokumentenverwaltungen unterstützt.

In der Projekteinsatzungsverfügung ist geregelt, dass das Staatsarchiv dieses Projekt und die Projekte in den Behörden und Ämtern bei der fachlichen Ausrichtung der elektronischen Dokumentenverwaltungen berät, den Beirat als Gremium für den behördenübergreifenden Informationsaustausch leitet und an der Weiterentwicklung des Softwareproduktes *teraDOC®* mitwirkt sowie in Zusammenarbeit mit der Finanzbehörde unter Beteiligung des Hamburgischen Datenschutzbeauftragten die Regelungen zum Umgang mit elektronischen Dokumentenverwaltungen erarbeitet.

Die aktive Mitgestaltung der Einführung eines Dokumentenmanagementsystems bietet Archiven die Möglichkeit, ihre Interessen frühzeitig mit denen der Schriftgut produzierenden Stellen abzustimmen und so die eigene Arbeit langfristig zu erleichtern. Darunter fallen die Mitwirkung bei der Erstellung von Aktenplänen und Richtlinien zur Aktenführung bei konventionellen Papierakten wie bei elektronisch gespeicherten Unterlagen. Einen weiteren Schwerpunkt hat das Staatsarchiv Hamburg aufgrund der Nachfrage auf die Fortbildung von Registratorinnen und Registratoren in Form von Grundkursen gelegt. Weitere Angebote für die Fachbereiche im Bereich der Schriftgutverwaltung in Aus- und Fortbildung wären angebracht.

Um die Sensibilität für die Probleme der Schriftverwaltung zu fördern, sind Gespräche mit Verwaltungsleitenden begonnen worden. Nicht nur im Bereich der Registraturen, sondern vor allem in den Fachbereichen ist es von grundlegender Bedeutung, das Bewusstsein für einen ordnungsgemäßen Geschäftsgang und eine entsprechende Aktenführung zu aktivieren, damit die Einführung elektronischer Speichersysteme erfolgreich sein kann – sonst wird man unbrauchbare Daten oder nicht mehr recherchierbare Dokumente und somit einschneidenden Informationsverlust nicht ausschließen können. Die seit gut drei Jahrzehnten vielfach zu konstatierenden Nachlässigkeiten im Umgang mit Schriftgut in den Verwaltungen müssen behoben werden, wenn neue Technologie erfolgreich implementiert werden soll.<sup>9</sup>

## 6. Schlussfolgerungen

Für Archive bedeutet der Komplex Digitalisierung, wie deutlich geworden ist, nicht nur die Lösung von technischen Fragen – Archive sind in diesem Kontext in den Bereich der Verwaltungsreform eingebunden und müssen auf unterschiedlichen Ebenen agieren. Ihnen kommt, wie das Beispiel des ELDORADO-Beirates in Hamburg zeigt, in diesem Prozess eine wesentliche strategisch-planerische Funktion zu. Sie heilen Brüche im Bereich der Schriftgutverwaltung, um die sich sonst keine andere Institution in der Verwaltung kümmern kann, dienen dem Transfer zwischen den verschiedenen Ebenen im Bereich der Verwaltungsreform und tragen dazu bei, die strategischen Ziele der E-Government-Konzepte in die konkrete Praxis vor Ort umzusetzen. Dies ist ein wesentlicher Beitrag nicht nur zur Überlieferungssicherung für die Zukunft, sondern auch zur Umsetzung der E-Government-Strategie für die Gegenwart. Daher nehmen Archive eine bedeutende Position ein im Rahmen der Verwaltungsreform als Antwort auf die gesamtgesellschaftlichen Veränderungen. Zugleich ist es die Chance für Archive, sich zwischen Vergangenheit und Gegenwart im Blick auf die Zukunft neu zu positionieren und so ihre Existenz als (öffentliche) Einrichtungen zu sichern.

---

9 Heinz Vogel: Die elektronische Dokumentenverwaltung für Hamburg. In: *Digitales Verwalten – Digitales Archivieren*. 8. Tagung des Arbeitskreises «Archivierung von Unterlagen aus digitalen Systemen» in Hamburg. Hrsg. von Rainer Hering; Udo Schäfer Hamburg 2004, S. 143-151; Ivy Gumprecht: Dem Informellen einen Rahmen geben. Die Einführung des digitalen Dokumentenmanagements unter besonderer Berücksichtigung der Kategorie des Informellen in Veränderungsprozessen. In: Ebd., S. 153-166; Hering, Change Management (wie Anm. 5), bes. S. 177-181.

# Digitale Archive<sup>1</sup>

Dr. Martin Warnke

Universität Lüneburg

## *Das Phantasma der Unvergänglichkeit*

Die Funktion des Archivs ist eine unmögliche. Der Aufgabe, die Archivalien zu bewahren, indem man sie vor dem zerstörerischen Zugriff ihrer Nutzer bewahrt, ist unvereinbar mit der Aufgabe, die eingelagerten Materialien allgemein zugänglich zu machen. Jeder Griff nach dem Manuskript hinterlässt Fingerabdrücke, das Anfertigen von Faksimiles beansprucht die brüchigen Kostbarkeiten, jeder Blick bleicht die Schätze.

So ganz anders und viel versprechend benehmen sich da digitale Speicherinhalte. Sie lassen sich mit Hilfe raffinierter Verfahren so gut gegen Abschreibefehler schützen, dass man sagen kann, sie ließen sich verlustfrei kopieren und vervielfältigen. Die Unterhaltungsindustrie weiß davon ihr klagend' Lied zu singen, denn sie hat seinerzeit in Form der Audio-CD Musik verkauft, die sich endlos und verlustfrei kopieren lässt, und zwar von Jeder und Jedem im Besitze eines CD-Brenners.

Digitale Datenhaltung könnte also die Lösung des archivarischen Dilemmas sein, könnte dem Metier, das so nach Staub und Moder duftet, neuen, modernen, effizienten Auftrieb geben, Speicherungsprobleme lösen, wie auf so vielen anderen Gebieten des spätmodernen Lebens.

Ein Grund für die Anrufung des Digitalen mag die Tatsache sein, dass die digitalen Medien nach und nach die alten, analogen abzulösen scheinen. Es gibt zwar Retro-Moden – die Leidenschaft für die gute alte Schallplatte aus Vinyl bei den jüngeren Musikkonsumenten etwa gehört dazu –, aber insgesamt scheint das Digitale handstreichartig eine analoge Bastion nach der anderen einzunehmen, auch die der Archive. Es wäre dann mithin normal, auch auf diesem Feld digitale Tendenzen zu erwarten.

