

Veröffentlichungen der Archivschule Marburg

Institut für Archivwissenschaft

Nr. 51

Veröffentlichungen der Archivschule Marburg

Institut für Archivwissenschaft

Nr. 51

Katrin Wenzel, Jan Jäckel (Hrsg.)

Retrokonversion, Austauschformate und Archivgutdigitalisierung

Beiträge zum Kolloquium aus Anlass des 60-jährigen Bestehens
der Archivschule Marburg

zugleich 14. Archivwissenschaftliches Kolloquium der Archivschule
Marburg am 1. und 2. Dezember 2009

Bibliografische Information der Deutschen Nationalbibliothek:
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der
Deutschen Nationalbibliografie; detaillierte bibliografische Daten
sind im Internet unter <http://dnb.ddb.de> abrufbar.

ISBN 978-3-923833-38-2

Gedruckt auf alterungsbeständigem Papier ∞

©Archivschule Marburg 2010

Bismarckstraße 32, D 35037 Marburg

Konzeption der Umschlaggestaltung: Designbüro Glas, Seeheim-
Jugenheim

Druck: Druckerei Schröder, Schuppertsgasse 2, 35083 Wetter

Inhalt

Vorwort

von Katrin Wenzel und Jan Jäckel 11

Retrokonversion

Aufgaben und Erfahrungen der DFG-geförderten Koordinierungsstelle Retrokonversion an der Archivschule Marburg

von Frank M. Bischoff 15

Evaluation der Koordinierungsstelle Retrokonversion

von Katrin Wenzel und Jan Jäckel..... 49

Komplexe Findbücher? Erfahrungen mit der Retro- konversion durch Dienstleister und im Eigenbetrieb

von Nils Brübach 63

Zum Stellenwert der Retrokonversion im Rahmen der Katastrophenbewältigung – das Beispiel des Stadtarchivs Köln

von Ulrich Fischer 79

Retrokonversion von bibliothekarisch erschlossenen Handschriftenkatalogen im Archiv am Beispiel des Bestandes J 1 im Hauptstaatsarchiv Stuttgart

von Jeannette Godau 109

Von der Eindimensionalität zum Netzwerk: Die Retrokonversion des Systematischen Katalogs im Deutschen Literaturarchiv von Karin Schmidgall	143
--	-----

Austauschformate

Archival Exchange Formats: Developments and Future Issues by Michael Fox.....	163
---	-----

Exchanging Archival Descriptions: The Role of EAD and EAC in the Construction of the SAN Portal in Italy by Stefano Vitali	183
--	-----

The DTD of the EAG <i>Censo-Guía de los Archivos de España e Iberoamérica</i>: An Electronic Guide of the Spanish and Ibero-American Archives by Blanca Desantes.	199
--	-----

Austauschformate bei der Retrokonversion und ihr Nutzen für die deutschen Archive von Sigrid Schieber.....	221
--	-----

Datenaustausch in der Praxis: Das archivsparten- übergreifende Portal „Archive in NRW“ nach dem Relaunch (2007) von Peter Worm	237
--	-----

Archivgutdigitalisierung

„Europeana“ – die europäische digitale Bibliothek und der Aufbau einer „Deutschen Digitalen Bibliothek“. Stand und Perspektiven von Gerald Maier	251
The Digitization of Archives in France: Projects and Perspectives by Claire Sibille-de Grimoüard.....	275
Digitalisierung und Onlinestellung von Archivgut im Bundesarchiv: Ziele, Verfahren und Werkzeuge von Angelika Menne-Haritz	291
Realisierung der technischen Integration von Digitalisaten unter der Verwendung des Austauschstandards METS von Paul Bantzer	319
Archivgutdigitalisierung im Rahmen der Bestandserhaltung. Die Praxis im Landesarchiv Nordrhein-Westfalen von Johannes Kistenich.....	339
Das virtuelle Urkundenarchiv „Monasterium“ – bayerische Urkunden im internationalen Kontext von Joachim Kemper	361
Autorenverzeichnis	379

Vorwort

von Katrin Wenzel und Jan Jäckel

Informationen bereitzustellen und Benutzung zu fördern ist eine der wichtigsten Aufgaben von Archiven und ihnen verwandten Gedächtnisorganisationen. Diesem Anliegen war das anlässlich des 60-jährigen Bestehens der Archivschule Marburg abgehaltene 14. Archivwissenschaftliche Kolloquium zum Thema „Retrokonversion, Austauschformate und Archivgutdigitalisierung“ gewidmet, das am 1. und 2. Dezember 2009 stattfand.¹ Das Thema bot sich auch deshalb an, da seit Sommer 2007 an der Archivschule die Koordinierungsstelle Retrokonversion etabliert ist, die von der Deutschen Forschungsgemeinschaft mit dem Auftrag zur Koordinierung und Durchführung der Digitalisierung von Findmitteln des deutschen Archivwesens gefördert wird. Ziel und Leitmotiv ist es dabei, den Zugang zu den Archivalien für die Forschung im Zeitalter des Internets zu erleichtern.

Nach nun zweijähriger Tätigkeit der Koordinierungsstelle wurden im Rahmen der Tagung die besonderen Herausforderungen archivischer Retrokonversionsprojekte beleuchtet, die Thematik aber auch durch Einbeziehung ähnlicher Programme in anderen Ländern oder in anderen Sparten befruchtet. Zu berücksichtigen waren neben den

¹ Vgl. den Tagungsbericht: Katrin Wenzel/Jan Jäckel, Retrokonversion, Austauschformate und Archivgutdigitalisierung – 14. Archivwissenschaftliches Kolloquium der Archivschule Marburg, in: Der Archivar 63, 2010, S. 187-188; auch abgedruckt in: Archivnachrichten aus Hessen 10, 2010, Heft 1, S. 11-13 und Forum – Newsletter der Archivschule Marburg 33, 2009, S. 7-13.

spezifischen Erfahrungen der Archive auch jene von Dienstleistern, denen im Rahmen der Tagung sowohl auf dem Podium als auch in Posterausstellungen Möglichkeit zur Darstellung ihrer Leistungen und Wünsche gegeben wurde.

Weil die Retrokonversion von Findmitteln in öffentlich zugänglichen Onlineangeboten münden muss, war in besonderem Maße über die Formate nachzudenken, in die retrokonvertiert werden soll. Die Kommunikationsfähigkeit von archivischen Informationen in elektronischer Form wird in Zukunft maßgeblich von der Einhaltung geeigneter Standards abhängig sein. Hier spielt die internationale Standardisierung der Codierung von Erschließungsinformationen eine ebenso zentrale Rolle wie die deutschen Erschließungstraditionen und Erfahrungen mit Austauschformaten. Die internationalen und nationalen Beiträge zu dieser archivwissenschaftlichen Kernthematik zeigten deshalb Perspektiven auf, wie die Sichtbarkeit für unterschiedlichste Zwecke – Austausch und Präsentation – gewährleistet werden kann.

Während noch die flächendeckende Digitalisierung der Findmittel des deutschen Archivwesens operativ umgesetzt wird, haben bereits einige Archive die Digitalisierung analog vorliegenden Archivguts in Angriff genommen. Wenn auch kein Zweifel besteht, dass die elektronischen Findmittel Voraussetzung für die digitalisierten Archivbestände darstellen, muss dieser nächste Schritt doch bereits jetzt in den Blick genommen und zukunftsweisend geplant werden. Auch die Digitalisierung des Archivguts selbst wurde daher in verschiedenen Beiträgen beleuchtet. Dabei wurden Erfahrungen aus dem Ausland

sowie Aktivitäten im Rahmen internationaler Programme gemeinsam mit Projekten deutscher Archive vorgestellt.

All diese Tagungsbeiträge werden in bewährter Tradition im vorliegenden Sammelband veröffentlicht.² Nicht mit in den Band aufgenommen wurden allerdings die Statements der Dienstleister. Andererseits konnten den Aufsätzen eine thematisch passende Transferarbeit der Archivschule Marburg sowie ein Artikel der beiden Herausgeber zur Evaluation der Koordinierungsstelle Retrokonversion hinzugefügt werden.

Zahlreiche Kollegen haben am Entstehen des Bandes mitgewirkt – Ihnen gilt dafür ein herzlicher Dank: Werner Engel, Volker Hirsch, Alexandra Lutz, Karsten Uhde und Wolfgang Krauth. Besonderen Dank sind wir Waltraud Noll für die Hilfe bei der Erstellung der Druckvorlage schuldig. Viel Dank gilt nicht zuletzt auch der DFG für einen finanziellen Zuschuss zur Drucklegung. Belohnung für die Unterstützung der zuletzt Genannten, für die Autoren und nicht zuletzt für die Herausgeber wäre es, wenn dieser Tagungsband viele Leser finden und so das große Vorhaben der Retrokonversion weiter befruchten und fördern könnte.

Marburg, im August 2010

Katrin Wenzel und Jan Jäckel

² Die Powerpoint-Präsentationen der Referenten sind einsehbar unter <http://www.archivschule.de/forschung/retrokonversion-252/veranstaltungen/vortraege-kolloquium-09.html> [23.8.2010].

Aufgaben und Erfahrungen der DFG-geförderten Koordinierungsstelle Retrokonversion an der Archivschule Marburg

von Frank M. Bischoff

1. Zielsetzung

Eine der grundlegenden Aufgaben von Archiven ist es, die in ihren Magazinen aufbewahrten historischen Dokumente der Öffentlichkeit zugänglich zu machen. Mit den deutschen Archivgesetzen ist der Zugang zum Archivgut *de facto* zu einem Recht für Jedermann geworden.¹ Die Realisierung dieses Zugangsrechts bedingt, dass nicht nur das Archivgut selbst den Benutzern zugänglich ist, sondern auch der Weg dorthin kundenfreundlich aufgezeigt wird. Erschließung von Archivalien und die Bereitstellung von Findmitteln für Benutzer sind Voraussetzungen, die die Benutzung von Archivgut erst ermöglichen. Vor wenigen Jahren ist anlässlich des 11. Archivwissenschaftlichen Kolloquiums der Archivschule Marburg zum Thema „Aktuelle Anforderungen an archivische Erschließung und Findmittel“ bereits ausgeführt worden, dass nicht erschlossenes Archivgut brach liegendes

¹ So hat etwa die gerade vollzogene Reform des nordrhein-westfälischen Archivgesetzes in § 6 Abs 1 geregelt: „Jeder hat nach Maßgabe dieses Gesetzes und der hierzu ergangenen Benutzungsordnung das Recht, Archivgut auf Antrag zu nutzen, soweit aufgrund anderer Rechtsvorschriften nichts anderes bestimmt wird.“ Die erste Gesetzesfassung von 1989 enthielt in § 7 Abs. 1 noch die – in der Praxis allerdings kaum relevante – Einschränkung: „Archivgut kann nach Ablauf der Sperrfristen (...) nutzen, wer ein berechtigtes Interesse an der Nutzung glaubhaft macht.“

historisches Erbe ist, das Kosten verursacht, ohne Nutzen und Erkenntnis stiften zu können.² Angesichts der Entwicklung des Onlinemarkts für historisch-kulturelle Informationen in den vergangenen Jahren darf man inzwischen weitergehen und postulieren, dass Archive, die mit ihren Erschließungsinformationen nicht in einer fachlich angemessenen und zugleich benutzerfreundlichen Form online recherchierbar sind, für eine große Zahl potentieller Kunden verborgen bleiben. In einer weitgehend auf die Wissensermittlung im World Wide Web fixierten Gesellschaft wird Archivgut, über das keine Informationen im Internet verfügbar sind, keine Beachtung finden, ungenutzt bleiben und damit nutzlos sein für die historische Erkenntnis.

In einer 2008 an der Archivschule verfassten Transferarbeit zur Frage der Integration von Beratungsangeboten in Online-Findmitteln stellte ein baden-württembergischer Referendar auf der Basis einer Auswertung des Internet-Nutzerverhaltens fest: „Internetbesucher von archivischen Homepages fahnden nach jenen exklusiven Informationen, die nur von den Archiven bereitgehalten werden. Ihre Interessenlage entspricht damit sehr genau jener des klassischen Archivbenutzers. Sie suchen Inhalte und kein kulturelles Entertainment.“³ Ähnlich wie die Bibliotheken in Deutschland schon vor Jah-

² Frank M. Bischoff, Vorwort, in: Frank M. Bischoff (Hrsg.), Benutzerfreundlich – rationell – standardisiert. Aktuelle Anforderung an archivische Erschließung und Findmittel. Beiträge zum 11. Archivwissenschaftlichen Kolloquium der Archivschule Marburg, Marburg 2007 (Veröffentlichungen der Archivschule Marburg 46), S. 11.

³ Christoph Volkmar, Service für den virtuellen Nutzer. Vorschläge zur Integration von Beratung in Online-Findmitteln, Marburg 2008, S. 16, <<http://www>.

ren ihre Kataloge online zugänglich gemacht haben, sollten auch Archive ihre Findmittel – unter Beachtung schutzrechtlicher Belange – ins Netz stellen. Es handelt sich hier nicht allein um den gerechtfertigten Anspruch der historischen Forschung auf zeitgemäße Vermittlung von Informationen durch jene Institutionen, die auf diesem Gebiet als Kompetenzzentren zu gelten haben. Vielmehr besteht hier zugleich eine ideelle Verpflichtung von Archiven im Rahmen ihres Charakters als Häuser der Geschichte, Gedächtnis der Gesellschaft und vor dem Hintergrund ihrer Mitverantwortung für die historisch-politische Bildung.⁴

Mit der seit der zweiten Hälfte der 1990er Jahren von vielen deutschen Archivarinnen und Archivaren grundsätzlich akzeptierten Zielsetzung, die Findmittel der Archive im großen Stil im Internet zugänglich zu machen, waren zwar die Weichen – mental – in die richtige Richtung gestellt, die praktischen Probleme aber noch nicht gelöst. Archivische Findmittel sind Generationenwerk und können in

landesarchiv-bw.de/sixcms/media.php/25/Transferarbeit_Volkmar.pdf. > [18.08.2008]. – Die Einschätzung von Volkmar wird im Übrigen durch Website-Ratingsysteme, wie <<http://www.seitwert.de>> oder <<http://bizinformation.org>>, gestützt, die Portalen mit umfangreichen Findmitteln, etwa <www.hadis.hessen.de>, <<http://archivportal.niedersachsen.de>> oder <www.archive.nrw.de>, einen vergleichsweise guten Rang innerhalb der Domäne .de bescheinigen.

⁴ Vgl. hier und im Folgenden auch Frank M. Bischoff/Udo Schäfer, Das Angebot der Archive in der digitalen Welt. Von der Retrokonversion zum Deutschlandportal, in: Rainer Hering/Jürgen Sarnowsky/Christoph Schäfer/Udo Schäfer (Hrsg.), Forschung in der digitalen Welt. Sicherung, Erschließung und Aufbereitung von Wissensbeständen. Tagung des Staatsarchivs Hamburg und des Zentrums „Geisteswissenschaften in der digitalen Welt“ an der Universität Hamburg am 10. und 11. April 2006, Hamburg 2006 (Veröffentlichungen aus dem Staatsarchiv der Freien und Hansestadt Hamburg 20), S. 169-182.

absehbarer Zeit nicht neu erstellt werden. Viele Archive sind zur Recherche in ihren Beständen auf Findmittel angewiesen, die weit über hundert Jahre alt sind. Diese Findbücher können aus Kapazitätsgründen nicht ohne weiteres durch neue ersetzt werden, abgesehen davon, dass manche darunter eine durchaus vorzügliche Qualität aufweisen. Eine Neuerschließung der Archivalien in den heute üblichen elektronischen Erschließungssystemen der Archive würde eine rasche Onlinestellung behindern. Selbst eine Nacherschließung auch nur von Teilen der Archivbestände oder eine inhaltliche Nachbesserung der vorhandenen Findbücher würde Jahre verschlingen, bevor onlinefähige Findmittel ins Netz eingestellt werden könnten.

Um die Masse archivischer Findmittel im Internet verfügbar zu machen, muss ihre Digitalisierung folglich in der vorliegenden Form erfolgen. Bereits die deutschen Bibliotheken hatten seit den 1980er Jahren auf die Strategie der Retrokonversion gesetzt. In Anlehnung an die bibliothekarische Terminologie wird im Folgenden unter Retrokonversion die Umsetzung älterer Findmittel in ein digitales Format verstanden. Die Übertragung erfolgt dabei im Wesentlichen unverändert, ist also nicht von nennenswerten Nach- oder Neuverzeichnungsarbeiten begleitet.⁵

⁵ Als Beispiel für die bibliothekarischen Bemühungen um die Retrokonversion sei auf folgenden Bericht verwiesen: Dietmar Buschey/Axel Halle/Reinhard Harms (Hrsg.), Zwanzig Jahre Retrokonversion an der SUB Göttingen: Ein Zwischenbericht, in: Bibliotheksdienst 35, 2001, S. 1089-1102. Deutlich wird u.a. die große Bedeutung der Nachnutzung konvertierter Titelaufnahmen für andere Bibliotheken, mit der im Bibliothekswesen Synergieeffekte erzielt werden können, die im Archivwesen undenkbar sind, da Archive im Wesentlichen Unikate verwahren, deren Erschließungsinformationen nicht für andere Archivalien nachgenutzt werden

2. Vorgeschichte der DFG-Förderlinie

In deutschen Archiven wurden nach der Jahrtausendwende unterschiedliche Ansätze zur Retrokonversion verfolgt. Das Hauptstaatsarchiv Düsseldorf, heute die Abteilung Rheinland des Landesarchivs Nordrhein-Westfalen, beantragte bereits im Jahr 2000 bei der DFG Mittel für ein Projekt zur „Entwicklung von Werkzeugen zur Retrokonversion archivischer Findmittel“.⁶ Das im Rahmen dieses Projekts entwickelte Tool ermöglichte, aufgrund von inhaltlichen und Layout-Merkmalen die in maschinengeschriebenen oder gedruckten Findmitteln enthaltenen Informationen zu identifizieren, zu zerlegen und strukturiert in einem XML-Dokument abzulegen.⁷ Allerdings mussten die Identifikationsparameter für die diversen Merkmale im Detail angepasst werden. Wegen der Unterschiedlichkeit der Findbücher

können. Der aktuelle Wikipedia-Bertrag zur Retrokonversion ist leider zu bibliotheksbezogen und im Hinblick auf die Archive sogar falsch, wenn dort in generalisierender Weise über die Vorzüge der Online-Konversion ausgeführt wird: „Da zahlreiche Bibliotheken, bzw. Archive Mitglied desselben Verbunds sind, können später hinzukommende Bibliotheken, Archive bereits vorhandene Datensätze mitverwenden (sogenannten [sic!] "Ansigelung" an bestehende Titelaufnahmen).“ Vgl. <<http://de.wikipedia.org/wiki/Retrokonversion>> [24.05.2010].

⁶ Vgl. Frank M. Bischoff, Das Projekt Retrokonversion von Findbüchern, in: Mechthild Black-Veldtrup/Ottfried Dascher/Axel Koppetsch (Hrsg.), Archive vor der Globalisierung ? Beiträge zum Symposium des Nordrhein-Westfälischen Hauptstaatsarchivs in Verbindung mit den Allgemeinen Reichsarchiven in Brüssel (Belgien) und Den Haag (Niederlande) vom 11. bis 13. September 2000 in Düsseldorf, Düsseldorf 2001 (Veröffentlichungen der staatlichen Archive des Landes Nordrhein-Westfalen Reihe E Nr.7), S. 139-155.

⁷ Vgl. den Abschlussbericht zum DFG-Projekt „Entwicklung von Werkzeugen zur Retrokonversion archivischer Findmittel“ vom 30.06.2004 unter <http://www.archive.nrw.de/LandesarchivNRW/BilderKartenLogosDateien/Abschlussbericht_Retrokonversionsprojekt.pdf> [24.05.2010].

war der Anpassungsaufwand hoch, da die Parameter nahezu für jedes Findbuch individuell definiert werden mussten.

Schon in den ausgehenden 1990er Jahren wurde in baden-württembergischen Staatsarchiven mit der Retrokonversion begonnen und zwar zunächst vor allem mit Inhouse-Lösungen, also eigenem Stammpersonal, ABM-Kräften und Praktikanten, die die Findbücher abschrieben.⁸ Das Bundesarchiv hat seit 2004 Erfahrungen in der Findmittelretrokonversion im Rahmen von Fremdaufträgen.⁹ Die Aufträge wurden zum Teil in Rumänien durch händische Erfassung in einem strukturierten Format abgewickelt. Die hessischen Staatsarchive schließlich haben mit eigenem Personal unterstützt durch Projekt-Hilfskräfte ihre maschinengeschriebenen Findmittel in den letzten zehn Jahren gescannt, mit Mustererkennung (OCR) bearbeitet, die elektronischen Texte dann händisch strukturiert und in ihre Datenbank HADIS eingespeist. In Wiesbaden und in Darmstadt sind die Arbeiten zum Abschluss gelangt.¹⁰

⁸ Vgl. Peter Müller, Vom handschriftlichen Repertorium zur Datei – Retrokonversion von Findmitteln im Staatsarchiv Wertheim, in: Archivnachrichten 23, 2001, S. 11. Monika Dirsch/Elke Koch, Retrokonversion archivischer Findbücher im Staatsarchiv Ludwigsburg, in: Archivnachrichten 25, 2002, S. 8f. hoben in Ihrem Beitrag bereits hervor, dass für jedes Findbuch eine Erfassungsanweisung erstellt werden musste.

⁹ Vgl. Petra Rauschenbach, Start des neuen Rahmenvertrags zur Retrokonversion von Findmitteln – Anlass für eine Zwischenbilanz, in: Mitteilungen aus dem Bundesarchiv 16, 2008, Heft 1, S. 102.

¹⁰ Nicola Wurthmann, Pilotphase zum DFG-Projekt „Retrokonversion archivischer Findmittel“ abgeschlossen. Recherchedaten zu 16 Stadtarchivbeständen des

Um von solchen punktuellen Ansätzen weg zu einer bundesweiten Förderlinie bei der Deutschen Forschungsgemeinschaft zu gelangen, wie sie in den 1980er Jahren für die Altbestände der Bibliotheken aufgelegt wurde,¹¹ mussten zunächst das Mengengerüst ermittelt und Umsetzungsszenarien entwickelt werden. Das Landesarchiv Nordrhein-Westfalen führte 2005 mit Unterstützung der Deutschen Forschungsgemeinschaft eine Vorstudie zur Retrokonversion archivischer Findmittel durch.¹² Der Abschlussbericht vom Juni 2006 geht von mehr als 60 Millionen zu digitalisierenden Titelaufnahmen in den öffentlichen Archiven Deutschlands aus.¹³ Die Vorstudie schlug eine Anschubfinanzierung durch die DFG vor, bei der ein Antragsteller die Hälfte der DFG-Fördersumme auch als Eigenanteil beisteuern sollte. Außerdem empfahl die Vorstudie die Einrichtung einer Geschäftsstelle, um die deutschen Archive zu beraten, die Retrokonversions-

Staatsarchivs Marburg online verfügbar, in: Archivnachrichten aus Hessen 8, 2008, Heft 2, S. 12f.

¹¹ Vgl. Altbestandserfassung in wissenschaftlichen Bibliotheken der Bundesrepublik Deutschland einschließlich Berlin (West): Positionspapier des Bibliotheksausschusses der Deutschen Forschungsgemeinschaft, in: Zeitschrift für Bibliothekswesen und Bibliographie 35, 1988, S. 51-60.

¹² Wilfried Reininghaus, Die Retrokonversion von Findmitteln in Deutschland – ein DFG-Vorprojekt, in: Der Archivar 58, 2005, S. 201.

¹³ Der Projektbericht von 2006 findet sich unter <http://www.archive.nrw.de/LandesarchivNRW/BilderKartenLogosDateien/RetrokonversionProjektbericht2006.pdf> [24.05.2010]. Vgl. auch Ulrich Fischer/Wilfried Reininghaus, DFG-Vorstudie "Retrokonversion archivischer Findmittel". Die wichtigsten Ergebnisse des Projekts, in: Der Archivar 4, 2006, S. 330.

arbeiten zu verstetigen und die Abwicklung der Retrokonversionsaufträge durch Dienstleister zu begleiten.¹⁴

Die Notwendigkeit einer breit angelegten Fördermaßnahme wurde von den Fachgutachtergremien der DFG befürwortet¹⁵ und mündete im Mai 2006 in ein DFG-Positionspapier zum Bereich „Wissenschaftliche Literaturversorgung und Informationssysteme“. Darin sind die Schwerpunkte der Förderung bis 2015 festgelegt. Unter den 17 hier genannten Einzelmaßnahmen findet sich die Digitalisierung der Findmittel des deutschen Archivwesens als eigenständiges Förderprogramm.¹⁶

Im Gefolge der Vorstudie schlugen Vertreter der DFG-Gutachtergremien und des deutschen Archivwesens – unter Letzteren Vertreter der staatlichen Archivverwaltungen, der Bundeskonferenz der Kommunalarchive und des Verbands deutscher Archivarinnen und Archivare – vor, dass die Archivschule Marburg als quasi neutrale Instanz ohne eigene Findmittel, aber ausgestattet mit ent-

¹⁴ Zur Frage der Einbindung von Dienstleistern vgl. auch Ulrich Fischer, Vergabe, Förderung, Eigenregie. Strategien zur Konversion handschriftlicher Findmittel in deutschen Archiven, in: Alexandra Lutz (Hrsg.), Neue Konzepte für die archivistische Praxis. Ausgewählte Transferarbeiten des 37. und 38. wissenschaftlichen Kurses an der Archivschule Marburg (Veröffentlichungen der Archivschule Marburg 44), S. 295-326.

¹⁵ Vgl. die bereits im Jahr 2004 formulierten Forderungen einer Gruppe von DFG-Gutachtern: Thekla Kluttig/Robert Kretzschmar/Karl-Ernst Lupprian/Wilfried Reininghaus/Udo Schäfer/Barbara Schneider-Kempf/Günther Wartenberg, Die deutschen Archive in der Informationsgesellschaft – Standortbestimmung und Perspektiven, in: Der Archivar 57, 2004, S. 28-36, bes. S. 30ff.

¹⁶ Vgl. <<http://www.dfg.de/download/pdf/foerderung/programme/lis/positionspapier.pdf>>, bes. S. 5 u. 8, [24.05.2010].

sprechender Fachkompetenz und einem satzungsgemäßen Forschungsauftrag, eine Koordinierungsfunktion übernehmen soll. Seitens der Archivschule wurde dieser Vorschlag aufgegriffen und in einem Förderantrag konkretisiert. Seit September 2007 ist die mit DFG-Förderung eingerichtete Koordinierungsstelle Retrokonversion an der Archivschule Marburg aktiv und bündelt die bundesweiten Retrokonversionsmaßnahmen.¹⁷

Dass die Deutsche Forschungsgemeinschaft ein nationales Förderprogramm zur Digitalisierung der Findmittel aufgelegt hat, darf durchaus als bemerkenswert eingestuft werden. Nach Kenntnisstand des Verfassers ist lediglich in Großbritannien mit dem *Access to Archives*-Portal (A2) ein ähnliches, dort aus Lotteriemitteln gefördertes Programm durchgeführt worden. A2A wurde 2008 abgeschlossen und hat immerhin die stolze Summe von rund 10 Mio. digitalisierten Verzeichnungseinheiten aus über 400 britischen Archiven vorzuweisen.¹⁸

3. Grundlage der Retrokonversion: Volltextauflösung, Austauschformate und Archivportal

Bevor die DFG-Förderlinie starten konnte, war eine Reihe von grundsätzlichen Fragen zu klären. Es musste Einvernehmen darüber beste-

¹⁷ Vgl. dazu die Mitteilung von Frank M. Bischoff/Sigrid Schieber, DFG-Förderprogramm zur Retrokonversion von Findbüchern, in: *Der Archivar* 61, 2008, S. 36-38.

¹⁸ Nach Aussage der National Archives handelt es sich dabei um rund 30 Prozent der Findmittel zu Archivbeständen in England und Wales; vgl. <<http://www.nationalarchives.gov.uk/a2a/faqs.aspx>> [eingesehen am 24.05.2010].

hen, welche Qualität die retrokonvertierten Findmittel aufweisen sollten, ob ein bestimmtes Zielformat anvisiert werden sollte und wie die Bereitstellung der Findmittel im Internet sichergestellt werden konnte. Man darf nicht verkennen, dass die letztgenannte Frage für die Deutsche Forschungsgemeinschaft entscheidend war, weil sie die Retrokonversion der Findmittel der deutschen Archive letztlich nur als Infrastrukturmaßnahme fördert, damit die Forschung davon profitieren kann.

3.1 Imaging vs. Volltextauflösung

Wie das Resultat einer Retrokonversion von analogen Findmitteln aussieht und welche Qualität der Daten erreicht werden soll, wurde bereits in früheren Projekten erarbeitet. Eine reine Image-Konversion, d.h. die Bereitstellung der Findmittel in einem Bildformat, galt im Archivwesen im Gegensatz zum Bibliothekswesen nie als zielführend. Zu den Ergebnissen des unter Federführung des Hauptstaatsarchivs Düsseldorf durchgeführten Projekts zur Erstellung eines Retrokonversions-Werkzeugs gehörte die Feststellung, dass die notwendige Recherchierbarkeit in einem solchen Verfahren nicht gegeben sei, trotzdem aber eine Reihe von Metadaten zusätzlich erfasst werden müssten, um die Images zu verwalten und zugänglich zu machen. In dem Abschlussbericht an die DFG wurde deshalb geurteilt: „Bewertet man die vorgestellten Konversionsmethoden und die dabei jeweils zu erwartenden Ergebnisse, so ist aus archivfachlicher Sicht das Imagingverfahren eindeutig nicht zu empfehlen, da der damit zu erreichende Erschließungs- und Informationsgrad deutlich hinter dem der anderen Verfahren zurücksteht. Dies gilt selbst wenn die Bildfolge in einer Klassifikation dargeboten und durch Metadaten

oder gar eine partielle Verschlagwortung angereichert wird.“¹⁹ Auf die Notwendigkeit einer Stichwortsuche und einer navigierenden Recherche für die angemessene Nutzung und Auswertung von Online-Findmitteln wird nochmals in der DFG-Vorstudie vom Juni 2006 hingewiesen.²⁰

Die vordergründig betrachtete kostengünstigere Variante des Imagingverfahrens ohne Texterkennung (OCR) kann die damit verknüpften Nachteile nicht aufwiegen. Die Retrokonversion mit Volltexterkennung lässt demgegenüber Volltextrecherchen bzw. verschiedene Varianten komplexer Suchmöglichkeiten zu, ermöglicht die flexible Darstellung in der jeweiligen Anwendung gemeinsam mit originär digitalen Findbüchern und kann den Import in beliebige Archivsoftware und die Weiterbearbeitung der Daten sicherstellen. Somit war von vornherein der Weg eines Imaging-Verfahrens, das ohnehin nur ein Zwischenprodukt geliefert hätte, als unwirtschaftlich ausgeschlossen. Als qualitative Mindestanforderung ist die Volltextauflösung gefordert.

Praktisch kann das Ziel auf zwei Wegen erreicht werden, nämlich mit Mustererkennung (OCR) und mit händischer Erfassung. Das OCR-Verfahren setzt maschinengeschriebene Findmittel voraus. Eine

¹⁹ Abschlussbericht zum DFG-Projekt „Entwicklung von Werkzeugen zur Retrokonversion archivischer Findmittel“ vom 30.06.2004, Anlage 1: Projektbericht „Retrokonversion handschriftlicher Findmittel“, S. 61, <http://www.archive.nrw.de/LandesarchivNRW/BilderKartenLogosDateien/Projektbericht_Retrokonversion_hs_Findmittel.pdf> [24.05.2010].

²⁰ DFG-Vorstudie „Retrokonversion archivischer Findmittel“. Projektbericht, wie Anm. 13, Kap. 7.3, S. 73ff.

Texterkennungsoftware mit oder ohne Layouterkennung „liest“ den Text aus den Bildscans ein. Danach erfolgen die Zuordnung der Elemente der Verzeichnungseinheiten und eine Fehlerkorrektur. Das manuelle Verfahren kann sowohl bei maschinengeschriebenen wie bei handgeschriebenen Findmitteln angewendet werden. Die Findmittel werden händisch erfasst, entweder durch eine Person mit entsprechend aufwendiger nachfolgender Fehlerkorrektur oder durch ein *double keying* Verfahren. Beim *double keying* werden die Daten zweimal durch zwei verschiedene Personen erfasst und die abgeschriebenen Texte anschließend automatisch verglichen. Beide Verfahren gelangen zum Einsatz.²¹ Vorgegeben ist lediglich eine Mindestgüte der Ergebnisse, nämlich 100 Prozent Fehlerfreiheit bei den Klassifikationsstufen und den Signaturen und 99,98 Prozent bei den sonstigen Texten. Lediglich im Falle handgeschriebener Findmittel

²¹ Das in unterschiedlichen Graden automatisierbare OCR-Verfahren scheint gegenüber der manuellen Erfassung auch bei maschinengeschriebenen Findbüchern bei geforderter Qualität nicht zwingend kostengünstiger zu sein. Insbesondere das manuelle double-keying-Verfahren liefert sehr gute Qualitäten, die aber auch mit einem hohen Personaleinsatz erkaufte werden müssen. Im schweizerischen Bundesarchiv, wo bereits in der ersten Hälfte der 1990er Jahre mit der Retrokonversion begonnen wurde, bestand bis 2003 aufgrund von Preis-Leistungs-Analysen eine deutliche Präferenz für eine hausinterne, händische Retrokonversion; vgl. Bärbel Förster, Retrokonversion archivischer Findmittel im Schweizerischen Bundesarchiv. Beitrag zur Tagung des Nordrhein-Westfälischen Hauptstaatsarchivs zum DFG-Projekt "Entwicklung von Werkzeugen zur Retrokonversion archivischer Findmittel", Düsseldorf, 22.-23.09.2003, unter <<http://www.archive.nrw.de/dok/tagung-retro/08-Foerster.pdf>> [24.05.2010]. Nach den Erfahrungen der Koordinierungsstelle Retrokonversion macht es jedoch auch unter Berücksichtigung des Preis-Leistungs-Verhältnisses keinen Sinn, OCR-basierte Verfahren bei der Retrokonversion maschinengeschriebener Findmitteln auszuschließen.

liegt die Fehlertoleranz höher; hier müssen nur 98 Prozent Fehlerfreiheit erreicht werden.

3.2 Austauschformate

Die Digitalisierung von Findmitteln in einer heterogenen Archivlandschaft setzt voraus, dass die in ein digitales Format konvertierten Findbücher mit relativ geringem Aufwand in verschiedene Zielsysteme portiert werden können. Will man die Kommunikation von Erschließungsinformationen zwischen unterschiedlichen Systemen sicherstellen, benötigt man dazu einen gemeinsamen Sprachcode, ein standardisiertes Austauschformat für Erschließungsinformationen.

Das Zielformat sollte verschiedene Eigenschaften aufweisen. Zu den formalen Anforderungen an ein Austauschformat zählt, dass es anerkannt, standardisiert, verbreitet, mit einfachen Mitteln verarbeitbar und nicht proprietär ist. Die strukturellen Anforderungen verlangen, dass die Findbuchstrukturen und die Kontextinformationen zu den einzelnen Findbucheinträgen verlustfrei abgebildet werden können. Die inhaltlichen Anforderungen schließlich entsprechen den inhaltlichen Anforderungen an herkömmliche, analoge Findmittel und bedingen, dass alle gebräuchlichen Elemente in Findmitteln adäquat dargestellt werden können.

Im Rahmen des Düsseldorfer Retrokonversionsprojekts von 2001-2004 wurde das sog. Standardaustauschformat (Saft-XML²²)

²² Vgl. Ziel und Austauschformat des Retrokonversionstools (AF-Findmittel_Version1.0.dtd – „Findbuch-DTD“) – Tag Library. Teilprojekt 1 des DFG-Projekts zur Entwicklung von Werkzeugen zur Retrokonversion archivischer Findmittel,

entwickelt, das sich an der deutschen Erschließungstradition inhaltlich und terminologisch orientiert und einen verlustarmen Austausch von Erschließungsinformationen zwischen verschiedenen Systemen gewährleisten soll. Die ISO-Norm *Encoded Archival Description* (EAD) findet inzwischen weltweit Beachtung, ist etwas weniger eng der deutschen Erschließungstradition verhaftet, eignet sich aber gleichermaßen als Format für den Transfer von Erschließungsinformationen.

Beide Formate sind als Zielformate für die Retrokonversion zugelassen, im Falle von EAD sogar lediglich eine Untermenge der EAD-Tag-Library. Zugleich ist jede Abweichung davon ausgeschlossen.²³ Diese strikte Vorgabe für die Teilnahme am DFG-Förderprogramm beinhaltet zugleich eine Chance für das deutsche Archivwesen. Hier wird ein hoffentlich nachhaltiger Standard für den Austausch und die Kommunikation von Erschließungsinformationen definiert, eine in der Praxis angewendete Norm geschaffen. Wenn Informationen vernetzt werden sollen, wird man von solchen Standards profitieren, Standards, die über das einzelne Archiv oder eine Archivsparte hinaus die technische Verständlichkeit der Erschließungsinformationen über-

unter <http://www.archive.nrw.de/LandesarchivNRW/BilderKartenLogosDateien/SAFT_TagLibrary.doc> [24.05.2010]. Die DTD findet sich veröffentlicht unter <<http://www.archive.nrw.de/LandesarchivNRW/BilderKartenLogosDateien/SAFT-XML.zip>> [24.05.2010].

²³ Auf Nachfrage für EAD kann hier angesichts der Vielzahl von einschlägigen Veröffentlichungen des Bundesarchivs verzichtet werden; vgl. etwa <<http://www.bundesarchiv.de/daofind/ead/>> [24.05.2010]. Entsprechende Informationen und Links finden sich im Onlineangebot der Archivschule: <<http://www.archivschule.de/forschung/retrokonversion-252/vorstudien-und-saft-xml/>> [24.05.2010].

haupt erst gewährleisten.²⁴ Es ist nicht ausgeschlossen, dass sich nach Abschluss der Retrokonversion der Findbücher deutscher Archive die Einführung der Standards für den Findmittelaustausch im Rückblick als wichtigste Weichenstellung darstellen wird.

3.3 *Archivportal*

Als Zielsysteme, in die die retrokonvertierten Verzeichnungseinheiten eingespeist werden, müssen die je eigenen Anwendungen der Archive angesehen werden. Das ist allein schon deshalb notwendig, weil viele archivrechtlichen Normen der Länder Findmittel unter einen ähnlich hohen Schutz stellen, wie das Archivgut selbst.

Dort, wo keine daten- bzw. persönlichkeitschutzrechtlichen Belange betroffen sind, dürfen und wollen die Archive Findmittel aber öffentlich zugänglich machen, auch in Fremddatenbanken. Die Deutsche Forschungsgemeinschaft hat als Bedingung für eine Retrokonversionsförderung den teilnehmenden Archiven auferlegt, ihre Findmittel nicht nur in eigenen Datenbanken, auf eigenen Homepages oder in regionalen Archivportalen zugänglich zu machen, sondern auch über zentrale Portale, die eine Bündelungsfunktion für das deutsche Archivwesen erfüllen. Diese Verpflichtung erstreckt sich

²⁴ Dass auch Standards wie EAD Interpretationsspielräume bieten und es ermöglichen, gleichartige Informationen in unterschiedlicher Weise zu taggen, sei an dieser Stelle nicht verhehlt. Es hat in diesem Zusammenhang auch Absprachen zwischen Archivschule, Bundesarchiv und anderen Archivverwaltungen gegeben, wie bestimmte Informationen zu taggen sind; vgl. etwa die Ergebnisse eines Arbeitstreffens am 7.11.2007 in Berlin unter <<http://www.archivschule.de/forschung/retrokonversion-252/newsretro/arbeitstreffen-zur-abstimmung-eines-gemeinsamen-ead-austauschformats.html?PHPSESSID=57b45ef866120f73949809ef0c0ab1d9>> [24.05.2010].

nicht nur auf die Findmittel, die mit DFG-Unterstützung retrokonvertiert wurden. Auch die in Eigenleistung retrokonvertierten oder bereits in digitalen Formaten vorliegenden Findmittel müssen in solchen institutionenübergreifenden Portalen zur Verfügung gestellt werden.

Hintergrund dieser Forderung der Deutschen Forschungsgemeinschaft ist wieder die Orientierung an den Interessen der Forschung. Man geht bei der DFG von der Philosophie eines One-Stop-Shop aus, in der ein zentrales Informationsportal die Wünsche der Forschung nach einem Maximum an Information am ehesten wird befriedigen können. Dezentrale und heterogene Informationen werden an einem Ort fachlich angemessen zusammengeführt, dem Forscher für beliebig komplexe Recherchen zur Auswertung zur Verfügung gestellt, um diesen von dort aus dann wieder zielgerichtet an die eigentliche Provenienzadresse zu verweisen.