Doch über solche Normalität eines digitaltechnischen Imperialismus hinaus gibt es allerorten signifikante Ewigkeitsvorstellungen zum Digitalen, die den Schluss von «digital» auf «Archiv» nahe legen. Es ist allerdings, um die Pointe vorwegzunehmen, ein Kurzschluss, ein Phantasma und kein technisches Faktum, um das es hier geht. So ist etwa die Rede von der Immaterialität der computerisierten Daten, es gelte «für

den digitalen Raum, dass in virtuellen Welten ... gerade die Möglichkeit des Defekts und Verfalls»<sup>2</sup> fehle, womit zwar der Mainstream-Diskurs über das Digitale richtig wiedergegeben wird, was aber dennoch technisch nicht zutrifft. Das Digitale, wir werden es noch sehen, ist ein archivarischer Albtraum, und zwar aus vielerlei Gründen, die noch zur Sprache kommen sollen.

## *Ein digitales Archiv von überhaupt allem*

Kommen wir aber zunächst zu Aspekten der reinen Quantität, zu technischen Randbedingungen digitaler Speicherung als Teil einer Archiv-Strategie: Was halten Sie davon, einmal grundsätzlich überhaupt alles archivieren zu wollen?

Fangen wir klein an, mit dem Gedächtnis eines Menschen etwa. Ein gewisser kognitiver Psychologe mit dem Namen Thomas Landauer aus Boulder, Colorado, hat geschätzt, dass der Informationsgehalt eines typischen menschlichen Langzeitgedächtnisses zwischen 150 und 225 MB umfasse.<sup>3</sup> Das ist immerhin der Text von etwa ein paar Hundert Paperbacks. Manchmal beschleicht mich das Gefühl, bei mir müsse es sehr viel weniger sein, aber auch dann, wenn das Gedächtnis gut sein sollte, bei Ihnen vielleicht: es passt ohne Probleme mittlerweile auf briefmarkengroße Speicher-Karten, wie ich sie in meine digitale Kamera schiebe, um Urlaubs-, also Erinnerungsfotos zu machen. Ein gängiger PC müsste nur etwa ein Hundertstel seines Festplatten-Speichers opfern, um das Gedächtnis einer typischen Kleinfamilie abzuspeichern. Sofern dieses Bild überhaupt stimmt, wohlgemerkt, sofern also Erinnerung und Speicher vergleichbar wären. Rechnen wir diesen Wert auf die gesamte Menschheit hoch, so ergeben sich 1.350 PetaByte, also kein Problem, wie wir gleich sehen werden. Auch was ein PetaByte ist, wird gleich geklärt.

Nehmen wir uns als nächstes vor, was – anders als ein menschliches Gedächtnis – ohnehin schon als Datenmasse externalisiert ist: die Texte, Bilder und Töne, die die Menschheit über technische Medien ständig absondert.

1 Vortrag, gehalten auf der 9. Tagung der AG Zeitungen im Forum Zeitschriften (GeSIG), Hamburg, 13./14.9. 2005

2 Wolfgang Ernst: *Das Rumoren der Archive*. Berlin: Merve 2002. S. 28

3 Landauer, Thomas: How Much Do People Remember? Some estimates of the quantity of learned information in long-term memory. In: *Cognitive Science*. 10.1986, S. 477-493.

Sehen wir einmal nach.<sup>4</sup>

Die Library of Congress hält 20 Millionen Bücher, macht 20 Tera-Byte (Tera ist  $10^{12}$ , eine Million Millionen), 13 Millionen Fotos, macht 13 TB, 4 Millionen Karten und Pläne, macht 200 TB, 500.000 Filme zu etwa 500 TB, 3,5 Millionen Klangdokumente: 2.000 TB. Zusammen ergibt das etwa 3.000 TB oder 3 PetaByte, Peta heißt  $10^{15}$ , eine Milliarde Millionen.



*Dr. Martin Warnke*  
Foto © HWWA – Olaf Hoppe

Nehmen wir noch mehr hinzu: alles Schriftliche, alle Fotos, alle neu ausgestrahlten Fernsehsendungen, der Hörfunk, alle publizierte Musik und vor allem: alle geführten Telefonate, die einige spezielle staatliche Stellen sowieso gern aufzeichnen würden, machen auf der ganzen Erde pro Jahr ca. 4.600 PB oder 4,6 ExaByte, Exa:  $10^{18}$ , eine Milliarde Milliarden. Eine ganze Menge. Aber: das ist weniger als der jährlich von der Industrie gefertigte und verkaufte elektronische Speicher.

Mithin: wenn wir wollten, könnten wir – alles speichern.

Wenngleich kritische Stimmen angesichts der enormen menschlichen Mitteilbarkeit, vor allem am Telefon, fragen: «So who wins the war here – a handful of cy-

brarian archivists, or the entire chattering human race?»<sup>5</sup> «Schnatternde menschliche Rasse» nennt man uns, das haben wir nun davon.

### *Das World Wide Web als Archiv*

Doch man muss gar nicht zu den Sternen, also zum gesamten Weltgeist greifen, um auf erhebliche Probleme zu stoßen: Nehmen wir uns vor, was ohnehin digital da und weltweit verfügbar ist: das Internet, speziell das World Wide Web!

Ist es selbst vielleicht schon ein digitales Archiv? Es liegt ja bitweise zugreifbar vor.

Technisch umfasst das Internet die Gesamtheit aller Datenleitungen, aller Computer, die Relais-Funktionen übernehmen, aller Datenpakete, die transportiert werden, und aller Server und Clients, die Informationen anbieten und abfordern.

Wären Internet-Pakete solche wie bei der Paketpost, so würden sie, zusammengenommen, den Bestand des Internet ausmachen. Doch: Internet-Pakete sind vergänglich. Damit alle die Daten, die zu übertragen sind und die gelegentlich verschiedene Wege ausprobieren müssen, um dann schließlich hoffentlich beim Empfänger anzukommen, damit diese Daten nicht sinnlos herumliegen und alles verstopfen, hat man ihnen einen Selbstzerstörungs-Mechanismus eingebaut, das TTL-Feld, das heißt «Time To Live». Normalerweise trägt diese digitale Lebenserwartung anfangs den Wert 255, und bei jeder Passage über eine Computer-Relais-Station, bei jedem *hop*, wird dieser Wert um Eins dekrementiert. Ist er Null, wird das Paket gelöscht.<sup>6</sup>

Das Internet ist offenbar eine Mischung aus reinem Transport und temporärer Speicherung, die bei den Paketen nie länger als ein paar Handvoll Sekunden dauert. Es kommt für die Frage nach der Lebensdauer von Internet-Dokumenten also auf die Endgeräte an, denn: die Übertragungspakete verschwinden von selbst. Was nicht mehr auf den Servern liegt, kann nicht mehr erreicht werden, es schlägt der berüchtigte «Error 404, document not found» zu.