Vor diesem Hintergrund hatte die Konferenz der Archivreferenten des Bundes und der Länder (ARK) an das Bundesarchiv das Anliegen herangetragen, ein zentrales Archivportal D mit Unterstützung der DFG zu entwickeln, das als Verbundfindmittel die digitalen Findbücher der deutschen Archive aufnehmen soll. Leider scheint dieses Portal ungeachtet der von der DFG dafür bewilligten Mittel nicht oder zumindest nicht zeitnah errichtet zu werden, so dass diese Lücke durch die bereits bestehenden und seit Jahren erfolgreich

arbeitenden regionalen Archivportale,²⁵ durch das BAM-Portal²⁶ oder durch die in der Entwicklung begriffene Deutschen Digitale Bibliothek²⁷ gefüllt werden muss.

4. Aufgaben, Zeitabläufe und Erfahrungen

Um die DFG-geförderte Retrokonversion rationell abwickeln zu können, wurde die Antragstellung in Phasen eingeteilt. Nach der Bewilligung der Einrichtung einer Koordinierungsstelle Retrokonversion im Sommer 2007 startete die Förderung mit einer „produktiven Pilotphase“, in der die Verfahrensabläufe für die weiteren Antragsphasen im Echtbetrieb erarbeitet wurden. An der Pilotphase beteiligten sich vier Archivverwaltungen, das Bundesarchiv, das Landesarchiv Baden-Württemberg, die Hessischen Staatsarchive und das Staatsarchiv Sachsen.²⁸

Nach dieser ersten Pilotphase begannen die regulären Antragsphasen, in der Regel zwei pro Jahr mit Antragsfristen für die beteiligten Archive jeweils im Mai/Juni und im November/Dezember eines

²⁵ Z.B. das nordrhein-westfälische Portal unter <http://www.archive.nrw.de> oder die hessische Recherchedatenbank Hadis unter <http://www.hadis.hessen.de/>. Für weitere Nachweise vgl. <http://www.archivschule.de/service/archive-im-internet/archive-in-deutschland/archivportale/> [24.05.2010].

²⁶ Vgl. <http://www.bam-portal.de/> [24.05.2010].

²⁷ Vgl. <http://www.deutsche-digitale-bibliothek.de/> [24.05.2010].

²⁸ Vgl. neben den in den Anm. 9 und 10 zitierten Beiträgen auch Jörg Ludwig, Retrokonversion von Findmitteln, in: Sächsisches Archivblatt, 2008, Heft 1, S. 10f. und Wolfgang Zimmermann, Neue Findmittel im Internet. Deutsche Forschungsgemeinschaft fördert das Landesarchiv Baden-Württemberg, in: Archivnachrichten 38, 2009, S. 34.

Jahres. In 2008 konnte wegen zu knapper Fristen nur eine reguläre Antragsphase aufgenommen werden. Im Dezember 2009 liefen die Arbeiten für die Antragsphase 1/2010, für die die Teilnahmeanmeldung der Archive bis zum 1. November 2009 und die Einreichung der fertigen Anträge bis zum 15. Dezember 2009 bei der Koordinierungsstelle an der Archivschule erfolgen musste. Die Bescheide der DFG für diese Antragsphase – Bewilligungen oder auch Ablehnungen – ergehen dann voraussichtlich im Juni/Juli 2010.

4.1 Aufgaben der Koordinierungsstelle

Bei den Aufgaben der Koordinierungsstelle sind übergreifende, die unabhängig von einzelnen Antragsphasen wahrgenommen werden, von den Aufgaben zu unterscheiden, die sich auf konkrete Antragsphasen beziehen. Zu den übergreifenden Aufgaben zählen

- Öffentlichkeitsarbeit,
- Berichterstattung gegenüber der DFG und diversen archivischen Gremien (z.B. im Beirat der Archivschule, in der Archivreferentenkonferenz des Bundes und der Länder und in der Bundeskonferenz der Kommunalarchive),
- die Entwicklung von Antragsvorlagen, Musterberichten usw.,
- die kontinuierliche Pflege der Standards für die Austauschformate

- sowie die Verwaltung der Koordinierungsstelle, hier vor allem die Verwaltung der großen Schriftgutmengen und der Schriftgutablage.²⁹

Der Aufwand für diese Aufgaben war in der Anfangszeit erheblich und band teilweise mehr als eine halbe Mitarbeiterstelle. Im weiteren Projektverlauf reduzierte sich die zeitliche Belastung, beträgt aber in der Regel immer noch mehr als ein Viertel der gesamten Arbeitszeit.

Die Aufgaben, die sich jeweils auf konkrete Antragsphasen beziehen, lassen sich in fünf Gruppen einteilen. In der Vorbereitung einer Antragsphase fallen folgende Arbeiten an:

1. Beratung und Antragstellung

- Beratung der Archive bei der Antragstellung,
- ggf. die Prüfung der Unterlagen und der ebenfalls mit einzureichenden Findbuch-Faksimiles durch die Koordinierungsstelle zur Einholung von unverbindlichen Kostenvoranschlägen,
- ggf. die Einholung der unverbindlichen Kostenvoranschläge bei Dienstleistern durch die Koordinierungsstelle,

²⁹ Im Zwischenbericht vom September 2008 wurde bereits darauf hingewiesen, dass allein von der Projektbearbeiterin im höheren Dienst im Laufe des ersten Jahres rund 1.100 schriftliche Ausgänge (in der Regel per E-Mail) erstellt wurden. Davon bezog sich der größte Teil auf die Abwicklung der Pilotprojekte, gefolgt von den Mails zur Vorbereitung der nächsten Antragsphase inklusive Einholung der Kostenvoranschläge und zur allgemeinen Beratung von interessierten Archiven. Hinzu kommen noch telefonische Auskünfte, deren Menge allerdings im Verhältnis deutlich geringer ausfiel. Bei 217 Arbeitstagen ergibt das durchschnittlich fünf ausgehende Schreiben je Arbeitstag.

- die Fertigstellung der Anträge durch die Archive und ihre Übersendung an die Koordinierungsstelle.
- 2. die Bündelung bzw. Erstellung eines Rahmenberichts durch die Koordinierungsstelle für die Gutachter und der Versand der Anträge mit Rahmenbericht an die DFG

Nach ausgesprochener Bewilligung von Fördermitteln erfolgt die Abwicklung der Antragsphasen.

3. Die Durchführung der Retrokonversion durch die Dienstleister muss begleitet werden und erfordert immer wieder Rückkopplungen zu den Archiven.
4. Die Kontrolle der Ergebnisse (Erstkontrolle) muss vorgenommen werden.
5. Der Datenimport in die Zielsysteme der Archive sowie die Prüfung auf Konformität der Austauschformate bzw. der Schnittstellen der Zielsysteme mit den Vorgaben muss durchgeführt werden. Hier besteht bei vielen Archiven Bedarf nach Beratung und konkreter Unterstützung bei der Umsetzung.

Als arbeitsaufwendigste Einzelmaßnahme darf die Kontrolle der Ergebnisse gelten, die in der Pilotphase deutlich mehr als ein Viertel der Arbeitszeit der Mitarbeiter beansprucht hat. Im Zeitaufwand an zweiter Stelle steht die Beratungsleistung für die Antragsphasen, die immerhin 18 Prozent der Arbeitszeit ausmacht, die sich in den kommenden Phasen aber aufgrund inzwischen geleisteter Vorarbeiten und vorhandener Erfahrungen etwas reduzieren wird.

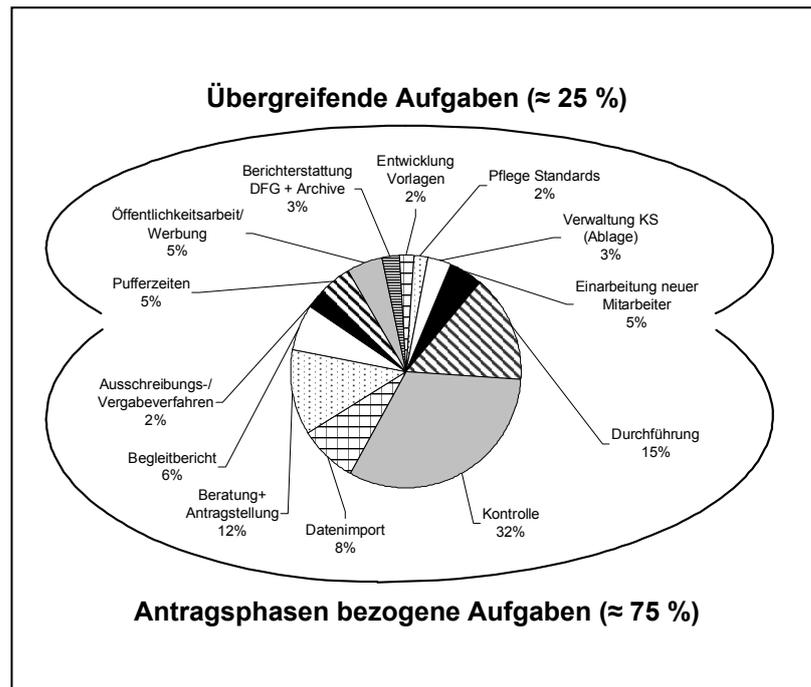


Abb. 1: Arbeitsaufwände in verschiedenen Arbeitsbereichen der Koordinierungsstelle Retrokonversion: Planung für 2009–2011 auf der Basis der Erfahrungswerte 2007-2009

Gegenüber den Anfangsphasen wurden Aufwände bei der Koordinierungsstelle bereits dadurch heruntergeschraubt, dass die Erfassungsanweisungen, die für jedes Findbuch einzeln erstellt werden müssen, nicht mehr von der Koordinierungsstelle, sondern von den Dienstleistern oder den Archiven zu erstellen sind. Die Erfassungsanweisungen geben vor, welche Elemente ein Findbuch enthält, wie diese Elemente konvertiert werden müssen und welche Besonderhei-

ten bei einem Findbuch zu beachten sind. Es handelt sich also um eine Art Leistungsbeschreibung der Retrokonversion für jedes Findbuch.

Für den Förderzeitraum von 2009-2011 dürften sich im Vergleich zum ersten Förderzeitraum in der Koordinierungsstelle Retrokonversion eine leichte Entlastung der übergreifenden Aufgaben und eine stärkere Gewichtung der auf Förderphasen bezogenen Aufgaben ergeben. Derzeit wird davon ausgegangen, dass die übergreifenden Aufgaben ca. 25 Prozent und die auf die Antragsphasen bezogenen Aufgaben etwa 75 Prozent des Personalaufwands ausmachen (vgl. Abb. 1).

4.2 Zeitliche Abläufe in den Antragsphasen

Die Dauer einer Antragsphase ist durchaus erheblich. Vorbereitung und Abwicklung einschließlich der für eine DFG-Begutachtung notwendigen Mindestfristen können bis zu 28 Monaten in Anspruch nehmen, sofern ein öffentliches Ausschreibungs- und Vergabeverfahren notwendig ist. Im Einzelnen wird heute mit folgenden Zeitabläufen kalkuliert:

- Die Vorbereitung von Antragsphasen nimmt bis zu 6 Monate in Anspruch.
- Für die Bearbeitung der Anträge bei der DFG, das Durchlaufen des Begutachtungsverfahrens bis hin zum Eingang des Bewilligungsschreibens und der Mittelbereitstellung bei den antragstellenden Archiven müssen erfahrungsgemäß mindestens vier Monate eingerechnet werden. Durch Rückfragen oder anderweitige, auch unvorhersehbare Verschiebungen kann sich dieser Zeitraum leicht ausdehnen.

- Für die Abwicklung der bewilligten Retrokonversionsprojekte muss man im Falle notwendiger öffentlicher Ausschreibungsverfahren ca. 18 Monate veranschlagen, wobei auch hier Unwägbarkeiten bestehen. Dort, wo bereits Rahmenverträge bestehen oder eine freihändige Vergabe erfolgen kann, wird sich dieser Zeitraum möglicherweise auf 12 Monate verkürzen. Allerdings kann er sich durch einen großen Projektumfang, unkalkulierbare Fristabweichungen bei Dienstleistern oder Auftraggebern erweitern. Im Einzelnen wird von folgendem Szenario ausgegangen:
 - Für die erfolgreiche Durchführung einer öffentlichen Ausschreibung bis zur Vergabe müssen 3-6 Monate veranschlagt werden.
 - Für die Erstellung der Erfassungsanweisungen, der Mappings, für die Durchführung der Retrokonversion durch den Dienstleister, für die Ergebniskontrolle, für evtl. notwendige Korrekturschritte des Dienstleisters, für die Abnahme und schließlich für die Onlinestellung muss mit ca. 12 Monaten gerechnet werden.
 - Die Fertigstellung des Eigenanteils wird in dieser Zeitplanung nicht berücksichtigt. Schöpft ein Antragsteller den maximal zulässigen Zeitrahmen aus, kann die Abwicklung insgesamt sogar 3 Jahre umfassen. Die Koordinierungsstelle hat auch hier den Auftrag, die Ergebnisse zu prüfen.

Die Übersicht zeigt, dass ab Sommer 2009 vier, ab Frühjahr 2010 sogar zeitweise fünf Antragsphasen parallel zu begleiten sind.

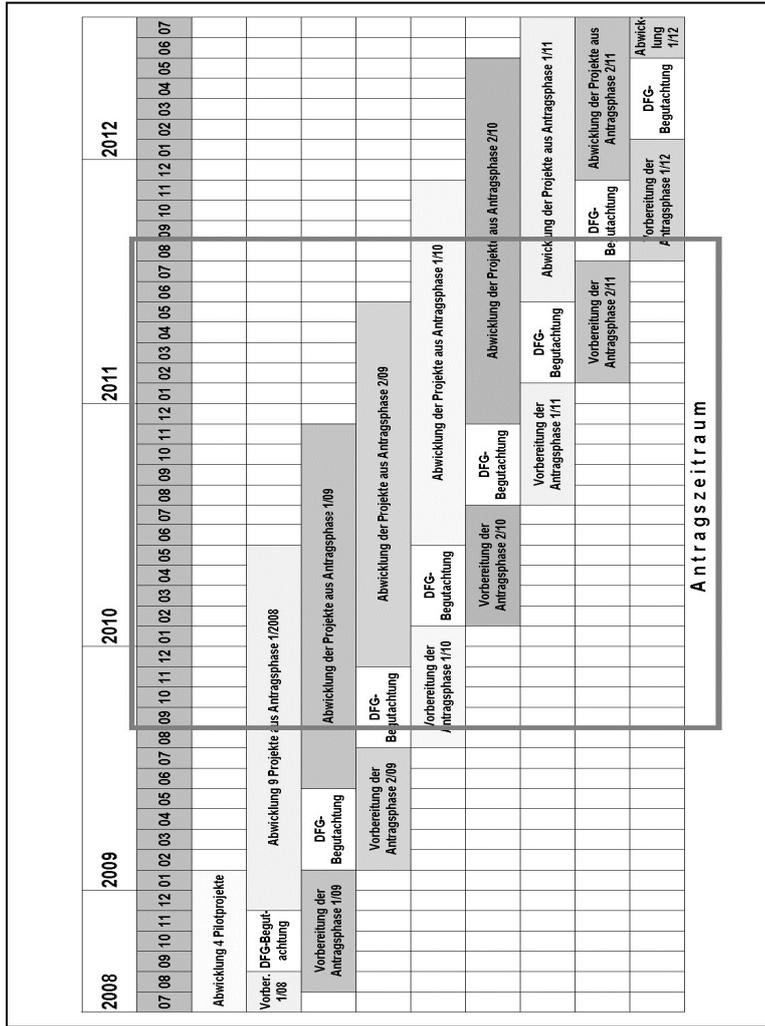


Abb. 2: Zeitliche Abläufe in den Antragsphasen (eingehämt: aktuelle DFG-Förderphase für die Koordinierungsstelle Retrokonversion)

4.3 Praktische Erfahrungen mit der Retrokonversion: Dienstleister vs. Eigenpersonal

Zu den grundlegenden Erfahrungen der Retrokonversion gehört die Feststellung, dass man in deutschen Archiven kaum zwei identisch strukturierte Findbücher findet. Fast jedes Findbuch weist eine Besonderheit auf. Folglich bedarf auch jedes Findbuch einer eigenen Erfassungsanweisung.

Neben den Erfassungsanweisungen binden die Endkontrollen erhebliche Arbeitskraft, zumindest in Fällen, in denen Aufträge an Dienstleister vergeben werden. Durch die Koordinierungsstelle wird eine Erstkontrolle der Ergebnisse auf der Basis einer Stichprobe vorgenommen. Die Endkontrolle muss aber vom Archiv durchgeführt werden und ist die Grundlage für die Abnahme und Vergütung der Leistung. Das dafür zur Verfügung stehende Zeitfenster ist relativ schmal, da die Dienstleister darauf angewiesen sind, sich in absehbarer Zeit refinanzieren zu können.

Die DFG stellt nicht allein Sachmittel zur Durchführung der Retrokonversion durch Dienstleister zur Verfügung, sondern gewährt ebenso Personalmittel, um die Digitalisierung der Findbücher von eigenen Mitarbeitern durchführen zu lassen. Obwohl die Retrokonversion mit eigenem Personal auch Qualitätssicherung erfordert, berichten die Archive in diesen Fällen von geringeren Belastungen. Demgegenüber haben die Erfahrungen der Abteilung Rheinland des Landesarchivs Nordrhein-Westfalen in der zweiten Hälfte des Jahres 2009 gelehrt, dass die Retrokonversion mittels Werkverträge mit Studenten, Historikern oder anderen, im Umgang mit Archivalien ausgewiesenen Personen auch erhebliche Arbeitskraft in den Archi-

ven bindet.³⁰ Die Einarbeitung oder Schulung dieser Personen und die Qualitätssicherung am Projektende – und ggf. zusätzlich noch zwischendurch – ist zeitraubend.

Im Verlauf der vergangenen drei Antragsphasen hat sich verschiedentlich gezeigt, dass im Vergleich zu den Kostenvoranschlägen von Dienstleistern einige Archive die Retrokonversion mit eigenem Projektpersonal zu gleichen Preisen oder günstiger durchführen können. Es handelt sich dabei nicht allein um komplex strukturierte Findmittel, bei denen das von vornherein zu erwarten war, sondern auch um Findmittel mit einfacher Struktur.

Eindeutig lassen sich die bisherigen Erfahrungen nur bezüglich der aus paläographischen oder strukturellen Gründen als schwierig einzustufenden Findmittel interpretieren. Hier scheint eine Retrokonversion im eigenen Haus und mit eigenem Personal in vielen Fällen wirtschaftlicher zu sein, weil nur auf diese Weise ein regelmäßiger und zeitnaher Kontakt mit dem fachlich versierten Personal des Archivs zur Klärung von Fragen sichergestellt werden kann.

³⁰ Die Abteilung Rheinland des Landesarchivs Nordrhein-Westfalen hat im Herbst 2009 über rund drei Monate hinweg im Rahmen von zehn Werkverträgen 58 Findbücher und Karteien retrokonvertieren lassen. Die Werkvertragskräfte mussten in die jeweiligen Findbücher eingeführt und im Umgang mit dem ebenfalls zur Verfügung gestellten Laptop und der darauf installierten Archivsoftware V.E.R.A. geschult werden. Da parallel dazu auch ein Auftrag zur Retrokonversion von 28 in obsoleten Datenformaten vorliegenden Findbüchern an ein Unternehmen erteilt wurde, konnte der Betreuungsaufwand unmittelbar verglichen werden. Gleichwohl ist zuzugestehen, dass die in obsoleten Datenformaten vorliegenden Findbücher aufgrund ihrer zumeist klaren Struktur auch einen deutlich niedrigeren Schwierigkeitsgrad aufwiesen, als der größere Teil der durch die Werkvertragskräfte konvertierten Findbücher.

Insbesondere in Fällen eines geringen oder mittleren Schwierigkeitsgrads von Findmitteln, ja selbst bei partiell anspruchsvollem Material, lässt sich ein Verzicht auf die Unterstützung durch Dienstleister derzeit allerdings kaum vorstellen. Für die Koordinierungsstelle Retrokonversion war die Zusammenarbeit mit Dienstleistern eine wertvolle Erfahrung, gerade auch bei der realistischen Einschätzung und Kalkulation von Aufwänden.

Ein Beispiel für diesen Erfahrungsgewinn stellt die Kategorisierung der Schwierigkeitsgrade von Findmitteln dar. Es war überraschend, dass die meisten Dienstleister sich auf weniger Schwierigkeitsgrade für die Kostenermittlung beschränkten, als die Koordinierungsstelle es ursprünglich vorsehen wollte. In der Praxis wird derzeit von sechs verschiedenen Schwierigkeitsgraden ausgegangen, unterschieden einerseits in (weitestgehend) maschinengeschriebene und (mindestens zum Teil) handgeschriebene Findmittel und andererseits nach Umfang der einzelnen Titelaufnahmen. Der Umfang der Titelaufnahmen wird dabei nur grob in drei Gruppen unterteilt, nämlich Titelaufnahmen bis 300 Zeichen, solche mit 300 bis 1000 Zeichen und die, die darüber hinausgehen. Nach den bisherigen Erfahrungen mit den Retrokonversionskosten scheinen diese sechs Schwierigkeitsgrade in einem relativ einfachen Kosten- und Aufwandsverhältnis von 1 : 2 : 3 : 3 : 6 : 9 zu stehen bzw. sich durch diese Faktoren näherungsweise berechnen zu lassen. Natürlich ist auch hier die Realität im Einzelfall komplexer, als Mittelwerte das erkennen lassen.

Qualität der Findmittel Anzahl Zeichen pro VZE	maschinengeschrieben	handschriftlich
< 300	1	3
> 300 und ≤ 1000	2	6
> 1000	3	9

Tab. 1: Kosten- und Aufwandsverhältnis unterschiedlicher Schwierigkeitsgrade von Findmitteln bei der Retrokonversion

5. Zwischenbilanz und Perspektiven

Eine Zwischenbilanz müsste an erster Stelle von dem Kompetenz- und Erfahrungsgewinn ausgehend hervorheben, dass nach fast dreijähriger Laufzeit des DFG-Förderprogramms heute sehr klare Kenntnisse über die notwendigen Arbeitsabläufe bei der Retrokonversion der Findmittel bestehen, die eine rationelle Abwicklung im großen Stil möglich machen. Die Koordinierungsstelle Retrokonversion verfügt heute über einen guten Überblick über den Dienstleistermarkt, über große Erfahrungen hinsichtlich potentieller Schwierigkeiten bei der Retrokonversion von archivischen Findmitteln, verfügt über umfangreiche Erfahrungen hinsichtlich von Angebotseinholungen und der Abwicklung öffentlicher Ausschreibungen, hat erhebliche Kompetenzen auf dem Gebiet der Austauschformate aufbauen und vermitteln können und hat mit den Abnahmeverfahren der vergangenen

Jahre inzwischen auch ein scharfes System der Qualitätskontrolle entwickelt.

Neben diesen qualitäts- und qualifikationsbezogenen Erträgen dürfen die eigentlichen Produkte, also die konvertierten Findmittel und Verzeichnungseinheiten in einer Bilanz nicht zu kurz kommen. Dies sei im Folgenden anhand eines knappen statistischen Überblicks dargelegt.

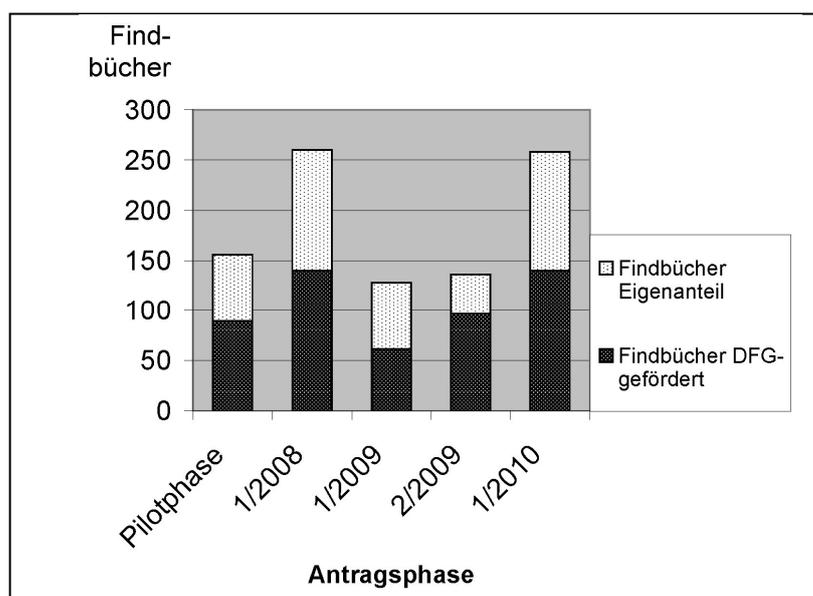


Abb. 3: Anzahl retrokonvertierter bzw. zu retrokonvertierender Findbücher der Pilotphase und der Förderphasen 1/2008-1/2010: DFG- und Eigenfinanzierung der Archive

Nimmt man den DFG-Förderanteil und den Eigenanteil der antragstellenden Archive zusammen, dann ist bis einschließlich der Antragsphase 1/2010 die Retrokonversion von 940 Findmitteln beantragt worden. Abgelehnt wurden bisher (Stand Dezember 2009) zwei Anträge, davon einer mit dem Hinweis, den Antrag erneut und unverändert vorzulegen, sobald das bemängelte Defizit behoben sei.

Den 940 Findbüchern entsprechen 1,87 Mio. Verzeichnungseinheiten. Das Antragsvolumen für die DFG-Förderung ohne Eigenanteil beziffert sich auf gut 1,6 Mio. Euro. Beteiligt haben sich bisher elf staatliche Archive, d. h. neben dem Bundesarchiv und dem Geheimen Staatsarchiv Preußischer Kulturbesitze neun Landesarchive, zehn Kommunalarchive und ein Forschungsinstitut.

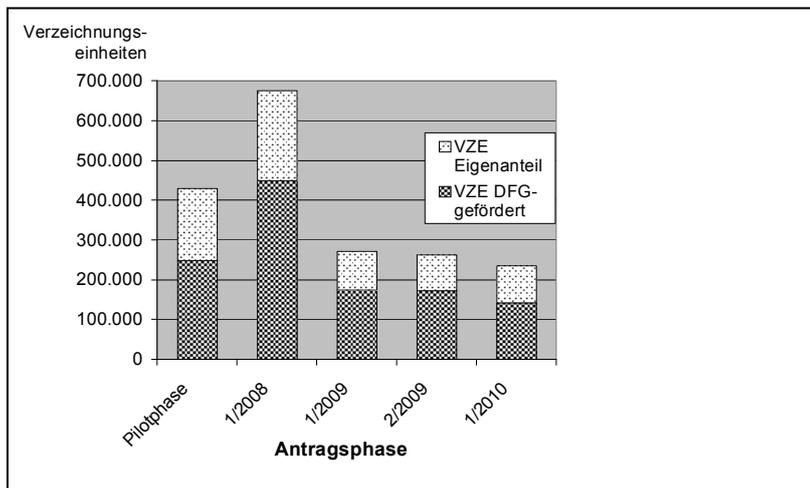


Abb. 4: Anzahl retrokonvertierter bzw. zu retrokonvertierender Verzeichnungseinheiten der Pilotphase und der Förderphasen 1/2008-1/2010: DFG- und Eigenfinanzierung

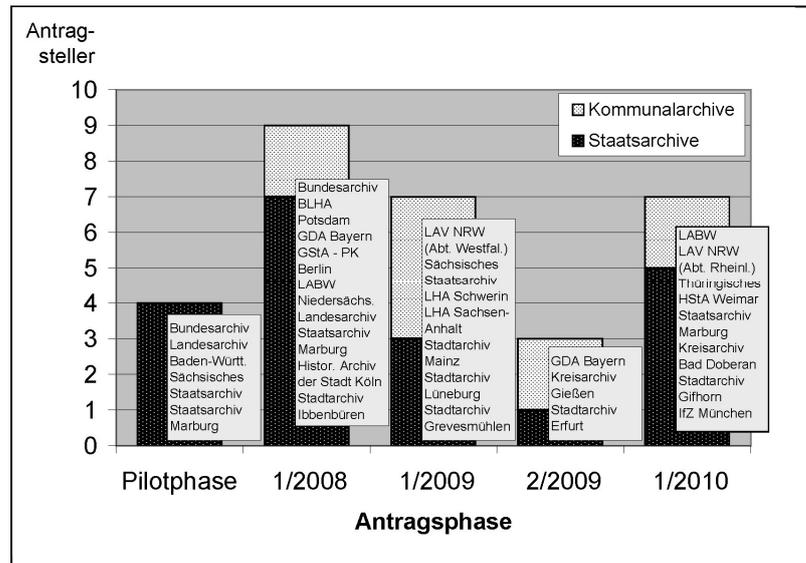


Abb. 5: Anzahl der Antragsteller in staatlicher und nicht-staatlicher Trägerschaft

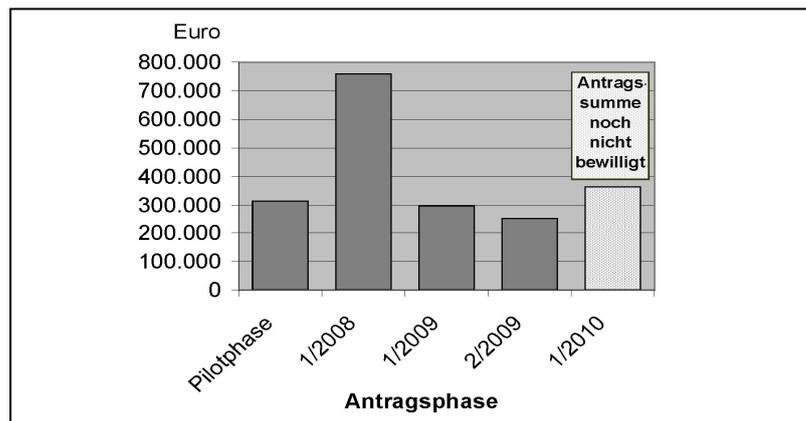


Abb. 6: Umfang der bewilligten bzw. beantragten Fördersumme

Einer von der Koordinierungsstelle Retrokonversion getätigten Umfrage zufolge wurden die Leistungen der Marburger Koordinierungsstelle von den beteiligten Archiven bisher im Wesentlichen positiv beurteilt. Beratungsleistungen, Formulare und Muster für die Antragstellung oder Berichterstattung und die Kommunikation zwischen der Koordinierungsstelle und den Dienstleistern waren als hilfreich eingeschätzt. Auch seitens der DFG kommen Hinweise, die auf eine wohlwollende Beurteilung der Leistungen der Koordinierungsstelle schließen lassen.

Wenn hier zum Abschluss ein knapper Blick auf die Perspektiven geworfen werden darf, sind dabei vor allem zwei Sachverhalte in den Blick zu nehmen.

1. Die DFG-Förderlinie Retrokonversion läuft im Prinzip bis zum Jahr 2015, de facto aber nur bis 2012, da die DFG grundsätzlich von einer bis zu drei Jahren dauernden Laufzeit der Projekte ausgeht. Im Jahr 2012 kann aber noch nicht ein Ergebnis erreicht sein, das man sich als angemessene Anschubfinanzierung erhoffen würde. Es ist bereits jetzt absehbar, dass der Förderzeitraum um wenigstens drei Jahr erweitert werden sollte.
2. Das Archivwissenschaftliche Kolloquium von 2009 hat sich auch mit der Digitalisierung von Archivgut befassen. Hier sind die Archive gefordert, rasch die Weichen in Richtung einer Digitalisierung von analogem Archivgut stellen, sowohl aus konservatorischen wie auch aus benutzer- und forschungsorientierten Gründen. Digitalisiertes Archivgut benötigt aber zwingend digitalisierte Erschließungsinformationen, damit Benutzer das Gesuchte auch

finden können. Auch aus diesem Grund sollte die Retrokonversion der Findmittel des deutschen Archivwesens verlängert werden.

Aus diesen Zielen lässt sich bereits herleiten, dass nach Auffassung des Verfassers einem bundesweiten Programm zur Retrokonversion und Onlinestellung von Findmitteln ein weiteres Programm zur Digitalisierung und Onlinestellung von analogem Archivgut folgen sollte. Es gibt in Deutschland bereits einige Archive die mit diesen Arbeiten jenseits klassischer Zimelienschauen begonnen haben und sich inzwischen auf gutem Weg befinden. Mit fachlicher Beratung, finanzieller Förderung und der Bereitstellung geeigneter Präsentationssysteme könnte dieser Prozess in der Fläche nachhaltig stimuliert werden.

Evaluation der Koordinierungsstelle Retrokonversion¹

von Katrin Wenzel und Jan Jäckel

1. Einleitung

Kurz vor Ablauf der ersten zweijährigen Projektphase der Koordinierungsstelle Retrokonversion (September 2007 – September 2009) erfolgte eine Evaluation der Geschäftsabläufe und der erzielten Resultate. Die Evaluation fußte nicht allein auf einer internen Analyse der Prozesse und evtl. aufgetretener Probleme, sondern auch auf einer Befragung der antragstellenden Archive. Dabei sollte insbesondere überprüft werden, ob die Aufgabe einer Koordinierungsstelle erfüllt wurde.

Im September 2009 wurde daher eine anonyme Befragung der bisher 20 Projektteilnehmer zur Bewertung der Tätigkeit der Koordinierungsstelle durchgeführt. Jedes Archiv erhielt einen Fragebogen, ganz gleich, ob es an mehreren Antragsphasen beteiligt war bzw. ist (bezüglich der bereits bewilligten Projekte betrifft dies alle Teilnehmer der Pilotphase: Bundesarchiv, Landesarchiv Baden-Württemberg, Sächsisches Staatsarchiv und das Staatsarchiv Marburg). Um die Anonymität zu gewährleisten, sollten die Teilnehmer den Fragebogen in einem weißen, unbeschrifteten Umschlag verstauen und diesen wiederum zum Versand in einen weiteren adressierten Briefumschlag

¹ Die Autoren danken Dr. Frank M. Bischoff, Dr. Sigrid Schieber und Ulrike Vogel M.A. für wertvolle Hinweise. Dieser Beitrag wird auch abgedruckt in: Forum – Newsletter der Archivschule Marburg 2010, Heft 34, S.11-13.

stecken. Von den angeschriebenen Archiv(verwaltungs)en haben 18 geantwortet.

Zu folgenden vier Komplexen wurden die Beurteilungen der Teilnehmer abgefragt: „Allgemeine Angaben“, „einzelne in Anspruch genommene Dienstleistungen“, „Vorlagen und Formulare“ sowie „sonstige Tätigkeit der Koordinierungsstelle“. In den meisten Fällen der insgesamt 23 Fragen standen fünf Antwortmöglichkeiten zur Auswahl: „sehr hilfreich“, „hilfreich“, „nicht hilfreich/verbesserungsbedürftig“, „nicht genutzt“ und „keine Antwort“. Nach Auszählung der Antworten wurden entsprechende Diagramme erstellt, wovon einige im Folgenden abgebildet sind. Bei zwei Fragen konnten mehrere Möglichkeiten angekreuzt werden. In einigen Fällen haben die Teilnehmer kurze Zusatzkommentare verfasst. Die hier vorliegende Auswertung ist den Befragten sowie der DFG bereits mitgeteilt worden.

Bei der Auswertung ist zu beachten, dass über die Anträge der Phase 2/2009 zum damaligen Zeitpunkt noch nicht entschieden worden war und die Archive der Antragsphase 1/2009 erst im Juli 2009 die Förderbewilligungen der DFG erhalten haben (zwei der ursprünglich neun Anträge wurden abgelehnt, über einen weiteren Antrag war damals noch nicht entschieden worden). Folglich haben diese Archive lediglich den Prozess der Antragstellung bewerten können. Bei Fragen zum späteren Projektverlauf wurde daher häufig „nicht genutzt“ oder „keine Antwort“ angekreuzt.

2. Allgemeine Angaben der Befragten

Von den Projektteilnehmern wurden zunächst allgemeine Angaben abgefragt. *Eigene Erfahrungen in der Retrokonversion von Findmitteln* mit eigenem oder zusätzlichem Personal besaß die Hälfte der Befragten (9 Archive). Zwei weitere haben entsprechende Projekte nach außen vergeben. Beide Varianten (mit eigenem oder zusätzlichem Personal und durch Vergabe nach außen) probierten bereits sechs Teilnehmer. Die restlichen beiden haben noch keine eigenen Retrokonversionsprojekte durchgeführt.

Die Anzahl der einzelnen Antragsteller verteilte sich folgendermaßen auf die bisherigen vier *Phasen* (die folgende Aufstellung orientiert sich nicht an den Antworten des Fragebogens, sondern an einer Übersicht der Koordinierungsstelle): vier Archive bzw. Archivverwaltungen in der Pilotphase, jeweils neun in der ersten regulären Antragsphase 1/08 (davon drei Zweitteilnahmen) und 1/09 (davon eine Zweitteilnahme) sowie drei in Antragsphase 2/09 (davon ebenfalls eine Zweitteilnahme).

Die *Unterstützung der Koordinierungsstelle* wurde in den einzelnen Antragsphasen vor allem auf drei Gebieten in Anspruch genommen (vgl. Diagramm 1). Die Archive konnten hier mehrere Antworten ankreuzen. Fast alle Befragten (17 von 18 Archiven) ließen sich bei der Antragstellung von der Koordinierungsstelle beraten, 14 bei der Formulierung der Leistungsbeschreibung und ungefähr die Hälfte (11) bei der Einholung der Kostenvoranschläge. Rückfragen während des Retrokonversionsprozesses stellten nur vier Teilnehmer. Jeweils drei setzten sich bezüglich der Vertragsgestaltung und der Vermittlung zwischen Dienstleistungsfirma und Archiv mit der Koordinie-

rungsstelle in Verbindung. Nur wenige Projekte waren zum Zeitpunkt der Befragung bereits mit der Erstkontrolle der Ergebnisse und der Gestaltung des EAD- oder SAFT-Imports (je 2) befasst. Aufgrund der Erstellung von Erfassungsanweisungen suchte nur ein Archiv den Kontakt, ein weiteres nahm keine der angebotenen Unterstützungsmöglichkeiten in Anspruch.

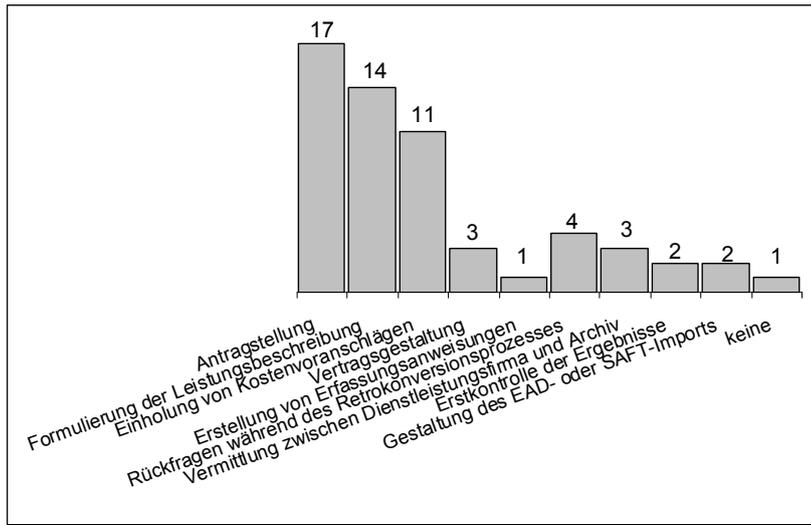


Abb. 1: In Anspruch genommene Unterstützung der Koordinierungsstelle nach Angeboten und Häufigkeit

3. Bewertung einzelner in Anspruch genommener Dienstleistungen

Im nächsten Abschnitt sollten einzelne Dienstleistungen der Koordinierungsstelle beurteilt werden. Die *Beratung bei der Antragstellung* wurde von fast allen Archiven (16) als sehr hilfreich eingeschätzt.

Die übrigen beiden Teilnehmer empfanden diese als hilfreich bzw. nicht hilfreich/verbesserungsbedürftig.

Die Hälfte der Antragsteller beurteilte die *Beratung bei der Formulierung der Leistungsbeschreibung* als sehr hilfreich (9), ein weiteres Drittel (5) als hilfreich. Ungenutzt blieb dieser Service von drei Archiven.

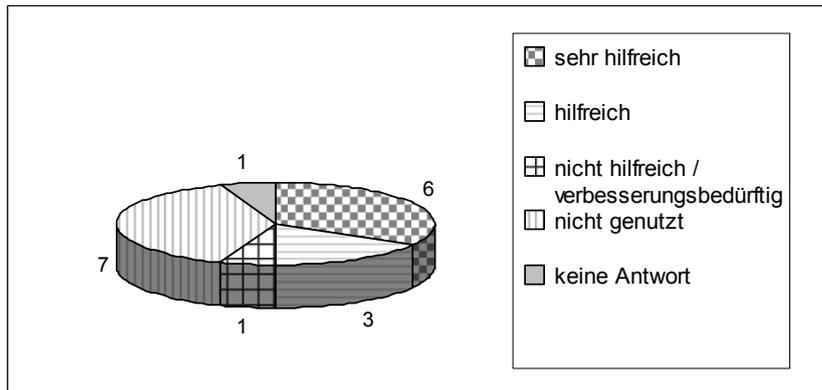


Abb. 2: Bewertung der Beratung oder Unterstützung bei der Einholung von Kostenvorschlägen nach Häufigkeit

Einem Drittel der Archive (6) erschien die *Beratung oder Unterstützung bei der Einholung von Kostenvorschlägen* sehr nützlich (vgl. Diagramm 2). Ca. ein weiteres Drittel der Teilnehmer (7) machte von diesem Angebot keinen Gebrauch. Als hilfreich schätzten es drei Archive ein, als nicht hilfreich/verbesserungsbedürftig bzw. ohne Antwort je einer der Teilnehmer.

Nur vier der 18 Archive nutzten die *Beratung bei der Vertragsgestaltung*. Teilweise liegt dies darin begründet, dass einige Teilnehmer der regulären Phasen bereits durch ihre Beteiligung an der Pilotphase entsprechende Erfahrungen besaßen. Lediglich ein Archiv empfand die Beratung bei der Vertragsgestaltung als sehr hilfreich, drei weitere als hilfreich.

Ein ähnliches Bild bietet sich bei der *Unterstützung zur Erstellung von Erfassungsanweisungen*: zwei Archive haben sie genutzt und erachteten sie als hilfreich (ein Befragter davon mit Einschränkungen). Die Erstellung von Erfassungsanweisungen ist nach den Erfahrungen aus der Pilotphase an die Dienstleister übertragen worden.

Doppelt so viele, nämlich vier Archive, nutzten die *Unterstützung bei Rückfragen während des Retrokonversionsprozesses* (vgl. Diagramm 3). Je zwei der Befragten beurteilten die Unterstützung als sehr hilfreich bzw. hilfreich.

Ein ähnliches Bild bieten die Antworten auf die *Kontaktvermittlung zwischen Dienstleistungsfirma und Archiv*: vier Archive nahmen sie in Anspruch. Die Ursache liegt zum einen darin, dass einige Teilnehmer Personalmittel beantragt haben und daher keine Firma benötigten. Zum anderen hatten die Archive z.T. schon durch frühere Retrokonversionsprojekte eigene Kontakte zu entsprechenden Dienstleistern geknüpft. Je ein Teilnehmer hielt die Kontaktvermittlung für sehr hilfreich bzw. nicht hilfreich/verbesserungsbedürftig. Zwei Befragte fanden die Vermittlung hilfreich.

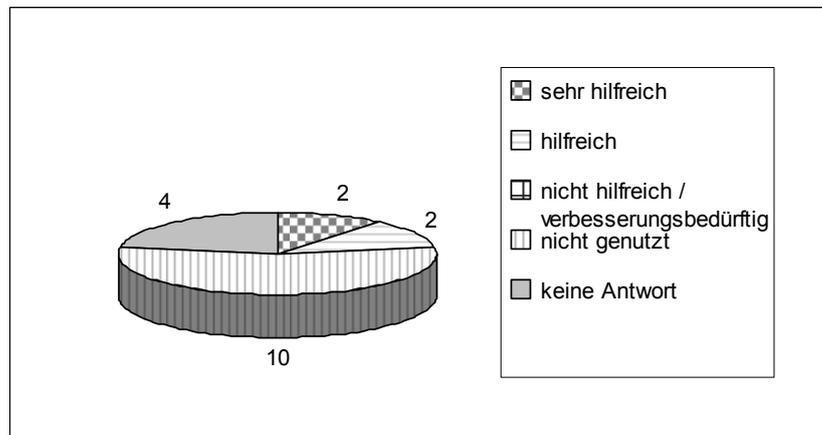


Abb. 3: Bewertung der Unterstützung bei Rückfragen während des Retrokonversionsprozesses nach Häufigkeit

Die *Erstkontrolle der Ergebnisse* konnte zum Zeitpunkt der Befragung erst in wenigen Fällen durchgeführt werden, da sich nur einige der Projekte schon in diesem fortgeschrittenen Stadium befanden. Dadurch lassen sich die sechs darauf antwortenden Archive erklären. Zwei davon schätzten diese Unterstützung als sehr hilfreich (ein Teilnehmer bemerkte „wird selbst vorgenommen und auch von Koordinierungsstelle“). Hilfreich fanden sie drei Teilnehmer, ein weiterer hingegen nicht hilfreich/ verbesserungsbedürftig.

Aufgrund des bereits genannten Grundes konnte auch die *Beratung bei der Gestaltung des EAD- oder SAFT-Imports* nicht in vollem Umfang genutzt werden: nur zwei von 18 Archiven hatten bisher Gebrauch von diesem Angebot gemacht und beantworteten diese Frage überhaupt. Beide Archive bewerteten die Beratung beim Datenimport als hilfreich.

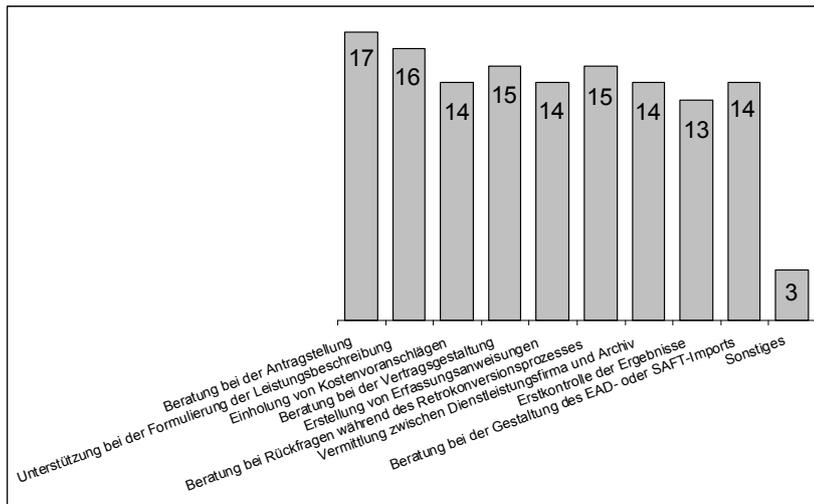


Abb. 4: Diese Dienstleistungen sollten neuen Antragstellern zur Verfügung gestellt werden (nach Häufigkeit)

Breite Zustimmung zu den bisherigen Dienstleistungen der Koordinierungsstelle ergab die Frage, *welche der Angebote auch neuen Antragstellern zur Verfügung gestellt werden sollten* (vgl. Diagramm 4). Auch hier konnten die Archive mehrere bzw. alle Antworten ankreuzen. Fast alle Befragten stimmten dabei für die Beratung bei der Antragstellung (17 Archive), die Unterstützung bei der Formulierung der Leistungsbeschreibung (16) und die Beratung bei der Vertragsgestaltung sowie bei Rückfragen während des Retrokonversionsprozesses (je 15). Weiterhin empfahlen je 14 Befragte die Einholung von Kostenvoranschlägen, die Erstellung von Erfassungsanweisungen (1 Teilnehmer/in merkte zusätzlich an, dies möglichst von einer Firma machen zu lassen), die Vermittlung zwischen Dienstleistungs-

firma und Archiv sowie die Beratung bei der Gestaltung des EAD- oder SAFT-Imports. Die Erstkontrolle der Ergebnisse soll laut Aussage von 13 Projektteilnehmern weitergeführt werden. Ein/e Befragte/r schlug weiterhin vor, die Ergebnisse ohne vorherige Erstkontrolle direkt ins Zielsystem zu importieren und dort abschließend zu kontrollieren. Sonstige Dienstleistungen empfahlen drei der Befragten. Eine interessante Anregung war dabei der Vorschlag, „im Anschluss an Projektpräsentationen auf Archivtagen etc. möglichen Interessenten eine Inhouse-Veranstaltung anzubieten, bei der konkrete Details erläutert werden“. Vor allem solle dabei die Realisierung des Eigenanteils, die Erstellung von Schnittstellen sowie Projektablauf und -aufwand im Archiv vorgestellt werden. „Die Präsentation [auf Archivtagen etc.] dient v.a. der Erstinformation. Danach bleiben aber viele Fragen zum eigenen Arbeitsaufwand offen, die vor dem Entschluss zur Teilnahme am Projekt geklärt werden müssen“, so diese/r Befragte weiter. Unter dem Punkt „Sonstiges“ wurde weiterhin die Unterstützung bei der Datenlieferung an das Archivportal D angegeben.

4. Bewertung der Vorlagen und Formulare

Die Einschätzung der über die Projekthomepage zur Verfügung gestellten Vorlagen und Formulare fiel ebenfalls weitgehend positiv aus.²

² Der Musterantrag sowie die Formulare 1, 2 und 3 sind abrufbar unter <http://www.archivschule.de/forschung/retrokonversion-252/antragsstellung/unterlagen-zur-antragstellung.html> [20.04.2010].

Dreiviertel der Archive (14) beurteilten den *Musterantrag* – den eigentlichen Kern der Antragsunterlagen – als sehr hilfreich (vgl. Diagramm 5). Weitere zwei Teilnehmer empfanden ihn als hilfreich, die restlichen beiden nutzen diese Vorlage nicht bzw. gaben darauf keine Antwort.

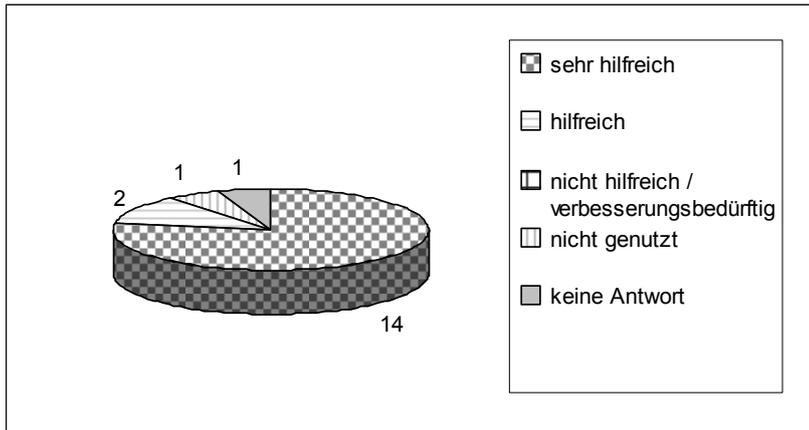


Abb. 5: Bewertung des Musterantrags nach Häufigkeit

Die im *Formular 1* abgefragten Informationen zum jeweiligen Antragsteller beurteilte ungefähr die Hälfte (7 Archive) als sehr hilfreich, elf als hilfreich (vgl. Diagramm 6). Jeweils neun Befragte empfanden das *Formular 2* (Informationen zu den einzelnen hand- und maschinenschriftlichen Findmitteln) als sehr hilfreich bzw. hilfreich. Ein/e Teilnehmer/in schlug zur Verbesserung eine Aufteilung der Informationen einerseits für die DFG und andererseits für das Dienstleistungsunternehmen vor. Beim Mappingschema für die Elemente

der Verzeichnungseinheiten (*Formular 3*) überwog die sehr gute Einschätzung: zehn Archive sahen es als sehr hilfreich, sieben als hilfreich an (keine Antwort: ein Archiv).

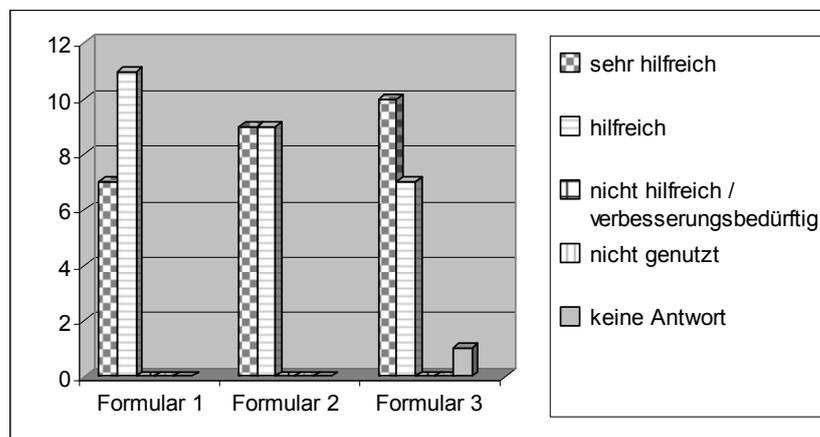


Abb. 6: Bewertung der Formulare 1, 2 und 3 nach Häufigkeit

Die Mehrheit der Antragsteller (10 Archive) beurteilte die *Musterleistungsbeschreibung* als sehr hilfreich. Zwei weitere sahen sie als hilfreich an. Von fünf Archiven blieb sie ungenutzt, ein weiteres antwortete nicht.

Da viele Antragsteller Fördermittel für die Einstellung von Personal beantragten, blieben die *Mustervergabeunterlagen* wenig genutzt: fünf Archive nutzten sie. Als sehr hilfreich beurteilten vier Teilnehmer diese Unterlagen, als hilfreich ein befragtes Archiv.

Der *Mustervertrag* als ein Teil der Mustervergabeunterlagen wurde ebenfalls aus bereits benanntem Grund nur in wenigen Fällen

benutzt: nur fünf Archive machten davon Gebrauch. Zwei davon empfanden ihn als sehr hilfreich. Als hilfreich erschien er drei der Befragten.

Ein ähnliches Bild bietet die Beurteilung des *Musterberichts*, welcher erst nach Abschluss des Retrokonversionsprojekts abzuliefern ist (also maximal drei Jahre nach der Bewilligung). Dadurch erklärt sich dessen Nutzung durch fünf der Befragten. Sehr hilfreich war diese Vorlage für vier, hilfreich für ein Archiv.

5. Bewertung der sonstigen Tätigkeit der Koordinierungsstelle

Die *Öffentlichkeitsarbeit* über die DFG-Förderung der Retrokonversion erschien für mehr als die Hälfte der Archive (10) als hilfreich, für ein Drittel der Befragten (6) sehr hilfreich (vgl. Diagramm 7).

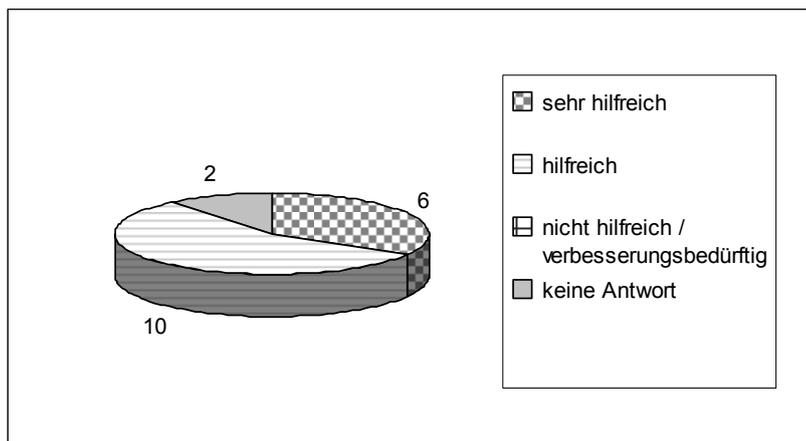


Abb. 7: Bewertung der Öffentlichkeitsarbeit nach Häufigkeit

Sehr wichtig und durchaus noch erweiterbar ist dieses Tätigkeitsfeld für einen Projektteilnehmer. Keinen Kommentar lieferten zwei Archive.

Die Antragsteller waren mit der *Sicherung der einheitlichen Verwendung der Austauschformate EAD und SAFT* zufrieden. Ein Drittel (6) empfand diese Arbeit als sehr hilfreich, für weitere sieben Archive war sie hilfreich (keine Antwort: 5).

Am Schluss des Fragebogens konnten die Projektteilnehmer mittels Freitext weitere Bewertungen vornehmen (was war besonders positiv/besonders negativ?) bzw. Verbesserungsvorschläge für die nächsten Antragsphasen liefern. Die Resonanz fiel überwiegend positiv aus. Gelobt wurde die schnelle und unkomplizierte Hilfestellung bei der Klärung von Fragen und Unklarheiten. Als freundlich, kompetent, engagiert und geduldig wurde die Unterstützung der Antragsteller durch die Koordinierungsstelle beschrieben. Eine/r der Projektteilnehmer/innen hob besonders den Kontakt in den Antragsphasen und die Hilfe beim Ausfüllen der Formulare hervor. Jedoch wurde auch bemerkt, dass durch die Teilnehmer der jüngeren Antragsphasen bisher lediglich die Beratung während der Antragstellung bewertet werden kann. Für die folgenden Phasen wurde die Empfehlung ausgesprochen, den Kontakt zum Bundesarchiv zu verstärken. Der Grund dafür ist die „Nutzung von MEX zum Import von XML-Dateien aus [sic] Standard-Archivsoftware nach dem EAD-XML/Profil des Bundesarchivs für die künftige Bereitstellung im Archivportal D“.

6. Zusammenfassung

In der Gesamtschau lässt sich ein überwiegend positiver Eindruck der Projektteilnehmer erkennen. Aus deren Sicht ist die Koordinierungsstelle in den allermeisten Fällen den ihr gestellten Aufgaben und Anforderungen gerecht geworden. Besonders zufrieden zeigten sich die Befragten mit der Unterstützung bei der Antragstellung, der Beratung bei der Formulierung der Leistungsbeschreibung, dem Mustervertrag sowie der Öffentlichkeitsarbeit. Die Koordinierungsstelle ist um die Umsetzung der Vorschläge sowie um die kontinuierliche Verbesserung und Erweiterung ihres Serviceangebots bemüht. Sie wird darüber auf ihrer Projekthomepage berichten.³

Das Projekt wurde im Juni 2009 von der DFG um weitere zwei Jahre verlängert. Am Ende dieser zweiten Förderphase bieten sich eine erneute Evaluation sowie deren Vergleich mit den vorliegenden Ergebnissen an. Bei einer erneuten Befragung sollte aber möglichst darauf geachtet werden, dass hinter die Antwortmöglichkeit „nicht hilfreich/ verbesserungsbedürftig ...“ die Projektteilnehmer mit dem Zusatz „... weil“ zur Angabe von Gründen animiert werden. Auch aus diesen Reaktionen kann die Koordinierungsstelle Verbesserungsvorschläge für ihre eigene Arbeit ableiten.

³ <<http://www.archivschule.de/retrokonversion>> [20.4.2010]

Komplexe Findbücher? Erfahrungen mit der Retrokonversion durch Dienstleister und im Eigenbetrieb

von Nils Brübach

1. Einleitung

Der folgende Beitrag skizziert Beobachtungen, Ergebnisse und Erfahrungen aus dem Hauptstaatsarchiv Dresden, die aus der Mitarbeit an der Konzeption und Betreuung von Retrokonversionsvorhaben sowie der Entwicklung, dem Test und der Implementierung von Verfahren zur Präsentation von Erschließungsergebnissen in Onlinefindbüchern stammen. Einen Anspruch auf Allgemeingültigkeit können und wollen sie nicht erheben. Es ist ein wesentliches Merkmal von Erschließungsvorhaben und damit auch von Projekten zur Konversion von Papierfindbüchern und Findkarteien in elektronische und online recherchierbare Hilfsmittel (Retrokonversion), dass sie von höchst individuellen, von Einrichtung zu Einrichtung unterschiedlichen Bedingungen auszugehen haben. Es macht einen fundamentalen Unterschied, ob – wie etwa im Stadtarchiv Köln – Retrokonversion eine Schlüsselfunktion im Rahmen des Disastermanagements erhält oder ob sie sich im ruhigen Fahrwasser als langfristig geplantes Projekt vollzieht. Ebenso spielt es eine Rolle, in welchem Maße die notwendigen IT-Voraussetzungen vorliegen. Zu ihnen gehören neben der Verwendung von Erschließungsdatenbanken als Produktionssysteme für die Erschließung auch Werkzeuge für die Aufbereitung und Bereitstellung von Erschließungsergebnissen im Internet. Diese Werkzeuge müssen den fachlichen Anforderungen an die Präsentation archivischer Erschließungsergebnisse entsprechen und sie müssen

zumindest eines der beiden Austauschformate SAFT oder EAD unterstützen. Mit Blick auf laufende Projekte zum Aufbau von Verbundfindmitteln auf nationaler und europäischer Ebene wird die Bedeutung von EAD als Austauschstandard zukünftig noch wachsen. Natürlich gehören zu den Ausgangsbedingungen auch entsprechende personelle Ressourcen, um Projekte zur Retrokonversion von Findmitteln vernünftig abwickeln zu können. Wie noch zu zeigen sein wird, spielt dabei die Nutzung externen Personals eine wichtige Rolle. Aber ganz ohne eigene Mittel wird es nicht gehen.

Neben den technischen und logistischen Voraussetzungen muss auch auf die archivfachlichen Ausgangsbedingungen geachtet werden. Im Hauptstaatsarchiv Dresden wurde und wird Retrokonversion als Teil einer weiter gefassten Erschließungskonzeption verstanden. Die Erarbeitung dieser Strategie begann im Jahr 2003 als Teil eines übergeordneten Prozesses.¹ Dabei wurde zunächst die Kernüberlieferung des Hauptstaatsarchivs in den Blick genommen und ein Ziel- und Maßnahmenplan definiert, der sich auf die folgenden vier Punkte konzentrierte:

1. Erschließung nicht erschlossener Bestände.

Ausgehend von der vollständigen Verpackung des Bestandes des Hauptstaatsarchivs in Vorbereitung auf den Neu- und Umbau

¹ Vgl. den Bericht des Sächsischen Staatsministeriums des Innern an den Sächsischen Landtag über die Umsetzung des Konzeptes zur Erhaltung des staatlichen Archivgutes und die Unterbringung der Archivverwaltung bis zum Jahr 2020, die Mengenentwicklung und die Bestandsgefährdung des staatlichen Archivgutes, hier Ziffer 5.4 und 5.5.; Vgl. auch Sächsischer Landtag, 3. Wahlperiode, Drs. 3/6494. Ziffer 5.

seiner Gebäude und der dazu erforderlichen kompletten Revision aller Bestände ergab sich rasch eine Übersicht über mehr oder weniger unerschlossene Bestandsreste. Für ihre Aufarbeitung erfolgte zunächst eine Konzentration auf die Kernüberlieferung, d.h. diejenigen Bestände, die für eine Vielzahl von Forschungsfragen herangezogen werden können und somit auch besonders häufig benutzt werden, für die Beantwortung von Anfragen besonders wichtig sind bzw. allgemein durch ihre herausgehobene Stellung im historischen Verwaltungsaufbau ihres Entstehungszeitraums oder ihre Funktion als herausragend gekennzeichnet sind.

2. Die Konversion von analogen Findbüchern und Findkarteien in das Archivdatenbanksystem verfolgte zwei Ziele. Zum einen sollte die erleichterte Benutzung für die Allgemeinheit erreicht werden. Unabhängig von dem Einsatz von Onlinefindbüchern führte diese Retrokonversion zu verbesserten Recherchemöglichkeiten in den im Lesesaal des Hauptstaatsarchivs nutzbaren Datenbanken. Zum anderen erbrachte sie aber auch eine Verbesserung der Recherchemöglichkeit zum Zwecke der Beauskunftung durch die Mitarbeiter(innen) des Archivs.
3. Die Verbesserung der Erschließungsqualität dient dazu, dem Benutzer hinreichende Informationen zum Inhalt des Archivgutes bereit zu stellen, so dass eine Einschätzung möglich ist, ob ein Archivale für ein bestimmtes Anliegen einschlägig oder zumindest von Interesse ist. Die Basis für die Verbesserung der Erschließungsqualität ist eine Zuordnung der Bestände in Erschließungsklassen, die nicht nur formalen, sondern auch inhaltlichen

Kriterien folgen. Diese Qualitätsverbesserung, die richtig betrieben, auch zu einer erheblichen Steigerung der Nutzungsqualität („Kundenzufriedenheit“) und zu einer Rationalisierung bei der Benutzung durch gezielte Bestellung führen kann, ist unbedingt eine Fachaufgabe, die nur eingeschränkt Dritten übertragen werden kann. Maßnahmen im Rahmen der Erschließungsverbesserung sind z.B. die inhaltliche Ergänzung vorhandener Erschließungsergebnisse auf der Ebene eines Verzeichnungsdatensatzes oder auch die Erstellung einer Klassifikation und die entsprechende Einordnung der Verzeichnungseinheiten, die dem Benutzer durch Navigation und Suche ein gezieltes Auffinden von Archivgut zu seinem Interessensgebiet in vielen Fällen erst ermöglicht.

4. Auch die Präsentation der Erschließungsergebnisse in Form von Onlinefindbüchern war im Rahmen der Strategie von Anfang an vorgesehen. Hierzu mussten allerdings zunächst geeignete Werkzeuge entwickelt werden, um Onlinefindbücher aus der Archivdatenbank als dem führenden System heraus zu erzeugen. Notwendig war es auch, zunächst eine systematische, strukturierte Präsentation der Bestandsinformation auf der Grundlage der Einheitstektonik des Sächsischen Staatsarchivs herzustellen, mit der die Onlinefindbücher anschließend systematisch verknüpft werden konnten.

Parallel erfolgten Retrokonversionsarbeiten, die sich v.a. auf die Ministerialüberlieferung des 19. und 20. Jahrhunderts sowie auf die Überlieferung des Geheimen Rates mit Akten vom 14. bis zum 19. Jahrhundert konzentrierten. Basis für diese Vorhaben waren für die Bestände aus der Zeit bis 1945 häufig Behördenrepertorien, Findkar-

teilen und in wenigen Fällen Findbücher aus den 1950er Jahren. Zum Einsatz kamen hierbei Aushilfskräfte. Ihre Auswahl erfolgte möglichst abgestimmt auf das jeweilige Retrokonversionsprojekt im Hinblick auf archivfachliche, paläographische und verwaltungsgeschichtliche Vorkenntnisse. Es zeigte sich bald, dass dieses differenzierte Vorgehen und der Einsatz von Aushilfskräften mit möglichst soliden Fach- und Vorkenntnissen sowie die Zuordnung des jeweiligen bestandsverantwortlichen Facharchivars zum Projektteam die richtige Strategie war. Es gelang auf diese Weise innerhalb von vier Jahren (2004-2008) die Mehrzahl der Findmittel derjenigen Bestände zu digitalisieren, die als Teil der Kernüberlieferung des Hauptstaatsarchivs identifiziert worden waren. Als Vorteile der Retrokonversion im Haus erwies sich dabei die Möglichkeit, in das laufende Verfahren jederzeit steuernd eingreifen zu können. Sachliche Korrekturen, das Eingehen auf inhaltliche Besonderheiten und andere Fragen, die erst im Laufe der Retrokonversion auffallen, konnten ohne Verzögerung geklärt werden. Diese steuernden Eingriffe durch die zuständigen Facharchivare reduzierten den zeitlichen Aufwand für eine Nachbearbeitung teilweise erheblich im Vergleich mit einer außer Haus beauftragten Retrokonversion. Zwischen den Facharchivaren und den Aushilfskräften fand eine Art Wissenstransfer statt: Nicht alles, was Archivare zu einem Bestand wissen und was ein Benutzer wissen sollte, steht schon in dem für die Retrokonversion vorgesehenen Findbuch. Häufig liegen zu einem Bestand mehrere parallele Findmittel vor, die zusammengenommen einen Informationsbestand bilden, der in der Retrokonversion berücksichtigt werden soll. Der somit notwendige fachliche Abstimmungsprozess ist leichter und mit weniger Aufwand zu leisten, wenn zwischen Projektkraft und Fach-

archivar auch eine kurzfristige Abstimmung möglich ist. Schließlich müssen in Verbindung mit den meisten Retrokonversionsprojekten auch Fragen der Bestandstrennung, der Nacherschließung von bislang nicht erschlossenen Bestandsresten und die Qualität der vorgefundenen Findbuchgliederung beachtet werden. Dies gilt natürlich für alle organisatorischen Varianten bei der Durchführung von Retrokonversionsprojekten. Allerdings ist auch hier wieder die Integration von Arbeitsergebnissen leichter, wenn vor Ort gearbeitet wird. Es hat sich gezeigt, dass häufig weniger die Komplexität des einzelnen Findmittels, z.B. im Hinblick auf die Anordnung von Erschließungsinformationen oder eine hand- bzw. maschinenschriftliche Vorlage dafür entscheidend sind, ob die Retrokonversion im Hause oder extern durchgeführt wird. Fragen der Qualitätssicherung und die Bereitstellung von Erschließungsinformationen in einer für den Benutzer hinreichenden Informationsqualität, die möglichst aus sich selbst heraus verständlich sind, spielen eine ebenso große Rolle. Das Ziel der Retrokonversion ist eben nicht nur die „Elektrifizierung“ der vorliegenden analogen Findmittel, sondern die Bereitstellung von aussagekräftigen Erschließungsinformationen in einer strukturierten, verständlichen und sinnvoll recherchierbaren Form. Nur wenn dies gewährleistet ist, können aus retrokonvertierten Findbüchern Metadaten gewonnen werden, die für eine nutzbare und verständliche Anbindung von Digitalisaten unverzichtbar sind. Findbücher sind keine Bibliothekskataloge, sie sind Ergebnisse inhaltlicher Erschließung und formaler Erschließung zugleich. Die Zuordnung einer Verzeichnungseinheit zu einer Klassifikationsgruppe ist nicht wahllos, denn sie bildet den Kontext, Entstehungs- bzw. Funktionszusammenhänge und Strukturen ab. Informationen in der Findbucheinleitung und in der

Beständeübersicht stellen den Bestand in den inhaltlichen Kontext der Gesamtüberlieferung des Archivs. Deswegen sind diese Verzeichnungselemente unverzichtbar und an sie sind ebenso Qualitätsanforderungen zu stellen, wie an die Verzeichnungseinheiten selbst. Daher wird auch bei Retrokonversionsprojekten immer eine inhaltliche Redaktion erforderlich sein, die durchaus über eine reine Qualitätskontrolle hinausgeht. Auch auf eine moderate Verbesserung der Erschließungsergebnisse sollte nicht verzichtet werden. Dies kann z.B. in Verbindung mit der Redaktion für die Onlinepräsentation erfolgen und auch durch die Bereitstellung von Benutzungshinweisen geschehen. Perfektionismus ist ebenso zu vermeiden wie „Laissez faire“. Eine kluge Nutzung von inhaltsbezogenen Erschließungsstandards (Content Standards) ist hierbei hilfreich.

2. Die Teilnahme des Sächsischen Staatsarchivs an der Pilotphase der DFG-geförderten Retrokonversion

Das Sächsische Staatsarchiv war in den Jahren 2008/2009 am Produktivpiloten des DFG-Projektes zur Retrokonversion analoger Findbücher beteiligt. Das Hauptstaatsarchiv steuerte sieben Findbücher bei, die mit insgesamt 21.500 Verzeichnungseinheiten den Löwenanteil des sächsischen Kontingents ausmachten. Die zur Konversion vorgesehenen ausschließlich maschinenschriftlichen Findbücher stammten bis auf eine Ausnahme aus den 1980er Jahren mehrheitlich aus der Zeit um 1960. Zusammen mit den anderen an der Pilotphase beteiligten Archiven (Bundesarchiv, Landesarchiv Baden-Württemberg, Hessisches Staatsarchiv Marburg) wurde die Nutzung von EAD als Austauschformat vereinbart, wobei das zwischen dem Bundesarchiv und der Archivschule Marburg vereinbarte Profil zu-

grunde gelegt wurde. Ziel des Produktivpiloten war es, den für das Projekt vereinbarten Arbeitsablauf im Zusammenspiel zwischen beauftragendem Archiv, externem Dienstleister und Projektgeschäftsstelle zu erproben und zu optimieren sowie die Abläufe zur eigentlichen Retrokonversion zu testen – von der Erstellung der Erfassungsanweisungen über das Scannen der Vorlagen, ihre Umsetzung nach EAD, Korrekturdurchläufe bis hin zur Qualitätssicherung beim beauftragenden Archiv. Hierzu gehörte auch die Bereitstellung der Onlinefindbücher und – außerhalb der DFG-Förderung – der Import der EAD-Dateien in die Verzeichnungsdatenbanken. Die Unterstützung durch die bei der Archivschule Marburg angesiedelte Koordinierungsstelle des DFG-Projektes nutzte das Hauptstaatsarchiv bei der Erstellung der Eingabeanweisungen, die nach Empfehlungen der Koordinierungsstelle durch das Archiv formuliert und mit dem Dienstleister ausführlich erörtert wurden. Dieser Prozess ist für den Erfolg und den Aufwand des Retrokonversionsvorhabens entscheidend: Erfolgt doch dabei die Umsetzung der fachlichen Anforderungen in präzise Arbeitsanweisungen an den Dienstleister. Findmittel eines bestimmten Entstehungszeitraumes haben meist typische Merkmale und Gemeinsamkeiten, bei der Erstellung der Eingabeanweisungen kommt es darauf an, die individuellen Merkmale jedes einzelnen Findbuches zu erfassen und ihre Umsetzung möglichst präzise abzusprechen. Stolpersteine waren anfangs dabei z.B. die Auflösung von Abkürzungen, die Einordnung handschriftlicher Ergänzungen und die Auflösung von Verweisungen bei Serientiteln („desgl.“) und Laufzeit („dito“). Der Zuordnung der Verzeichnungsangaben zu den einschlägigen Tags des EAD-Schemas sollte auch Beachtung geschenkt werden. Weniger ist hier häufig mehr – eindeutige Zuordnungen sind zwin-

gend, eine klare sowie möglichst einfache Strukturierung kann ein erfolgskritischer Faktor sein. Wie oben bereits ausgeführt, kommt es bei der Zusammenarbeit mit externen Dienstleistern darauf an, das zu konvertierende Findmittel als eine Einheit zu betrachten und nicht nur die Verzeichnungseinheiten im Blick zu haben. Komplexe Fragen sind dabei zu beachten, z.B. wie soll mit der Findbucheinleitung verfahren werden, sollen Indizes mit konvertiert werden (Onlinefindbücher bieten eine Freitextsuche!), können Teile des Findbuches bei der Konversion weg gelassen werden. Die Klärung dieser Fragen und die Erarbeitung von Lösungen nebst einer eindeutigen Vermittlung an den Dienstleister ist ohne archivarische Fachkenntnis und v.a. Bestandskenntnis nicht leistbar. Es war von Vorteil, dass sowohl in der Geschäftsstelle wie auch beim Dienstleister Facharchivarinnen als Ansprechpartner zur Verfügung standen. Je umsichtiger, sorgfältiger und genauer die Eingabeweisungen erarbeitet werden, desto geringer muss im laufenden Projekt nachgesteuert werden – und v.a., desto geringer ist der Aufwand für Nacharbeiten, die nicht aufgeschoben werden können, wenn man nicht die Termine des Projektlaufes über den Haufen werfen will. Im Rahmen des Produktivpiloten erwies sich die abschließende Phase der Qualitätssicherung aufwändiger als erwartet. Zum einen musste eine große Ergebnismenge innerhalb kurzer Zeit geprüft und in Absprache zunächst durch den Dienstleister und schließlich im Archiv korrigiert werden. Zum anderen war die vereinbarte Korrekturstichprobe nicht ausreichend. Aber hier kann durch entsprechendes Projektmanagement und durch angepasste Abnahmekriterien leicht Abhilfe geschafft werden. Insgesamt war die Pilotphase für das Hauptstaatsarchiv ein Erfolg und hat das Haus zur Teilnahme an den folgenden Antragsphasen ermutigt.

Die hervorragende Unterstützung durch die Projektgeschäftsstelle und der engagierte Einsatz des Dienstleisters haben dazu beigetragen.

Eine Retrokonversion einer großen Anzahl von Findmitteln ohne eine Möglichkeit zur Bereitstellung dieser Information für Benutzer online zur einen Archivbesuch vorbereitenden Recherche und auch im Lesesaal schied für das Hauptstaatsarchiv aus. Daher wurde schon vor dem Beginn der DFG-Förderung seit 2004 nach einem Weg gesucht, Findbücher online zu präsentieren. Das Hauptstaatsarchiv unternahm eine Reihe von Tests verschiedener Verfahren, die alle die Erschließungsdatenbank als führendes System nutzten und darauf ausgerichtet waren, mit möglichst geringem technischen Aufwand, ohne tiefgehende IT-Kenntnisse auf Seiten der Archivare und unter Nutzung der bestehenden Internetplattform des Sächsischen Staatsarchivs zu Ergebnissen zu kommen. Da das in Verbindung mit der im Sächsischen Staatsarchiv eingesetzten Verzeichnungssoftware angebotene Verfahren fachlich nicht vollständig befriedigte und seine Nutzung im Rahmen der bestehenden IT-Architektur auf Schwierigkeiten stieß, wurde eine Alternative erforderlich. Sie fand sich in der Nutzung der unter Federführung der Archivschule Marburg entwickelten Programmpakete Midosa-Online und MidosaXML. Insbesondere MidosaXML erwies sich für die gedachten Zwecke als hervorragend geeignet, bot es doch als Findbucheditor auch noch die Möglichkeit einer der Internetpräsentation gerecht werdenden Endredaktion des Onlinefindbuches. Und beide Programmversionen ermöglichen die Nutzung von EAD als Austauschformat. Somit bestand hier die Möglichkeit zur Nachnutzung von Erschließungsinformationen z.B. in Portalen, aber auch ein Im-

port von Findbüchern im passenden EAD-Profil. Das Hauptstaatsarchiv orientierte sich bei diesem Verfahren am Bundesarchiv und nutzte auch das dort entwickelte EAD-Profil. Dies erwies sich von Anfang an als übersichtlich und gut beherrschbar. Es bildete in seiner Struktur alle notwendigen Funktionen ab, die zum Austausch und zur Präsentation der überwiegenden Mehrheit von Findbüchern benötigt werden. Es ist primär auf die Erschließung von Akten ausgerichtet, mit geringfügigen Anpassungen aber auch für andere Schriftgutformen adaptierbar. Diese Vorgehensweise verletzt übrigens in keiner Weise den Normcharakter von EAD. Angepasste EAD-Profile werden auch in anderen europäischen Ländern, wie zum Beispiel in Italien und Frankreich, verwendet – mit der derselben Intention. Sie bilden ganz pragmatisch ab, wie EAD als Standard vor dem Hintergrund nationaltypischer Verzeichnungstraditionen gut angewandt werden kann. Ein „Richtig“ oder „Falsch“ kann es in diesem Zusammenhang nicht geben, sondern nur die Frage nach dem Grad der Nutzbarkeit. Standards wie EAD bilden einen Werkzeugkasten, aus dem heraus nach den immanenten Regeln die Werkzeuge entnommen werden, die für den jeweiligen Zweck erforderlich sind. Für die Weiterentwicklung eines Standards ist die Beobachtung dieser Anwendungsformen und der Art seiner Implementierung sehr wichtig. Die Diskussionen hierzu erfolgen auf internationaler Ebene in einem offenen Prozess.

Für das Hauptstaatsarchiv war die Nutzung von EAD von Anfang an ein Glücksfall, denn dadurch wurde die Beteiligung am DFG-Retrokonversionsprojekt erleichtert, da dieses ja ebenfalls EAD als ein Austauschformat nutzt. Die einzelnen Verfahrensschritte, die erforderlich sind, um aus der Erschließungsdatenbank Onlinefindbücher her-

zustellen, sind an anderer Stelle ausführlich beschrieben.² Zentrales Werkzeug ist dabei ein Konverter, der automatisch Datenbanken im MS-Office-Format in XML-Dateien umformt, die anschließend mit MidosaXML weiter bearbeitet werden können.

Mit der Entwicklung eines Verfahrens zur Präsentation der Erschließungsinformationen aus dem Produktivsystem, ihrer Redaktion und Onlinestellung war auch der letzte Baustein der eingangs skizzierten Erschließungsstrategie im Hauptstaatsarchiv Dresden umgesetzt. Das Verfahren wurde im Sächsischen Staatsarchiv mittlerweile allgemein eingeführt und hat sich als praxisgerecht und flexibel erwiesen, auch wenn es nicht jede Besonderheit einzelner Findmittel zu berücksichtigen vermag. In jüngster Zeit wurden zur Weiterentwicklung Tests mit dem im Rahmen des beim Bundesarchiv angesiedelten DFG-Projekt „Ausbau des Portals Netzwerk SED-Archivgut zu einer Referenzanwendung für ein Archivportal Deutschland“ entwickelten MIDEX-Editors unternommen. Sein Einsatz soll die leichte Verknüpfung von in EAD vorliegenden Erschließungsinformationen mit Digitalisaten erlauben. Von den derzeit rund 250 Onlinefindbüchern des Hauptstaatsarchivs ist etwa ein Drittel aus Retrokonversionsprojekten entstanden.

² Vgl. Nils Brübach, Mit Midosa-XML-Findbüchern ins Internet, in: Sächsisches Archivblatt, 2008, Heft 1, S. 8 ff. und die ausführliche Präsentation des Workflows unter <<http://www.archivgut-online.de>>. Der Zugriff auf alle im Beitrag genannten Webseiten erfolgte am 21.6.2010.