Das Netz selbst ist also offenbar als Archiv untauglich.

Lassen Sie uns nachsehen, wie es um die Daten auf den Servern selbst bestellt ist.

Ein Blick auf die Statistik eines typischen Servers – hier desjenigen der eigenen Universität – besagt, dass während des überwachten Zeitraums rund 2,5% aller Anforderungen an WWW-Seiten nicht bedient werden

4 nach Hartmut Krech: Der Weltgeist : 1350 Petabyte, In: *Die Zeit*, 53.1998,46(5.Nov.)

5 Dead Media Working Note 26.6

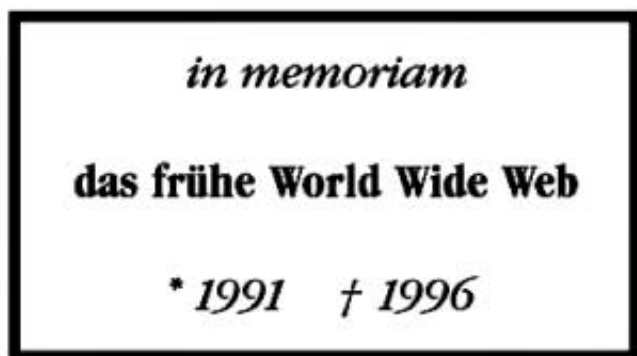
6 Tanenbaum, Andrew S. T: *Computernetzwerke*. - München : Addison-Wesley, 2000. S. 438

konnten, weil es sie nicht mehr gab. Sie sind also verschwunden. Gelöscht. Nicht mehr für wichtig oder erhaltenswert befunden, vielleicht durch Neuere ersetzt.

Das wirft die Frage auf, wie hoch die Lebenserwartung einer WWW-Seite ist, bis sie von den Servern verschwunden ist. Tage, Monate, Jahre?

Brewster Kahle weiß, wie viele WebSites stehen, und er weiß auch, wie lange sie am Himmelszelt des Internet glänzen: noch Ende 1996 sollten es 400.000 WebSites mit insgesamt 1,5 TeraByte Umfang sein, im März 2000 waren es 13,8 TB<sup>7</sup>, also wie vorhersehbar etwa 10 Mal so viele. Und wie lange bleibt so ein Dokument typischerweise auf seinem Server erreichbar? Auch hier weiß Brewster Kahle Antwort. Seine Verfahren registrieren, wenn sich etwas auf einem Server verändert, wenn Daten verschwinden. Die Veränderungsrate der Dateien 1996 von 600 GB im Monat lässt den Schluss zu, dass die mittlere Lebenserwartung eines Dokuments im WWW ganze 75 Tage beträgt, denn bei 600 GB pro Monat muss man zweieinhalb Monate oder 75 Tage warten, bis im Mittel 1,5 TeraByte, also alles, weg ist. Nach 75 Tagen sind die meisten Seiten also nicht mehr dort, wo sie einmal vorzufinden waren. Referenzen, Zitate, Bezüge auf sie liefern dann den «Error 404, document not found». Zweieinhalb Monate sind wahrhaftig nicht viel, vergleicht man das mit der Langlebigkeit analoger Medien, etwa den viele Tausend Jahre alten frühesten Texten der Menschheit, den Proto-Keilschrift-Tafeln aus Mesopotamien. Selbst Bücher auf säurehaltigem Papier kommen gut dagegen weg.

Was heißt das für die Inkunablen des Internet, die «Wiegendrucke», Erstlinge, Erstauflagen?



Sie sind verschwunden. Die Frühzeit des WWW ist verloren. Es gibt die Dokumente nicht mehr –, weil niemand sie bewahrt oder archiviert. Ständig verschwinden Dokumente im digitalen Nirwana, und niemand scheint das aufhalten zu können.

Fast niemand. Brewster Kahle versucht es. Seine Organisation heißt «The Internet Archive», und sie macht Schnappschüsse des gesamten WWW, weshalb er auch die Kenndaten des WWW besitzt. In Form einer Skulptur ist ein solcher seiner Schnappschüsse der Library of Congress als Geschenk übergeben worden.<sup>8</sup> Sie zeigt auf den Monitoren die digitale Momentaufnahme des Web von 1997, etwa 2 TeraByte Daten. Man kann nur hoffen, dass diese zwei TeraByte nicht die ganze Zeit draußen im Freien herumstehen, sie werden es ohnehin schon schwer genug haben, die nächste Zukunft zu überdauern.

Man kann den Internet-Archiv-Service auch als normaler *user* in Anspruch nehmen: unter dem Namen «Wayback Machine»<sup>9</sup> kann man sich HomePages aus vergangenen Tagen anzeigen lassen. Besonders ergiebig ist dieser Dienst übrigens nicht, oft gibt es nur die Leitseiten, manchmal auch diese nicht mehr.

Zu den üblichen Verlusten kommt hinzu, dass vieles im Web erst bei der Abfrage entsteht, etwa die Bahn- oder die Telefonauskunft oder aktuelle Preis- und Produktlisten oder alles was mit Content-Management-Systemen gemacht wird, weil die Seiten dynamisch erst zur Laufzeit aus Datenbanken erzeugt werden. Diese sind den Suchrobotern ohnehin unzugänglich und können von ihnen nicht archiviert werden.

Fazit: das WWW taugt bei einer Dokument-Lebensdauer von zweieinhalb Monaten nicht zum Archiv. Das Web dennoch zu archivieren ist ein heroischer Akt, der seine Frühzeit ohnehin nicht mehr retten kann. Und auch die meisten der zeitgenössischen Websites lassen sich so nicht dem Vergessen entreißen, sie gehen ständig und unwiederbringlich verloren.

#### *Lebensdauer digitaler Speichermedien*

Nehmen wir einmal an, alles, was archiviert werden sollte, wäre tatsächlich schon von irgendeinem Cyberspace-Librarian, einem «cybrarian», digital gespeichert. Sind die Kisten voller Bänder, Platten und CDs dann ein sanftes Ruhekissen für den guten Menschen, dem der Erhalt der Kulturgüter so am Herzen liegt?