3. Ausblick

Seitdem sich seit dem Beginn der 1990er Jahre im deutschen Archivwesen Verzeichnungsdatenbanken durchzusetzen begannen, war die Erörterung von Methoden zur Umsetzung bestehender analoger Findbücher immer ein Diskussionspunkt. Die Zukunft des Findbuches als dem Standardzugangsinstrument wurde zeitweise in Frage gestellt; der Einsatz von Datenbanken als Ersatz wurde ebenso diskutiert wie der Ersatz der archivischen Erschließung durch eine Indizierung. Erst durch die Revolution des Internets trat nicht nur das Findbuch als für eine Navigation im Webbrowser besonders geeignetes Zugangsinstrument nun in seiner Form als Onlinefindbuch wieder stärker in den Mittelpunkt, auch die Bereitstellung großer Mengen von Erschließungsinformationen für den Benutzer war sofort auf der Agenda. Die aufkommende Diskussion um die Repräsentation archivistischer Erschließung wurde durch die neuen technischen Möglichkeiten geradezu befeuert. Anwendungen wie das regionale Archivportal "Archive in NRW",³ aber auch die Entwicklung von Werkzeugen wie das 1997 vorgestellte Softwareprogramm MidosaOnline oder die nur unwesentlich jüngere Datenbanklösung "Findbuch.net"⁴ der Firma AUGIAS trugen dazu bei, dass den theoretischen Überlegungen ganz schnell die Praxis folgen konnte. Anders als im Bibliotheksbereich fehlt im Archivwesen der Impetus, Katalogdaten über die Bereitstellung im Verbundkatalog für andere nachnutzbar zu machen. Der unikale Charakter von Archivgut erklärt, warum dieser wichtige Mo-

³ Vgl. dazu den Beitrag von Peter Worm im vorliegenden Band.

⁴ Vgl. <<http://www.findbuch.net>>

tor für die Konversion archivischer Findmittel im Vergleich zur Katalogkonversion zunächst fehlte. Vielleicht wird man auch sagen können, dass im archivischen Bereich der deutlich zu beobachtende Trend zur Nutzung von offline-Datenbanken seine Ursache darin hatte, dass so zumindest die interne Recherche zu Zwecken der Beauskunftung deutlich verbessert und erleichtert werden konnte – ein Aspekt, der im bibliothekarischen Sektor vielleicht nicht ganz die Bedeutung wie im Archivwesen hat. Die Archivverwaltungen einiger Bundesländer, wie z.B. Hessen oder Niedersachsen, konnten so einen erheblichen Vorsprung herausarbeiten.

Umso bemerkenswerter ist daher, dass innerhalb weniger Jahre andere Archive, wie z.B. die baden-württembergische Archivverwaltung oder das Bundesarchiv, aufholten und für die Ausgestaltung und Weiterentwicklung von archivischen Internetplattformen sowie einer Vernetzung mit bibliothekarischen und musealen Angeboten Vorreiter werden konnten. Der Motor dieser Entwicklung war die online-Bereitstellung der Erschließungsinformationen, die Nutzung offener Standards, insbesondere von EAD, die Verbesserung des Serviceangebots und die Erhöhung des Nutzungskomforts. Rasch zeigte sich dabei, dass nicht nur die Bereitstellung einer „kritischen Masse“, sondern möglichst aller sinnvoll nutzbaren Erschließungsinformationen das Ziel sein muss. Dabei weist – wie oben am Dresdner Beispiel bereits gezeigt – das Archivwesen die Besonderheit auf, dass mit Beständeübersichten und Onlinefindbüchern zwei Kategorien von Metadaten systemen bereit gestellt werden müssen, um Benutzern ein sinnvolles Verständnis und einen verständlichen Zu-

gang zu ermöglichen. Werkzeuge wie das BAM-Portal,⁵ die Europeana⁶ und aktuelle Planungen wie etwa die „Deutsche Digitale Bibliothek“ müssen dies berücksichtigen, damit die Einbindung von Digitalisaten aus dem archivischen Bereich gut gelingt. In das Bewusstsein der Benutzer tritt nur, was ohne größeren Aufwand gefunden werden kann und v.a. aus sich heraus verständlich ist.

Der Aufbau von archivischen Verbundfindmitteln sollte daher beschleunigt und konzentriert werden. Sie bieten Gewähr dafür, dass von archivischer Seite selbst Erschließungsinformationen „archivgut-affin“ bereitgestellt und gut mit den Angeboten aus dem bibliothekarischen und musealen Bereich verknüpft werden können. Der Vergleich mit dem Bibliothekswesen bietet sich im Übrigen auch für das Thema Retrokonversion an.⁷ Vor allem bei Projekten, die die Konversion älterer komplexer Katalogsysteme zum Ziel haben, ähneln sich die Erfahrungen von Archiven und Bibliotheken.⁸ Auch hier gilt: eine möglichst präzise Analyse des Kataloges ist entscheidend für den Aufwand und den Projekterfolg. Der Einsatz von Fachpersonal ist

⁵ Vgl. <<http://www.bam-portal.de>>

⁶ Vgl. <<http://www.europeana.eu>> sowie den Beitrag von Gerald Maier im vorliegenden Band. Zu hoffen ist, dass auch das Projekt eines deutschen Archivportals weiter fortgesetzt wird, zum Prototypen des Verbundfindmittels siehe <<http://www.archivgut-online.de>>

⁷ Vgl. zu Retrokonversionsprojekten in Bibliotheken den Beitrag von Karin Schmidgall im vorliegenden Band.

⁸ Vgl. Claudia Fabian/Werner Holbach, Die Konversion des Quartkataloges der Bayerischen Staatsbibliothek – Schlussstein auf dem langen Weg zu einem vollständigen maschinenlesbaren Bestandskatalog?, in: Zeitschrift für Bibliothekswesen und Bibliographie 54, 2007, S. 103–119.

hier ebenso unverzichtbar wie für die Qualitätskontrolle und die -sicherung. Ein wesentlicher Unterschied ist die Nutzung von bereits vorhandenen Katalogdaten auch bei der Katalogkonversion, die im archivischen Bereich wegen des unikaligen Charakters des Archivguts nicht gegeben ist. Aber gerade hierin liegt eine große Chance für die Archive.

Findbücher sind historisch gewachsene, komplexe Informationssysteme, die einen Zugang zu einmaligem Kulturgut gewährleisten – sie sind die Ariadnefäden unseres historischen Gedächtnisses. Retrokonversion ist ein Instrument, mit dessen Hilfe der in der Vergangenheit erworbene und bis in die Gegenwart gepflegte Reichtum an Wissen und Erkenntnis in eine zeitgemäße Form des Nachweises überführt werden kann. Pflege und Verbesserung der Datenqualität unter Nutzung der zeitgegebenen und zeitgebundenen technischen Entwicklung sind eine Grundvoraussetzung, um die Bedürfnisse bei der – insbesondere wissenschaftlichen – Nutzung zu decken und auch mit den Nutzern in einen Dialog zu treten. Und schließlich: Retrokonversion ist die notwendige Ausgangsbasis für die Einbindung von digitalisiertem Archivgut.⁹

⁹ Vgl. dazu die Beiträge von Angelika Menne-Haritz und Paul Bantzer im vorliegenden Band.

Zum Stellenwert der Retrokonversion im Rahmen der Katastrophenbewältigung – das Beispiel des Stadtarchivs Köln

von Ulrich Fischer

1. Retrokonversion und die Katastrophe von Köln – wie passt das zusammen?

Wenn bislang Retrokonversionsmaßnahmen einem Fachpublikum vorgestellt wurden, dann meist als eine nach den Regeln der Zunft mit ausreichendem Vorlauf sach- und fachgerecht durchgeführte Arbeit an einem festgelegten Ausschnitt der Findmittel. Diskutiert wurden Korrekturbedarf, die Rolle von Dienstleistern und die „richtige“ Nutzung des gewählten Austauschformates, über die sich bekanntlich trefflich streiten lässt.

Was aber passiert, wenn mit einem Mal alle verfügbaren Findmittel eines großen Archivs konvertiert werden sollen, und das in einer Situation, in der alle Kräfte für die Bergung und Erstversorgung des Archivgutes angespannt werden und in der für Vorstudien und Fachkonzepte keine Zeit ist? In dieser Situation befand sich das Historische Archiv der Stadt Köln bekanntlich nach dem 3. März 2009.¹

¹ Bettina Schmidt-Czaia/Ulrich Fischer/Max Plassmann, Zum Einsturz des Historischen Archivs der Stadt Köln, in: *Der Archivar* 62, 2009, S. 148-152; Ulrich Fischer/Max Plassmann/Nadine Thiel, Die Katastrophe von Köln: Bergung – Erstversorgung – Zwischenbilanz, in: *Journal of Paper Conservation - IADA Reports* 10, 2009, Heft 2, S. 8-14; Johannes Kistenich, Phasen der Bergung und Erstversorgung des Archivguts aus dem Historischen Archiv der Stadt Kölner, in: *Der Archivar* 62, 2009, S. 305-313; Max Plassmann/Andrea Wendenburg, Histori-

Dabei war von Anfang an deutlich, dass die Retrokonversion aller Findmittel ein zentraler Aspekt jeglicher Überlegungen zur Katastrophenbewältigung sein würde.

Dies schon allein deshalb, weil durch den Einsturz die archivische Ordnung und die Kontexte bei großen Teilen der Bestände verloren gegangen waren. Damit wurde schnell deutlich, dass ein Überblick über die betroffenen Bestände nur über die Findhilfsmittel zu gewinnen war, Findmittel, die in großen Teilen nur analog vorhanden und daher für eine dezentrale Aufarbeitung der geborgenen Bestände untauglich waren. Die Retrokonversion der Findmittel war also im Kölner Fall sofort zu einem integralen Teil des Prozesses geworden, der mittelfristig die Arbeitsfähigkeit des Kölner Archivs wiederherstellen soll. Ihre Durchführung unter Katastrophenbedingungen war zwar für das Kölner Stadtarchiv hinsichtlich Umfang und Verfahren ohne Beispiel, allerdings hat das Stadtarchiv nicht erst mit der Katastrophe mit der Digitalisierung der eigenen Findmittel begonnen. Es soll also im Folgenden ein knapper Überblick über die bisherigen, durchaus als repräsentativ für ein großes deutsches Kommunalarchiv zu sehenden Konversionsmaßnahmen in Köln geboten werden, dann die Retrokonversion nach dem 3.3. geschildert und schließlich der Versuch eines Fazits unternommen werden.

Die Retrokonversion von Findmitteln ist ein vielgestaltiger Prozess. Diese Erfahrung machten vor den Archiven bekanntlich die

sches Archiv der Stadt Köln – ein halbes Jahr nach dem Einsturz, in: *Der Archivar* 62, 2009, S. 399-403; Ulrich Fischer, *Einsturz, Bergung, Perspektiven. Ansichten und Einsichten*, in: Bettina Schmidt-Czaia/Ulrich S. Soénius (Hrsg.), *Gedächtnisort. Das Historische Archiv der Stadt Köln*, Köln u.a. 2010, S. 39-65.

Bibliotheken.² Und auch in den Archiven sind viele Formen der Findmittelkonversion zur Anwendung gelangt.³ Schon seit fast einem Jahrzehnt nimmt sich die Deutsche Forschungsgemeinschaft der Problematik an und hat bereits einige Strukturen von bleibender Bedeutung geschaffen.⁴ Eine Vorstudie und dann die Umsetzung der Ergebnisse in Gestalt eines Förderprogrammes mit eigener Koordinierungsstelle an der Archivschule Marburg haben schließlich dafür gesorgt, dass nun in Deutschland zentral Fördermöglichkeiten für alle

² Vgl. dazu neben einer Vielzahl von Einzeldokumentationen in der bibliothekarischen Fachliteratur die Synopse aus Sicht der Deutschen Forschungsgemeinschaft als hauptsächlichem Geldgeber unter http://www.dfg.de/download/pdf/foerderung/programme/lis/retro_digitalisierung_eval_050406.pdf [21.07.2010]

³ Vgl. zu nationalen Retrokonversionsmaßnahmen in Deutschland und darüber hinaus u.a. Angelika Menne-Haritz (Hrsg.), *Online-Findbücher, Suchmaschinen und Portale*. Beiträge des 6. Archivwissenschaftlichen Kolloquiums der Archivschule Marburg, Marburg 2002 (Veröffentlichungen der Archivschule Marburg 25), und zusammenfassend Ulrich Fischer, *Imaging versus händische Eingabe – eine Wirtschaftlicheitsanalyse zum Problem der Retrokonversion handschriftlicher Findmittel* Marburg 2005 (Archivschule Marburg, Transferarbeit), unter <http://www.archive.nrw.de/dok/Transferarbeiten/Fischer/Transfer%20Text.pdf> [21.07.2010], bes. S. 4, sowie neuerdings auch Jeannette Godau, *Retrokonversion von bibliothekarisch erschlossenen Handschriftenkatalogen im Archiv am Beispiel des Bestandes J 1 im Hauptstaatsarchiv Stuttgart*, Marburg 2009 (Archivschule Marburg: Transferarbeit) unter http://www.landesarchiv-bw.de/sixcms/media.php/120/49559/Retrokonversion_Godau.pdf [21.07.2010]. Die in vielerlei Hinsicht einflussreiche Digitalisierungsstrategie des Bundesarchivs aus dem Jahr 2006 ist unter http://www.bundesarchiv.de/imperia/md/content/bundesarchiv_de/fachinformation/digitalisierungsstrategie.pdf [21.07.2010] zu finden.

⁴ Vgl. dazu besonders die Beiträge der Düsseldorfer Tagung „Entwicklung von Werkzeugen zur Retrokonversion archivischer Findmittel“ 2003 unter <http://www.archive.nrw.de/dok/tagung-retro/Inhalt.html> [21.07.2010].

Archive in öffentlicher Trägerschaft bereit stehen. Vor allem aber wird hier das *know-how* in Sachen Retrokonversion gebündelt und – so ist zu hoffen – durch die Macht des Faktischen und der finanziellen Förderung Standards zum Durchbruch verholfen, die in Zukunft zu mehr Interoperabilität, weniger doppelter Datenhaltung und mehr Nutzungskomfort führen können.⁵ Das Ausland ist Deutschland – gerade auch in diesen Fragen – weit voraus.⁶

Jedenfalls ist auch am Kölner Beispiel einmal mehr das deutlich geworden, was schon seit dem 68. Deutschen Archivtag 1997 in Ulm (unter dem Motto „Vom Findbuch zum Internet“) in der Fachwelt akzeptiert ist. Retrokonversion von Findmitteln ist eine notwendige Vorarbeit für die Archivgutdigitalisierung. Archivgutdigitalisate ohne Bestandskontexte werden schnell zu optisch ansprechenden Ikonen, jedenfalls laden sie nicht ein zu einer Auseinandersetzung mit historischer Information in ihrem jeweiligen Entstehungskontext.⁷ Der Verzicht auf eine Einbettung des Digitalisates in seinen Be-

⁵ Die Vorstudie findet sich unter <<http://www.archive.nrw.de/LandesarchivNRW/BilderKartenLogosDateien/RetrokonversionProjektbericht2006.pdf>> [21.07.2010]. Vgl. dazu den Beitrag von Frank M. Bischoff in diesem Band und Ulrich Fischer/Wilfried Reininghaus, DFG-Vorstudie „Retrokonversion archivischer Findmittel“. Die wichtigsten Ergebnisse des Projekts, in: *Der Archivar* 59, 2006, S. 329-333.

⁶ Vgl. auch schon Ulrich Fischer/Wilfried Reininghaus, „Keeping it simple“ – das englische Retrokonversionsprogramm A2A tritt in seine 4. Projektphase. Arbeitsbesuch bei The National Archives, Kew, in: *Der Archivar* 59, 2006, S. 198-200.

⁷ Besonders groß ist diese Gefahr bei Portalen wie der EUROPEANA, die zuvorderst dokumentbasiert operieren und bei denen ein Konnex zwischen Einzelobjekten wenn überhaupt, dann über den „original context“ des jeweiligen Einstellers außerhalb des Portals erfolgt.

stands- und Provenienzkontext erschwert die forschende Beschäftigung mit historischen Sachverhalten und ihrem schriftlichen Niederschlag – ja, er mag sie gar in einigen Fällen verhindern. Dies insbesondere dann, wenn der gegenwärtige Trend in der Forschung wie im täglichen Leben weitergeht, Dinge nicht mehr wahrzunehmen, die online nicht zumindest referenziert werden.⁸

Schließlich ist der enge Zusammenhang mit der informationstechnologischen Umsetzung der Findmittelinformation zu betonen, insbesondere mit der Verwendung von Austauschformaten. Zwar bezeichnet der Begriff der Retrokonversion an sich lediglich die Überführung von analog vorliegenden Findmittelinformationen in irgendein digitales Zielsystem; tatsächlich aber hat sich überall dort, wo Archivinstitutionen gemeinsam in die Konversion von Findmitteln eingestiegen sind, die Verwendung eines Austauschformates durchgesetzt.⁹ Dies ist insofern nachvollziehbar, als nur die gemeinsame Präsentation, die gemeinsame Durchsuchbarkeit von archivischer Metainformation für den forschenden Nutzer sinnvoll ist und die Wahrnehmbarkeit einzelner beitragender Archive erhöht.¹⁰ Portal-

⁸ Vgl. Ulrich Fischer, Findbuch – Kartei – Datenbank. Praktische Beispiele für den Umgang mit analogen Findmitteln im digitalen Zeitalter, in: Marcus Stumpf/Katharina Tiemann (Hrsg.), Aufbruch ins digitale Zeitalter – Kommunalarchive zwischen Vorfeldarbeit und Nutzerorientierung, Münster 2008 (Texte und Untersuchungen zur Archivpflege 21), S. 57-75.

⁹ Vgl. für ein sehr effektives ausländisches Beispiel das Projekt A2A und sein Nachfolgeprojekt Access4All bei den National Archives in Kew bei London, unter <<http://www.nationalarchives.gov.uk/a2a>> [21.07.2010].

¹⁰ Vgl. Frank M. Bischoff (Hrsg.), Benutzerfreundlich – rationell – standardisiert. Aktuelle Anforderungen an archivische Erschließung und Findmittel. Beiträge zum

lösungen werden seit langer Zeit in der Archivwelt angestrebt und mit Erfolg umgesetzt; sie funktionieren aber nur dort, wo alle Teilnehmer Ihre Informationen in einer Form aufbereiten, die wechselseitig verständlich ist.¹¹ International ist EAD der Standard zur systemübergreifenden Codierung von Archivinformationen, national hat sich dazu noch der deutsche Verzeichnungstraditionen genauer abbildende Standard SAFT-XML etabliert.¹²

2. Retrokonversion vor dem 3. März 2009

Schon vor dem 3. März 2009 wurde am Historischen Archiv der Stadt Köln Retrokonversion betrieben, und zwar auf verschiedenen Wegen und mit verschiedenen Mitteln. War das Kölner Stadtarchiv einerseits ein digitaler „Spätstarter“ – immerhin wurde erst im Jahr 2007 eine echte Archivsoftware eingeführt – so hatte andererseits die Retrokonversion früh begonnen.

11. Archivwissenschaftlichen Kolloquium der Archivschule Marburg, Marburg 2007 (Veröffentlichungen der Archivschule Marburg 46).

¹¹ Vgl. etwa den Beitrag von Michael Fox in diesem Band. Die weitergehenden Fragen von mehrfacher Datenhaltung (und der Notwendigkeit jeweils häufiger Aktualisierungen) werden dabei in der Zukunft besonders wichtig werden, vgl. auch den Beitrag von Stefano Vitali in diesem Band.

¹² Ulrich Fischer, Austauschen, sichern, präsentieren – Ansprüche an ein archivisches Standardaustauschformat, in: Frank M. Bischoff (Hrsg.), Benutzerfreundlich – rationell – standardisiert. Aktuelle Anforderungen an archivische Erschließung und Findmittel. Beiträge zum 11. Archivwissenschaftlichen Kolloquium der Archivschule Marburg, Marburg 2007 (Veröffentlichungen der Archivschule Marburg 46), S.177-202.

Seit 2001 wurden in Zusammenarbeit mit dem Berliner Dienstleister *ArchivInForm* Findmittel vor allem der Alten Abteilung (Depositalebestände des Landes NRW zu den Kölner Klöstern und Stiften) konvertiert – in Word-Dokumente und Excel-Tabellen. Etwa gleichzeitig begann ein Kollege, in Eigenarbeit und mittels OCR-Texterkennung vor allem die unbenutzten, auf mehrere Bände der Reihe *Mitteilungen aus dem Kölner Stadtarchiv* verteilten Findmittel zu großen Urkundenbeständen in Excel-Tabellen zu überführen. Insbesondere die Arbeiten am Bestand 1 (Haupturkundenarchiv) mit allein 22.000 Nummern nahmen dabei Jahre in Anspruch, gingen aber auch mit einer inhaltlichen Verbesserung der Erschließung einher. Gleiches gilt für den Bestand 210 (Dom).¹³ Kleinere Bestände wurden parallel mit bearbeitet.

2006 wurde Retrokonversion – gemeinsam mit der Aufarbeitung von Verzeichnungsrückständen – als ein zentraler Aspekt der weiteren Archivstrategie identifiziert. Es gelang, Mittel für diese Arbeiten in den Haushalt einzustellen, so dass die Zusammenarbeit mit Dienstleistern fortgeführt und intensiviert werden konnte. Weiterhin lag der Fokus auf den Beständen der Kölner Klöster und Stifte, um diesen Block möglichst zeitnah vollständig digital vorlegen zu können. Bis 2008 wurden nahezu alle handschriftlichen Findmittel aus diesem Bereich bearbeitet. Zudem wurden mit Inbetriebnahme der Archivsoftware durch Mitarbeiterinnen und Mitarbeiter händisch Bestände eingepflegt, unter anderem aus der umfangreichen Nach-

¹³ Beide Bestände wurden zwischenzeitlich in einer vorläufigen Version in die Datenbank eingestellt.

lassabteilung.¹⁴ Im Haushalt 2008 wurde eine zusätzliche halbe Stelle für die Retrokonversion bewilligt. Die Arbeiten der betreffenden Kollegin konzentrierten sich auf Bestände zu Kölner Schulen, die von Bürgern intensiv nachgefragt werden. Zudem wurde verstärkt an semiautomatischen Konversionsverfahren (OCR, Layouterkennung) gearbeitet und die Ergebnisse in die Archivsoftware übernommen.

Das Kölner Stadtarchiv nahm außerdem mit einem Paket von Findbüchern aus allen Abteilungen an der ersten Förderphase 2008 des DFG-Projektes *Retrokonversion von Findmitteln* teil. Insgesamt 16209 Einheiten aus den Findbüchern zu zehn Beständen wurden gemeldet. Die Auswahl der betreffenden Bestände folgte den DFG-Kriterien (zentrale Referenzbestände), bezog aber zusätzlich die Nutzungsintensität und die inhaltlichen Zusammenhänge mit ein.¹⁵ Die erforderlichen Eigenanteile stellten Konversionsarbeiten im Haus dar,

¹⁴ Neben den Arbeiten in der Nachlassabteilung wurden aber auch die Beschreibungen der bislang nicht im Druck oder auf <http://www.manuscripta-mediaevalia.de> [21.07.2010] beschriebenen Handschriften des Archivs manuell in die Archivsoftware konvertiert.

¹⁵ So wurden einerseits mit den im Druck vorliegenden Beständen 214 (Georg) und 251 (Mariengraden) sowie dem handschriftlichen Repertorium zum Kloster St. Heribert in Deutz (Bestand 208) die letzten Kölner geistlichen Institutionen digital verfügbar. Gleichzeitig wurden als zentrale Bestände die Französische Verwaltung (Best. 350, handschriftliches Repertorium mit 6987 Nummern) und die Bestände Oberbürgermeister und Oberstadtdirektor (Acc. 2 und Acc. 5) aus der Nachkriegszeit identifiziert, die beide gleichzeitig im Rahmen der Landesinitiative Substanzerhalt entsäuert und aufgearbeitet worden waren. Diesen beigegeben wurde mit den Nachlässen der Oberbürgermeister Schwering (Best. 904) und Pünder (Best. 1304) die gleichsam persönliche Seite der Amtsinhaber. Zusätzlich wurde der bedeutende (Korrespondenz-)Nachlass des Kunstsammlers Sulpiz Boisserée (Best. 1018) angemeldet.

für die Mitarbeiterinnen und Mitarbeiter freigestellt wurden. Nach der Genehmigung und der Beauftragung der Dienstleister wurde diese Maßnahme – allerdings schon unter den Vorzeichen des Einsturzes im Herbst 2009 – fertig gestellt.

Um die Jahreswende 2008/2009 waren damit einige Retrokonversionsprojekte angelaufen, die dazu führen sollten, dass zum Jahresende 2009 für mindestens 15% der Bestände des Stadtarchivs digitale, online verfügbare Erschließungsinformationen vorliegen sollten. Bei einer Grundgesamtheit von etwas mehr als 3000 Beständen bedeutete dies online-Findmittel für mindestens 450 Bestände.¹⁶ Parallel zu den Retrokonversionsbestrebungen wurden gleichzeitig Erschließungsarbeiten durchgeführt, sowohl langfristig angedachte, wie zum Bestand 403 (Oberbürgermeister) als auch projektweise und *ad hoc* durchgeführte wie die Schnellerfassung der Kölner Personenstandsregister (Best. 541) im Januar und Februar 2009.

Die Konversion der Findmittel wurde und wird im Historischen Archiv der Stadt Köln zweifach verstanden – als ein Mittel zur Effektivierung der Arbeitsabläufe im Haus, und als ein Teil einer Gesamtstrategie zur Verbesserung der Sicht- und Nutzbarkeit archivischer Informationen für alle Nutzergruppen. Im Kontext dieser Strategie stand die Beteiligung des Kölner Stadtarchivs an der Weiterentwicklung des Portals *www.archive.nrw.de* und des Austauschformates SAFT-XML, aber auch erste Projekte, die auf eine Digitalisierung von Archivgut im Kontext ihrer (verbesserten) Erschließung

¹⁶ In Anbetracht der großen Erschließungsrückstände bedeutete diese Zahl einen Anteil von deutlich über 25% an den erschlossenen Beständen.

abzielten. Hier ist unter anderem die projektierte Kooperation mit der Universitätsbibliothek Köln bei der Digitalisierung eines Plakatbestandes aus der Zeit der deutschen Besatzung in Belgien und Nordfrankreich zu nennen, ebenso wie die Zusammenarbeit mit der Universität zu Köln bei einem Regesten-/Editionsprojekt zum Urkundenbestand der Kölner Deutschordenskommande St. Katharina.¹⁷

Technisch wurden diese und andere Projekte bestimmt vom Bestreben, möglichst flache Strukturen bei der Bewältigung aller Aufgaben im digitalen Bereich zu schaffen. Als eine kleine Dienststelle bestand für das Historische Archiv der Stadt Köln weder die Möglichkeit noch das Bestreben, eigenständige IT-Strukturen zu entwickeln. Stattdessen setzten wir vor allem auf eine intensive Kooperation mit dem Amt für Informationstechnologie und orientierten uns an der dort vorgehaltenen Technik. Damit waren etwa die zu verwendenden Speichertechnik und das Speichermanagement nicht verhandelbar, während anders herum die archivischen Anforderungen die technischen Lösungen im Bereich von Austausch- und Archivierungsformaten beeinflussten. Das archivische Informationssystem ACTApro spielt dabei die Rolle einer „Spinne“ im Datennetz der Stadt Köln, zumindest was das Management der archivischen Daten angeht.

¹⁷ Daneben fanden weitere, kleinere Projekte statt, etwa die Digitalisierung der Filmplakate aus dem Bestand 1475 (Tabertshofer).

3. Retrokonversion nach dem 3. März 2009

Dies war der Stand, als der Einsturz des Historischen Archivs alle weiteren Planungen zunächst obsolet machte. Es war selbstredend entscheidend, die materielle Bergung des Archivgutes und seine Erstversorgung sicher zu stellen.¹⁸ Dennoch gehörte es zu den ersten Bergungsmaßnahmen, die Erstexemplare der analogen Findmittel zu bergen, die glücklicherweise den Einsturz überlebt hatten. Dies war vor allem deshalb entscheidend, weil die Duplikate nicht flächendeckend mit Nachträgen ausgestattet worden waren. Gerade diese Nachträge sind allerdings in vielen Fällen entscheidend, können doch nur durch sie Um- und Neusignierungen, Überführungen in andere Bestände, Neuerwerbungen und Verluste nachvollzogen werden – ganz zu schweigen von Hinweisen auf Parallelüberlieferungen, neuerer Forschungsliteratur usw. Einige wenige Findmittel waren gar Unikate, selbst als Zeugnis von archivischer Arbeit oder Registraturpraxis vom Charakter her Archivgut oder doch zumindest mit Blick auf die archivische Praxis hochwillkommene Konkordanzen, Übersichten oder Zusammenfassungen.¹⁹

¹⁸ Vgl. dazu nun auch Bettina Schmidt-Czaia, Das Historische Archiv der Stadt Köln. Geschichte – Bestände – Konzeption Bürgerarchiv, in: Bettina Schmidt-Czaia/Ulrich S. Soénius (Hrsg.), Gedächtnisort. Das Historische Archiv der Stadt Köln, Köln u.a. 2010, S. 10-38, bes. S. 23f. und Ulrich Fischer, Einsturz – Bergung – Perspektiven. Ansichten und Einsichten, ebd., S. 39.

¹⁹ Beispiele: Registraturbuch (19./20. Jh.) der zentralen städtischen Registratur als Findmittel für Bestand 403; handschriftliche Konkordanz der Handschriftenbestände (Best. 7002, 7004, 7008, 7010 und 7020) mit Verweis auf die Beschreibungen; Zusammenfassung der verstreut in den Mitteilungen aus dem Stadtarchiv von Köln publizierten Regesten zum Best. 1 (Haupturkundenarchiv).

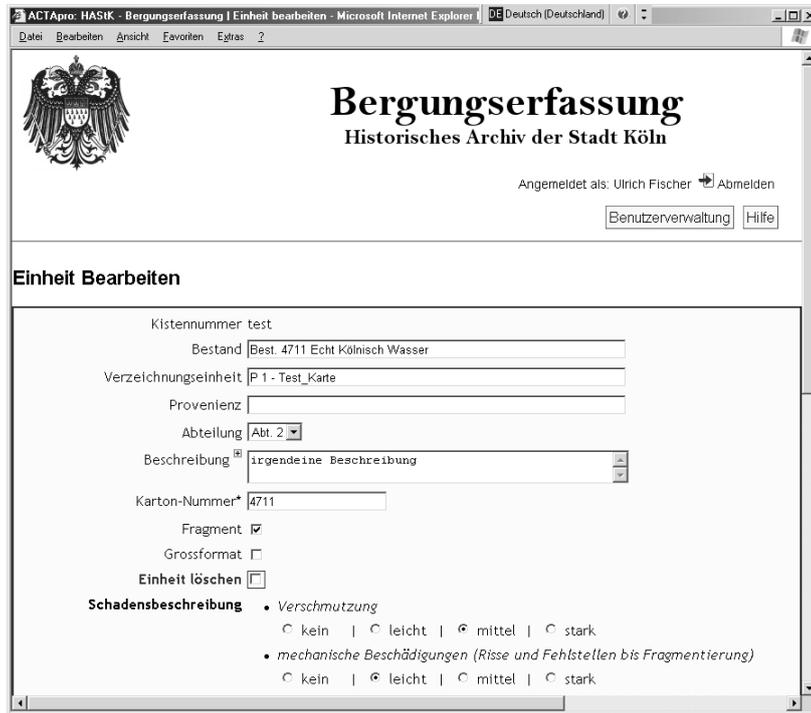
Bereits in den ersten Tagen der Bergungsarbeiten an der Unglücksstelle wurde deutlich, dass eine vollständige Durchmischung der Bestände stattgefunden hatte. Gleichzeitig stellte sich heraus, dass mitnichten alle gefundenen Stücke weitgehend intakt und mit einer genauen und auch für Laien verständlichen Signatur versehen wieder aus dem Schutt auftauchten. Zu dem bekannten Problem beträchtlicher Erschließungsrückstände (mit weit mehr als 40% unerschlossener Bestände) gesellte sich also ein weiteres, nämlich, dass auch verzeichnetes Archivgut nicht unmittelbar identifizierbar sein würde.

Es wurde deshalb noch in der ersten Märzhälfte gemeinsam mit dem Hersteller der Archivsoftware nach Lösungen gesucht, um die Zuordnung und Identifikation geborgenen Materials zu erleichtern und möglichst schnell wieder die Bestandskontexte zu etablieren. Das Resultat dieser Bemühungen war die sogenannte „Bergungserfassung“, ein Modul, das es zunächst erlaubte, für größere Blöcke Bergungsgut (die „blauen Wannen“ aus dem Erstversorgungszentrum bzw. die Gitterboxen mit nassem Material) summarische Inhaltsbeschreibungen zu hinterlegen und den Verbleib des Materials in einem der 19 Asylarchive festzuhalten.

Schnell wurde deutlich, dass die Erfassungslisten noch keine punktgenaue Identifikation der einzelnen Stücke, geschweige denn einen Abgleich gegen die in den Findmitteln beschriebenen Verzeichnungseinheiten, ja nicht einmal eine bestandsweise Zuordnung erlaubten. Für eine Weiterentwicklung des Moduls Bergungserfassung wurden also folgende Anforderungen definiert:

1. Die Verknüpfung eines geborgenen Einzelstückes (einer „Bergungseinheit“) mit einer Verzeichnungseinheit muss live auf der ACTApro-Datenbank festgehalten werden können.
2. Ist keine Verknüpfung mit einer spezifischen Verzeichnungseinheit möglich, so muss entweder die Zugehörigkeit zu einem Bestand, einer Provenienz oder zumindest zum Bestandskomplex einer „Abteilung“ des Stadtarchivs (Alte Abteilung, Neue Abteilung, Nachlässe und Sammlungen) hinterlegt werden können.
3. Jede Bergungseinheit muss leicht und fehlerfrei zu finden, zu identifizieren und in einen anderen räumlichen Kontext oder einen anderen Prozessschritt (Restaurierung, physische Bestandszusammenführung) zu überführen sein.
4. Jede Bergungseinheit muss leicht und von Laien hinsichtlich ihres Schadensbildes beschrieben werden können und einer Schadensklasse zuzuordnen sein.
5. Es müssen Digitalfotos zu Bergungseinheiten hochgeladen werden können, mit Hilfe derer entweder Schadensbilder illustriert oder durch ortsfern agierende Experten Identifikationsversuche unternommen werden können.
6. Die geborgenen Einheiten müssen nach verschiedenen Parametern zu durchsuchen und zu filtern sein.

Mittlerweile verfügt das Modul „Bergungserfassung“ über diese Funktionen – und es wird für die Kontrolle der verschiedenen Workflows mit einem Modul „Umzugsmanagement“ ergänzt.



ACTApro: HASIK - Bergungserfassung | Einheit bearbeiten - Microsoft Internet Explorer | Deutsch (Deutschland)

Datei Bearbeiten Ansicht Favoriten Extras 2

 **Bergungserfassung**
Historisches Archiv der Stadt Köln

Angemeldet als: Ulrich Fischer [Abmelden](#)

[Benutzerverwaltung](#) [Hilfe](#)

Einheit Bearbeiten

Kistennummer test

Bestand

Verzeichnungseinheit

Provenienz

Abteilung

Beschreibung

Karton-Nummer*

Fragment

Grossformat

Einheit löschen

Schadensbeschreibung

- *Verschmutzung*
 - kein | leicht | mittel | stark
- *mechanische Beschädigungen (Risse und Fehlstellen bis Fragmentierung)*
 - kein | leicht | mittel | stark

Abb.1: Modul „Bergungserfassung“ der Archivsoftware ActaPro, oberer Teil

The screenshot shows a web browser window titled "ACTapro: HASTK - Bergungserfassung | Einheit bearbeiten - Microsoft Internet Explorer". The page content is as follows:

Schadensbeschreibung

- Verschmutzung
 kein | leicht | mittel | stark
- mechanische Beschädigungen (Risse und Fehlstellen bis Fragmentierung)
 kein | leicht | mittel | stark
- Knicke/Stauchungen
 kein | leicht | mittel | stark
- Deformierung
 kein | leicht | mittel | stark
- Microbenbefall
 kein | leicht | mittel | stark
- Wasserränder
 kein | leicht | mittel | stark
- Schäden an Text- und Schriftbild (Beeinträchtigung der Lesbarkeit)
 kein | leicht | mittel | stark

Gesamtschadensklasse 2

Sonstiges

Barcode

Bild

Bilddatei

Titel

Nach dem Speichern: Zurück zur Einheitenübersicht | neue Einheit anlegen | neue Einheit anlegen, Daten dieser Einheit übernehmen

Abb.2: Modul „Bergungserfassung“ der Archivsoftware ActaPro, unterer Teil

Damit diese Softwareunterstützung allerdings wirkungsvoll genutzt werden konnte, mussten zunächst sämtliche verfügbaren Findmittel digital vorliegen und in die Datenbank der Archivsoftware ACTapro eingespielt werden. In unserem Fall bedeutete dies zusätzlich zum bereits Erreichten die Digitalisierung von weiteren mehr als 1000 Findhilfsmitteln verschiedener Art in möglichst kurzer Zeit. Diese Arbeiten konnten – zumal während des laufenden Bergungsbetriebes – nicht mit Bordmitteln umgesetzt werden. Genausowenig

kamen angesichts von Gesamtumfang und knapper Zeit Fachdienstleister in Frage – an diese wurden lediglich einige komplexe Spezialfälle vergeben, und die noch ausstehenden Arbeiten aus dem regulären DFG-Retrokonversionsprojekt abgearbeitet.²⁰

Ein Unternehmen, das sich allgemein mit Verwaltungsdigitalisierung beschäftigt und in diesem Zusammenhang nicht nur DMS-Lösungen anbietet, sondern auch als Scandienstleister tätig ist, bot an, diese Arbeiten in der knappen Zeit durchzuführen. In einigen Gesprächen wurden die Parameter von Angebot und angeforderter Leistung wechselseitig konkretisiert und schließlich ein annahmefähiges Angebot vorgelegt.

Blieb die Frage der Finanzierung. Zwar ist das Verfahren in der Zusammenarbeit mit dem ausgewählten Dienstleister – auch aufgrund der zu realisierenden Skalenerträge – recht günstig, in der Summe aber blieb die Aufgabe und damit der Finanzbedarf beträchtlich. In der Zwischenzeit hatte die Deutsche Forschungsgemeinschaft die Sicherung der Erschließungsinformationen durch Konversion der Findmittel als einen zentralen Aspekt ihrer Katastrophenhilfe identifiziert. Damit standen ausreichende Mittel zur Verfügung, um alle Findmittel zu konvertieren.²¹

²⁰ Z.B. ein Konvolut von insgesamt sechs teilweise handschriftlich erschlossenen Findmitteln zu komplexeren Pertinenzbeständen aus der Alten Abteilung (Best. 33 (Militaria), Best. 45 (Reformation), Best. 82 (Hanse II), Best. 290 (Clerus secundarius) und Best. 303 (Mülheim, Älteres Stadtarchiv)).

²¹ Insgesamt wurden dem Historischen Archiv der Stadt Köln 600.000,- EUR zur Verfügung gestellt. Diejenigen Mittel, die nicht für die Retrokonversion aufgewandt werden, werden zusammen mit Landesmitteln für die Digitalisierung der mikroverfilmten Bestände verwendet.

In der Zusammenarbeit mit dem ausgewählten Dienstleister, der Fa. Paperless wurde ein Verfahren entwickelt, das konsequent dem Grundsatz größtmöglichen Durchsatzes folgte. Teil dieses Verfahrens war zunächst die bewusste Konzentration auf wenig tief strukturierte Findmittel, die möglichst schnell und einfach zu konvertieren waren. Früh wurden also einfache Findmittel vor allem für Behördenschriftgut des 20. Jahrhunderts bearbeitet, mit einem Schwerpunkt auf Massenschriftgut (Fallakten u.a.m.). Später wurde mit der Bearbeitung von Sammlungs- und Nachlassbeständen, älteren Beständen, Erschließungswerken zu Karten- und Planbeständen und den Handschriftensammlungen auch komplexere Findmittel auf diesem Wege konvertiert. Eingeschlossen in die Retrokonversion mit diesem Dienstleister sind allerdings auch die nun laufenden komplexen Erfassungsarbeiten der Karteikarten der Zeitgeschichtlichen Sammlung Bild und des Bibliothekskataloges.

Die Erfassung beim Dienstleister erfolgt teilweise automatisch und über OCR, wird aber immer durch eigens dafür angestellte Mitarbeiter kontrolliert. Je nach Anforderung und gewähltem Verfahren verändern sich der Anteil der manuellen Arbeit und die Anforderungen an die Mitarbeiter beim Dienstleister. Für die jetzt laufende Konversion des Bibliothekskataloges ist extra eine bibliothekarische Fachkraft eingestellt worden.

Für eine effiziente Bearbeitung der „Massenfälle“ brachte das Historische Archiv den Softwarehersteller direkt mit dem ausgewählten Retrokonversionsdienstleister in Kontakt. Vereinbart wurde eine direkte Erstellung von Findmitteln im Zielformat, dem XML-Dialekt der Software. Dies gelang nach einigen wenigen Sitzungen zufrie-

denstellend – am Komplexesten blieb die Festlegung der zu verwendenden IDs (sog. „DocKeys“), um einzelne Verzeichnungseinheiten oder Klassifikationspunkte eindeutig zu identifizieren und zu referenzieren. Die konvertierten Findmittel waren oft in der Klassifikation reduziert – in einigen Fällen war (und bleibt teilweise bis jetzt) auch in der Struktur Nacharbeit erforderlich.²² Mit Blick auf Fehlertoleranzen und der Korrektheit der XML-Instanzen gab es über die ganze Phase bislang keine Beanstandungen. Die Konzentration auf „einfache“ Findmittel und die enge und direkte Zusammenarbeit zwischen dem für die Konversion zuständigen Dienstleister und dem sehr archiverfahrenen Softwarehersteller ermöglichten zudem einen weitgehenden Verzicht auf die Erstellung individueller Erfassungsanweisungen. Stattdessen lieferte das Archiv in bislang etwa einem Dutzend Sitzungen Antworten auf entstandene Fragen und stellte möglichst homogene Gruppen von Findmitteln zusammen.