Lassen wir Augenzeugen berichten (ich übersetze aus dem Amerikanischen):

«In Taiwan habe ich Disketten gesehen, die voller Pilze und Schimmel waren (grün und haarig). In Missouri habe ich Disketten-Hüllen gesehen, die von der Hitze im Inneren eines Autos völlig verzogen waren. In Utah habe ich Disketten voller Flugsand gesehen. Der Be-

<sup>7</sup> <http://www.archive.org>

<sup>8</sup> <http://www.loc.gov/loc/lcib/9811/alexa.html>

<sup>9</sup> <http://www.archive.org/web/web.php>

sitzer sagte mir, es habe ein seltsames Kratzgeräusch gegeben, als sie seinem Laufwerk den Lesekopf zu Schrott geschliffen haben.»<sup>10</sup>

Aber selbst vorsichtigeren Zeitgenossen, ja selbst den sagenhaften Raketenwissenschaftlern bei der NASA, mit Technologie doch auf Du und Du, ist schon Schlimmes widerfahren:

«Der Inhalt von 1,2 Millionen Magnetbändern, die drei Jahrzehnte amerikanische Raumfahrt dokumentieren, ist hinüber», so schreibt Dr. Michael Friedewald vom Fraunhofer-Institut für Systemtechnik und Innovationsforschung in Karlsruhe.» Die Aufbewahrung in Lagerhäusern hat den wertvollen Tapes nicht gut getan. Die Trägerfolie löst sich auf, die Bänder zersetzen sich.<sup>11</sup> Bei Audio-Bändern kennt man das als Quietschen beim Abspielen, als ein Zeichen dafür, dass bald alles vorbei sein wird.

Wie lange Medien halten, hängt davon ab, wie sorgsam sie aufbewahrt werden. Aber auch bei größter Vor- und Umsicht ist ihnen nur eine gewisse Spanne beschieden. Etwa so:

Bänder halten zwischen 2 und 30 Jahren. CD-ROM (die silbrigen, industriell gefertigten) 5 bis 100 Jahre, ebenso wie magneto-optische Platten. Ähnliches kann von anderen Medien gesagt werden, die nur ein Mal beschrieben werden.

Kurz und knapp: «Computerbänder, Videobänder und Tonbänder halten ungefähr so lange wie ein Chevy oder ein Pudel.»<sup>12</sup> Dasselbe kann von digitalen Medien im Allgemeinen behauptet werden. Dann schlägt *data rot* zu, die kalte Datenrotte. Für Menschen mit starken Nerven gibt es eine WebSite<sup>13</sup>, die den Sound von «dying disks» – «sterbenden» Festplatten – zu Gehör bringt. Wer solches hört, ohne dass ihr oder ihm das Blut in den Adern gefriert, hat noch nie einen Platencrash erlitten.

Und doch ist das noch nicht einmal der Schlimmste, denn nicht nur die Träger altern, auch die Lesegeräte kommen in die Jahre und sterben einfach aus. Haben Sie noch ein 5 1/4-Zoll-Laufwerk? Wo kann ich meine Lochkarten von vor zwanzig Jahren einlesen lassen? Kennt jemand noch das Format eines Schneider-Schreibcomputers?

Schätzt man somit die Lebensdauer von Datenformaten anhand der Lesegeräte, die damit etwas anfan-

gen können, so landet man bei noch sehr viel niedrigeren Werten, die bei fünf bis zehn Jahren liegen. Danach hilft nur noch ein Computermuseum mit geschickten Technikerinnen oder Technikern, die ohne Ersatzteile, die die Industrie natürlich nicht mehr liefern kann, durch Basteln die alten Geräte am Laufen halten.

Glücklicherweise ist *retro computing* zum Sport einiger Unverzagter geworden, die Spaß daran haben, mit Computern zu spielen, in die man noch hineinblicken kann.<sup>14</sup>

Eine Interessengruppe namens «Dead Media» bringt ihre «Dead Media Working Notes»<sup>15</sup> heraus, in denen natürlich nur von toten Medien die Rede ist, etwa der pneumatischen Post in Paris, Hummels Telediagrammen, den «Peek-a-Boo-Index-Cards», aber auch den obsoleten Computern, die nach Gordon Moores Gesetz alle eineinhalb Jahre durch ihre schnelleren Nachfolger ersetzt werden. Dort, wo tatsächlich alte Formate und die Anmutung der alten Hardware gebraucht werden, hilft nur noch eines: man muss das alte Zeug auf neuem Gerät simulieren, oder, wie Informatiker sagen: emulieren.

Dead-Media-Aktivist Bruce Sterling merkt an (übersetzt aus dem Amerikanischen): «Diese Entwicklung ist für Dead Media Studies interessant, weil die rasche Folge, durch die elektronische Komponenten obsolet werden, immer ein Kainsmal der elektronischen Medien war. Simulation und Emulation toter Hardware wird weiter an Bedeutung zunehmen, solange der Friedhof toter Multimedien nur so wimmelt von Opfern des Mooreschen Gesetzes.»<sup>16</sup>

So sieht sich die amerikanische Air Force auch gezwungen, spezielle Vorsorge zu treffen, um beim Generationswechsel elektronischer Schaltungen, mit denen die modernen Kampffjets ja vollgestopft sind, nicht auch gleich neue Flugzeuge bauen zu müssen. Denn die Hardware und die darauf implementierte Software müssen einwandfrei laufen, damit der Vogel am Himmel bleibt, auch wenn die Chips schon längst nicht mehr hergestellt werden. Also gibt es Ersatzteil-Probleme.

«Die Teileknappheit rührt größtenteils von der kurzen kommerziellen Lebensspanne digitaler elektronischer Komponenten, verglichen mit dem langen Wartungsleben von Waffensystemen. Eine digitale Komponente z.B. mag eine Lebenszeit von 18 Monaten haben, wäh-

10 <http://www.phlab.missouri.edu/~ccgreg/tapes.html>

11 Digital-Alzheimer. In: *Macmagazin*. 2000,10, S. 134

12 <http://www.phlab.missouri.edu/~ccgreg/tapes.html>

13 <http://kiza.kcore.de/technology/harddisks.shtml>

14 Borchers, Detlef: Der Glanz von Gestern. In: *Süddeutsche Zeitung*. 56.2000,233(10.10.), S. V2/15.

15 <http://www.well.com/user/jonl/deadmedia/NOTES26-28.txt>

16 ebd.

rend ein Waffensystem, das diese Komponente verwendet, oft Jahrzehnte im Einsatz ist.»<sup>17</sup> Das kommt sehr teuer. Und es führt uns wieder zurück zu unserem eigentlichen Thema, den digitalen Archiven, die ohne heftigste Anstrengungen auf dem Feld einer aufwendigen Daten-Archäologie sehr schnell digitalem Vergessen anheim fallen.

### *Daten-Archäologie*

Jeff Rothenburg ist durch ein Diktum bekannt geworden, das da lautet: «Digital documents last forever – or five years, whichever comes first.»<sup>18</sup> Und er weiß, wovon er spricht, denn sein in die Ewigkeit, das heißt in die nächsten fünf Jahre, greifendes Urteil ist Frucht einer aus- ausführlichen und sehr überzeugenden Studie zum Thema digitaler Dokumentarchivierung. Das Resultat lautet: «... there is – at present, no way to guarantee the preservation of digital information.»

Wenngleich Garantien nicht abzugeben sind, so gibt es doch eine Strategie, die, wenn verfolgt, Hilfe verspricht, und von der auch schon die Rede war:

«The best way to satisfy the criteria for a solution is to run the original software under emulation on future computers.» Die Originalsoftware unter einer Emulation hoffnungslos veralteter Betriebssysteme längst verrotteter Hardware muss immer wieder zum Laufen gebracht werden, um Funktion und Anmutung obsoleter digitaler Dokumente wiederherzustellen.

Zunächst ist jedoch über die Jahrzehnte der Bitstrom der digitalen Daten zu erhalten, umzukopieren auf je neue Speichermedien, als Maßnahme gegen den Verschleiß von Trägermaterial und Gerätschaft, zu ergreifen, etwa alle ein bis zwei Jahre. Anschließend hat man dafür zu sorgen, dass die Daten auch korrekt interpretiert werden. Wenn man nicht weiß, wie der Inhalt eines Mediums zu interpretieren ist, ist man noch nicht viel weiter. Neben dem *data rot* war auch fehlende Beschriftung eine der Ursachen für die massiven Datenverluste der NASA. Metadaten sind anzubringen. Sie beschreiben, was wie zu interpretieren ist. Schlägt man die Emulations-Strategie ein, müssen die Metadaten beschreiben, unter welchem Betriebssystem und auf welcher Hardware die Software lief, die die Daten einstmals interpretierte. Rothenburg schreibt dazu: «This point cannot be overstated: in a very real sense, digital documents exist only by virtue of software that understands how to access and display them; they come into existence only by virtue of running this software.»

17 ebd.

18 Rothenberg, Jeff: Avoiding technological quicksand. 1998. Nach: <http://www.clir.org/pubs/reports/rothenberg/>

19 Derrida, J.: *Dem Archiv verschrieben – Eine Freudsche Impression*. (Übers. Von Gondek, H.-D. und Naumann, H.) Berlin: Brinkmann + Bose 1997. Original: *Mal d'Archive*. S. 11

20 ebd., S. 35

Anzulegen ist also auch ein Archiv von Betriebssystem-Emulationen in allen relevanten Versionen und eine Sammlung von Software, die die Dokumente interpretieren kann. Ständig frisch umkopierte Dokumente könnten so auf neuesten Computern unter Betriebssystem-Emulationen von Originalsoftware angezeigt, mithin archiviert und verwendet werden.

Nur dann, wenn wir diesen Aufwand treiben, werden digitale Dokumente archiv-fähig. Sie sehen, das ist nichts mehr für Privatleute, hier sind staatliche Institutionen gefragt, die eine solche außerordentlich aufwändige Arbeit kontinuierlich leisten. Der Archiv-Begriff, der ja auf das Amtshaus des Archonten zurückgeht, der die Macht über die Regierungsdokumente ausübt, zeigt seine ursprüngliche Bedeutung.

Derrida schreibt in *Mal d'Archive*:

«... <archive>, sein einziger Sinn, vom griechischen *archeion*: zuerst ein Haus, ein Wohnsitz, eine Adresse, die Wohnung der höheren Magistratsangehörigen, die *archontes*, diejenigen, die geboten. Jenen Bürgern, die auf diese Weise politische Macht innehatten und bedeuteten, erkannte man das Recht zu, das Gesetz geltend zu machen oder darzustellen. Ihrer so öffentlich anerkannten Autorität wegen deponierte man zu jener Zeit bei ihnen zuhause, an eben jenem Ort, der ihr Haus ist (ein privates Haus, Haus der Familie oder Diensthaus), die offiziellen Dokumente. Die Archonten sind zunächst Bewahrer. Sie stellen nicht nur die physische Sicherheit des Depots und des Trägers sicher. Man erkennt ihnen auch das Recht und die Kompetenz der Auslegung zu. Sie haben die Macht, die Archive zu interpretieren.»<sup>19</sup>

Und weiter:

«... die technische Struktur des archivierenden Archivs bestimmt auch die Struktur des archivierbaren Inhalts schon in seiner Entstehung und in seiner Beziehung zur Zukunft. Die Archivierung bringt das Ereignis im gleichen Maße hervor, wie sie es aufzeichnet. Das ist auch unsere politische Erfahrung mit den so genannten Informationsmedien.»<sup>20</sup>

Insbesondere wird das die Erfahrung mit digitalen Archiven sein. Nur Macht und Geld können sie vor dem Verfall retten, die so viel anfälliger sind als ihre analogen Vorläufer. Und wer die überkommenen Dokumente so unter seiner Ägide hat, kann sie nach Belieben einsetzen, interpretieren, vorenthalten.

Sehr real vernichtet der technische Fortschritt, der unabdingbar ist, um immer mehr Dokumente in digitale



Archive einstellen zu können, ganz real also vernichtet genau dieser Fortschritt das Archiv selbst: digitale Archive als Schauplätze eines *mal d'archive*, eines digitalen Archiv-Übels.

### *XPliztheit*

Ein wenig lässt sich das technische Problem entschärfen, indem man digitale Dokumente möglichst in solchen Formaten abspeichert, dass der Inhalt noch lange interpretierbar bleibt, dass also schon durch den Erhalt des Bitstroms der Daten wesentliches gerettet wird. Dies versäumt zu haben, war der Fehler der NASA-Leute, die ihrer Nachwelt nur unverständliche Bitfolgen auf ihren Bändern hinterließen.