Diese Besprechungen befassten sich stets mit den Problemen der neu anzugehenden Findmitteltypen. Am Anfang der gemeinsamen Arbeit stand die Darstellung genereller Prinzipien (Wie sind Laufzeiten aufzunehmen? Welche häufigen Abkürzungen sind wie aufzulösen?) am Beispiel von Aktenfindmitteln und Abgabelisten, da diese als „einfachste“ Findbuchgattung identifiziert wurden. Weitere Absprachen waren dann für Findmittel für Großformate und Fotografien, Handschriften und den Bibliothekskatalog erforderlich. Als letz-

²² Es ist insgesamt geplant, die Findmittel abteilungsweise einer Revision zu unterziehen; hier werden einige Nacharbeiten erforderlich sein.

tes werden einige wenige noch nicht konvertierte Urkundenfindmittel bearbeitet werden.²³

Die organisatorische Schwierigkeit besteht wie immer in der Retrokonversion in der Verbindung von technischem Verständnis und dauerhafter Verfügbarkeit von archivischem Fachwissen. Die Katastrophenbewältigung, die für alle Beteiligten an erster Stelle stand, band eigentlich alle verfügbaren Kräfte. Dennoch gelang es, zu den jeweils entscheidenden Zeitpunkten die Fachkompetenz für die Vorbereitung der Erfassungsarbeiten sicher zu stellen. Der Import in die Archivdatenbank wurde als prioritäre Aufgabe identifiziert und als solche stetig von einem Kollegen durchgeführt.

Kontrollen, händische Konversionsmaßnahmen und teilweise erforderliche Umarbeiten an den importierten Daten werden von Kollegen immer dann erledigt, wenn – etwa zwischen Phasen des externen Einsatzes – dazu die Möglichkeit besteht. Auch durch hausinterne Arbeiten wurden komplexere Findmittel konvertiert, so etwa die 1864 umfangreich erschlossenen Reichskammergerichtsakten (Bestand 310).

Insgesamt wurden bislang mehr als 200.000 Einheiten extern konvertiert. Die Kosten dafür betragen bislang etwa 150.000,- EUR, also pro Verzeichnungseinheit rund 0,75 EUR.²⁴ Dieser im Vergleich

²³ Diese sind insofern nicht prioritär, als die Urkundenbestände gesondert (und weitgehend schadenfrei) geborgen worden sind und in großen Teilen nicht zeitnah für die Bergungserfassung benötigt werden.

²⁴ Nach dem Ende des Projektes werden diese Berechnungen abschließend konkretisiert. Die Preise variieren auch beim gewählten Modell deutlich je nach Komplexität des zu bearbeitenden Materials.

sehr günstige Preis konnte aufgrund von zwei Vorgaben erreicht werden. Zum Ersten ermöglichten die großen Mengen dem Anbieter das Aufsetzen einer eigenen Infrastruktur – der große und schnelle Durchsatz teilweise ähnlicher Findmittel ermöglichte zudem Skalenerträge. Zum anderen aber wurde bewusst ein gegenüber dem Verfahren mit spezialisierten Retrokonversionsdienstleistern reduzierter Aufwand in Kauf genommen. Für den Kunden ergibt sich damit zwar das Erfordernis von Nacharbeiten (insbesondere im Bereich der Klassifikationen), andererseits aber war der gewünschte hohe Durchsatz bei geringer Beanspruchung der anderweitig gebundenen Mitarbeiter sichergestellt. Dadurch wurde es dem Historischen Archiv letztendlich möglich, bereits im Herbst mit der Beständeidentifikation in den Asylarchiven zu beginnen.

Zwei weitere, spezielle Aufgaben der Konversion stehen parallel zu den letzten „klassischen“ Konversionen an, die Konversion der Findkartei zur Zeitgeschichtlichen Sammlung Bild (ZSB, Best. 7350-7360) und die Digitalisierung des Bibliothekskataloges. Erstere dient der Verbesserung der Erfassungsarbeiten in den Asylarchiven und dient als Basis für eine digitale Verfügbarkeit zumindest des meistgenutzten (und vollständig erhaltenen) Bildbestandes des Historischen Archivs. Zweitere wird einerseits den Nachweis und damit die Benutzbarkeit der umfangreichen Fachbibliothek des Kölner Stadtarchivs mit ihren teilweise singulären Beständen zur Kölner und Rheini-

schen Geschichte verbessern und andererseits einfach den teilweise verlorenen analogen Katalog sichern.²⁵

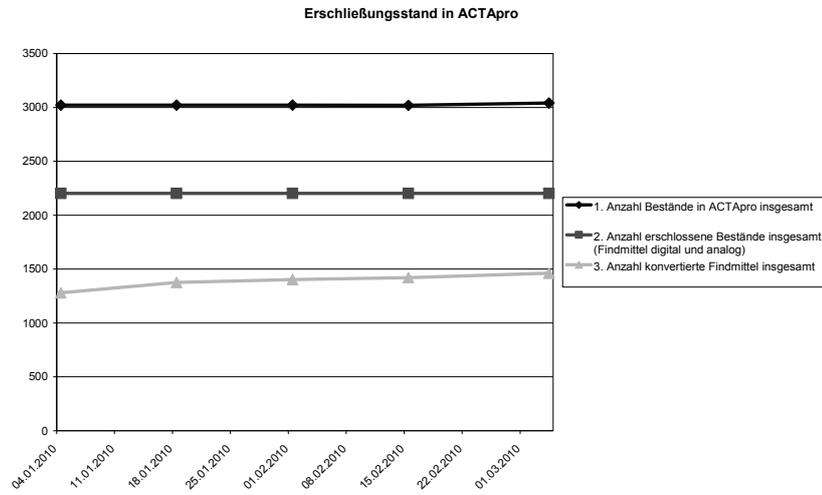
Die Lieferung der fertigen Dateien erfolgte auf Datenträger bzw. per FTP-Import. Eine rudimentäre Erstkontrolle wurde zunächst durch den Hersteller der Archivsoftware durchgeführt, seit einigen Monaten aber im Archiv. Es hatte sich schnell herausgestellt, dass der damit befasste Mitarbeiter im Archiv mehrere Korrektur- und Umarbeitungsarbeitsschritte in einem Durchgang durchführen und damit effizienter arbeiten kann. Insgesamt ist allerdings festzuhalten, dass diese Nacharbeiten selbst bei gut abgelieferten Findmitteln beträchtliche Arbeitszeiten bindet – selbst wenn keine inhaltlichen Verbesserungen oder Kontrollen stattfinden (siehe Tab.1, nächste Seite).

Flankiert wurden diese großen Konversionsmaßnahmen mit Aufbereitungsarbeiten innerhalb des Archivs. Hier wurden vor allem spezielle Findmittel in die Archivsoftware übernommen, ebenso wie rudimentäre Erfassungen – etwa Inhaltszusammenfassungen für Sammlungen und Nachlässe (siehe Tab. 2, nächste Seite).²⁶

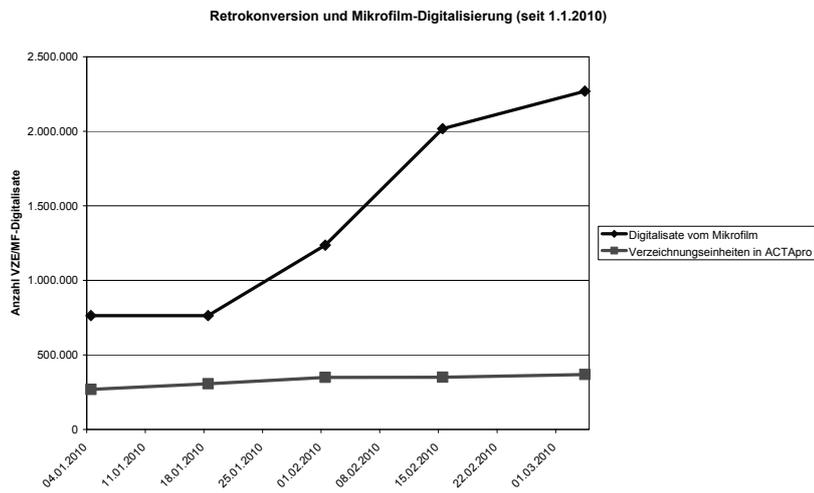
Die vollständige Erschließung der Bestände (zumindest soweit Findmittel vorhanden sind) dient weiterhin als Basis für die Schadenskatastrierung, die Priorisierung für Restaurierungsmaßnahmen

²⁵ Die Zweitschrift des alphabetischen Kataloges befand sich im Keller des eingestürzten Magazinbaus, die Erstschrift konnte in der Zwischenzeit (fast) vollständig wieder zusammengetragen werden.

²⁶ Dies schließt auch die in alten Datenbanken, Excel- und Textdateien vorhandenen Erschließungsarbeiten ein, etwa die 7.307 in einem obsoleten Faust-System erfassten Einheiten des Best. 1671 (René König).



Tab.1: Fortschritt der Retrokonversions- und Digitalisierungsarbeiten seit dem 1.1.2010



Tab.2: Fortschritt der Bestandserschließung und Retrokonversion

und schließlich für die Steuerung, Durchführung und Dokumentation der Restaurierungsmaßnahmen selbst. Hierfür wird in der ersten Jahreshälfte 2010 ein Modul der Archivsoftware beauftragt, das Massenrestaurierungen und deren Dokumentation stückbezogen unterstützen kann.

Nach Ihrer Übernahme in die Archivsoftware werden die Daten aus dem Retrokonversionsprojekt nach und nach im Standardformat SAFT-XML ausgegeben und im Portal <<http://www.archive.nrw.de>> eingestellt. Dabei wird gleichzeitig ein „Fundstatus“ mit ausgegeben, d.h. für alle Einheiten wird angegeben, ob das Stück schon in einem der Asylarchive identifiziert wurde. Standardmäßig wird zunächst von einem Verlust am 3. März ausgegangen. Sobald das entsprechende Archivale aufgefunden ist, wird dieses „Tag“ gelöscht. In der Archivsoftware ist daher ein tagesaktueller Stand über die aufgefundenen Stücke möglich, im nordrhein-westfälischen Archivportal wird die Aktualisierung voraussichtlich im Dreimonatsrhythmus vorgenommen. Gegenwärtig sind etwa 50.000 VZE in www.archive.nrw.de nachgewiesen, Tendenz steigend.

Weiterhin werden – zukünftig automatisiert – freigegebene Erschließungsinformationen aus der Archivsoftware an den Auftritt des Archivs unter <<http://www.historischesarchivkoeln.de>> weitergegeben.²⁷ Sie dienen hier nicht zuletzt als strukturelle Grundlage für

²⁷ Vgl. zur Digitalisierungsstrategie des Historischen Archivs jetzt auch Andreas Berger, Digitalisierung – Zukunft des Archivs?, in: Bettina Schmidt-Czaia/Ulrich S. Soénius (Hrsg.), Gedächtnisort. Das Historische Archiv der Stadt Köln, Köln u.a. 2010, S. 84-95.

die Einbringung der Digitalisate, die zukünftig nicht nur von Nutzern, sondern auch vom Archiv im Rahmen der umfassenden Mikrofilm-digitalisierung eingestellt werden.²⁸ Diese wird mit Mitteln des Landes NRW und der Deutschen Forschungsgemeinschaft vorangetrieben, um Öffentlichkeit und Forschung wenigstens die Benutzung der sicherungsverfilmten Bestände zeitnah zu ermöglichen – die Benutzung der Originale ist bekanntlich auf Jahre hinaus unmöglich. Damit ist die Retrokonversion aller analogen Findhilfsmittel eine unabdingbare Voraussetzung für jede Benutzung.

Aber mehr noch – auf der Basis der Technik des Auftritts <<http://www.historischesarchivkoeln.de>> werden zukünftig auch Lesesaalbetrieb, Verwertung und Reproduktionserstellung gesteuert. Sowohl der Betrieb des *analogen* Lesesaals (zunächst als Provisorium im Rahmen des Wiederaufbaus der Abteilung für Restaurierung und Digitalisierung) als auch der *digitale* Lesesaal (im zukünftigen Haupthaus am Kölner Heumarkt) sowie schließlich der *virtuelle* Lesesaal im Internet nutzen aber nicht nur diese Technik, sondern bedürfen auch eines möglichst vollständigen Bestands an Erfassungsinformationen. Nutzer – einerlei ob sie an einer Nutzung von Archivgutdigitalisaten oder an den Originalen interessiert sind – erwarten heute eine weitgehende Digitalisierung der Findmittelinformationen und einen leichten, schnellen, zeit- und ortsungebundenen Zugriff auf diese.

²⁸ Stand Anf. März sind mehr als 2 Mio. *images* erzeugt worden, die nun hinsichtlich der Metadaten aufgearbeitet und bis zum Sommer in das „Digitale Historische Archiv“ hochgeladen werden.

4. Fazit

Vorbereitung und Durchführung der Retrokonversionsmaßnahmen am Kölner Stadtarchiv machten deutlich, dass auch und gerade nach der Katastrophe die allgemein erkannten Gründe für eine Digitalisierung der Findmittel weitergelten. Gerade nach dem Einsturz verlangen Nutzer – und besonders diejenigen mit akademischem Hintergrund – nach Informationen über die Bestände. Neu hinzugekommen ist als Schwerpunkt die Frage nach Erhaltungszustand und Benutzbarkeit von Beständen oder einzelnen Archivalien. Da eine Autopsie des Kölner Archivguts in vielen Fällen auf absehbare Zeit unmöglich sein wird, ist die Gefahr eines Vergessens oder Übersehens der nicht online recherchierbaren Bestände ungleich größer. Das Szenario einer digitalen Amnesie im Sinne eines Verlustes des Wissens über Kölner Bestände war nie realer – und eine umfassende Präsenz von digitalen Informationen zu Beständen und einzelnen Objekten damit nie notwendiger als jetzt.

Gleichzeitig wird durch die auf Jahre hinaus unmögliche reale Beständezusammenführung eine virtuelle Übersicht über das Gesamtarchiv das wichtigste Arbeitsinstrument von Archivaren und Restauratoren sein. Die Bearbeitung von Rechercheanfragen aus Wissenschaft, Bürgerschaft und Verwaltung, die Weiterentwicklung von Übernahmestrategien und Bewertungsmodellen, die Abschätzung von Raumbedarfen oder die Suche nach Ersatzüberlieferungen für verlorene Bestände oder Bestandsteile – all dies ist ohne digitale Findmittel nicht zu leisten. Die Retrokonversion des analogen Findmittelbestandes ist damit ein unverzichtbarer Schritt hin zu effektiver archivischer Arbeit auch in Zukunft.

In der Katastrophe von Köln wurde allerdings deutlich, dass die Findmittelkonversion mehr ist als nur eine Reaktion auf ein geändertes Nutzerverhalten oder ein Mittel zur Effizienzsteigerung archivischer Arbeitsprozesse. Richtig verstanden und umgesetzt ist sie als Teilaufgabe der Erschließung ein notwendiger, vielleicht entscheidender Aspekt jeder vorbeugenden Bestandserhaltung.²⁹

Die Bewältigung der Kölner Katastrophe bedurfte von Anfang an digitaler Erschließungsinformationen. Diese wären schon bei der Bergung aus den zunächst überlebenden Gebäudeteilen hilfreich gewesen. Als eigentlich unverzichtbar stellten sie sich bei den Überlegungen zur Kalkulation von Schadenssummen und zur Einsturzgeometrie heraus. Vollständig digital verzeichnete Bestände hätten möglicherweise zu einem anderen, letztlich einstufigen Verfahren bei der Erstversorgung von Archivgut führen können und damit viel Arbeit ersparen können, die nun in den Asylarchiven dezentral zu verrichten ist.³⁰

Die digitalisierten Findmittel dienen in vielerlei Hinsicht als Basis für die weiteren Arbeiten zum Wiederaufbau des Kölner Stadtarchivs. Sie sind die Grundlage für die Beständeidentifikation, und

²⁹ Das Folgende zusammengefasst auch schon in: Wilfried Reininghaus/Andreas Pilger (Hrsg.), *Lehren aus Köln. Dokumentation zur Expertenanhörung „Der Kölner Archiveinsturz und die Konsequenzen“*, Düsseldorf 2009 (Veröffentlichungen des Landesarchivs NRW 25), S. 34-36.

³⁰ Fairerweise ist hierbei allerdings zu bedenken, dass es die Kombination aus beträchtlichen Verzeichnungsrückständen mit der nur analogen Verfügbarkeit der Findmittel war, die zu dem gewählten Verfahren bei Erstversorgung und Bestandszusammenführung geführt hat.

ebenso für die Planung, Priorisierung und Steuerung von Restaurierungsmaßnahmen. Sie dienen auch als Grundraster für die Erstellung und die Einordnung der digitalisierten Schutzmedien, um möglichst zeitnah der Forschung einen komfortablen Zugang zu allen Beständen zu gewähren, die verfilmt oder digitalisiert worden sind. Schließlich sind sie die Basisinformation für Nutzer und Mitarbeiter, um nachzuhalten, welche Stücke erhalten und welche verloren sind.

Vor diesen Erfahrungen ist es erstaunlich, dass Retrokonversion und die Sicherstellung der erforderlichen Erschließungstiefe auf der auf dem Deutschen Archivtag in Regensburg durch den Verband deutscher Archivare verabschiedeten *Kölner Erklärung* fehlen.³¹ Hätten Kölner Kollegen bereits in der Entstehungsphase an dem Dokument mitgewirkt, wäre sicherlich die Schwerpunktsetzung eine deutlich andere geworden. Denn in der Katastrophe zeigte sich deutlich, welche Fragen für Bestandserhaltung und Notfallprävention tatsächlich entscheidend sind. Neben der zu recht angemahnten konservatorisch adäquaten Verpackung ist dies in allererster Linie die Forderung nach einer vollständigen und umfassenden digitalen Erschließung der Bestände.

Würden Kölner noch einmal eine Kölner Erklärung formulieren, so enthielte diese folgende fünf Forderungen:

1. Schaffung von *lokalen* Verbänden einschließlich intensivem fachlichen Austausch zur gemeinsamen Arbeit im Notfall: Die ge-

³¹ Resolution der Mitgliederversammlung des VdA – Verband deutscher Archivarinnen und Archive e.V. am 24. September 2009 in Regensburg – Kölner Erklärung zur Sicherung und Erhaltung der Archivbestände, in: *Der Archivar* 62, 2009, S. 453.

naue Kenntnis von Haus und Beständen durch die Kölner Kolleginnen und Kollegen hat sich in unserem Fall extrem vorteilhaft ausgewirkt.³²

2. Enge und regelmäßige Zusammenarbeit mit den *Einsatzkräften vor Ort*, so dass diese Liegenschaft und Kulturgut aus eigener Anschauung kennen und nicht allein auf abstrakte Angaben in einem Online-Register angewiesen sind.³³
3. Bereitstellung von ausreichenden Mitteln für adäquate, konservatorisch einwandfreie, auch *bei Feuchte farbbeständige* und schützende Verpackungsmaterialien.³⁴
4. Bereitstellung von Mitteln für die Erstellung und Erhaltung von *Sicherungsmedien*. Hierbei treten die kostentechnischen Vorteile für Mikroformen zunehmend zurück; aus Nutzersicht ist ohnehin Digitalisaten der Vorzug zu geben.³⁵

³² „Regionenübergreifende Notfallverbände“, wie unter Grundsatz 4 der Kölner Erklärung gefordert, entfalten dagegen ihre Wirksamkeit erst mit zeitlicher Verzögerung. Damit sollten sie erst nach Eintritt der Konsolidierungsphase im Gefolge eines Notfalls bedarfsgerecht verhandelt werden.

³³ Der Eintrag im Notfallregister für Archive (NORA) beim Bundesarchiv kann eine sinnvolle Zusatzmaßnahme sein.

³⁴ Nach den Kölner Erfahrungen erfüllen selbst DIN-zertifizierte Verpackungsmaterialien die Grundanforderung nach Farbechtheit nicht immer.

³⁵ Die Verpflichtung der Archive zur Langzeiterhaltung digitalen Archivgutes verlangt ohnehin nach Einrichtung und Betrieb eines sicheren elektronischen Archivs, in das mit geringem Zusatzaufwand auch digitale Schutzmedien aufgenommen werden können. Vor diesem Hintergrund ist die einseitige Konzentration auf den Mikrofilm als Mittel zur Informationssicherung (so noch Kölner Erklärung, Grundsatz 5) weder aus Wirtschaftlichkeits- noch aus Nachhaltigkeitsgründen

5. Und eben: Bereitstellung von Mitteln und Personalressourcen, um alles Archivgut zeitnah nach der Ankunft im Archiv inhaltlich und formal ausreichend und digital zu *erschließen* und alle analogen Findmittel sach- und fachgerecht zu *retrokonvertieren*.³⁶

zeitgemäß. Für die digitalen Findmittel ist die o.a. Verpflichtung zur Erhaltung der Digitalisate ohnehin schon akzeptiert.

³⁶ Zudem ist die Findmittelkonversion unverzichtbare Grundlage für ein auch in der Katastrophe sinnvolles Archivkataster, wie es in der Kölner Erklärung, Grundsatz 2, gefordert wird.

Retrokonversion von bibliothekarisch erschlossenen Handschriftenkatalogen im Archiv am Beispiel des Bestandes J 1 im Hauptstaatsarchiv Stuttgart¹

von Jeannette Godau

1. Einleitung

1.1 Die Handschriften der Sammlung J 1 und ihre bibliothekarische Katalogisierung

Der Begriff „Handschrift“ bezeichnet in Bibliotheken ein in sich abgeschlossenes, ungedrucktes Buch und damit eine von anderen Beständen abgrenzbare Gruppe von Bibliotheksgut.² Im archivischen Bereich ist der Terminus nicht eindeutig bzw. wenig geeignet, um eine definitorische Abgrenzung einer bestimmten Beständegruppe

¹ Bei vorliegendem Beitrag handelt es sich um die überarbeitete und gekürzte Fassung der gleichnamigen Transferarbeit der Autorin, welche im Rahmen der Laufbahnprüfung für den höheren Archivdienst an der Archivschule Marburg (42. Wissenschaftlicher Kurs) am 1. April 2009 eingereicht wurde. Betreuer der Arbeit waren Dr. Alexandra Lutz (Archivschule Marburg) und Dr. Peter Bohl (Hauptstaatsarchiv Stuttgart). Neben den Gutachtern danke ich Dr. Regina Keyler vom Hauptstaatsarchiv Stuttgart für wertvolle Hinweise. Der Beitrag geht nicht auf Strategien zur Index-Retrokonversion ein. Diese können im betreffenden Kapitel der online-Fassung der Transferarbeit nachgelesen werden, vgl. <http://www.landesarchiv-bw.de/sixcms/media.php/120/49559/Retrokonversion_Godau.pdf>. Alle im Beitrag erwähnten Webseiten wurden am 19. Januar 2010 aufgerufen.

² Vgl. Michael Klein, Die Handschriften der Sammlung J 1 im Hauptstaatsarchiv Stuttgart, Wiesbaden 1980 (Die Handschriften der Staatsarchive in Baden-Württemberg 1), S. 11 mit weiteren Literaturverweisen.

von anderen Materialien zu erzielen, da diese Begriffsbestimmung kein Alleinstellungsmerkmal beinhaltet. Die Provenienz (Herkunft) ist das vorherrschende Kriterium für die Abgrenzung von archivischen Beständen. Doch in Form von Sammlungen verwahren auch Archive Unterlagen, die den bibliothekarischen Terminus „Handschriften“ rechtfertigen. Bei dem Bestand J 1 im Hauptstaatsarchiv Stuttgart handelt es sich um eine landesgeschichtliche Handschriftensammlung, die Bestandteile unterschiedlichster Provenienz in sich vereint.

Der Bestand ist über einen Zeitraum von fast 300 Jahren gewachsen, über 190 Vorbesitzer lassen sich feststellen.³ Er umfasst vornehmlich „Württembergica“ und bildete ursprünglich eine „Handbibliothek“ für Archivare.⁴ Entstanden ist die Sammlung vornehmlich im 17. Jahrhundert, als viele chronikalische Berichte und Dokumente zur Geschichte Württembergs zusammengetragen wurden.⁵

Die älteste Handschrift reicht zurück in das 15. Jahrhundert, die jüngsten sind Abschriften aus dem 20. Jahrhundert.⁶ Es handelt sich

³ Vgl. ebd., S. 32.

⁴ Vgl. zum Folgenden ebd., S. 21f., 28, 57.

⁵ Sammelschwerpunkt war die Geschichte des württembergischen Hauses und die südwestdeutsche Adelsgeschichte. Formal umfasst die Sammlung Chroniken und Landesbeschreibungen, Landbücher sowie Handschriften mit Wappen- oder Siegelabbildungen, Urkundenabschriften, juristische Abhandlungen, Tagebücher und Briefsammlungen, Komödien und Singspiele.

⁶ Die älteste Handschrift des Bestandes ist das Stuttgarter Wappenbuch, J 1 Nr. 289, das zwischen 1439 und 1450 datiert. Laut Michael Klein, ebd., S. 60, erfolgt das Zitat einer Handschriftensignatur aus J 1 folgendermaßen: „HStAst J 1 Nr. 48g“. Innerhalb des Kataloges verweist er aber nicht mit „Nr.“, sondern mit „Hs“, z.B. J 1 Nr. 155a „gleichen Codex wie Hs 48n“. Im vorliegenden Beitrag

größtenteils um nachmittelalterliche Handschriften, der Buchdruck hatte sich bereits durchgesetzt, als sie entstanden.⁷ Einige der in den Handschriften überlieferten Texte sind Vorarbeiten zu Drucken, andere haben vornehmlich persönlichen Charakter, weshalb sie nicht dem teuren Druckverfahren unterzogen wurden, wie Tagebücher, Stammbücher, Briefsammlungen und Vorlesungsmanuskripte. Oder die nachmittelalterlichen Handschriften sind literarische Texte, die nur einem kleinen Teilnehmerkreis gewidmet waren, wie Predigt-handschriften und Liederbücher. Weitere Handschriften, wie die Wappenbücher, sind künstlerisch ausgeschmückt, was im Druck schwer wiedergegeben werden konnte. Zudem ist eine Vielzahl der zum Druck vorgesehenen Handschriften ungedruckt geblieben, weil ihre Verbreitung untersagt wurde.⁸

Michael Klein definiert nach theoretischen Vorüberlegungen die in der Sammlung J 1 befindlichen Handschriften als ungedruckte literarische Erzeugnisse, die durch Autoren individuell geprägt sind.⁹ Diese Definition, so Klein, beschreibt auch die Texte, die in den Handschriftenabteilungen von Bibliotheken aufbewahrt werden. Einzelne Handschriften wurden in der Vergangenheit aus der Sammlung

wird eine gekürzte Form der von ihm vorgeschlagenen Zitierweise verwendet: „J 1 Nr. 48g“.

⁷ Vgl. zum Folgenden Klein, Handschriften (wie Anm. 2), S. 42f.

⁸ Wenn z.B. eine Arbeit zur württembergischen Ortsgeschichte ungeklärte Besitzverhältnisse offenbarte, konnte sie mit Zensur belegt und die Handschrift dem Archiv einverleibt werden. Vgl. ebd., S. 52.

⁹ Vgl. ebd., S. 12.

herausgelöst.¹⁰ Auch wenn viele Nutzer diese Art von Schriftgut eher in Bibliotheken vermuten, ist die Sammlung im Archiv verblieben und nicht etwa als Ganzes an eine Bibliothek abgegeben worden. Dafür spricht Klein zufolge nicht nur der Entstehungszusammenhang der Sammlung – sie ist im Archiv entstanden und gewachsen – sondern auch die enge Verzahnung mit weiteren im Archiv verwahrten Beständen an genuinem Geschäftsschriftgut.¹¹

Doch der enge Zusammenhang des Bestandes mit der Überlieferung in der Württembergischen Landesbibliothek veranlasste die Archivverwaltung dazu, eine Anregung der WLB aufzugreifen und J 1 nach dem Vorbild der Handschriftenkataloge der WLB zu erschließen.¹²

¹⁰ Abgaben an die Württembergische Landesbibliothek (WLB), später auch im Rahmen eines Beständeaustauschs, erfolgten seit der Gründung der Bibliothek 1765 wiederholt bis zum Jahr 1924. Vgl. Günther Haselier, Vorwort, in: Klein, *Handschriften* (wie Anm. 2), S. 9 und die Einführung von Klein, S. 29 mit Anm. 71 und S. 30f. Die ursprünglich für die abgegebenen Handschriften vergebenen Signaturen sind im vorliegenden Katalog unter Angabe der neuen Signaturen in der WLB aufgeführt. Vgl. dazu unten Kapitel 4.4 und Anhang 5 der Transferarbeit in der online-Fassung (URL in Anm. 1).

¹¹ Vgl. Klein, *Handschriften* (wie Anm. 2), S. 13.

¹² Diese basieren auf: Richtlinien Handschriftenkatalogisierung Deutsche Forschungsgemeinschaft. Unterausschuss für Handschriftenkatalogisierung, Bonn Bad-Godesberg 1973. Vgl. Johanne Autenrieth, *Die Handschriften der ehemaligen Hofbibliothek Stuttgart Bd. 3, Codices iuridici et politici* (HB VI 1-139). *Patres* (HB VII 1-71), Wiesbaden 1963 (*Die Handschriften der Württembergischen Landesbibliothek Stuttgart Zweite Reihe*), das Beschreibungsschema auf S. XV-XVIII.

1.2 Problemlage und Zielstellung

Durch die Nutzung der Informationstechnologie ergeben sich neue Möglichkeiten, um Archivgut zu verbreiten und zugänglich zu machen. Gerade um die Assoziation von Handschriften mit Bibliotheken zu umgehen, ist eine Präsentation von Erschließungsergebnissen im Internet ratsam. Daher soll nachfolgend ein Konzept für die Retrokonversion des Katalogs J 1 erarbeitet werden.

Der Begriff Retrokonversion bezeichnet die Digitalisierung von Findmitteln *ohne* eine inhaltliche Veränderung, Anreicherung oder Bearbeitung, ohne Nachbesserungen.¹³ Die Voraussetzungen und Vorzüge, aber möglicherweise auch Einschränkungen, die dieses Vorgehen unter Verwendung archivischer Software für die Sammlung J 1 impliziert, werden im Folgenden erörtert.

Dafür sind zunächst die nach bibliothekarischen Regeln erfolgte Tiefenerschließung der Handschriften und deren Strukturierung nachzuvollziehen. Anschließend wird mit der *Manuscripta mediaevalia* ein bibliothekarisches Retrokonversionsprojekt für Hand-

¹³ Vgl. Frank M. Bischoff, DFG-Förderprogramm zur Retrokonversion von Findbüchern, in: Forum – Newsletter der Archivschule Marburg, 2007, Heft 28, S. 5-14, hier S. 6, online: <<http://www.archivschule.de/uploads/Publikation/forum28.pdf>>. Der Begriff der Konversion stammt aus dem bibliothekarischen Bereich und bezog sich in der Anfangszeit vor allem auf die Digitalisierung von Zettelkatalogen in Bibliotheken. Nach und nach wurde er ausgeweitet auf alle analogen Kataloge und – im archivischen Bereich – Findbücher. Vgl. Sigrid Schieber, Die Retrokonversion archivischer Findmittel – mit Hilfe der DFG Findmittel digitalisieren und online anbieten, in: Archivpflege in Westfalen-Lippe 68, 2008, S. 28-30.

schriftenkataloge vorgestellt. Nachfolgend sind mögliche Retrokonversionsverfahren für den vorliegenden Katalog gegeneinander abzuwägen. Dabei werden auch Probleme der Überführung analoger Erschließungsdaten in ein digitales Medium erörtert. Im vierten Abschnitt werden konkrete Anforderungen für die Retrokonversion des Katalogs formuliert.

2. Standards und Praxis bei vergleichbaren Projekten

2.1 Verzeichnungsrichtlinien der DFG und der Württembergischen Landesbibliothek

Die Richtlinien Handschriftenkatalogisierung der DFG in der Fassung von 1973, an denen Klein seine Katalogisierung ausgerichtet hat, behandeln „Mittelalterliche Handschriften“ und „Nachmittelalterliche Buchhandschriften“ in zwei getrennten Abschnitten.¹⁴ Die Mehrzahl der Handschriften der Sammlung J 1 entstand nach dem Mittelalter, jedoch hat sich der Bearbeiter für die detailliertere Beschreibung nach den Vorgaben für mittelalterliche Handschriften entschieden und erfasst auch Formalangaben wie Lagen, alte Folierung, fehlende Blätter, Technik bei Stempelbänden und Initien, die nach den Richtlinien für die nachmittelalterlichen Buchhandschriften fakultativ sind. Eine ausführliche Vorstellung der in den Richtlinien verbindlich festgelegten Beschreibungselemente ist an dieser Stelle nicht notwendig, da der Katalog diesen in fast allen Punkten entspricht.¹⁵

¹⁴ Vgl. Richtlinien (wie Anm. 12).

¹⁵ Einen tabellarischen Überblick über die Elemente enthält Anhang 7 der Transferarbeit in der online-Fassung (URL in Anm. 1). Darin werden die von Klein beschriebenen Elemente mit denen der DFG-Richtlinien gemappt und so können

Während die DFG-Richtlinie strukturelle Anforderungen formuliert, dienten die Handschriftenkataloge der WLB als praktisches Vorbild für Kleins Erschließung.¹⁶

Die Erschließungstiefe von Handschriften in Bibliotheken kann mit der von Urkunden in Archiven verglichen werden, wie die detaillierte Erschließungsmaske in scopeArchiv aufzeigt.¹⁷ Die Handschriften in Archiven werden größtenteils nach inhaltlichen Gesichtspunkten ausgewertet, denn die vorliegende Sammlung enthält wertvolle historische Quellen zur Geschichte Württembergs. Nachgefragt werden Handschriften aber auch aus kunsthistorischem oder philologischem Interesse, was in Bibliotheken häufiger vorkommt. Für die Ermittlung von Schreibschulen, Werkstätten oder Buchbindereien ist die kodikologische Beschreibung von Handschriften unentbehrlich.¹⁸

kleinere Abweichungen, die aber nicht von Relevanz für die Retrokonversion sind, nachvollzogen werden.

¹⁶ Dies wird vor allem drucktechnisch durch den Satz deutlich. Z.B. werden im Abschnitt „Beschreibung des Inhalts“ die Hervorhebung des Verfassers durch Kapitälchendruck, die des Titels durch Kapitälchenminuskel sowie die Kennzeichnung von erschlossenen Angaben durch bestimmte Klammerformen und die Auszeichnung wörtlicher Zitate analog zum WLB-Katalog gestaltet. Vgl. Autenrieth, Handschriften (wie Anm. 12), S. XVII und Klein, Handschriften (wie Anm. 2), S. 61.

¹⁷ Vgl. Landesarchiv Baden-Württemberg, Beiblätter für Erschließungsrichtlinien (AZ 7511.5-312/16/O/Fz, Beiblatt Nr. 4 Erschließung von Urkunden mit scopeArchiv, Formular: Urkunden (ID 10028), Autor: Dr. Th. Fritz (Abt. 1), Stand: 30.08.2006 (unveröffentlicht).

¹⁸ Vgl. Otto Mazal, Zur Praxis des Handschriftenbearbeiters. Mit einem Kapitel zur Textherstellung, Wiesbaden 1987 (Elemente des Buch- und Bibliothekswesens 11), S. 4ff. Vgl. auch Hermann Knaus, Studien zur Handschriftenkunde. Ausgewählte Aufsätze, München u.a. 1992.

Daher werden die Handschriften in Bibliotheken ebenso ausführlich erschlossen wie in Archiven die Urkunden. Die Katalogisierung berücksichtigt gleichermaßen inhaltliche wie formale Merkmale.

2.2 Retrokonversion und Präsentation von Handschriftenkatalogen in der Datenbank „Manuscripta mediaevalia“

Die DFG fördert seit den 1960er Jahren die wissenschaftliche Erschließung von Handschriften nach inhaltlichen und äußeren Gesichtspunkten.¹⁹ Ursprünglich wurde aus den mit Förderung der DFG entstandenen Katalogen der „Gesamtindex mittelalterlicher Handschriftenkataloge“ erstellt, welcher alle Initien- und Sachregistereinträge zusammenführte.²⁰ Dieser war seit 1993 in Form der DBI-LINK-Datenbank als Auskunftssystem für den „Gesamtindex“ recherchierbar.²¹

¹⁹ Grundlage sind die Richtlinien Handschriftenkatalogisierung der DFG (wie Anm. 12). Vgl. Robert Giel, „Manuscripta mediaevalia“. Möglichkeiten und Grenzen eines Auskunftssystems für die mediävistische Forschung, in: Hannelore Benkert/Burkard Rosenberger/Wolfgang Dittrich (Hrsg.), Die Bibliothek zwischen Autor und Leser. 92. Deutscher Bibliothekartag in Augsburg 2002, Frankfurt/M. 2003 (Zeitschrift für Bibliothekswesen und Bibliographie, Sonderheft 84), S. 223-231, hier S. 224.

²⁰ Vgl. Robert Giel, „Manuscripta mediaevalia“. Ein Internet-Angebot für die Handschriftenforschung, in: Mittellateinisches Jahrbuch 38, 2003, S. 291-293, hier S. 291.

²¹ Vgl. Bernd Michael, Elektronische Wege zu den mittelalterlichen Handschriften in Deutschland. Der „Gesamtindex mittelalterlicher Handschriftenkataloge“ und die „DBI-LINK“-Datenbank „Handschriften des Mittelalters“, in: Rita Schlusemann (Hrsg.), Sources for the History of Medieval Books and Libraries, Groningen 1999, S. 353-363, hier S. 357.

Seit 1999 wurden diese „Findehilfen“²² zu den gedruckten Handschriftenkatalogen abgeschlossen und die DFG förderte den Aufbau des Webportals *Manuscripta mediaevalia*.²³ Die Registerinformationen aus dem Gesamtindex zu etwa 30.000 Handschriftensignaturen gingen in die Datenbank der *Manuscripta mediaevalia* ein und werden fortlaufend mit Registerdaten aus weiteren Katalogen oder durch eine direkte Katalogisierung in der Datenbank ergänzt.²⁴ Die *Manuscripta mediaevalia* wie ihr Vorgängerprojekt, der Gesamtindex, verfolgen das Konzept der Tiefenerschließung der Handschriften, um den Bedürfnissen der Handschriftenforschung adäquat entgegen zu kommen.²⁵

²² So die Bezeichnung bei Bernd Michael, vgl. ebd., S. 355.

²³ <<http://www.manuscripta-mediaevalia.de>>

²⁴ Vgl. zum Folgenden Robert Giel, „Manuscripta mediaevalia“. Handschriften aus deutschen Bibliotheken im Internet, in: Gazette du livre médiéval 39, 2001, S. 34-40.

²⁵ Vgl. Christian Bracht, „Manuscripta mediaevalia“. Ergebnisse der Handschriftenkatalogisierung im Internet, in: Gazette du livre médiéval 47, 2005, S. 39-42, hier S. 39. Dass die der Retrokonversion zugrunde liegenden Kataloge mit der Berücksichtigung kodikologischer Informationen, dem Nachweis der einzelnen Texte in einem Kodex und mit der überlieferungsgeschichtlichen Einordnung von Handschriften die im Wissenschaftsalltag entstehenden Fragen inhaltlich angemessen beantworten können, bestätigt Hermann Leskien. Als Vorsitzender der Arbeitsgruppe des Bibliotheksausschusses der DFG zur Erarbeitung eines datenbankgestützten Verfahrens der Handschriftenkatalogisierung in der Nachfolge des „Gesamtindex“ überprüfte er zunächst die Forschungstauglichkeit der nach den DFG-Richtlinien erstellten Kataloge. Vgl. Hermann Leskien, Handschriftendatenbank – Überlegungen im Rahmen des DFG-Förderprogramms Handschriftenkatalogisierung, in: Sabine Wefers (Hrsg.), Die Herausforderung der Bibliotheken durch elektronische Medien und neue Organisationsformen. 85. Deutscher Bibliothekar-

Die *Manuscripta mediaevalia* vereint inzwischen drei Informationsebenen: erstens die Datenbankrecherche in den Registerinformationen, zweitens einen Zugriff auf die Images der retrokonvertierten Kataloge und drittens einen Zugang zu digitalisierten Handschriften. Derzeit sind hier über 70.500 Handschriftenbeschreibungen aus 290 digitalisierten Handschriftenkatalogbänden und -inventaren verfügbar.²⁶

Bei der Übernahme der Daten aus dem Gesamtindex und deren Ergänzung mit Daten aus den Registern weiterer retrokonvertierter Handschriftenkataloge wurden diese aufgearbeitet und verbessert. Die Kreuzregistereinträge²⁷ wurden bei der Retrokonversion mit halbmaschinellen und manuellen Verfahren Kategorien wie Person, Körperschaft, Ort oder Ikonographie zugewiesen, in denen nun recherchiert werden kann.²⁸ Für jedes der so gebildeten 23 Suchfelder ist außerdem eine Liste der vorhandenen Einträge abrufbar.²⁹ Während der Retrokonversion wurden Begriffe normiert und z.B. Personennamen mit vorhandenen Synonymen oder Verweisformen verknüpft. Mit der Personennamendatei PND steht den Bibliotheken

tag in Göttingen 1995, Frankfurt/M. 1996 (Zeitschrift für Bibliothekswesen und Bibliographie, Sonderheft 63), S. 262-272, hier S. 263.