Das wichtigste dieser Formate heißt XML, eXtensible Markup Language, und es hat den Vorteil, völlig vom Erscheinungsbild der Dokumente und der Funktionalität der anzeigenden Programme abzusehen, damit unabhängig zu machen von Software und Hardware, weshalb man diese auch nicht über die Jahrzehnte und Jahrhunderte zu retten braucht. XML kodiert sehr explizit Inhalt, Struktur und Semantik der Daten, es ist lesbar von Menschen und von Programmen. Schafft manes, den Bitstrom zu erhalten, indem man immer wieder auf neue Speichermedien umkopiert, hat man die Chance, durch direkten Augenschein und durch Programme die Daten immer wieder interpretieren zu können, aber auch, wenn die Darstellungssoftware nicht mehr läuft, neu algorithmisch interpretieren zu müssen.

Man fügt den Daten Metadaten zu, kleine Schildchen gleichsam, die alle Datenatome etikettieren. Das World-Wide-Web-Consortium, das XML betreut, gibt in seinem Einführungskursus<sup>21</sup> folgendes einführende Beispiel:

```
<note>
<to>Tove</to>
<from>Jani</from>
<heading>Reminder</heading>
<body>Don't forget me this weekend!</body>
</note>
```

Diese eigentlich wahrlich zu Herzen gehende Bitte Janis an Tove verliert durch ihre Expliztheit massiv an der ansonsten in Liebesdingen erforderlichen Zweideutigkeit, und das ist für unsere Archiv-Zwecke auch gut so. Absender und Empfänger bleiben genau so wenig

im Dunkeln wie der Charakter des Schreibens und seine Unterteilung in Kopf und Körper. Jedem Fitzelchen sein Schildchen, sein *tag*, ersichtlich an den spitzen Klammern.

Welchen Fortschritt in Hinblick auf spätere Verstehbarkeit diese extensive und explizite Etikettierung hat, wird vielleicht noch deutlicher an Daten, deren Semantik sich nicht unmittelbar erschließt.

So führt eine niederländische Forscherinnen-Gruppe<sup>22</sup> folgendes Beispiel an, zunächst im dürren ASCII-Code,

```
26502 Martensz Matheeus Kruidenier Antwerpen 19-05-1586 B 35
```

danach in XML,

```
<record>
<row>
<persID> 26502 </persID>
<family_name> Martensz </family_name>
<first name> Matheeus </first_name>
<profession> Kruidenier </pfeession>
<origin> Antwerpen </origin>
<date_of_entry> 19-05-1586 </date_of_entry>
<entry_number> B </entry_number>
<entry_page> 35 </entry_page>
</row>
</record>
```

was sich doch, das muss man zugeben, entschieden klarer liest.

Diese *tags* darf man selbst erfinden, und so kann jeder Archivarin, jeder Archivar ihren und seinen eigenen Satz von Metadaten erfinden. Dieser Umstand schafft Freiheit und somit Probleme, und dieser versuchen die *cybrarians* durch die Standardisierung von Metadaten Herrin und Herr zu werden.

Bei den Metadaten-Standards, die Identifizierung und Suche erleichtern, gibt es gute Vorschläge, die prominentesten lauten: Dublin Core und die *Open Archive Initiative*<sup>23</sup>. Diese Konventionen könnten, wenn sie weite Verbreitung fänden, die Suche und das Auffinden digitaler Dokumente erheblich vereinfachen. Ihre Uni-Bibliothek und hoffentlich auch Ihr Rechenzentrum wissen, worum es geht, wenn Sie sie nach Näherem fragen. Die Verwendung solcher Standards macht

21 [http://www.w3schools.com/xml/xml\\_what\\_is.asp](http://www.w3schools.com/xml/xml_what_is.asp)

22 Nispen, Annelies van; Kramer, Rutger; Horik, René van: The eXtensible past : the relevance of the XML data format for access to historical datasets and a strategy for digital preservation. In: *D-Lib Magazine*, 11.2005,2. <http://www.dlib.org/dlib/february05/vannispenn/02vannispenn.html>

23 <http://www.openarchives.org/>

Dokumentarchivierung damit noch nicht zu einem Kinderspiel, aber doch wenigstens zu einem, bei dem auch akademische Institutionen mittun können. Doch machen wir uns nichts vor: die Archivierung digitaler Daten erfordert – genau wie bei ihren analogen Vorläufern – ständige Pflege, einen großen Aufwand und: viel Geld.

Für Textdokumente kann XML eine Lösung sein, aber für Bilder und für Klänge, für Filme und alles, was gerade kein Text ist, bleibt die Kluft, dass die maschinenlesbaren Kodierungen gerade für Menschen unverständlich sind, das das *mal d'archive* auch durch Zaubersprüche in XML nicht gebannt werden kann.

In einem fünfjährigen Projekt haben meine Kolleginnen und Kollegen und ich eine Erschließungs- und Archivierungsarbeit an der Kunst Anna Oppermanns gemacht, deren Datenformat natürlich XML lautet, deren Bildbestand aus Dateien im JPEG-Format besteht, das zwar wenigstens ein offener Standard ist, von dem wir aber noch nicht wissen, wie lange er hält.<sup>24</sup>

Was auf den allerersten Blick wie ein Versprechen der modernen Digitaltechnik aussieht – sehr viel speichern zu können – entpuppt sich am Ende als ein Anlass zu umfänglichster Regelung und Verwaltung. Einen knappen Einblick in den Verhauf einer Zertifizierung digitaler Repositorien gewährt etwa die Checkliste<sup>25</sup> der Research Libraries Group<sup>26</sup>, die unter ihren vier Kategorien zwar auch eine technische hat, aber sehr viel mehr Augenmerk der Organisation und der Finanzierung, den Geschäftsprozessen und der Benutzbarkeit der Dokumentbewahrung widmet.

### *Das Genom als Archiv*

Gestatten Sie mir einen kurzen Exkurs in die Human-genetik, denn wenn es um große digitale Dokumente geht, muss einem auch das menschliche Genom in den Sinn kommen.

Das Genom ist zweifellos der Träger digital kodierter Information. Das Alphabet des Kodes besteht aus diskreten Zeichen, die mit den Basen Adenin, Cytosin, Guanin und Thymin identifizierbar sind. Die Länge der Zeichenkette entspricht etwa 5.000 Buch-Bänden, ei-

ner recht stattlichen Privatbibliothek, die wir in jeder unserer Zellen mit uns her herumtragen.<sup>27</sup>

Doch natürlich stimmt die Text- und damit auch die Buch-Metapher nicht. Denn das meiste, was in der Zeichenkette steht, ist Schrott, Datenmüll. Nur ca. 3% aller Zeichen kodieren die Gene, unsere Erbanlagen, auf die es eigentlich ankommt. Beim Umkopieren der Daten auf jeweils frische Datenträger, unseren Kindern, – einem Vorgang, der, wie wir wissen, mindestens zum Erhalt einer digitalen Informationssammlung erforderlich ist – bei diesem Umkopieren verändert sich das Genom, und es schleichen sich Fehler ein, die evolutionären Fortschritt, aber auch Krankheit bedeuten können. Das Umkopieren geschieht auch bei jeder Zell-Neubildung in einem Organismus, wobei es einen dem *data rot* entsprechenden Prozess gibt: mit zunehmendem Alter und unter ungünstigen äußeren Bedingungen erhöht sich die Zahl der Kopierfehler, und was dabei entstehen kann, heißt bei Mensch und Tier: Krebs. Will man den Vergleich mit dem Internet wagen, dann wären das Zellen mit einer TTL, einer TimeToLive, von unendlich: eben eine bösartige Wucherung.

Ein weiterer Grund, warum die Schrift-Metapher in die Irre führt, liegt darin, dass die Zeichen und ihre Anordnung zwar durchaus für den Menschen lesbar sind, nämlich für die Ribosomen, die Erzeuger der Proteine, für die das Genom die Bauanleitung ist, aber durchaus nicht für das menschliche Bewusstsein, das doch ansonsten für Textinterpretation zuständig ist. Die vier Basen und ihre Kombinationen bilden keine Symbolschrift, wie wir sie aus Texten gewohnt sind. Sie wirken nur durch das Leben und Sterben selbst, nicht über Symbol-Interpretation. Denn, was vor allem dazu fehlt, das sind die Metadaten: nirgends steht geschrieben, welche Bedeutung einzelne Abschnitte der DANN haben und wie sie zu interpretieren sind, kein XML-tag markiert, wo genau die Augenfarbe, der Intelligenzquotient, die Länge der Wimpern beschrieben stehen. Und wahrscheinlich ginge schon die Frage nach dem Ort solcher Inschrift in die Irre. Wir stehen im Moment, der ja als der gefeiert wird, zu dem wir 99% der DNA nachbuchstabieren können, wir stehen jetzt vor der Situation, dass wir einen Dokumenttext ohne

24 Terstegge, Christian; Warnke, Martin; Wedemeyer, Carmen: PETAL-: a proposal an XML Standard for the Visual Arts, Vito Cappellini, James Hemsley und Gerd Stanke. – Florenz : Pitagora Editrice Bologna, 2002. In: *Tagung EVA*. 2002, Florence. - ISBN 88-371-1312-9. – URL: [http://kulturinformatik.uni-lueneburg.de/warnke/Petal\\_EVA\\_2002\\_Florence.doc.pdf](http://kulturinformatik.uni-lueneburg.de/warnke/Petal_EVA_2002_Florence.doc.pdf).

Warnke, Martin: Daten und Metadaten. In: *Zeitenblicke*, 2.2003,1. - ISSN 1619-0459. - <http://www.zeitenblicke.historicum.net/2003/01/warnke/index.html>.

Schnede, Uwe; Warnke; Martin (Hrsg.): *Anna Oppermann in der Hamburger Kunsthalle*. - Hamburg : Hamburger Kunsthalle 2004. – ISBN 3-922 909-91-3. - Mit einer DVD von Martin Warnke, Carmen Wedemeyer und Christian Terstegge.

25 <http://www.rlg.org/en/pdfs/rlgnara-repositorieschecklist.pdf>

26 <http://www.rlg.org/>

27 Ridley, M.: *Alphabet des Lebens*. München : Claasen 2000.

Dokumentation haben, der unter unbekannter Software auf einer nur schlecht bekannten Hardware läuft, um einmal eine andere, sicher auch sehr schlechte Metapher zu wählen.

Jedenfalls sind wir weit davon entfernt, den Bitstrom der Daten etwa verstehen zu können, und bei der Größe und Komplexität dieses Problems bin ich auch eher verzagt, dass jemals erwarten zu können. Mich erinnert diese Situation eher an den Daten-GAU bei der NASA: ein Haufen von Daten, das meiste davon Schrott, alles unbeschriftet, entzifferbar nur per *trial and error*.

Die DNA ist ein digitales Archiv des Lebens, eines, das in je neuen Versionen die Evolution dokumentiert, das aber nicht von Menschen lesbar zu sein scheint, ein Geheimarchiv größter Bedeutung, aber ohne Zutritt für uns Sterbliche, was dafür aber seine Integrität noch ein kleines Weilchen sichern kann.

wie wir wissen, *memory*, das Gedächtnis eines Künstlichen Gehirns.

Dabei ist der Neurophysiologie noch gar nicht klar, wie das Gedächtnis eines Lebewesens mit Zentralnervensystem funktioniert. Es hat etwas mit Hirnmaterie zu tun, wie die Folgen von Gehirnverletzungen zeigen, aber niemand kann unter dem Mikroskop irgendwelche Speicherplätze zeigen, an denen Gedächtnisinhalte zu lokalisieren wären. Gerhard Roth schreibt: «Das Gedächtnis ist ... unser wichtigstes <Sinnesorgan>. Es ist zugleich aber ... nur ein Glied im Kreisprozess von Wahrnehmung, Gedächtnis, Aufmerksamkeit, Erkennen, Handeln und Bewerten.»<sup>28</sup> Hier ist er einig mit dem erkenntnistheoretischen Konstruktivismus eines Heinz von Foerster, aber auch mit dem späten Wittgenstein, der polemisiert: «Ein Ereignis lässt eine Spur im Gedächtnis, das denkt man sich manchmal. ... Der Organismus mit einer Diktaphonrolle verglichen; der Ein-



Foto © HWWA – Olaf Hoppe

### *Gedächtnis vs. Speicher*

Ich möchte mit einigen Überlegungen zum Verhältnis von Speicher und Gedächtnis so langsam zum Ende kommen.