²⁶ <<http://www.manuscripta-mediaevalia.de/hs/hs-einfuehrung.htm>>. Letzte Aktualisierung: 3.12.2008.

²⁷ Ein Kreuzregister ist eine Zusammenfassung von Sach- und Namenregister in einem einzigen Index.

²⁸ Die Recherche im Gesamtindex erlaubte nur die Suche im Initien- oder alternativ im Kreuzregister. Vgl. Giel, Möglichkeiten (wie Anm. 19), S. 225.

²⁹ Vgl. Bracht, Manuscripta (wie Anm. 25), S. 41.

dabei ein wirkmächtiges Instrumentarium zur Verfügung.³⁰ Über die in Bibliotheken verwendete Software PICA und durch ein Redaktionssystem unter Leitung der *Deutschen Nationalbibliothek* ist den großen Bibliotheken jederzeit das Abrufen und Verknüpfen sowie die Neuanlage von Namen in der Personendatenbank möglich. Da Archive in dieses Redaktionssystem nicht eingebunden sind, können sie zwar vorhandene Namen und die dazugehörige PND-Identnummer über das Internet abrufen, sie haben jedoch keinen schreibenden Zugriff auf die Datei, weshalb deren Verwendung bisher eher mit mehr Arbeitsaufwand als Nutzen für die Archive verbunden war.³¹

Durch die Möglichkeit der Kombination verschiedener Suchabfragen in Bezug auf eine oder mehrere Signaturen ist mit der *Manuscripta mediaevalia* ein mehrschichtiges und differenziertes Recherchewerkzeug entstanden. Aufgrund der umfangreichen Verbesserungen und Normierungstätigkeiten ist die oben gegebene enge Definition von Retrokonversion nicht mehr zutreffend.³² Diese Maßnahmen sind beim Internetnutzer ohne Frage willkommen, eine Aufbereitung der Daten hängt jedoch von dem vorhandenen Budget und von

³⁰ <<http://www.d-nb.de/standardisierung/normdateien/pnd.htm>>. Die Form der Namensansetzungen spielt auch in der Nachlassverzeichnung in Bibliotheken eine große Rolle. Vgl. Gerhard Schmid (Hrsg.), Bestandserschließung im Literaturarchiv. Arbeitsgrundsätze des Goethe- und Schiller-Archivs in Weimar, München u.a. 1996 (Literatur und Archiv 7), bes. S. 79-85.

³¹ Das Landesarchiv Baden-Württemberg führte als Partner im Projekt „Kalliope II“ von Mai 2003 bis Mai 2004 die normierte Ansetzung von Personendaten für alle Bestandsbeschreibungen durch. Projektinformationen unter: <<http://www.landesarchiv-bw.de/web/47340>> und <<http://www.kalliope-portal.de>>.

³² Vgl. die Definition bei Anm. 12.

personellen Ressourcen ab, die für die Retrokonversion des Handschriftenkatalogs J 1 nicht zusätzlich verfügbar sind. Die Nachbearbeitung der Register erfolgte mit finanzieller Unterstützung der DFG. Die *Manuscripta mediaevalia* ist seit 2002 in dauerhafter Trägerschaft einer Zentralredaktion bei der Staatsbibliothek zu Berlin mit Dependancen in der Bayerischen Staatsbibliothek und dem Bildarchiv Foto Marburg, die jeweils Stellen dafür bereithalten. Abgesehen von wenigen Ausnahmen hat die *Manuscripta mediaevalia* den Fokus auf Bibliotheken.³³ Der zeitliche Schwerpunkt der retrokonvertierten Kataloge liegt auf mittelalterlichen Handschriften,³⁴ so dass der von Klein erstellte Katalog mit der Mehrzahl frühneuzeitlicher und späterer Handschriften aus diesem Rahmen herausfällt. Da es sich inhaltlich bei dem Bestand vor allem um Handschriften mit inhaltlichem Schwerpunkt auf „Württembergica“ handelt, ist es von Vorteil, wenn sie auch über das Internetangebot des Landesarchivs recherchierbar sind, unabhängig von einer möglichen späteren Übernahme

³³ Laut Bracht ist es das Ziel „den gesamten Handschriftenbestand in deutschen Bibliotheken nachzuweisen“. Vgl. Bracht, *Manuscripta* (wie Anm. 25), S. 39. Giel verweist auf die Liste der „Bibliotheksorte“, in der nach dort erstellten Katalogen zu recherchieren ist. Vgl. Giel, *Handschriften* (wie Anm. 24), S. 38. – Ausnahmen sind Handschriftkataloge aus dem Stadtarchiv Köln, Spitalarchiv Biberach und Landeshauptarchiv Koblenz. Handschriften aus Archiven in Hannover, Hildesheim und Osnabrück sind durch die Reihe „Mittelalterliche Handschriften in Niedersachsen“ in Katalogen erfasst und in die *Manuscripta Mediaevalia* aufgenommen worden.

³⁴ So ergibt eine Expertensuche in der Datenbank nach Handschriften, die zwischen 1600 und 1800 entstanden sind, nur 2.906 Treffer, nicht einmal 5 % der Gesamtmenge von 70.500 Handschriften (Abfrage am 29.01.2010). Legt man die 123.348 Fundstellen für den Oberbegriff „Handschriften“ zugrunde, sinkt der Anteil weiter.

in die *Manuscripta mediaevalia*. Daher sollen hier weiter Möglichkeiten der Retrokonversion in ihrer engen Definition, das heißt ohne Bearbeitung und Anreicherung, verfolgt werden.

3. Verfahren der Retrokonversion und Probleme der Überführung analoger Medien in eine digitale Umgebung

3.1 OCR oder Image-Retrokonversion?

Bei der Retrokonversion werden manuelle oder maschinelle Verfahren eingesetzt.³⁵ Maschinelle Verfahren sind das Erfassen eines gescannten Textes mit Hilfe von OCR (Optical Character Recognition) oder das Einscannen einer Vorlage ohne Texterkennung und Ablage der Informationen als „Non Coded Information“, d.h. als Bildinformation (Image), deren Inhalt nicht indexiert werden kann und die durch zusätzliche Metadaten erschlossen werden muss.

Die Wahl des Verfahrens hängt von verschiedenen Faktoren ab: der Qualität der Vorlage, vom Strukturierungsgrad der Informationen (als Voraussetzung für eine OCR-Erkennung)³⁶ und nicht zuletzt vom Zweck der Retrokonversion, das heißt vom beabsichtigten Zielformat und welche Recherchemöglichkeiten es offerieren soll.

³⁵ Vgl. Werner Holbach, Retrokonversion und Retrokatalogisierung in den Bayerischen Staatlichen Bibliotheken, in: Bibliotheksforum Bayern 34, 2006, Heft 3, S. 206-228, hier S. 218. Manuelle Verfahren sind das einfache Abschreiben des Textes oder das so genannte „Double Keying“, für das ein Text zwei Mal abgetippt und dann mittels maschineller Verfahren Abweichungen der beiden Versionen ermittelt werden, um so die Fehlerquote rationell zu minimieren.

³⁶ Vgl. ebd., S. 220.

Ziel einer Retrokonversion von archivischen Findbüchern sollte es sein, die vormals analogen Daten in die bereits vorhandene Softwareumgebung von datenbanktechnisch erschlossenen Daten zu integrieren.³⁷ Die DFG-Vorstudie „Retrokonversion archivischer Findmittel“ entschied sich aus nachvollziehbaren Gründen gegen eine Image-Konversion, da eine Recherchierbarkeit in den Daten dann nicht gegeben wäre und Metadaten für Recherche und Verwaltung der Images darüber hinaus trotzdem erfasst werden müssten.³⁸ Im Zuge der Vorstudie wurde ein eigenes Software-Tool „DocWorks-Findbuch“ entwickelt, welches über OCR und Layouterkennung eine (semi-)automatische Konversion maschinenschriftlicher Findmittel erlaubt.³⁹ Für die Retrokonversion des Handschriftenkatalogs ist dieses Tool, das auf die Spezifika archivischer Findbücher ausgerichtet ist, jedoch nicht einsetzbar.

Eine Anforderung für eine möglichst geringe Fehlerquote beim Einsatz von OCR erfüllt der Katalog an sich: er liegt in Druckqualität

³⁷ So auch die Forderung der DFG für die Vorstudie. Vgl. DFG-Vorstudie „Retrokonversion archivischer Findmittel“. Projektbericht. Eingereicht für das Landesarchiv Nordrhein-Westfalen von Prof. Dr. Wilfried Reininghaus, Düsseldorf 2006, hier S. 35. Vgl. <<http://www.archive.nrw.de/LandesarchivNRW/BilderKartenLogosDateien/RetrokonversionProjektbericht2006.pdf>>. Vgl. Ulrich Fischer / Wilfried Reininghaus, DFG-Vorstudie „Retrokonversion archivischer Findmittel“. Die wichtigsten Ergebnisse des Projekts, in: *Der Archivar* 59, 2006, S. 329-333, hier S. 330.

³⁸ Vgl. Bischoff, DFG-Förderprogramm (wie Anm. 13), S. 5.

³⁹ Vgl. DFG-Vorstudie „Retrokonversion archivischer Findmittel“ (wie Anm. 37), S. 9.

vor.⁴⁰ Für das automatische Einlesen von Textbestandteilen in Datenbankfelder ist jedoch eine gleich bleibende Strukturierung der Beschreibungen Voraussetzung, wie sie etwa in Zettelkatalogen von Bibliotheken vorhanden ist. Klein entwirft zwar ein Schema für die Handschriftenbeschreibungen und benennt die Elemente, die in den Beschreibungen auftauchen können.⁴¹ Die Ausführlichkeit der tatsächlichen Beschreibungen richtet sich aber nach dem Alter und der Bedeutung der jeweiligen Handschrift, so dass nicht immer alle Elemente belegt sind. Schon der eigentlich obligatorische Kurztitel ist nicht immer besetzt, denn für Teilbestände wird nur die Signatur (z.B. „78a“, „78b“) angegeben, ohne den Kurztitel der Oberaufnahme (hier: „Johann Gottlieb Breyer“) zu wiederholen. Auch die nachfolgenden Abschnitte wie die „Geschichte der Handschrift“ (Zweiter Absatz in kleinem Schriftgrad) und „Literatur“ (Dritter Absatz in kleinem Schriftgrad) sind nicht immer belegt. Für den effizienten Einsatz von OCR-Software ist damit die Voraussetzung einer eindeutigen und konsequent befolgten Anordnung von Textteilen oder das Wiederkehren von Signalworten (wie „Enthält“) nicht erfüllt.⁴²

⁴⁰ Vgl. Thomas Fricke/Gerald Maier, Automatische Texterkennung bei digitalisiertem Archiv- und Bibliotheksgut, in: Hartmut Weber/Gerald Maier (Hrsg.), Digitale Archive und Bibliotheken. Neue Zugangsmöglichkeiten und Nutzungsqualitäten, Stuttgart/Berlin/Köln 2000 (Werkhefte der Staatlichen Archivverwaltung Baden-Württemberg, Serie A Landesarchivdirektion, Heft 15), S. 201-221, hier S. 208, online: <http://www.landearchiv-bw.de/sixcms/media.php/120/Werkheft_Staatl_Archiv.pdf>

⁴¹ Vgl. Klein, Handschriften (wie Anm. 2), S. 60-62.

⁴² Vgl. Abschlussbericht zum DFG-Projekt „Entwicklung von Werkzeugen zur Retrokonversion archivischer Findmittel“. Landesarchiv NRW, Hauptstaatsarchiv Düsseldorf. 30.06.2004, S. 31.

Schwierigkeiten bei der Zeichenerkennung kann auch ein Wechsel verschiedener Schrifttypen bereiten.⁴³ Klein kennzeichnet die Elemente seiner Beschreibung mit Kursivdruck, mit Kapitälchen- und Kapitälchenminuskel, graphische Zeichen signalisieren die Trennung von Beschreibungselementen und Zeichen für Quadrangel geben den exakten Fundort an (z.B. J 1 Nr. 291). Er verwendet hochgestellte Buchstaben für „recto“ und „verso“, Bruchzahlen sowie diverse Klammerformen. Es ist möglich, OCR-Software auf die Erkennung solcher Formate zu „trainieren“, jedoch scheint bei der Einzigartigkeit des Katalogs⁴⁴ dieses Verfahren wenig wirtschaftlich. Außerdem könnten linguistische Verfahren zur Rechtschreibprüfung nicht zum Einsatz kommen, da der gewinnbringende Einsatz eines Wörterbuchs zur Unterstützung der Texterkennung die Verwendung einer einzigen Sprache voraussetzt, der Katalog aber mindestens drei Sprachen aufweist.⁴⁵

Aus diesen Gründen ist die Anwendung von OCR auf den Beschreibungsteil der Handschriftenkatalogisierung wenig Erfolg versprechend oder wäre mit einem unverhältnismäßig hohen Aufwand für die Nachkorrektur verbunden.

⁴³ Vgl. Fricke/Maier, *Texterkennung* (wie Anm. 40), S. 209.

⁴⁴ Ein weiterer Katalog liegt für das GLA Karlsruhe vor. Vgl. Michael Klein, *Die Handschriften 65/1-1200 im Generallandesarchiv Karlsruhe*, Wiesbaden 1987 (*Die Handschriften der Staatsarchive in Baden-Württemberg* 2).

⁴⁵ Deutsch, Latein, Französisch. Zur Empfehlung des Einsatzes eines OCR-Readers mit hinterlegter und durch den Nutzer erweiterbare Wortliste bei umfangreicheren und längerfristigen Projekten vgl. Abschlussbericht zum DFG-Projekt (wie Anm. 42), S. 25. Zu linguistischen Verfahren vgl. Fricke/Maier, *Texterkennung* (wie Anm. 40), S. 202.

3.2 Image-Retrokonversion oder händische Erfassung?

Da eine OCR-Erkennung aus den oben erwähnten Gründen für den Katalog J 1 ausscheidet, ist nun zu prüfen, ob eine Image-Retrokonversion nach dem Vorbild der *Manuscripta mediaevalia* oder aber eine händische Erfassung der Handschriftenbeschreibungen vorzuziehen ist. Die Transferarbeit von Ulrich Fischer untersucht für die Retrokonversion handschriftlicher Findmittel die Qualität und die Effizienz eines erweiterten Imaging-Verfahrens (mit Anreicherung durch Metadaten) im Vergleich zu einer händischen Eingabe.⁴⁶ Wenngleich handschriftliche Findmittel noch zusätzliche Anforderungen an die Retrokonversion stellen als der hier vorliegende gedruckte Katalog, sind Fischers Erkenntnisse dennoch hierfür relevant. Die Daten wurden direkt in die im Staatsarchiv Münster eingesetzte Archivsoftware VERA eingegeben, für die Präsentation der Images wurde eine zusätzliche Maske eingerichtet, um die Image-Dateien in die Verzeichnungseinheiten einzubinden.⁴⁷ Ein solches Bildbetrachtungsmodul liegt für scopeArchiv bzw. das Präsentationssystem des Landesarchivs Baden-Württemberg OLF 21 bereits vor. Hinsichtlich der Metadaten, die für die Anzeige und Verwaltung der Images erhoben werden müssen, benennt Fischer Felder, die auch für J 1 relevant sind: „Archivsignatur“, „Altsignatur“ und „Laufzeit“.⁴⁸

⁴⁶ Vgl. Ulrich Fischer, Imaging versus händische Eingabe – Eine Wirtschaftlichkeitsanalyse zum Problem der Retrokonversion handschriftlicher Findmittel, Marburg 2005 (Archivschule Marburg, Transferarbeit) unter <<http://www.archive.nrw.de/dok/Transferarbeiten/Fischer/Transfer%20Text.pdf>>

⁴⁷ Vgl. ebd., S. 6.

⁴⁸ Vgl. ebd., S. 8.

Fischer kommt nach Auswertung der Erfassungsgeschwindigkeiten und Bearbeitungszeiten zu dem Schluss: „Je komplexer und umfangreicher die handschriftliche Erschließung eines Bestandes, desto größer der Vorteil der Hybriderfassung“ (d.h. des Imaging-Verfahrens).⁴⁹ Diese Schlussfolgerung zieht er mit Blick auf ein retrokonvertiertes Urkundenbuch, das durch Hervorhebungen, lateinische Buchstabenformen und Verweise gekennzeichnet ist und damit der Strukturierung des vorliegenden Katalogs nahe kommt.

Für eine Image-Retrokonversion hat man sich im Referenzprojekt *Manuscripta mediaevalia* entschieden, um „die ursprünglichen Arbeitsergebnisse der Katalogbearbeiter zu erhalten“, so Robert Giel.⁵⁰ In den Bibliotheken wird meistens das Einstellen von Images mit einer strukturierten Datenerfassung kombiniert, so dass in den alten Katalogaufnahmen geblättert und in den erfassten Daten recherchiert werden kann.⁵¹

Die *Manuscripta mediaevalia* bietet zwei Zugriffsmöglichkeiten auf die Image-Kataloge: entweder über das von der analogen Nutzung inspirierte Blättern in den digitalisierten Katalogseiten oder über eine Datenbankabfrage. Im ersten Fall wird zunächst ein Katalog aus einer Liste ausgewählt. Als Einstieg werden dann die mit Hyperlinks hinterlegten Seitenzahlen des Katalogs eingeblendet. Durch das

⁴⁹ Vgl. ebd., S. 23.

⁵⁰ Giel, Handschriften (wie Anm. 24), S. 36.

⁵¹ Dies trifft auch für die Forschungsdokumentation zu den Handschriften der Bayerischen Staatsbibliothek zu. Vgl. Holbach, Retrokonversion (wie Anm. 35), S. 213. <<http://www.bsb-muenchen.de/Forschungsdokumentation.172.0.html>>

Auswählen und Anklicken einer Seitenzahl wird das Image aufgerufen und kann gelesen werden. Eine Recherche in den Inhalten ist innerhalb dieser PDF-Dateien nicht möglich.

Die zweite Möglichkeit ist die Recherche in der Datenbank. Nach der Eingabe eines Suchbegriffs und der Auswahl eines Treffers aus der Ergebnisliste werden zunächst die Metadaten des Dokuments angezeigt und ein Link „Dokument öffnen“ führt zur dazugehörigen eingescannten Katalogseite. Es wird auf die Seite verlinkt, auf der die Katalogisierung der Handschrift mit Angabe der Signatur beginnt. Wenn eine Handschriftenbeschreibung also mehrere Seiten umfasst, springt der Nutzer nicht sofort auf die Seite, auf der sich der im Index verzeichnete Eintrag befindet.⁵² Es ist keine Highlighting-Funktion implementiert und der indizierte Begriff muss, wie in der analogen Vorlage, im Image selbst gesucht werden. Auch Robert Giel ist sich im Hinblick auf die Retrokonversion der Kataloge als Images für die *Manuscripta mediaevalia* der Grenzen des Verfahrens bewusst. „Was in der Vergangenheit für nicht registerwürdig befunden worden ist, steht also auch heute nicht für eine differenzierte Datenbankabfrage zur Verfügung“.⁵³

⁵² Z.B. „Georg Wolf von Kotzau“ im Bestand der WLB HB VI 8. Eine Suche nach dem Namen verweist auf den Bestand der WLB. Bei Auswahl des Treffers wird dann die Signaturangabe spezifiziert mit „Cons. 57“, was für eine Untergliederung der Handschrift in „Consilia“ steht. Die Handschriftenbeschreibung HB VI 8 beginnt auf S. 10, der Namensnachweis steht auf S. 11. Da die Handschriftenbeschreibung die Einzelabschnitte der Handschrift und deren Gliederung in „Cons.“ berücksichtigt und auf S. 10 mit „Cons. 56“ endet, ist dem Nutzer klar, dass er auf die nächste Imageseite des Katalogs blättern muss.

⁵³ Vgl. Giel, Möglichkeiten (wie Anm. 19), S. 230.

Dieser Einschränkung muss man sich bei der Wahl dieses Verfahrens bewusst bleiben. Dennoch scheint bei allen Vorüberlegungen folgender Schluss nahe zu liegen: Die Erschließung des Bestandes J 1 erfolgte nach bibliothekarischen Richtlinien, so dass auch für die Retrokonversion das bibliothekarische Vorgehen als Orientierung dienen kann. Analog zum Verfahren der *Manuscripta mediaevalia* empfiehlt sich für den Beschreibungsteil des Katalogs eine Image-Retrokonversion.⁵⁴

Sollte die Digitalisierung in Eigenleistung erfolgen oder an einen Dienstleister vergeben werden? Nach einer detaillierten Kostenaufstellung favorisiert Ulrich Fischer aus finanziellen und ressourcetechnischen Erwägungen eine Retrokonversion in Eigenleistung.⁵⁵ Bärbel Förster ermittelte anhand von Fallbeispielen bei der Retrokonversion archivischer Findmittel im Schweizerischen Bundesarchiv, dass bei der externen Erfassung der interne Aufwand der Kontrolle und Nachbearbeitung im Durchschnitt höher lag als bei der manuellen internen Erfassung. Letztere sei daher die beste Variante hinsichtlich ihres Kosten-Nutzen-Verhältnisses.⁵⁶

⁵⁴ Mit dem umfangreichen Index liegt ein potentiell wirkungsmächtiges Rechercheinstrument vor, das mit unterschiedlich hohem Aufwand (und damit Nutzen) retrokonvertiert werden kann. Vgl. dazu Kapitel 3.2 der Transferarbeit in der online-Fassung (URL in Anm. 1).

⁵⁵ Vgl. Fischer, Imaging (wie Anm. 46), S. 31.

⁵⁶ Vgl. Bärbel Förster, Retrokonversion archivischer Findmittel im Schweizerischen Bundesarchiv. Vortrag auf der Tagung des Nordrhein-Westfälischen Hauptstaatsarchivs zum DFG-Projekt „Entwicklung von Werkzeugen zur Retrokonversion archivischer Findmittel“. Düsseldorf, 22.-23.9.2003, hier S. 3, unter <<http://www.archive.nrw.de/dok/tagung-retro/08-Foerster.pdf>>

4. Retrokonversion von J 1 – Probleme und Lösungsansätze

4.1 Allgemeine Grundsätze für die Digitalisierung

Bevor auf die Retrokonversion der Handschriftenbeschreibungen eingegangen wird, ist kurz das Verfahren für die Digitalisierung der umfangreichen und in mehrere Abschnitte untergliederten „Einführung“ in Kleins Katalog zu erörtern.⁵⁷ Für diese schien mit Blick auf die Felddokumentation für Bestandsbeschreibungen in scopeArchiv zunächst eine händische Erfassung naheliegend.⁵⁸ Kleins Einführung ist in drei größere Abschnitte untergliedert, die den scopeArchiv-Datenfeldern für „Überlieferungsgeschichte“, „Inhalt / Bewertung / Bearbeitungshinweise“ und „Benutzungshinweise“ entsprechen. Darüber hinaus gibt es jedoch analoge Inhalte, denen keine Datenfelder zugeordnet werden können. Zunächst beschreibt Klein im Abschnitt I die „Handschriften im Hauptstaatsarchiv Stuttgart“, aufgeteilt in allgemeine Vorüberlegungen zur archivischen Verzeichnung von Handschriften, gefolgt von Ausführungen über Handschriften außerhalb der Sammlung J 1. Diese Informationen gehören nicht unmittelbar zum beschriebenen Bestand und damit auch nicht zur „Überlieferungsgeschichte“, weshalb kein adäquates Feld dafür in der Erschließungsmaske vorgesehen ist.

⁵⁷ Sie umfasst 55 Katalogseiten. Vgl. Klein, Handschriften (wie Anm. 2), S. 11-65.

⁵⁸ Vgl. Landesarchiv Baden-Württemberg, Beiblätter für Erschließungsrichtlinien (AZ 7511.5-312/16/O/Fz, Beiblatt Nr. 11 Erfassung von (Teil-)Bestandsbeschreibungen mit scopeArchiv, Formular: Bestände und Teilbestände (ID 10025), Autor: Dr. Th. Fritz (Abt. 1), Stand: 07.12.2006 (unveröffentlicht).

Der Abschnitt III über „Aufbau der Handschriftenbeschreibungen und der Indizes“ (S. 60-65) könnte dem Feld „Benutzungshinweise“ eindeutig zugeordnet werden. Doch außer der schwierigen Feldzuordnung im Falle des Abschnittes I sprechen weitere Überlegungen gegen eine feldgetreue Erfassung der Einleitung. Nicht nur ihr großer Umfang würde die Übersichtlichkeit beim Lesen stören. Die Einführung hat zahlreiche Fußnoten, die nur mit großem Aufwand in Datenbanktechnologie umzusetzen sind. Außerdem wären die in der Literatur eingeführten Verweise auf die Seitenzahlen des Katalogs durch die Übernahme in neue Felder nicht mehr gegeben. Deshalb bietet sich eine Image-Retrokonversion der Einleitung an. Zusätzlich dazu sind die in einer Ergänzungslieferung gegebenen Nachträge und Verbesserungen einzuscannen, die für die Einleitung vor allem Literaturangaben in den Anmerkungen ergänzt.⁵⁹

Der erste Schritt ist die Digitalisierung aller Katalogseiten und ihre Speicherung im PDF/A-Format. Dabei muss der Scan von Einzelseiten erfolgen, um später als erstes Image die Seite anzuzeigen, auf der die Beschreibung der Handschrift beginnt. Die Vergabe der Dateinamen sollte für die verschiedenen Abschnitte des Katalogs getrennt erfolgen. Für die Katalogseiten mit den Handschriftenbeschreibungen werden die Signaturen in den Dateinamen übernommen, um eine spätere automatische Zuordnung der Images zu den

⁵⁹ Vgl. Michael Klein, Die Handschriften der Sammlung J 1 im Hauptstaatsarchiv Stuttgart, Wiesbaden 1987 (Die Handschriften der Staatsarchive in Baden-Württemberg 1) Nachträge und Verbesserungen (S. 563-584), Ergänzungen zur Einleitung, S. 565-567. Neben der Einleitung beziehen sich die Korrekturen und Ergänzungen auf die Handschriftenbeschreibungen und den Index.

Verzeichnungseinheiten zu gewährleisten. An die Signaturen können die Seitenzahlen angehängt werden. Wenn auf einer Seite die Signaturen J 1 Nr. 77 und Nr. 78 beschrieben sind, lautete der Dateiname: „077_078_S_145.pdf“.

Für die Abschnitte „Einführung“ und „Abkürzungen“ genügt die Übernahme der Katalogseite in den Dateinamen (z.B. S_047.pdf). Da sowohl Seitenzahlen als auch Signaturen dreistellig sein können, müssen für eine korrekte Sortierung der Dateien bei ein- bzw. zweistelligen Zahlen zwei bzw. eine führende Null im Dateinamen hinzugefügt werden.

Als Vorlage für eine Anzeige der Images im Internet kann die Präsentation des Bestandes J 56 Nachlass Christian von Martens dienen, dessen Tagebücher und Zeichnungen vollständig digitalisiert wurden.⁶⁰ Die Entscheidung, die Images des Handschriftenkatalogs auf die gleiche Art und Weise anzuzeigen wie einen digitalisierten Bestand, bedeutet freilich einen Bruch in der bisherigen Präsentation von Erschließungsergebnissen und von digitalisiertem Archivgut mit scopeArchiv. Die Images des Katalogs reproduzieren nicht das Archiv- bzw. Sammlungsgut, sondern zeigen die Erschließungsdaten in der Form, in der sonst digitalisierte Archivalien angezeigt werden.

4.2 Metadaten für die Präsentation der Images und Klassifikation des Bestandes

Für die Präsentation der Images im Online-Findmittel-System (OLF21) des Landesarchivs Baden-Württemberg, das über eine Schnittstelle

⁶⁰ Vgl. <<https://www2.landesarchiv-bw.de/ofs21/olf/startbild.php?bestand=5528>>.

mit dem Produktivsystem scopeArchiv verbunden ist, muss jede Handschriftenbeschreibung einer Verzeichnungseinheit zugewiesen werden. Pflichtfelder in scopeArchiv sind die Signatur und das Titelfeld, die in jeder Erschließungsmaske von scopeArchiv vorgesehen sind. Für eine eindeutige Identifikation der Handschrift sollten zudem die fünf Elemente in Felder überführt werden, die dem Kurztitel in der darunter stehenden Schlagzeile folgen: „Frühere Signatur“, „Umfang“, „Format“, „Entstehungsort“ und „Datierung“, wobei frühere Signaturen nur sehr selten erfasst sind. Um vorhandene Strukturen und Datenfelder nachnutzen zu können, bietet sich die Verwendung der Urkundenmaske an.⁶¹ Das einzige Feld, das in der Urkundenmaske fehlt, ist das Feld „Umfang“. Es ist jedoch Bestandteil anderer Erschließungsmasken in scopeArchiv. Da aus dem Onlinefindbuch kein gedrucktes Findmittel erstellt werden soll, entsteht ein Programmierungsaufwand nur für die Anzeige der Felder im Internet. Die Basis dafür bildete die Urkundenmaske unter Ausschaltung der nicht benötigten Felder, ergänzt um das Feld „Umfang“. Eine Modifikation der Programmierung ist zudem bei der zulässigen Zeichenanzahl vonnöten, denn das Feld „Umfang“ ist laut scope-Erschließungsrichtlinie auf 120 Zeichen beschränkt. Es konnte zumindest ein Fall ermittelt werden, in dem im Katalog diese Höchstzahl überschritten wird (J 1 Nr. 180). Die Handschrift umfasst mehrere Sammelbände, die ohne Untergliederung in eigene Signaturen summarisch erschlos-

⁶¹ Vgl. Landesarchiv Baden-Württemberg, Beiblatt Nr. 4 (wie Anm. 17). Ein Mapping der Elemente des Handschriftenkatalogs mit den entsprechenden Feldern aus scopeArchiv enthält Anhang 7 der Transferarbeit in der online-Fassung (URL in Anm. 1).

sen sind und daher im Abschnitt „Anzahl Bl.“ Angaben für alle Bände enthält. Deshalb müsste das Feld „Umfang“ um 80 auf 200 Zeichen erweitert werden.

Zwei weitere Felder der Urkundenmaske weichen von der scope-Erschließungsrichtlinie ab. Bei der Übernahme der Orte in das Feld „Entstehungsort“ bzw. „Ausstellungsort“ muss die Kennzeichnung von fraglichen Ortszuordnungen mit dem Zusatz eines „(?)“ erfolgen. Für das Feld „Format“ ist in der Urkundenmaske der Eintrag in der Form „ZAHL x ZAHL cm“ vorgesehen. Im Katalog treten jedoch auch Formatangaben wie „etwa 20 x 16,5 außer 102-108“ (J 1 Nr. 76b) auf, die zumindest in einem Bemerkungsfeld möglich sein sollten.

Die Präsentation der Images des retrokonvertierten Katalogs hätte folgende Struktur:

☐	Vorwort (S. 9-10)
☐	Einführung
☐	I. Handschriften im Hauptstaatsarchiv Stuttgart (S. 11-19)
☐	II. Geschichte, Aufgaben und thematische Schwerpunkte der Sammlung J 1 (S. 20-60)
☐	III. Aufbau der Handschriftenbeschreibungen und der Indizes (S. 60-65)
☐	Abkürzungen und Sigeln (S. 66-72)
☐	Beschreibungen der Handschriften 1 bis 419
	1 Simon Studion
	1 a Simon Studion
	1 b Sebastian Ebinger
	...
	419 Altes Repertorium
☐	Versanfänge (Initien) (S. 311-314)
☐	Personen-, Orts- und Sachindex (S. 315-560)

Abb. 1: Klassifikation der retrokonvertierten Katalogseiten

Im Navigationsbaum müsste unter Wiedergabe der Untergliederung des vorliegenden Katalogs in Einführung, Beschreibungsteil und Index⁶² eine mehrgliedrige Klassifikation anwählbar sein. Nur im Menüpunkt „Beschreibungen der Handschriften 1 bis 419“ wären die Images mit bestellbaren Einheiten verknüpft, denn jeder Signatur entspricht – abgesehen von den unten geschilderten Ausnahmefällen – eine Handschrift. Für die übrigen Klassifikationspunkte muss die Bestellfunktion deaktiviert werden.

Damit für jede Handschrift die Metadaten wie „Titel“, „Umfang“, „Format“ und „Datierung“ angezeigt werden können, muss für jede Signatur eine eigene Verzeichnungseinheit angelegt werden. Diese Verzeichnungseinheiten wären dann unter dem Klassifikationspunkt „Beschreibungen der Handschriften 1 bis 419“ eingebunden.

In einer Nachlieferung „Nachträge und Verbesserungen“ gibt Michael Klein Korrekturen für die Handschriftenbeschreibungen, die signaturbezogen nachgeführt werden müssen.⁶³ Das Image der Seite mit Nachträgen muss dafür an die Seite der Handschriftenbeschreibung des Hauptkatalogs angehängt werden.

⁶² Auch wenn für den Index möglicherweise eine strukturierte händische Erfassung als Retrokonversionsverfahren gewählt wird, sollten parallel dazu die Katalogseiten des Index online verfügbar sein. Vgl. Kapitel 3.2 der Transferarbeit in der online-Fassung (URL in Anm. 1).

⁶³ Es handelt sich um neue Literatur zu den Handschriften, neu ermittelte Signaturen verwandter Bestände oder Nachweise verloren geglaubter Handschriften. Vgl. Klein, Handschriften 1987 (wie Anm. 59), S. 567-572.

4.3 Signaturformate, Sammlungen und Serien

„Wann durch die Vergabe der noch heute gültigen Signaturen der Bestand endgültig formiert wurde, ist nicht exakt feststellbar“, hält Michael Klein fest.⁶⁴ Vermutlich wurden sie noch in der ersten Hälfte des 18. Jahrhunderts vergeben, wobei vorhandene Zwischennummern dafür sprechen, dass der Bestand nicht in einem Zug verzeichnet worden ist. Dies hat Folgen für die Retrokonversion. Zwischennummern wurden vergeben, wenn eine Handschrift aus mehreren Teilen bestand oder wenn inhaltlich ähnliche, später erworbene Handschriften einer vorhandenen Signatur durch Zusatzbuchstaben angegliedert wurden (vgl. z.B. Zwischennummern in J 1 Nr. 48a bis 48z). Dabei spiegelt sich in den Signaturen zum Teil die Erwerbungs-geschichte. Für den ersten Teil (Nr. 1-100), welchen Karl Friedrich Pfaff am Beginn des 19. Jahrhunderts verzeichnete, lässt sich Klein zufolge kein Ordnungsprinzip erkennen. Bei späteren Erwerbungen ab J 1 Nr. 101 erhielten Handschriften ohne inhaltlichen Bezug zu bereits Vorhandenem fortlaufend neue Signaturen, diejenigen mit inhaltlicher Verwandtschaft wurden der bereits vergebenen Signatur unter Zusatz eines Buchstabens beigeordnet.⁶⁵

Die verwendeten Buchstabenformate sind jedoch nicht eindeutig. Das Format lässt nicht immer Rückschlüsse auf den formalen Zusammenhang der Handschriften zu, da die Signierung historisch gewachsen ist. So gibt es Handschriften, bei denen die laufende Nummer ohne angefügten Buchstaben gar nicht besetzt ist (J 1 Nr. 20a,

⁶⁴ Vgl. Klein, Handschriften (wie Anm. 2), S. 25.

⁶⁵ Vgl. ebd., S. 28 mit Anm. 67.

20b), es gibt Signaturen, für die die „a“-Nummer nicht vergeben ist (J 1 Nr. 25, Nr. 25b), obwohl die b-Nummer eine spätere Erwerbung war, die der vorhandenen Nummer angefügt wurde. Manchmal bezeichnen die mit dem Buchstaben angefügten Handschriftensignaturen echte Teile der zugeordneten Handschrift (J 1 Nr. 21 und Nr. 21a), manchmal grenzen die Buchstaben unterschiedliche Entstehungsstufen derselben Handschrift voneinander ab (J 1 Nr. 126a, Nr. 126b, Nr. 126c), in anderen Fällen liegt gar kein formaler Zusammenhang vor (J 1 Nr. 121a, Nr. 121b). Es ist auch möglich, dass einer Untergliederung in Buchstaben keine weitere Detaillierschließung folgt (J 1 Nr. 129a-b).

Doch die *inhaltliche* Bedeutung der Zusatzbuchstaben hat kaum Relevanz für die Retrokonversion. Im Katalog sind die Signaturen und Beschreibungen der Unteraufnahmen in gleicher Schriftgröße gesetzt wie die Hauptaufnahmen. Diese Struktur sollte auch nach der Retrokonversion bestehen bleiben. Das bedeutet, dass jede Signatur, gleich ob Ober- oder Unteraufnahme, auf derselben Ebene klassifiziert wird.

Mehrteilige Handschriftensignaturen weisen jedoch im Beschreibungsteil und beim Kurztitel Eigenheiten auf, die bei der Retrokonversion zu beachten sind.⁶⁶ Ein eigener Kurztitel wurde für einige

⁶⁶ Die Signaturformate und mehrteiligen Handschriftenbeschreibungen mit Auswirkungen auf die Retrokonversion sind im Anhang 2.2 der Transferarbeit in der online-Fassung (URL in Anm. 1) aufgeführt.

Untersignaturen vergeben, für andere nicht.⁶⁷ In beiden Fällen muss der Titel der Obergabe nachgeführt werden.⁶⁸

Beim Zuweisen von Katalogseiten zu Signaturen ist bei mehrteiligen Handschriften zu beachten, dass häufig Beschreibungen, die sich auf den gesamten Handschriftenkorpus beziehen, nur ein Mal in den Abschnitten „Beschreibung des Äußeren“, „Geschichte der Handschrift“ oder „Beschreibung des Inhalts“ in der Obergabe erfolgt sind. Nur hier wurden umfangreichere Werktitel vollständig wiedergegeben. Klein entsprach damit den Erfassungsanweisungen der DFG, die festlegen, dass bei zusammengesetzten Handschriften und Sammelhandschriften alles Gemeinsame am Anfang behandelt wird, die Beschreibung jedes einzelnen Teils dann getrennt erfolgen soll.⁶⁹

In den Beschreibungen der Untergaben werden der Werktitel oder auf alle Teile zutreffende ausführliche Inhaltsangaben nur noch in Kurzform wiederholt. Wenn der Nutzer sich das Image einer Untersignatur aufruft, fehlen der vollständige Werktitel und die allgemeinen Informationen. Deshalb muss bei Signaturen, bei denen die Obergabe auf einer anderen Seite als die Untergabe steht, das Image der Obergabe zusätzlich eingestellt werden.

⁶⁷ Z.B. Obergabe J 1 Nr. 76 a-b "Wilhelm Schickard (Sammelmappe)", Beschreibungen in Nr. 76a und 76b ohne Kurztitel.

⁶⁸ Bei den Kurztiteln für Untergaben handelt es sich häufig nur um Bandzahlen, die bei der Retrokonversion durch den Titel der Obergabe ergänzt werden müssen. Z.B. Obergabe J 1 Nr. 3-7a „Oswald Gabelkover“, Untergabe Nr. 3 „Buch 1“, Nr. 4a „Buch 1 und 2“, Nr. 4b „Buch 2“.

⁶⁹ Vgl. Richtlinien (wie Anm. 12), S. 10.

4.4 Umsignierung, Verluste und Verweise auf heute in der WLB, der Dienstbibliothek oder in anderen HStAS-Beständen befindlichen Handschriften

Der Katalog weist auch Handschriften nach, die sich ehemals im Bestand befanden, dann aber durch Auslagerung, Feuer oder andere Umstände verloren gingen bzw. zerstört wurden (z.B. J 1 Nr. 54). Falls sie durch Umlagerungen oder Abgaben an eine andere Institution gelangten, wird dies durch Verweise angegeben, z.B. „jetzt LBS“ (heute: WLB).⁷⁰ Im Jahr 1977 wurden alle Texte des späten 19. und 20. Jahrhunderts, die nicht als genuine Handschriften einzustufen waren, aus J 1 entnommen und in andere Bestände oder die Dienstbibliothek eingefügt.⁷¹ Deshalb sind von der Neusignierung vor allem Signaturen mit höheren Nummern betroffen.

Wie sind diese Verweise zu retrokonvertieren? Üblicherweise weist ein Findmittel die Erschließungseinheiten aus, die im Archiv vorhanden sind. Verweise auf verwandte oder von der Provenienz her ehemals zugehörige Handschriften liegen jedoch im Nutzerinteresse. Für die früher in der Sammlung befindlichen Handschriften ist es nicht notwendig, dass der Nutzer erst ein Image öffnen muss, um diese Information zu erhalten. Deshalb erscheint es ratsam, die neu-

⁷⁰ Eine Zusammenstellung aller an die WLB abgegebenen Signaturen enthält die erste Tabelle im Anhang 5 der Transferarbeit in der online-Fassung (URL in Anm. 1).

⁷¹ Vgl. Klein, Handschriften (wie Anm. 2), S. 32. Eine Zusammenstellung aller in die Dienstbibliothek aufgenommenen Signaturen enthält Anhang 3 der Transferarbeit in der online-Fassung (URL in Anm. 1), die Signaturen der Handschriften in anderen HStAS-Beständen vgl. Anhang 4 der online-Fassung.

en Bestandssignaturen im HStAS, die Bibliothekssignatur oder die WLB-Signatur in das Feld „Titel“ abzuschreiben und dem Nutzer sofort von der Strukturansicht aus diese Information zur Verfügung zu stellen. Da der Katalog für keine der abgegebenen oder neu zugeordneten Handschriften einen Titel ausweist, ist so zugleich der scope-Erschließungsrichtlinie Genüge getan, die den Titel als Pflichtfeld definiert.

Weitere Sonderfälle mit Relevanz für die Retrokonversion betreffen irrtümlich nicht vergebene Nummern oder nicht mehr nachweisbare Handschriften.⁷²

5. Fazit und Ausblick

Die Wiedergabe von Erschließungsinformationen in Papierform basiert auf gänzlich anderen Mechanismen als ihre digitale Präsentation. Verschiedene Drucktypen, Absätze und Schriftgrößen bestimmen im Katalog die Orientierung, eine Erschließung in der Software erfordert eine durch Felder strukturierte Verzeichnung. Auch wenn die enge Definition von Retrokonversion eine Nachbearbeitung des

⁷² Vgl. Anhang 5, Tabelle 2 der Transferarbeit in der online-Fassung (URL in Anm. 1). Bei nicht besetzten Nummern ist im Titel „Irrtümlich nicht vergeben“ einzutragen, ohne Imageanzeige. Ein weiterer Sonderfall stellt J 1 Nr. 18 dar, das ohne Image den Verweis „Alte Bezeichnung von Hs 12a“ im Titel führen muss. In Fällen von verbrannten oder nicht mehr nachweisbaren Handschriften übernahm Klein die Erschließungsinformationen aus dem alten Repertorium, z.B. J 1 Nr. 57: „Druck: Johann Georg Walz, Fürstlich württembergischer Stamm- und Namensquell, Stuttgart 1657. 1788 an die LBS abgegeben, dort wohl 1944 verbrannt.“ Diese Angaben sollten verfügbar bleiben und daher das Image der Katalogseite eingestellt und im Titel jeweils „Nicht mehr nachweisbar“ oder „Verloren“ eingetragen werden. Weitere, für die Retrokonversion unerhebliche Teilverluste, vgl. Klein, Handschriften (wie Anm. 2), S. 32 Anm. 92f.

Materials ausschließt, ist doch hinsichtlich der Strukturierung und Auszeichnung der vorhandenen Elemente ein Umdenken erforderlich.

Ziel war die Erarbeitung eines Retrokonversionsverfahrens, das im Hauptstaatsarchiv Stuttgart des Landesarchivs Baden-Württemberg durchgeführt werden kann – ohne Involvierung von Dienstleistern, ohne aufwendige technische Programmierarbeiten – unter Verwendung vorhandener Erschließungssoftware. Die Entscheidung, die Handschriftenbeschreibungen in Form von Images im Internet zugänglich zu machen, bedeutet zwar eine Entscheidung gegen die Durchsuchbarkeit der Informationen, die jedoch durch eine strukturierte Übernahme des Index kompensiert werden kann.

Mit dem Bildbetrachtungsmodul des Online-Findmittel-Systems von scopeArchiv können die Images präsentiert werden. Auch wenn es bedeutet, dass zukünftig digitalisierte Archivalien und digitalisierte Erschließungsinformationen gemeinsam verwaltet werden müssen. Dies stellt einen konzeptionellen und praktischen Umbruch für die Anwendung von Erschließungssoftware dar. Für die Frage der Langzeitarchivierung und Verfügbarkeit von Dateiformaten ist es zunächst irrelevant, welche Informationen die Dateien enthalten. Für die Langzeitarchivierungsstrategie muss ohnehin eine konsistente, permanente Verknüpfung von Images und Erschließungsinformationen bzw. zwischen Images und Metadaten sowie die Lesbarkeit der vorhandenen Dateiformate sichergestellt werden. Mit der Entscheidung für eine Image-Retrokonversion der Handschriftenbeschreibungen ist der Vorteil verbunden, keine eigene Datenbank aufzubauen, die gesonderte Überlegungen für eine langfristige Archivierung erfordert. So können das Bestreben, mehr Erschließungsinformationen online

zugänglich zu machen und dabei rationell und kostengünstig zu verfahren, miteinander in Einklang gebracht werden.

Um die erörterten Probleme an die Praxis zurückzubinden, wären vor allem für die Ermittlung technischer Parameter Testreihen, Probescans und die Prüfung neuer Software notwendig, da diese ständigen Weiterentwicklungen unterworfen ist. Für die Suche nach praktischen Lösungen ist es von Vorteil, wenn, wie bei den von Fricke und Maier durchgeführten Testreihen oder bei der Erarbeitung der Transferarbeit von Ulrich Fischer konkrete Projekte bereits bestehen und sowohl Personal als auch technische Ausstattung wie Hard- und Software vorhanden sind.

Eine abschließende Bemerkung weist über die konkrete Bearbeitung der Fragestellung hinaus und betrifft die Zusammenarbeit von Bibliotheken und Archiven. Nicht nur die Genese der Handschriftensammlung in J 1 als Handbibliothek für Archivare offenbart bereits, dass der Versuch einer strengen Abgrenzung des Bestands entweder als Archiv- oder als Bibliotheksgut scheitern muss. Der vorhandene Katalog ist nach bibliothekarischen Richtlinien erschlossen, bildet aber einen Teil archivischer Überlieferung. Die Erfahrungen der Bibliotheken im Bereich der Retrokonversion und Standardisierung sind enorm, da durch bereits in den 1990er Jahren realisierte Verbundkonzepte die normierte Erschließung bundesländerübergreifend und basierend auf einheitlichen Richtlinien erprobt werden konnte. Die Archivlandschaft ist bis heute durch Heterogenität geprägt, wenngleich sich Standards und einheitliche Austauschformate langsam durchsetzen. Die Archive können in dieser Hinsicht von den Erfahrungen in Bibliotheken profitieren, denn in vielen Punkten über-

schneiden sich die Arbeitsfelder, nicht nur bei der Retrokonversion, sondern auch in Fragen der Bestandserhaltung, bei der Verwaltung und Langzeiterhaltung elektronischer Dokumente sowie der Präsentation von Erschließungsergebnissen in Portalen. Angesichts immer knapper werdender Ressourcen sollte trotz unterschiedlicher Traditionen die Chance genutzt werden, vorhandenes Know-how nachzunutzen und Erfahrungen zwischen Archiven und Bibliotheken auszutauschen.

Von der Eindimensionalität zum Netzwerk: Die Retrokonversion des Systematischen Katalogs im Deutschen Literaturarchiv

von Karin Schmidgall

1. „Buch, Bibliothek und geisteswissenschaftliche Forschung“

In dem Wort Retrokonversion steckt per se die Aufforderung in die Vergangenheit zu schauen. Als Initialzündung einer ganzen Reihe von Aktivitäten im Bereich Altbestandserfassung gilt Bernhard Fabians 1983 erschienenes Werk „Buch, Bibliothek und geisteswissenschaftliche Forschung“. Fabian benennt Probleme der Literaturversorgung in den Geisteswissenschaften und macht Lösungsvorschläge. Er sah damals schon voraus, was heute Tatsache ist: nur überregional – also heute im Netz – nachgewiesene Bestände werden von der Forschung wahrgenommen. Die DFG reagierte u.a. mit einem Förderprogramm zur Altbestandserfassung in wissenschaftlichen Bibliotheken der Bundesrepublik Deutschland, das während der Laufzeit (Beginn 1988) um fünf ostdeutsche Bibliotheken erweitert wurde. In 15 Einzelprojekten wurden konventionelle Titelaufnahmen für monographische Werke mit Erscheinungsjahren zwischen 1501 und 1850 in maschinenlesbare Daten konvertiert, um sie für den lokalen und regionalen Nachweis zur Verfügung zu stellen.¹

¹ Klaus Haller/Ekkehard Henschke/Reinhard Rutz (Hrsg.), Altbestandserschließung in wissenschaftlichen Bibliotheken. Ein Förderprogramm der Deutschen Forschungsgemeinschaft, Berlin 1995 (Dbi-Materialien 143), S. 17.

Fabian betont in seiner Studie, dass die „Geisteswissenschaften [...] auf große Literaturreserve angewiesen [sind], die sowohl die Quellenliteratur wie die moderne Forschungsliteratur umfassen. [...] Die Bibliothek behält den Charakter eines Archivs der modernen wissenschaftlichen Literatur, wird aber zugleich zu einem der Wissenschaft dienenden Archiv der kulturellen Überlieferung. Diese Doppelfunktion erfüllt eine Bibliothek dann in optimaler Weise, wenn sie dem wissenschaftlichen Benutzer eine umfassende Quellensammlung zusammen mit einem differenzierten Instrumentarium zu deren Erschließung bietet“.² In besonderem Maße kann diese Aussage auf die Bibliothek des Deutschen Literaturarchivs Marbach (DLA) und ihren Systematischen Kartenkatalog übertragen werden.

Die Bibliothek des DLA ist mit weit über 700.000 Bänden und einer umfassenden Mediendokumentation die größte Spezialbibliothek³ für neuere deutschsprachige Literatur sowie Literaturwissenschaft und wird von Forschern aus aller Welt frequentiert. Zu ihren Besonderheiten zählen über 130 Autoren-, Gelehrten- und Sammlerbibliotheken, Verlagsarchiv-Bibliotheken und gattungsspezifische Sammlungen. Die Sammlung literarischer Zeitschriften der Moderne seit 1880 gehört zu den umfangreichsten in der Bundesrepublik. Seit 1959 werden literarische und literaturwissenschaftliche Zeitschrif-

² Bernhard Fabian, *Buch, Bibliothek und geisteswissenschaftliche Forschung. Zu Problemen der Literaturversorgung und der Literaturproduktion in der Bundesrepublik Deutschland*, Göttingen 1983, S. 199 u. S. 27.

³ Jutta Bendt, *Die Bibliothek des Deutschen Literaturarchivs in Marbach*, in: Antonius Jammers/Dietger Pforte/Winfried Sühlo (Hrsg.), *Die besondere Bibliothek oder: Die Faszination von Büchersammlungen*, München 2002, S.111-126.

ten, Jahrbücher und Sammelwerke retrospektiv von 1880 beginnend in einem aktuellen Programm ausgewertet. Bei Monographien und unselbständig erschienenen Beiträgen erfolgt eine bibliographische, dokumentarische und sachliche Tiefenerschließung, welche eine entsprechend differenzierte Recherche ermöglicht. Ton- und Bildträger sowie Hörfunk- und Fernsehmanuskripte werden katalogisiert und inhaltlich ausgewertet. Der überregionale Nachweis findet im Südwestverbund⁴ und in der Zeitschriftendatenbank⁵ statt.

Der Bestand der Marbacher Spezialbibliothek ist bis einschließlich 1998 mit Hilfe verschiedener Kartenkataloge erschlossen worden. Das Herzstück und viel genutzte Arbeitsinstrument ist der systematisch gegliederte Sachkatalog (kurz SK). Er wurde nach Gründung des Literaturarchivs 1955 unter der Leitung von Paul Raabe nach dem Vorbild des Katalogs im Kieler Institut für Weltwirtschaft aufgebaut und vereint selbständig sowie unselbständig Erschienenes.⁶ Auf diese Weise werden vielfältige Aspekte der deutschen Literaturgeschichte dokumentiert: der Einfluss skandinavischer Autoren auf die Literatur der Jahrhundertwende, die Vielfalt expressionistischer Strömungen, die rege publizistische Tätigkeit der Autoren während der Weimarer Republik und die untergründigen Szenen im literarischen Leben der DDR – um nur wenige Beispiele zu nennen.

⁴ <<http://swb.bsz-bw.de>> [08.01.2010].

⁵ <<http://dispatch.opac.d-nb.de/DB=1.1>> [08.01.2010].

⁶ Paul Raabe, Die Bibliothek des Deutschen Literaturarchivs, in: Zeitschrift für Bibliothekswesen und Bibliographie 10, 1963, S. 213-222.

Der Systematische Katalog enthielt bei seinem Abbruch ca. 1,2 Millionen Nachweise (davon ca. 380.000 Kartenkopien für den Nachweis an verschiedenen Systematikstellen), von denen etwa 800.000 allein die literarischen Autoren seit dem 18. Jahrhundert betreffen. In diesem Bereich hat der Katalog personalbibliographische Ausmaße und wird für entsprechende Publikationen als Referenzorgan benutzt. Namen, Pseudonyme und biographische Daten der Autoren wurden seit jeher in einer Kartei gepflegt. Mit 43.000 Namen bildete sie den Grundstock für die seit 1999 abteilungsübergreifend gespeiste Namensnormdatei in der Erschließungsdatenbank Kallias.⁷

Die Fachsystematik des Katalogs basiert auf dem Dezimalsystem, die Ordnung der Karten funktioniert nach einem Leitkartenprinzip. Herausragender Kern sind die Gruppen „Epochen der Deutschen Literatur und Literaturwissenschaft“. Sie sind mit über 960.000 Titeln Karten der mit Abstand umfangreichste und am differenziertesten geordnete Komplex. Für die Literatur von und über Autoren gibt es eine eigene Untersystematik ("Autorenschema"), der Benutzer findet hier Primär- und Sekundärliteratur. Die Nachweise zu Autoren wie Goethe (23.400) und Schiller (15.000) beispielsweise haben den Umfang des Bestands einer kleinen Stadtbücherei. Das Autorenschema hat im Wesentlichen die Funktion, große Treffermengen überschaubar zu machen und Forschungen zu Einzelaspekten nachzuweisen.

Der Systematische Katalog ist ein historisch gewachsenes Nachweisinstrument, das an vielen Punkten die im Online-Katalog

⁷ <http://www.dla-marbach.de/opac_kallias/index.html> [08.01.2010].

obligatorische Standardisierung durch Notationen vermissen lässt. Allerdings ist diese Stringenz in einem Kartenkatalog, der ein ganz anderes, blätternes Suchverhalten ermöglicht, auch nicht zwingend.

2. Methodik der Retrokonversion⁸

Die verschiedenen Konversionsmethoden lassen sich typologisieren. Die geeignete Methode für das jeweilige konkrete Projekt muss mit Hilfe einer genauen Analyse der vorliegenden Katalogsituation gefunden werden. Dabei spielen mehrere Faktoren eine Rolle: die Menge der zu konvertierenden Titelaufnahmen, der Zustand des Katalogs (äußere Form, Qualität der Aufnahmen, Katalogabbrüche etc.), nachnutzbare Vorarbeiten, Zahl und Qualität der zur Verfügung stehenden Fremddaten, Einsatz technischer Verfahren, Formen der Arbeitsorganisation, Zeitvorgaben, technische, organisatorische und räumliche Rahmenbedingungen, Qualitätsanforderungen an das Ergebnis und das Wichtigste zum Schluß: der Kostenrahmen.⁹

⁸ Der Begriff Retrokonversion wurde 1989 von der Arbeitsgruppe für Retrospektive Katalogisierung des Europarats definiert: „Retrospektive Katalogkonversion (Retrokonversion) bedeutet die Umformung von bereits bestehenden Katalogen von der konventionellen in maschinenlesbare Form. Dieser Vorgang unterscheidet sich von der Retrospektiven Katalogisierung, die die Erstkatalogisierung von älteren Bibliotheksbeständen meint, die bisher noch nicht katalogisiert wurden.“ Vgl. Claudia Fabian, *Altbestandskatalogisierung als Gemeinschaftsaufgabe Europas. Bericht über die International Conference on Retrospective Cataloguing in Europe, 15th to 19th century materials, vom 28. bis 30. November 1990 in München*, in: *Zeitschrift für Bibliothekswesen und Bibliographie* 38, 1991, S. 227-235, hier S. 227.

⁹ Für einen Überblick sind immer noch die verschiedenen Dbi-Materialien hilfreich, die der 1992 eingerichtete Beratungsdienst „Retrospektive Konversion“ als ständige Dienstleistung des Deutschen Bibliotheksinstituts Berlin in den 1990er

Grundsätzlich haben sich zwei Modelle der Retrokonversion durchgesetzt: die strukturierte Erfassung und Einbindung der Daten in ein bestehendes Verbund- oder Lokalsystem oder der Aufbau eines eigenen Imagekatalogs (siehe Abb. 1).

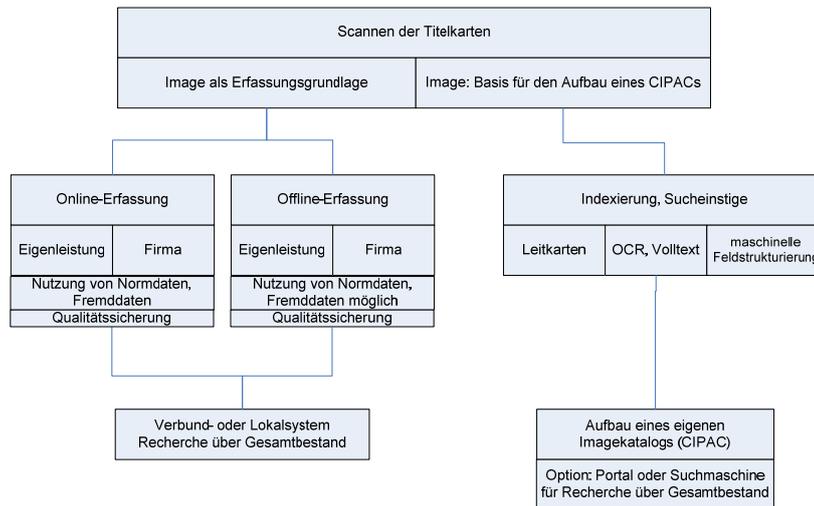


Abb. 1: Typologie der Konversionsmethoden

Für die strukturierte Erfassung muss zunächst ein Kategorienformat definiert werden. Die Erfassungsarbeiten können in Eigenleistung erfolgen oder an einen Dienstleister (Konversionsfirma) vergeben werden, in beiden Fällen ist die Qualitätssicherung eine Schlüs-

Jahren veröffentlicht hat. Insbesondere: Retrokonversion. Konversion von Zettelkatalogen in deutschen Hochschulbibliotheken. Methoden, Verfahren, Kosten, Berlin 1993 (Dbi-Materialien 128).

selmaßnahme. Die Erfassungsgrundlage ist das Original oder ein zuvor hergestellter Titeltkartenscan. Beim Online-Verfahren werden dann die Daten direkt über den Verbundclient oder das Lokalsystem unter Verwendung von Normdaten eingegeben und vorhandene Fremddatenquellen genutzt.

Bei der Offline-Erfassung werden die Daten zuerst in einem eigenen System erfasst – oft werden die Erfassungsarbeiten in Billiglohnländern durchgeführt – und anschließend in einem Standardaustauschformat (z.B. MAB2 oder MARC21) bereitgestellt und in das Verbund- oder Lokalsystem importiert.¹⁰ Auch bei einem Offline-Verfahren ist der Einsatz von Normdaten und die Nutzung von Fremddaten möglich, allerdings müssen die Daten für das Projekt zur Verfügung gestellt und synchronisiert werden.

Als zweites Konversionsmodell hat sich der Aufbau von Imagekatalogen etabliert: der sogenannte Card-Image Public Access Catalogue (CIPAC)¹¹ ist ein originalgetreues Abbild des Kartenkataloges

¹⁰ Die Grenzen und Möglichkeiten von Online- und Offline-Verfahren sowie eines kombinierten Verfahrens zur Online-Erfassung und Offline-Einspeicherung zeigt Werner Holbach auf: Retrokonversion und Retrokatalogisierung in den Bayerischen Staatlichen Bibliotheken, in: Bibliotheksforum Bayern 34, 2006, Heft 3, S. 206-223.

¹¹ Eine Übersicht der CIPACs hat Otto Carl Oberhauser im Rahmen seiner Dissertation „Card-Image Public Access Catalogues (CIPACs). A Critical Consideration of a Cost-Effective Alternative to Full Retrospective Catalogue Conversion. Univ. of Central England, Birmingham, UK, 2002“ ins Netz gestellt und weitergepflegt: <http://www.ub.tuwien.ac.at/cipacs/c-i.html> [19.11.2009]. Das im Text folgende Zitat stammt von dort.

im Netz. Der Zugang zu den Images kann über drei verschiedene Wege erreicht werden:

1. Die Leitkarten werden erfasst und als Einstiegspunkte zum Browsen zur Verfügung gestellt.
2. Der Text einer Titelaufnahme wird mit OCR ausgelesen und für eine Volltextsuche aufbereitet.
3. Anhand zuvor definierter Schlüsselwörter und Algorithmen erfolgt eine maschinelle Feldstrukturierung des mit OCR erkannten Textes. Damit wird neben der Volltextsuche eine feldweise Recherche, z.B. nach Autor und/oder Titel ermöglicht.

Während Imagekataloge bei Geldgebern als kostengünstige und schnelle Konversionsmethode bevorzugt werden und sich im deutschsprachigen Raum auffällig stark verbreitet haben, werden sie bei Benutzern und Bibliothekaren nur als Zwischenstation angesehen. Die Ergebnisse seiner Umfrage zum Thema CIPACs im Jahr 2002 fasste Otto Carl Oberhauser folgendermaßen zusammen: „[...] they are at best acceptable as short or medium-term, but not as permanent alternatives to ‚real‘ OPACs“.

3. Die Retrokonversion des Systematischen Katalogs im Deutschen Literaturarchiv

Ziel des Marbacher Projekts ist die komplette Überführung des konventionellen Sachkatalogs und damit die Anreicherung der leistungsstarken lokalen Datenbank Kallías: quantitativ um mehr als eine Million Nachweise literarischer und wissenschaftlicher, selbständig und unselbständig erschienener Literatur, die von der Gegenwart bis ins 18. Jahrhundert zurückreichen, qualitativ um die in den zurücklie-

genden 50 Jahren gewonnenen Ergebnisse der sachlichen Feinerschließung.

Kallias ist ein für bibliothekarisch-archivarische Strukturen optimiertes Datenbanksystem, das über eine Schnittstelle zum Südwestverbund verfügt. Seine Aufgabe ist der integrierte Nachweis aller Bestände des Deutschen Literaturarchivs. Im Zentrum stehen die Normdateien für Personen, Körperschaften, Werktitel, Sachschlagwörter sowie für die Systematik.

Bei der Konversion des Systematischen Katalogs sollten die Normdaten aus Kallias weitestgehend produktiv gemacht werden. Der Nutzer soll in einer Suchanfrage auf einen konsistent erschlossenen Gesamtbestand zugreifen können.

Als Methode wurde ein erstmals in der Bibliothek des Westdeutschen Rundfunks 1995 angewandtes „Roll-Down-Verfahren“ optimiert.¹² Mit Hilfe von Konkordanzen wurden normierte Informationen, d.h. Personen in der Relation „Von“ und „Über“ sowie Systematikstellen mit ihren Identnummern als Barcodeinformation auf Katalogkarten ausgedruckt und in den Kartenkatalog eingelegt. Dabei handelt es sich zum großen Teil um einen „Austausch“ der Leitkarten und zu einem anderen Teil um das Nachführen fehlender Leitkarten, um vorhandene, aber nicht markierte Unterteilungen für das Konversionsverfahren sichtbar zu machen.

¹² Thomas Lehmann, Retrokonversion in der Bibliothek des Westdeutschen Rundfunks Köln, in: Retrokonversionsprojekte – Planung und Durchführung, Berlin 1997 (Dbi-Materialien 155).

Das Roll-Down-Verfahren gewährleistet beim Scan-Vorgang die Übertragung der Barcodeinformation auf die nachfolgenden Titeltkarten bis zur jeweils nächsten Barcodekarte. Als Ergebnis trägt jede gescannte Titeltkarte im Header der Bilddatei die im Barcode verschlüsselte normierte Ansetzungsform des Verfassers bzw. die Sacherschließungsinformation. Der lesbare Text auf der Karte dient nur zur Information und zur Erleichterung des Einstellens in den Katalog. Insgesamt wurden 100.000 Barcode-Karten innerhalb eines Jahres in den Katalog eingestellt.

Mit den Erfassungsarbeiten, welche auf der Grundlage eines kompletten Katalogscans durchgeführt werden, ist die Firma Bibliographische Dienste GmbH (München) beauftragt worden. Das Angebot der Firma BDG beruhte auf unterschiedlichen Erfassungsmodellen für selbständig (Monographien, Zeitschriften, mehrbändige Werke) und unselbständig erschienene Publikationen (Beiträge aus Zeitschriften und Sammelwerken). BDG erkannte, dass Monographien und Einzelbeiträge unterschiedliche Anforderungen an die Erfassung stellen: die Konvertierung der Bücher erfordert den Einsatz bibliothekarischer Fachkräfte. Die Struktur der Beiträge dagegen ist einfach, die Erfassung kann durch Schreibkräfte kostengünstig erfolgen – zumal der Katalog doppelt so viele Beiträge wie Büchernachweise enthält.

Vorab wurden von der Firma alle nicht zu erfassenden Karten wie z.B. Leitkarten und Barcode-Karten aussortiert sowie die Titeltkarten für Beiträge von denjenigen für Monographien und Zeitschriften getrennt. Monographien und Zeitschriften werden von Bibliothekaren über Telearbeitsplätze direkt in Kallias nach festgelegten Stan-

dards erfasst. In großem Umfang können Fremddaten genutzt werden. Durch die Übertragung der Barcode-Information wird beim Speichern des Datensatzes automatisch die Relation zu den Normdaten in Kallias hergestellt. Die Verknüpfung von Stücktiteln mit der zugehörigen Reihenaufnahme oder die Verknüpfung weiterer Personen und Körperschaften, die nicht durch einen Barcode vorab codiert werden konnten, erfolgt durch die Bibliothekare. Über eine zusätzliche Verarbeitung wird nach dem Speichern im Dienstsysteem automatisch ein Link zum Original-Image hergestellt, anhand dessen Unklarheiten und Fragen rasch beantwortet werden können. Das Online-Verfahren läuft seit November 2007 in der Produktivphase.

Die Erfassung der Beiträge erfolgt in China. Zur Vermeidung von Tippfehlern werden die Informationen im Double-Keying-Verfahren erfasst.¹³ Um in einem eindimensionalen Zettelkatalog dem Benutzer mehrere Sucheinstiege zu ermöglichen, war es notwendig, Titeltkarten zu kopieren und je nach Dokumentinhalt unter verschiedenen Systematikstellen einzulegen. Dies ist in einem Online-Katalog nicht mehr notwendig, hier möchte man die Sacherschließungsinformationen in einem Datensatz vereinigt haben. Dafür wurden die Titeltkartenkopien mit ihren Barcodes über ein Match&Merge-Verfahren¹⁴ zusammengeführt. Die Verknüpfung mit den bereitgestellten Norm-

¹³ Beim Double-Keying-Verfahren werden Texte i.d.R. zweimal abgeschrieben und anschließend maschinell miteinander verglichen. Die Schreibfehlerquote wird dadurch erheblich gemindert.

¹⁴ Das Verfahren ermöglicht die maschinelle Zusammenführung von Titeldubletten, dafür werden mit Hilfe festgelegter Attribute (wie z.B. Erscheinungsjahr, Verfasser, Quelle der Publikation etc.) Algorithmen formuliert und eingesetzt.

daten für Personen und Quellen, d.h. zur übergeordneten Zeitschrift oder zum Sammelwerk wird halbautomatisch hergestellt. Die Ergebnisse werden im standardisierten Bibliotheksformat (MAB2) geliefert und anschließend in Kallías importiert. Da bei einem Offline-Verfahren die Erfassungssysteme (BDG-Datenbank und Kallías) nicht gekoppelt sind und die Normdateien sich unabhängig weiterentwickeln, müssen die Normdaten nach jeder Lieferung abgeglichen werden. Weiterhin wird eine Aktualisierung der BDG-Datenbank notwendig. Mit diesem Offline-Verfahren können gegenüber herkömmlichen Online-Erfassungsmethoden bei etwa 500.000 Beiträgen 40% der Kosten eingespart werden und eine zügige Konversion wird ermöglicht. Das Offline-Verfahren läuft seit Juni 2008 in Produktion, die Lieferungen umfassen jeweils ca. 100.000 Sätze.

Da die Abrechnung auf Stücktitelbasis erfolgt und unterschiedliche Preise für Neuaufnahmen, Fremddatenübernahmen, Mehrfachexemplare, Aussortierungen oder Barcode-Zutragungen gelten, stand am Anfang die Mathematik und eine genaue Kataloganalyse. Man sollte hier möglichst viel Aufwand hineinstecken und erprobte statistische Verfahren anwenden.

Ebenso wichtig wie das Zählen am Anfang ist eine Qualitätskontrolle. Eine gemeinsam mit dem Referat Wissenschaftliche Datenverarbeitung entwickelte datenbankgestützte Prüf- und Dokumentationsanwendung auf Web-Basis ermöglicht es, die Lieferungen formal lückenlos sowie interaktiv nach den geforderten Qualitätskriterien zu überprüfen und den Gesamtüberblick über den Bearbeitungsstand zu behalten.

Die Erfassungsqualität der Firma BDG ist hoch, dennoch bedarf es regelmäßiger Kontrollarbeiten nach Stichproben und einer ständigen lokalen Ansprechperson für die Firma in Erschließungsfragen. Diese Aufgaben werden von zwei Bibliothekarinnen neben ihrer Alltagsarbeit wahrgenommen. Für den Online-Teil sind die Bereinigungs- und Korrekturarbeiten in den normalen Monographien-Geschäftsgang integriert und somit auf viele Schultern verteilt. Dazu gehören insbesondere auch Autopsien am Bestand. Vor allem exemplarbezogene Informationen, wie z.B. Signaturänderungen, Mehrfachexemplare wurden auf den Karten nicht konsequent nachgetragen. Letztendlich ist jedes Retrokonversionsprojekt auch ein halbes Revisionsprojekt. In ihren Katalogen erzählt eine Bibliothek immer auch ihre eigene Geschichte.

4. Mitteleinwerbung für Retrokonversion – warum es im Bibliotheksbereich heute so schwierig ist

Klaus Haller schrieb vor 20 Jahren „[...] Retrospektive Katalogkonversion ist das Zauberwort, das uns seit 1980 in der Bundesrepublik fasziniert und motiviert“¹⁵ und nennt in einem Atemzug mit „Altbestandskatalogisierung“ die Schlagworte „Fremddatennutzung“, „Online-Katalog“, „Datenformat“, „MAB-Schnittstelle“, „Normdateien“, „Internationale Katalogisierungsregeln“ und „Freitextretrieval“ als pauschale Forderungen und Wünsche. Die Anfänge der überregionalen Normdateien, der Austauschformate und der Ruf nach Internationalisierung sind eng verknüpft mit den Anfängen der Retrokon-

¹⁵ Klaus Haller, Altbestandserfassung oder Altbestandskatalogisierung?, in: Bibliothek 13, 1989, S. 3-10, hier S. 3.

version in Europa. Viele dieser Anforderungen sind heute erreicht, manches wieder aktuell in der Diskussion. Dennoch ist die Anzahl der in Deutschland zu konvertierenden Altdaten immer noch fast so hoch wie 1988. Bis heute gibt es in den Onlinekatalogen für die Erscheinungsjahre 1850 bis 1970 eine Nachweislücke, wie eine aktuelle Statistik des Südwestdeutschen Bibliotheksverbundes anschaulich belegt.¹⁶ Was sind die Gründe?

1988 hielt es der Wissenschaftsrat in seinen Empfehlungen zur retrospektiven Katalogisierung an wissenschaftlichen Bibliotheken für dringend geboten, „nicht nur die historischen Altbestände, sondern auch die Katalogdaten nach 1850, insbesondere nach 1945, zu konvertieren. Da in den wissenschaftlichen Bibliotheken nachweislich die neuere, nach dem Zweiten Weltkrieg erschienene Literatur die höchste Benutzungsfrequenz aufweist (mehr als 80% der Gesamtausleihe) [...]“.¹⁷ Bund, Länder und DFG waren aufgefordert, die Mittel für eine möglichst rasche und flächendeckende Einführung von EDV-Katalogen und die dafür notwendige retrospektive Katalogisierung bereitzustellen. Der Kostenrahmen betrug 150 Millionen DM für rund 30 Millionen nach 1945 erfasste Titel.¹⁸

¹⁶ Bibliotheksservice Zentrum Baden-Württemberg (Hrsg.), BSZ-Kompakt 2009, Konstanz 2009, S. 15.

¹⁷ Empfehlungen des Wissenschaftsrates zur retrospektiven Katalogisierung an wissenschaftlichen Bibliotheken: Konvertierung vorhandener Katalogaufnahmen in maschinenlesbare Daten, in: Zeitschrift für Bibliothekswesen und Bibliographie 35, 1988, S. 423-437.

¹⁸ Die Kostenschätzung wurde auf der Grundlage der Deutschen Bibliotheksstatistik erstellt und berücksichtigt nur Universitätsbibliotheken (22,3 Mio. Titel) sowie die Zentralen Fachbibliotheken, die Deutsche Bibliothek, die Staatsbiblio-

1991 konstatiert Hansjürgen Maurer in der Zeitschrift für Bibliothekswesen und Bibliographie: „Wie wir alle wissen, ist trotzdem bis heute unser Traum immer noch nicht Wirklichkeit geworden, denn die Empfehlungen des Wissenschaftsrates trafen uns mehr oder weniger konzeptionslos an.“¹⁹ Zuletzt 2006 musste Werner Holbach von der Bayerischen Staatsbibliothek feststellen: „[...] nach wie vor gilt [...]: Die Retrospektive Konversion [...] stellt neben den neuesten technischen Entwicklungen [...] eine der größten Herausforderungen dar, der sich Bibliotheken, insbesondere die Bibliothekare, noch stellen müssen.“²⁰

Spätestens seit dem DFG-Positionspapier „Ziele und Struktur des Förderungsprogramms Kulturelle Überlieferung“ von 2002 heißt das Zauberwort für Drittmittelinwerbung in Bibliotheken nicht mehr Retrokonversion sondern Retrodigitalisierung.²¹ Braucht man angesichts von Google Books und Worldcat überhaupt noch eigene bibliographische Metadaten? Die Antwort der Bayerischen Staatsbibliothek, die mit beiden Giganten kooperiert, lautet eindeutig ja: „Nur ein

theek München und fünf ausgewählte Bibliotheken (7,44 Mio. Titel). Man ging von einem Richtwert von 5 DM pro konvertierter Katalogaufnahme aus. Vgl. Empfehlungen des Wissenschaftsrates (wie Anm. 17), S. 436.

¹⁹ Hansjürgen Maurer, Retrospektive Katalogkonversion in einem Verbundsystem, in: Zeitschrift für Bibliothekswesen und Bibliographie 38, 1991, S. 109-128, hier S. 111.

²⁰ Holbach, Retrokonversion (wie Anm. 10), S. 207.

²¹ Vorschläge des Unterausschusses für Kulturelle Überlieferung; durch den Bibliotheksausschuss der Deutschen Forschungsgemeinschaft im Oktober 2002

nach bibliotheksfachlichen Mindeststandards erschlossener Bestand erlaubt vernünftige Aussagen z.B. über etwaige Priorisierungen von Teilbeständen bei der Digitalisierung und insbesondere fundierte Schätzungen der dabei anfallenden Kosten. Ein Kernanliegen der Bibliotheken muss [...] daher sein, dass ihre Bestände [...] möglichst rasch in Datenbanken nachgewiesen und recherchierbar sind.“²²

Die Empfehlungen des Wissenschaftsrats von 1988 trafen auch die Bibliothek des Deutschen Literaturarchivs „konzeptionslos“ an. Nach einer langen Planungsphase wurde die Katalogisierung in den sammelnden Abteilungen des DLA im Jahre 1999 auf EDV umgestellt. Erst damit war die Voraussetzung für die Konversion der Kataloge geschaffen. 2004 stellte die Bibliothek bei der Deutschen Forschungsgemeinschaft einen Antrag zur Konversion des Marbacher Sachkatalogs. Ziel war der Ausbau der germanistischen Fachdatenbank Kallias zu einem Kernelement der Virtuellen Fachbibliothek Germanistik, zu deren Kooperationspartnern das DLA gehörte. Die DFG finanziert ein Fünftel des Projektvolumens, weitere Geldgeber sind die Robert Bosch Stiftung, der Bund und die Hertie-Stiftung. Den höchsten Anteil trägt die Deutschen Schillergesellschaft als Trägerverein des DLA. Die vollständige Mitteleinwerbung hat insgesamt vier Jahre gedauert. Das Bibliotheksservice-Zentrum Baden-Württemberg hat das Projekt durch Einbindung weiterer Fremddatenquellen wesentlich unterstützt.

verabschiedet. DFG-Positionspapier: Ziele und Struktur des Förderungsprogramms Kulturelle Überlieferung, [Bonn 2002].

²² Holbach, Retrokonversion (wie Anm. 10), S. 207.

Bei einem Projekt dieser Größenordnung muss für die Vergabe öffentlicher Aufträge eine EU-Ausschreibung durchgeführt werden. Der Aufwand für die Erstellung eines Pflichtenhefts, sodann Ausschreibung, Angebotssichtung- und -bewertung bis zur Vergabeentscheidung und rechtlich abgesicherte Vertragsschließung darf besonders in der Zeitplanung nicht unterschätzt werden. Dass sich der Aufwand jedoch lohnt und Retrokonversion ein wichtiger und unerlässlicher Schritt zu einer zukunftsfähigen, nutzerorientierten Präsentation des Gesamtbestands einer Bibliothek ist, soll zum Schluss deutlich gemacht werden.

5. Ausblick unter dem Aspekt FRBR

Aus der Nostalgie für alte Kartenkataloge können Zukunftskonzepte und neue Regelwerke entwickelt werden. Ein Beispiel stellen die Functional Requirements for Bibliographic Records (kurz FRBR)²³ dar, die das Grundgerüst für die neue Regelwerksgeneration Resource Description and Access (RDA)²⁴ bilden.

Die FRBR sind ein Versuch, das bibliographische Universum abzubilden. Ausgehend vom „Werk an sich“ (Work) sollen die einzelnen Ausgaben in verschiedenen Sprachen (Expressions), in ihren unterschiedlichen Veröffentlichungsformen (Manifestations) bis hin zu den Einzelexemplaren (Items) mit ihren Relationen zu Personen (Persons)

²³ <<http://www.ifla.org/en/publications/functional-requirements-for-bibliographic-records>> [17.12.2009].

²⁴ <<http://www.rda-jsc.org/rda.html#drafts>> [17.12.2009].

und Körperschaften (Corporate Bodies) für die Nutzung übersichtlich präsentiert werden.

Übertragen auf das Selbstverständnis des Deutschen Literaturarchivs und bildlicher ausgedrückt hieße das: das Deutsche Literaturarchiv bildet in seinen Sammlungen das Netzwerk des literarischen Lebens in all seinen Facetten ab. Es dokumentiert die Literatur von der Handschrift, also dem Entstehungsprozess eines Werkes, über die verschiedenen Ausgaben dieses Werkes sowie dessen Rezeption in Rezensionen und Zeitungsartikeln; die zum Werk geschaffenen Graphiken, seine dramaturgischen Umsetzungen in Hörfunk- und Fernsehinszenierungen, im Film oder auf Tonträger bis hin zum individuellen Exemplar, das durch Herkunft und Gebrauch ausgezeichnet sein kann.

Im alten Systematischen Katalog ist die Idee von "Work" und "Expression" in Ansätzen verwirklicht: die Primärliteratur eines Autors ist nach dem Erstveröffentlichungsjahr der Werke geordnet, die einzelnen Ausgaben folgen in der Chronologie ihres Erscheinens, im Anschluss und im Alphabet der Sprachen dann die Übersetzungen eines Werkes. Auch in der Magazinaufstellung spiegeln sich diese Preußischen Katalogisierungsregeln wider: Werkausgaben, Werkauszüge, Teilsammlungen, Briefe und Gedichte stehen jeweils beisammen. Im Anschluss an die Primärliteratur folgt die Sekundärliteratur zu einem Autor. Vorm Regal stehend überblickt man so die gesamte monographische Werk- und Wirkungsgeschichte eines Autors.

Die Marbacher Kartenkataloge reichen näher an die Idee der Functional Requirements for Bibliographic Records heran als moderne

Onlinekataloge. Aber ihnen fehlt ein entscheidendes Kriterium der FRBR: die Relationen – das Beziehungsgeflecht.