Die Geschichte der Computertechnik ist die Geschichte von schrägen Metaphern und Anthropomorphismen: Charles Babbage nannte Bestandteile seiner gebauten und geplanten Maschinen in Anlehnung an die Landwirtschaft noch *mill* und *store*, aber aus dem an ein Getreidesilo *store* wurde im öffentlichen Sprachgebrauch,

druck, die Spur, ist die Veränderung, die die Stimme auf der Rolle zurücklässt. Kann man sagen, das Diktaphon (oder die Rolle) erinnere sich wieder des Gesprochenen, wenn es das Aufgenommene wiedergibt?»<sup>29</sup>

Gedächtnis ist mithin nicht zu isolieren, schon gar nicht technischkonstruktiv. Gedächtnis ist weniger ein Ding als vielmehr ein Prozess, der sich als lebendiger vollzieht und sich dabei möglicherweise auf irgendwelche systemischen Zustandswechsel stützt, die im Vollzug des Erinnerns eine kodifizierende Rolle spielen könnten.

28 Roth, G.: *Das Gehirn und seine Wirklichkeit*. Frankfurt/M. : Suhrkamp 1996. S. 241

29 Wittgenstein, Ludwig: *Bemerkungen über die Philosophie der Psychologie*. Bd. 7, I. S. 220

Alan Turing hat ja seine Maschine als abstraktes Modell von Computern auch mit Zuständen und Symbolen bestückt, um die Arbeit eines rechnenden Menschen zu simulieren. Dabei spielen die Symbole noch die Rolle von «Gedächtnisstütze[n]»<sup>30</sup> Die eigentliche Anthropomorphisierung erfolgte erst später in der Künstliche-Intelligenz-Forschung, deren Credo die *symbol systems hypothesis* ist: alles, was die Welt ausmacht, sei durch Symbole kodifizierbar, die nach festen Regeln manipulierbar seien, wodurch das menschliche Denken inklusive Gedächtnis nachzubilden wäre. Hier gibt es nun keinen Unterschied mehr zwischen Speicher und Gedächtnis.

Mit der tatsächlichen Situation, in der auch und gerade digital kodierte Daten dem Verfall anheim gegeben sind, kann diese KI-Metapher nicht umgehen. Weltwissen und Fakten, ihre Pendants zu Erfahrung und Gedächtnis, können zwar eventuell durch spätere Ableitungs-Prozesse obsolet und widerlegt werden, aber ein einfaches Verschwinden aufgrund von *data rot* ist für dieses Wissen und solche Fakten ebenso wenig vorgesehen wie die Einbettung des Gewussten in Handlung, Wahrnehmung und Bewertung.

#### Das Paradox der digitalen Archive

Die Einsicht in die Zeitlichkeit digitaler Daten, die Notwendigkeit, digitale Archive mit hohem Aufwand über Jahrzehnte hinweg zu retten, schlägt mit Gerhard Roth, den Konstruktivisten oder dem späten Wittgenstein in dieselbe Kerbe: ein Speicher mag über lange Zeit hinweg intakt bleiben, ohne zu vergehen; Gedächtnis und Erinnerung bedürfen aber genau so wie interpretierbare Daten ständiger tätiger Erneuerung.

Archive, digitale zumal, überdauern nur, wenn sie ständig benutzt werden, wenn eine erhaltende Instanz

sie stets neu kodifiziert, interpretiert und bewertet, sich ihre Dokumente handelnd aneignet, sie herausgibt oder verheimlicht, damit Wissen ermöglicht und strukturiert, Handlungen provoziert oder zu unterdrücken trachtet. Nur so überstehen digitale Archive die Jahrzehnte.

Sie leben so lange, wie eine Macht sie trägt und ihren informationellen Stoffwechsel aufrecht erhält. Danach werden sie bestenfalls Mausoleen, in deren Innerem man nichts Brauchbares mehr finden wird, deren Deckel besser geschlossen bleiben, weil ihr Inhalt ohnehin Moores Gesetz oder dem zweiten Hauptsatz der Thermodynamik zum Opfer gefallen sein wird: der Entropie, der negativen Information, zu Datenstaub werdend, an das Vergessen vergessen.

Entgegen einer geläufigen Auffassung von Immaterialität des Digitalen taugen also informatische Archivierungsverfahren nicht so ohne Weiteres für die Ewigkeit: Medien verrotten, Festplatten scheinen zu «sterben», Chip- und ganze Rechnergenerationen lösen sich ab, Formate verschwinden so schnell, wie sie aufgetaucht sind.

Es gibt zwar Strategien, trotz *data rot* und technischer Veralterung ein Mindestmaß an Dauerhaftigkeit zu gewährleisten, aber eines funktioniert nicht mehr: das Liegenlassen und Wegschließen von Archivalien ist unter Bedingungen der Digitalität kein Schutz vor Abnutzung mehr, sondern ihr schlichtes Todesurteil.

Das Paradox der digitalen Archive, das da lautet: «Bei analogen Archivalien bleicht jeder Blick die Schätze, digitale Archivalien wollen Aufmerksamkeit um jeden Preis» oder, etwas salopper: «Rührmichnichtan trifft Betriebsnudel», dieses Paradox ist nicht mehr aus der Welt zu schaffen.

30 Turing, A.M.: On computable numbers, with an application to the Entscheidungsproblem. In: *Proceedings of the London Mathematical Society*. 2. Ser. 42.1937. – Dt. Übers. Dotzler und Kittler. 1987: 17-60, S. 21).

#### II E- Adressen

##### *Anschrift der Abteilung:*

Staatsbibliothek zu Berlin – Preußischer Kulturbesitz, Abt. II E, 10772 Berlin

##### *Sekretariat:*

Tel. (030) 266-2337, Fax (030) 266-2378, E-mail: [abt.ueberreg.dienste@sbb.spk-berlin.de](mailto:abt.ueberreg.dienste@sbb.spk-berlin.de)

##### *Telefonnummern:* (030) 266-

<i>Abt.-Leiter:</i>	Dr. Hartmut Walravens	- 2498
<i>Stellvertr.:</i>	Ulrike Junger	- 2478
<i>ZDB:</i>	Barbara Sigrist	- 2483
<i>ZRT:</i>	Barbara Pagel	- 2345
<i>GKD:</i>	Elisabeth Pitz	- 2343
<i>ZRZ:</i>	Bernd Augustin	- 2314
<i>ISBN, ISMN:</i>	Dr. Hartmut Walravens	- 2498

##### *Web sites:*

URL der Abteilung: [http://ueberregionale\\_dienste.staatsbibliothek-berlin.de](http://ueberregionale_dienste.staatsbibliothek-berlin.de)  
 ZDB: <http://www.zeitschriftendatenbank.de>  
 Sigelstelle: <http://www.sigel.staatsbibliothek-berlin.de>  
 ISBN: <http://isbn-international.org>  
 ISMN: <http://ismn-international.org>