Bei der Konzeption des Lokalsystems Kallías ging es darum, das Netzwerk der Bestände sichtbar zu machen. Im Bereich der Personen ist diese Integration gelungen. Über den Einstieg "Namen" erhält man als Ergebnis unter dem Namen eines Autors oder einer Autorin die verzeichneten Bestände der sammelnden Abteilungen in der Reihenfolge Bibliothek/Mediendokumentation, Handschriftensammlung, Bilder und Objekte gemeinsam präsentiert, gegliedert nach den Relationen "Von", "An", "Über", "Unter". Jeder 4. Internetbenutzer wählt diesen Einstieg. Wünschenswert ist ein entsprechender Einzeilen-Einstieg auch für Werke. Zwar gibt es auch für Werke einen Normdatensatz, der Kurzinformationen und weitere Titel enthält, aber dieser Werktitelsatz wird nur in der Relation „Über“ verwendet. Die Nachweise der Primärliteratur werden nicht mit diesem Werktitelsatz verknüpft.

Wie in den meisten Katalogen wird auch in Kallías vor allem mit Titel und Person gesucht. Genau an diesem Punkt, dem Suchverhalten, wird die Idee der FRBR für die Präsentation der Literaturarchivbestände interessant: nicht als Ersatz für bisherige Suchmöglichkeiten, sondern als zusätzliches Angebot. Die Bestände des DLA könnten wie in einem biobibliographischen Nachschlagewerk präsentiert werden – ein Gedanke, der auch der Konzeption des Systematischen Katalogs vor über 50 Jahren zugrunde lag. Damals wollte man die Nachfolge von Goedekes Grundriss der deutschen Dichtung aus ihren Quellen antreten. Eine Verlinkung zu anderen Online-Nachschlage-

werken, die Werke der Weltliteratur einzeln nachweisen, wie z.B. Wikipedia, wäre problemlos möglich.

Eine Aufgabe der nächsten Jahre besteht darin, die intellektuelle Leistung der Kartenkataloge für den Onlinekatalog weiter produktiv zu machen, z.B. die Nachweise der Primärliteratur um die Verknüpfung zum Werktitel durch halbautomatische Verfahren rückwirkend anzureichern.

Anfang 2011 wird die Überführung des konventionellen Sachkatalogs und die Anreicherung der Datenbank Kallías abgeschlossen sein. Dann ist einer der umfangreichsten und bedeutendsten Fachkataloge der Germanistik komplett auch im Rahmen des Fachportals "Germanistik-im-Netz"²⁵ nach dem Prinzip des offenen Zugangs für Wissenschaftler, Studenten, Kulturschaffende und literarisch Interessierte zugänglich.

²⁵ <<http://www.germanistik-im-netz.de>> [17.12.2009].

Archival Exchange Formats: Developments and Future Issues

by Michael Fox

1. Introduction

This article focuses on my remarks about the ways in which developments in the archival information landscape are affecting the exchange standards that we use to populate those systems. My comments fall into in three parts.

They begin by describing briefly the reasons why I believe we need to reexamine our current exchange standards. Next I will offer a more detailed analysis of the conceptual models of the archival systems that those standards are intended to support. Those models have evolved since they were first conceptualized as the result of the new theories about the nature of archival information and as the consequence of practical experience gained in the implementation of new systems, networks and portals. What changes do we have to make to these formats to support the exchange of information that facilitates the work of researchers? Finally, I will outline a number of specific modifications to existing standards that might enhance their usefulness in other ways.

In this presentation, I am assuming that you are acquainted with at the least the following basic ideas.

- a) There are international standards that prescribe the structure, content, and exchange formats for archival information as illustrated by this chart.

<i>Data Type</i>	<i>Data Fields and Content</i>	<i>Exchange Format</i>
Records	ISAD(G)	EAD
Creators	ISAAR (CPF)	EAC-CPF
Functions	ISDF	EAC-F?
Repositories	ISDIAH	EAG or EAC-IAH

- b) The ICA has defined four standards that prescribe the data elements, their structure and contents for archival information.
- General International Standard Archival Description – *ISAD(G)* – which prescribes the content of 26 core elements for archival finding aids.¹
 - International Standard Archival Authority Record for Corporate Bodies, Persons and Families – *ISAAR(CPF)* – which prescribes the data elements needed to describe persons and organizations associated with the creation or use of archival records.²

¹ International Council on Archives, ISAD (G) General International Standard Archival Description. 2a edition, Madrid 2000.

² International Council on Archives, ISAAR (CPF) International Standard Archival Authority Record for Corporate Bodies, Persons and Families, Paris ²2004.

- International Standard for Describing Functions – *ISDF* – which prescribes the data elements needed to describe the administrative functions that generate archival materials.³
 - International Standard for Describing Institutions with Archival Holdings – *ISDIAH* – which defines a set of elements that provide information about archival repositories.⁴
- c) There are also national standards in many countries that address the content of archival descriptions and archival descriptive systems.
- d) For three of these content standards, there are parallel internationally recognized protocols for the exchange of data. All are based on Extensible Markup Language (XML).

Encoded Archival Description (EAD) was developed to enable the standardized electronic representations of finding aids.

Encoded Archival Content (EAC-CPF) is an XML schema for the representation of information about individuals, families and corporate bodies that are associated with records in a variety of ways. Under development for many years, the final version should be released soon by its maintenance agency, the Staatsbibliothek zu Berlin.

³ International Council on Archives, *ISDF International Standard for Describing Functions*, Paris 2008.

⁴ International Council on Archives, *ISDIAH International Standard for Describing Institutions with Archival Holdings*, Paris 2008.

Encoded Archival Guide (EAG) is a protocol developed in Spain for recording information about archival repositories themselves. It is related to but not entirely parallel to ISDIAH.

There is no standard protocol for the exchange of information about functions. It remains to be seen whether a new standard will be developed or EAC-CPF will be extended to include the unique data elements for the description of functions.

2. What developments are affecting these three exchange standards?

EAC is just about to be released so it will be some time before we have sufficient practical experience to identify any problems. For information on repositories, we await the alignment of EAG with ISDIAH. Whether this process results in a major revision to EAG or can be accommodated by mapping elements between the content protocol (ISDIAH) and the exchange format (EAG) remains to be seen.

On the other hand, there are several developments that have the potential to affect EAD and to which the EAD community needs to react. They include the following issues.

- a) EAD is being used in increasingly larger systems. No longer limited to local and regional content, these networks are often national and international in coverage. They are also becoming more diverse in the nature of the data they include. Both factors result in greater complexity.
- b) There is a need for consensus on the way in which EAD will interact with other exchange formats, EAC and EAG to be sure,

but also with METS, the standard for managing digital objects. We must also consider the library standard MARC21, which has been adopted by German libraries as the data exchange format for bibliographic information.

- c) Certain technical choices in the EAC schema, such as the minimization of mixed content, reflect an evolving understanding of encoded information. Should these developments be incorporated into EAD as well?
- d) Several open source applications for the management of archival processes have appeared in the last few years. These include Archon and the Archivist Toolkit in the U.S. and ATOM, Access to Memory, being developed in Canada for the ICA. They all import and export data in EAD and EAC. But are there particular requirements of such systems that we need to account for?

3. Current models

What are the current models for archival information exchange? How are they evolving? What are the implications for our exchange formats? To describe these models, I will use the term archival information package. Let us call it an AIP. An AIP is a set of informational elements that are bundled together and shared with others as a unit.

In the library world, the catalog record is such an information package, containing data elements such as author, title, and publisher. Libraries have a second type of information package, the authority record. Some library systems share authority data with the public as information of interest in its own right. In other systems, it

is only available to staff, being seen as secondary and useful principally to facilitate the cataloging process.

For the most part, archives have historically had a single information package, whether we call it an inventory, finding aid, register, or a findbuch. While the data elements found in finding aids have varied over time and between different archival traditions, these single information packages typically contained some or all of the following types of information about the context and content of records from a variety of perspectives.

Repository: Name, address, contact information,

Contextual: Name of creator, biography/history, functions,

Content: A narrative description of the scope and nature of the records, and references to places, persons/ organizations, documentary forms, topics, and time periods associated with the materials,

Control: Administrative data,

Digital: References and/or links to electronic surrogates of original materials or born digital archives.

This commonality of tradition was codified, though not exhaustively, in ISAD(G). The data formulated according to this standard was expected to be contained in a single print or online document, a single archival information package that would be presented to the researcher.

EAD assumed this approach, though it provides for the representation of many additional data elements not found in ISAD(G). An

EAD instance is nothing more than the representation of a single finding aid in electronic form.

But before the ink was even dry on the first draft of ISAD(G), there was a call for a second standard for authority data. ISAAR(CPF) followed shortly. In this model, there would be two archival information packages, one describing archival materials and a second with rich content about their creators. Systems would link the two.

In the same vein, the relevant ICA committees have further parsed out portions of the traditional finding aid into two other information packages for function and repositories: ISDF and ISDIAH. This would result in an even more complex environment with content created according to four standards that would be linked together.

But if we look at all the data that might be found in traditional finding aids, we quickly see that there are many other groupings of data that could be divided up in the same way. One could manage multiple systems for data about repositories, functions, record content, documentary types, places, chronology, topics and digital surrogates and exchange information among all of them.

Which, if any, of these additional data types should be managed as separate information packages? Let me suggest two considerations in answering that question. First, do any of these sets of data elements constitute a rich and useful resource that provides important information for researchers on their own, apart from the context of the finding aid? In the case of information about record creators, the answer clearly was yes. Consider how much more

information about a record creator is found in an ISAAR(CPF) record than in a record created according to ISAD(G) or many national descriptive standards.

Information about creators in ISAD(G)/EAD

- Name of Creator
- Administrative/Biographical History
- Archival history

Information about creators in ISAAR(CPF)/EAC-CPR

- Multiple versions of names
- Related entities
- Dates of existence
- History
- Places
- Legal status
- Functions, occupations, and activities
- Mandates/Sources of authority
- Internal structures/Genealogy
- General context

An authority file with such rich content might well be consulted by a researcher for its own value.

The second consideration is one of control and use. Are the contents of these data elements likely to be created and maintained by archivists or are they under the control of other disciplines. Often the data in a finding aid comes from authority lists or thesauri from

which the archivist extracts terms in the same way he or she might consult a dictionary to identify the proper word or verify its spelling. In these situations, it seems unlikely that archivists will manage those sources ourselves.

Since the description of the records themselves is central to the archival descriptive enterprise, let us consider how the exchange of information might work from an EAD perspective, from the point of view of the finding aid, with respect to the four content standards defined by ICA. Basically, we can exchange either single or multiple information packages. Multiple scenarios are possible with either choice.

3.1. Single information package

With the single information package, we extract rich content from multiple sources, embed it in a finding aid, and share it through networks and portals. If we wish to continue to create and exchange a single information package, namely the comprehensive and self-contained finding aid, there are at least three ways we can do it with EAD.

a) Transcription

One choice is simply to transcribe the content of the data element from the external source into the finding aid. This requires no modification to EAD.

This solution appears to work well where the external sources are thesauri such as lists of place names, topical subjects, documentary types, or time periods managed by some other entity. The external source might also be an authority file system. In the following

example, a name has been transcribed from the United States Library of Congress Name Authority File into the Personal Name element of an EAD finding aid.

`<persname>`

Michael J. Fox (Michael James), 1946-

`</persname>`

b) Transcription and links

We know that the external source of such data may contain additional information of use to the researcher beyond the data that is transcribed. In such instances, do we want to go beyond simple transcription and provide links to such external resources?

Additionally, this data may be of interest to the archival information system itself for other data manipulation. Often the external data element exists in an exchange format of its own, such as an EAC record in an archival authority record or a MARC21 record in a library authority file. There might be utility in retaining the structure of those exchange formats within the EAD document to facilitate transfer of data between descriptive and authority systems. Data might flow not only into the EAD finding aid but also back from the finding aid into the external authority file system. Data about a record creator in an EAC formatted record might end up in an EAD-structured document but at some point might also be contributed from a finding aid to an EAC structured authority file. There are two obvious ways to accomplish this. Either technique would require modification to EAD.

In the first option, we could *extend EAD* by adding elements from other exchange formats to the EAD element set. In this example, the simple EAD personal name element is enriched by extending its content model to include the full structure of a MARC authority record.

```
<persname >
  <marctag="100" ind1="1" ind2="" >
    <subfield code="a">Michael J. Fox
  </subfield >
    <subfield code="q">Michael James
  </subfield >
    <subfield code="d">1946
  </subfield >
  </marctag >
</persname >
```

This would make the EAD structure larger and more complicated and subject to regular revision to keep it in synchronization with the latest versions of other standards.

The second option involves the use of the *XML namespace* convention. Data in the syntax of one XML namespace, MARC XML in this example, could be embedded directly into another XML namespace, here EAD.

```
<ead:persname >
  <marc:data-field tag = "100" ind1 = "1" ind2 = "" >
    <marc:subfield code = "a" > Michael J. Fox
  </marc:subfield >
    <marc:subfield code = "q" > Michael James
  </marc:subfield >
    <marc:subfield code = "d" > 1946
  </marc:subfield >
</marc:data-field >
</ead:persname >
```

No major alteration to EAD would be required beyond adding the general feature of enabling the insertion of other namespaces. Since this would be optional, it would be a backwardly compatible change. There are very vocal proponents of this approach as it requires less maintenance to keep EAD current with other XML schemas. But there are technical limitations in the lack of software support for multiple namespaces within a single XML file. Moreover, the possibility of having data from various namespaces embedded at different places within an EAD-based document would seem to create potential, possibly significant processing complications for systems.

3.2. Multiple Information Packages

Exchanging multiple information packages is our second major option for system design. Information exchange would rely on the linking capabilities of systems, including the web itself, to bring together data from multiple sources into a single package that would

be presented to the researcher. Again there are multiple ways to achieve this.

One choice is a hybrid to transcribe data from external sources such as a name authority file into the finding aid but also include links to those sources as in this example (*Transcribe and Link*). One could imagine archives exchanging as well as linking to at least four types of information packages that we create and manage – regarding creators, functions, and repositories, and descriptions of records themselves. This example illustrates how a name from an external authority file might be embedded in an EAD document which also provides a link to that authority file system.

<persname>

Michael J. Fox (Michael James), 1946-

<extref linktype = "simple"

href = "www.loc.gov/lcnaf/96021296"

actuate = "onrequest"

show = "new" >

More information is available about this individual at this location.

</extref >

</persname>

Archivists could be exchanging data with multiple external information systems and also creating links to them for the user. There is considerable appeal to this approach since such links are built into

the ICA content standards ISAAR (CPF), ISDF, and ISDIAH. Adding them to ISAD(G) seems a logical next step. The technical mechanism for linking is built into EAC but EAD would have to be modified to actually accomplish the linking illustrated above. But of course links have limitations as well since they must be created manually and maintained continuously into the future.

The concept of the *semantic web* is a possibility that would theoretically eliminate the need for explicit links. As I understand this model, we would not exchange data between systems but rather create a comprehensive search engine that understands the relationships between syntaxes of multiple information packages. For example, it would understand that the syntax of the data element `<datafield tag="100">` in a MARC authority record in the following example

```
<collection
xsi:schemaLocation="http://www.loc.gov/MARC21/slim
http://www.loc.gov/standards/./schema/MARC21slim.xsd">
  <record>
    <datafield tag="010" ind1=" " ind2=" ">
      <subfield code="a">96021296 </subfield>
    </datafield>
    <datafield tag="100" ind1="1" ind2=" ">
      <subfield code="a">Michael J. Fox</subfield>
      <subfield code="q">Michael James </subfield>
      <subfield code="d">1946</subfield>
    </datafield> ...
```

is the equivalent of the Personal Name <persname> element in the following excerpt from an EAD document.

```
<ead> ....  
  <persname source = "LCNAF" authfilenumber = "96021296" >  
    Michael J. Fox (Michael James), 1946-  
  </persname> ...
```

A powerful search engine would then assemble this data from multiple, independent systems and present the integrated results of that search.

Theoretically, such a system would minimize the human intervention required to create and maintain linkages across information packages. The principal challenge seems to be one of mapping the relationships of data elements across systems. While there are small-scale applications of this concept, many remain skeptical of the ability of such systems to scale up to large volumes of heterogeneous data.

4. Further changes to EAD

Many of these delivery options would require changes to EAD. It is my believe that there are also number of other possible modifications that would improve the usability of EAD regardless of the particular information delivery and exchange systems we employ. They are mentioned here only briefly as their implications probably will be fully understood only by those who are familiar with the details of EAD.

Several of these ideas reflect a widespread interest in simplifying the structure of EAD. Frankly, there were a number of features incorporated into EAD when it was first developed that seemed to be good ideas at the time, reflected the interests of certain constituencies, or were responses to the existing SGML technology. EAD was designed as a big tent that could accommodate a wide range of existing practice, both local and international. But now circumstances have changed. Some features have not proved to be so useful as imagined, technology has changed, and flexibility has created problems with training and implementation. On the other hand, one must exercise great caution with changes that are not backwardly compatible lest they create problems for the installed base of users. The following ideas are presented, not because I believe that they all should be adopted but because they need a fair discussion and final resolution as part of the next update cycle for EAD.

Minimize mixed content. One of the interesting features of EAC-CPF has been the almost complete elimination of mixed content, that is, elements that can have both text and other elements as children. Apparently, there is only a single element in EAC that contributes to mixed content and then only for a single purpose, namely the textual manipulation for presentation that is unrelated to the underlying data structure. The chief argument against mixed content seems to be the complications it creates for database applications. The contrary argument is that the technical systems should support the nature of the data and not the other way around.

Minimize content designation. Much mixed content could be avoided simply by eliminating many of the children of the descriptive

identification <did> element. The <unittitle> element is particularly rich with child elements that are probably seldom if ever used. Additionally, it may be time to eliminate <unitdate> as a child of <unittitle> all together.

Eliminate some wrapper elements. Whether is inherent in the archival personality or derives from our training, this profession likes to organize and categorize information into units of similar content. Aspects of EAD reflect this proclivity. Sometimes, it seems to be carried to excess. With the revision in EAD2000, the administrative information and adjunct descriptive data elements were eliminated as wrapper elements due to a lack on consensus as to their structure or utility. There was no discernable loss of functionality. In addition to the elements mentioned in the previous section, the description of subordinate components bears scrutiny. It was created simply as a solution to a particular limitation of the SGML stylesheet languages then available but is completely anachronistic both conceptually and technically in the age of XML and XSLT functionality.

Extend certain content designation. This idea may seem to contradict my previous suggestion to minimize content designation. But there are several situations where the universal availability of additional content designation would be beneficial. However, I am suggesting that these be added as optional attributes and not as data elements. The first would be an attribute to specify the language of any given element. This would facilitate the presentation of multilingual finding aids.

The second idea would be the addition of the label attribute to every relevant element. It can be argued that the head element is

simply a stylistic convention, a matter of local presentational preference that should not be considered part of the archival description. There is also the confusion caused by the use of the label attribute that accomplishes the same result but which appears only within the descriptive identification element. One could argue that the label attribute be added to every element. Use of the head element could be deprecated over time.

The third idea would be some mechanism for indicating for every element not only an appropriate encoding analog but also the standard to which the analog relates. As it is currently, one can specify only a single standard per finding aid. The encoding analog and its parent standard also need to be repeatable in some way within a particular element. Different elements within a finding aid may relate to different standards and a single element may have equivalencies in multiple standards. For example, a biography or administrative history element may correlate to both a MARC field and an EAC element.

I need to reiterate at this point that I am not so much suggesting answers as posing questions that others ought to consider with the next revision of EAD.

What else lies ahead? One clear trend is the international interest in integrating access to the holdings of libraries, archives and museums. Memory organizations like ours will participate inevitably in a richer, more diverse, and more complex resource access environment that includes metadata and digital representations of materials from archives of every sort and with every form of documentary resource, from the bibliographic world, and from the collections

of museums of all types. Best wishes to all you digital natives in expanding our horizons.

Exchanging Archival Descriptions: The Role of EAD and EAC in the Construction of the SAN Portal in Italy

by Stefano Vitali

1. The National Archives System

During the second half of the 20th century, in Italy as well as in many other countries of the developed world, the archival landscape underwent deep changes that helped reshape its territorial organization and the balance between its various institutional components. At the base of those changes are some typical phenomena of contemporary society on which archive literature has widely dwelled:

1. the advent of new historiographic trends, addressed at documentary sources that were generally neglected or undervalued up to that time;
2. the interest in recovering individual and collective memories, as well as local and territorial traditions; consequently, the new attention in archives as vehicles of memories and traditions and the rise of new expectations in them from an ever-growing, culturally and socially heterogeneous public;
3. the increasing importance and activism of local and regional institutions and of various cultural and social organizations who headed up a crisis in the traditional balance between central and local powers, favoring the second as opposed to the first.

In Italy, the outcome of these changes was the gradual exhaustion of the historical records concentration and keeping model

established after the unity (1861), which was based on the centrality of a network of provincial State Archives, controlled by a central archival administration in Rome. A different model, gradually affirmed for the past few decades, is based on the coexistence of many archival holders (archives of municipalities, provinces, regions, cultural institutions, research centers, business archives, etc.) and on the development of various initiatives for collecting, describing and promoting archives, supported by a large number of public and private bodies.

This dissemination of archival institutions and initiatives presents many problems of connection and coordination and urges archivists and state, regional and local administrators to establish new forms of collaboration, which are particularly important in a phase like the current one where financial resources continue to decrease and the risk of dispersing them without lasting outcomes is increasing. Therefore, as claimed by many parts, it is urgent to develop sustainable models of archival polycentrism capable of improve its positive features, while limiting the most critical, not to say negative, ones, such as an excessive dispersion and fragmentation of archival institutions and projects, which probably users of archives find unreasonable and which, at the same time, risk to put in danger the preservation of the extremely rich and precious Italian archival heritage.

The Second National Conference of Archives, held in Bologna from November 19 to 21, 2009, was dedicated to the need to "make a system" for governing the archival polycentrism. It was an important moment for debates and proposals where representatives

from State archives administration, regions, local authorities, cultural institutions and other public and private organizations discussed the need for coordination and cooperation on different fronts: from the construction of common repositories for archives to common strategies for digital preservation, from archival education to initiatives for promoting and publicizing archives. The result was the design of a National Archival System (SAN) to be built in the near future with its own national government bodies, regional coordination committees and common repositories and archival services at the local level.¹

2. The National Archives Portal and the Archives Resource Catalog

The core component of the National Archival System will be a web portal, which has momentarily been given the name Portale Archivistico Nazionale (National Archival Portal) or PAN, which should present itself as the integrated access point to national archival resources, irrespective of the juridical status and affiliation of the institutions or organizations that developed them. The portal, the planning and realization of which began in 2009 by the Directorate General for Archives of the Ministry for cultural heritage and activities, will be divided into multiple sections and will carry a lot of complex contents, such as:

¹ See the Conference Website <<http://www.conferenzanazionalearchivi.beniculturali.it/index.php?it/1/home>>, where all the preparatory documents and the final one can be accessed. All mentioned Websites have been accessed on 22/02/2010.

- an “Archipedia”, which is a sort of encyclopedia of archive definitions and concepts compiled in cooperation with the portal’s users and including a glossary which will provide simple and concise explanations of archival technical terms for novices and unskilled users;
- a database of bibliographic resources related to Italian archives;
- research guides, virtual tutorials and other materials for different typology of audience (for example: teachers, students, genealogists, historians, etc.), including novices and unskilled ones;
- specific thematic sections or sub-portals dedicated to describe and make available different typologies of records (business, genealogical, cartographical and fashion archives etc.);
- editorial and multimedia contents, including news, virtual exhibits, photo galleries etc. for illustrating the multiple aspects of the world of archives and records;
- a digital archive that allows access to digital reproductions of fonds and series published on Websites of local, regional or national archival institutions;
- Web 2.0 tools for communicating with the users of the portal, allowing them to collaborate in creating its contents and offering them the possibility to build communities on specific topics and research projects.

The central component of the PAN will be the Catalogo delle Risorse archivistiche (Archival Resource Catalog) or CAT that proposes to be a coordinating and integrated access point to archival descriptions

stored on archival databases and systems developed at the regional and local level, respecting their autonomy and specificity.

In fact, one way that the archival polycentrism mentioned above has been manifested in Italy, was the development of many software and systems for producing digital descriptions of archives and for publishing those descriptions on the Web. Systems have been developed at the national level by the state archives administration such as the General Guide to the state archives,² the State Archives Information System (SIAS),³ the Unified System of Archive Supervising Agencies (SIUSA),⁴ and the Mediterranean Historical Multimedia Archives,⁵ by some provincial State Archives (the State Archives of Florence,⁶ Milan,⁷ Bologna,⁸ Naples,⁹

² See its first version <<http://www.maas.ccr.it/h3/h3.exe/aguida/findex>> and the new one, published online in 2009 <<http://guidagenerale.maas.ccr.it/>>.

³ See <<http://www.archivi-sias.it/>>.

⁴ See <<http://sius.archivi.beniculturali.it/>>.

⁵ See <<http://www.archividelmediterraneo.org/portal/faces/public/guest/>>.

⁶ See the Florence State Archives Information System (SiASFI) <<http://www.archiviodistato.firenze.it/siasfi/>>.

⁷ See the Online Guide to the State Archives of Milan <<http://www.archiviodistatomilano.it/patrimonio/guida-on-line/>>.

⁸ See the Archival heritage of the State Archives of Bologna at <<http://patrimonio.archiviodistatobologna.it/asbo-xdams/>>.

⁹ See the Archival heritage of the State Archives of Naples <<http://patrimonio.archiviodistatonapoli.it/xdams-asna/>>.

Venice¹⁰); by some regions (like Lombardy,¹¹ Emilia-Romagna,¹² Piedmont¹³ or Umbria¹⁴); by other territorial entities (like the Province of Trento Historical Archives);¹⁵ by individual cultural or political institutions (like the Piedmont Institute for History of the Resistance and the “Giorgio Agosti” Contemporary Society,¹⁶ the Giangiacomo Feltrinelli Foundation,¹⁷ The Senate of the Republic,¹⁸ The Chamber

¹⁰ See the Online Guide to the State Archives of Venice (SiASVe) <<http://www.archiviodistatovenezia.it/siasve/cgi-bin/pagina.pl>> .

¹¹ See the section dedicated to the description of archives on the cultural heritage portal of Lombardy Region <<http://www.lombardiabeniculturali.it/archivi/>> .

¹² See the Ibc Archives Portal <<http://archivi.ibc.regione.emilia-romagna.it/ibc-cms/>> . Ibc (Istituto per i beni culturali) is the Institute for cultural heritage.

¹³ See Guarini Web for archives <<http://www.regione.piemonte.it/guaw/MenuAction.do>> .

¹⁴ See DOC-information System for the archives of the Umbria Region <<http://www.piau.regioneumbria.eu/default.aspx>> .

¹⁵ See the online inventories on the Website of the Historical Archives of Trento Province <http://www.trentinocultura.net/catalogo/cat_fondi_arch/cat_inventari_h.asp> .

¹⁶ See ArchOS. Integrated System for the archival Catalogs <<http://metarchivi.istoreto.it/>> .

¹⁷ See the Foundation online archives <http://www.fondazionefeltrinelli.it/feltrinelli-cms/cms.find?flagfind=quickAccess&type=1&munu_str=0_6_0&numDoc=95> .

¹⁸ See the Online Archives Project of the Senate of the Republic which includes descriptions and digital reproductions of archival fonds held also by other institutions <<http://www.archivionline.senato.it/>> .

of Deputies¹⁹ and many others) or by groups of 'federated' cultural institutes (the Institute of the Resistance Network²⁰ or the project Archives of the 1900's).²¹

The presence on the Web of these multiple systems is not only the result of the way in which historical archives have been computerized in Italy, nor is it just the result of the archival polycentrism mentioned above. Actually, the dialectic between 'local' and 'national' systems of archival descriptions reflects a deeper logic that has to do with the double meaning associated with archives today. On the one hand, archives are products of specific historical and geographical contexts and, hence, sources for knowledge of their history and vehicles of specific memories and identities. On the other hand, they are *tout court* cultural heritage, thus carriers of universal significance and values that cannot be closed within restricted territorial areas, but must acquire a national, and possibly international, visibility. Therefore, not only is this multiplicity of systems not casual, nor does it constitute a sort of limits to quickly overcome, possibly through their centralization into one single system, but contrarily, it represents an undoubtedly valuable resource. Nevertheless, it is true that in the previous years there has been a growing need to establish connections, data

¹⁹ See the Website of the Historical Archives of The Chamber of Deputies <<http://archivio.camera.it/archivio/public/home.jsp?f=10371>>

²⁰ See the Guide to Historical Archives of the Resistance Institutes <<http://beniculturali.ilc.cnr.it/insmli/guida.HTM>>.

²¹ See Archives of 1900's – Memory on the Net <<http://www.archividelnovecento.it/archivinovecento/>>.

exchanges, increasing levels of interoperability between local, regional and national archival systems so that a great deal of reflections and discussions has taken place on how to create fruitful forms of collaboration.²²

By developing integrated access tools and offering essential information on the nation's archival heritage, CAT wants to build an answer to such a need and represent a tool for joining many existing systems together without substituting them. On the contrary, it gives them greater visibility and enhances their specific characteristics. An operation like this is possible – even with the diversity of the software tools used and of some aspects of descriptive formats – since the systems developed in recent years share the same conceptual model and a common adoption of the international standards of archival description. These systems are generally based on an architecture that, besides the descriptions of archival materials created in conformity with ISAD (G), include the separate description of creators (corporate bodies, families, persons) in accordance with ISAAR (CPF) and the description of custodians of archives (archival institutions, but sometimes other entities or families and persons as well), according to the International Standard for the Description of Institutions with Archival Holdings (ISDIAH).

Therefore, the CAT will sketch a general map of the national archival heritage capable of providing an initial orientation to

²² See for example Stefano Vitali, *Verso un Sistema Archivistico Nazionale? Un' introduzione*, in: *Archivi e Computer* 13, 2004, 2 (special issue), p.7-15.

researchers and guide them towards more informative resources available in the systems that will participate in the National Archival Portal. It will contain descriptive records of custodians, of fonds or archival aggregations, finding aids and creators. It will be populated and updated through procedures that privilege procedures of data *harvesting* based on the protocol for metadata harvesting of the Open Archives Initiative (OAI-PMH). Other ways of importing will not be excluded, as for example the *upload* of XML files in a specific area of the portal or the direct entering data into the CAT database, using an ad-hoc on-line interface. The purpose of these multiple implementation procedures is to let even the less technologically equipped holder of archives participate in the project.

Custodians of archival materials will be univocally identified and essential information on each of them will be provided in order to produce an authority list of all entities, institutional and non-institutional, which hold and provide access to historical archives in Italy. In addition to the essential identification data of each custodian (name, location) and a brief description of its history and current status, the availability of a reading room with regular opening hours and a reference service provided by skilled archivists will be indicated. Description and identification data, if available, will be acquired directly by the systems that participate in the PAN. The portal's editorial staff will correct, update and standardize the descriptions, if necessary, in order to provide users with reliable and up-to-date information. The CAT record for each custodian will contain, along with a hypertextual link to its Website, links to other descriptions which can be found in any of the systems participating in the PAN.

Regarding *archival aggregations*, the CAT will include the highest levels of description of each archive (fonds, or groups of fonds). It will also include lower levels of description such as sub-fonds or even series if created by specific creators, different from those of the fonds which they belong to.

The CAT will provide also concise information on the availability and main features of finding aids, existing on paper or in digital formats, for the archival aggregations described. An appropriate hypertextual link will address users to the digital ones directly accessible on the Web.

The selection of the information elements to be included in the CAT database for describing archival aggregation, creators, finding aids was prompted by a sort of principle of subsidiarity and economy and aimed to identify just those elements which are really essential. For wider and deeper descriptions, users will be addressed to the harvested systems. Consequently, only the elements considered mandatory according to the international descriptive standards together with few others regarded as such in Italian archival tradition have been included. A maximum number of characters in open-text fields will be provided in order to avoid redundant information. The description imported into CAT database from the provider systems will be published without any correction or modification.

Each CAT record will contain a direct link to the corresponding record in the original system, from which the data has been imported. Following the link, the user will access the complete description of the archival aggregation, creator or the whole finding

aid – if it exists – in its original context. The descriptions of archival aggregations will also be linked to relevant digital reproduction projects which will be made available on the portal.

Each imported record is to be connected to a CAT record describing the original system in order to provide information about the provenance of the data. Moreover, since it cannot be excluded that the same archival aggregation, finding aid or creator is described in more than one of the archival systems exporting their data into the CAT database, provision of information about original systems will help users to correctly interpret and contextualize such multiple descriptions of the same entity.

Nevertheless, in order to provide users with information of increasing quality and for adding value to records imported into the CAT database, the descriptions of creators will be linked to an *authority file* of creators which will be progressively implemented by the central staff of the portal. This authority file should not only be the primary access point for research and navigation in the CAT, as well as the tool for connecting records imported from the provider systems, but also a national point of reference for identifying corporate bodies, persons and families and recording their authority names and descriptions. Therefore, in the future, local systems will not have to produce their own description of creators, but, if they would like, they could directly refer to the national authority file for creators published on the portal. Finally, this authority file will represent a bridge towards analogous authority files developed in catalogs and descriptive systems of other cultural heritage domain, such as the National Library System.

3. SAN standards, definitions and characters

The CAT architecture, the descriptive elements to be included and the exchange formats between the CAT itself and the systems which will provide the data have been developed by some ad-hoc workgroups appointed at the beginning of 2009 by the State-Regions Joint Technical Commissions for defining archival standards. Representatives from different archival institutions and administrations as well as from various regions took part in the discussions. Besides metadata sets for the various entities described in CAT,²³ the workgroups defined export-import formats, protocols and procedures and developed standards as well as methodologies for the preparation of the descriptions of corporate bodies, persons and families to be included in the authority file of creators. During the first few months of 2010, metadata sets will also be released for digital resources which will be made available on PAN, according to the same approach used in CAT. In the portal's digital archives, so-called thumbnails of the images and essential information, including hypertextual links to the original systems, will be imported, allowing users to research among the available digital resources, make a preliminary selection and then be directly addressed to the harvested systems for quality viewing of the digital reproductions of archival documents.

²³ See Sottocommissione tecnica per la definizione dei metadati [...], Tracciati descrittivi del CAT: soggetti, <http://ims.dei.unipd.it/data/san/metadati/docs/2009-04-17_Documento-conclusivo-sui-tracciati-CAT.pdf> (provisional URL).

For exporting the descriptions from the existing systems to the CAT, an XML exchange format has been developed. It has been named "SAN exchange format". It is based on three schemas. Each one includes a subset of elements of the Encoded Archival Description (EAD) or of the recently released Encoded Archival Context (Corporate Bodies, Persons, Families; EAC).²⁴

In particular, the schema for archival aggregations is based on EAD and includes only elements from the <did> element. Due to some specific requirements of the SAN exchange format, many of the EAD elements used had to be partially adapted and modified. The schema for finding aids is based on EAD and includes only elements from the <header> element. Also, in this case some of the elements have been adapted to meet the specific requirements of the SAN exchange format. The schema for creators is based on EAC. In this case, the standard worked much better and no specific adaptation was required to develop a schema completely compliant with the characteristics of the SAN exchange format. Finally, a fourth ad-hoc schema has also been developed for importing some essential information on custodians from the systems which can provide such information. As mentioned above, these descriptions – after being imported – will be amended, standardized and updated by the editorial staff of the portal.²⁵

²⁴ See relevant documentation on the Website of the project <<http://eac.staatsbibliothek-berlin.de/>>.

²⁵ See Sottocommissione tecnica per la definizione dei metadati [...], Tracciati descrittivi e schema XML di esportazione-importazione dai sistemi aderenti al catalogo delle risorse archivistiche (CAT) Versione 1.1, (ottobre 2009),

In developing the schemas, an approach was adopted which can be defined as “recordcentric”. Every item description exported will generate just one XML record, uniquely identified by the identity code of the description in the original system. Relationships between items – even those of hierarchical nature between archival aggregations – will be explicitly recorded, making reference to the identity code of the description/s of the related item/s.

The XML schemas for the four entities (custodians, creators, archival aggregations, finding aids), integrated with control information necessary to correctly implement the import procedures, have been accommodated into a whole export-import format. It is available to all those persons responsible for the existing archive descriptive systems, who wish to export their data into the National Archival Portal and contribute to the effort in constructing a single access point to Italian archival resources on the Web.²⁶

4. Conclusions: Why the PAN experience is important

The realization of the National Archival Portal and its Catalog of archival resources is destined to represent an important turning point in the ways with which the Italian archival community has until now conceived and constructed its own presence on the Web. For the first time, the community will have at its disposal a system built from the ground up through data exported from various providers. A

<http://ims.dei.unipd.it/data/san/metadati/docs/2009-11-08_tracciati_export_import_1-1.doc> (provisional URL).

²⁶ The overall schema and relevant documentation are provisionally accessible *<<http://gilgamesh.unipv.it/cat-import/cat-import.xsd>> and <<http://gilgamesh.unipv.it/cat-import/cat-import.html> - id6> .*

system which wants to be completely placed on the inside of the Web's new horizons based on cooperation, interoperability and data reusability.

The portal is going to radically change the Italian archival landscape, with significant impacts on the existing systems. The need to adapt, even if slightly, their data to make them fully compliant with the SAN exchange format will have inevitable fallbacks on the same systems that will be stimulated towards more elevated levels of homogeneity and standardization in order to allow wider and wider data reuse in contexts different from the original.

On the other hand, the open and cooperative model adopted for developing the SAN standards has largely already fulfilled satisfying and important results. Further consolidation of these results, which must derive from their wide use, will constitute an important premise for taking further steps forward in developing other more complex national exchange formats related to other aspects and components of archival description.

It is worth remembering that if standards and metadata are important because they make it possible for systems to exchange data, they are even more important because they force people and institutions to communicate.

The DTD of the EAG *Censo-Guía de los Archivos de España e Iberoamérica*: An Electronic Guide of the Spanish and Ibero-American Archives

by Blanca Desantes

1. Introduction

In this article the platform *Censo Guía de Archivos de España e Iberoamérica* as well as the data structure model and focus on the DTD of EAG (Encoded Archival Guide) is explained.¹

This platform is an electronic guide to archives for Spain and Ibero America, providing online information relating to: archival institutions responsible for the custody, management and service of documentation. From this view, this electronic guide is a directory of archives (where?). Additionally, the guide provides a description of

¹ The Project Censo Guía de los Archivos de España e Iberoamérica can be consulted at the following address: <<http://censoarchivos.mcu.es/CensoGuia/portada.htm>>. Also of interest for knowing about the projects and activities of the Spanish State Archives is the address <<http://www.mcu.es/archivos/index.html>> and specially PARES (Spanish Archives Web Portal), which started in May 2007. Actually, it spread 18 million digitized pictures of documents linked to 4,9 million archival records with access in the Internet from an archival net composed of 11 state archives. It has developed modules for description, spreading with single professional searchers, a tool to know the informatized records of every archive and its news. At the moment, PARES is developing modules for the internal electronical management of searchers access, searchers rooms and reprographic services: <<http://pares.mcu.e/>>. All mentioned Websites have been accessed on 09.02.2010.

the documentation held at each archival institution at a fonds and collection(s) level (what?) and a description of the creators (who?).

GOBIERNO DE ESPAÑA MINISTERIO DE CULTURA

CENSO-GUÍA de Archivos de España e Iberoamérica

Presentación

- Información general
- Directorio de Archivos
- Fondos Documentales
- Fichero de Autoridades
- Enlaces de Interés

ACCESO COLABORADORES

- Usuario y contraseña
- Contacte con el Censo-Guía de Archivos

Búsqueda por Directorio de Archivos.
Información de los archivos y otras instituciones que custodian documentos de archivo. Responde a la pregunta **Dónde** »

Búsqueda por Fondos documentales.
Información de los fondos y colecciones documentales que conserva cada archivo, descritos según la norma ISAD(G). Responde a la pregunta **Qué** »

Búsqueda por Autoridades.
Información de los organismos productores de los documentos, descritos según la norma ISAAR(CPF). Responde a la pregunta **Quién** »

UNESCO Archives Portal

PARES Portal de Archivos Españoles

DIRECCIÓN Y CONTACTO

Centro de Información Documental de Archivos (CIDA)
Paseo de Aguadores, 2-planta baja
Alcalá de Henares (Madrid)
28804
Teléfonos: 918838539-918836165
Fax: 918829470
correo: censoguia@mcu.es

© Ministerio de Cultura
Aviso Legal || Guía de Navegación || Accesibilidad

WC WAI-AA WEBA 1.1 Titulares RSS disponibles

Fig. 1: Homepage of the Censo Guía de Archivos de España e Iberoamérica

The objectives of this project are manifold, for instance, to provide an overall view of the fonds and/or collections held in the various archives included in this project. Further aims are to describe only higher levels of description and to provide summarized information regarding creators (corporate bodies, persons and family);

also, to provide general information on the archival institutions holding documents and the services they offer citizens. The main objective is to provide on-line information which traditionally has been gathered in the finding aids called Guides on the Web and electronically.

The electronic guide has two important qualities: the *cooperation* and the *standardization*. One of the most important qualities of this tool is that it is the result of an international scope deriving from cooperation, not only international (with other countries in Latin America), but also inside Spain in cooperation with the different autonomous regions.

At the beginning in 1979, the objective of this project was focused on Spanish archives. That scope has been expanded to other archives in Ibero-American countries through collaboration agreements.² At the present, more than 50,000 archival institutions are registered.

It is work for dissemination and an instrument of the promotion of cultural contents as regards archives on the Web. It is also a framework and information portal interrelated with the archive centers and documental fonds that offer a global vision of the Hispanic archival situation, facilitating more detailed queries. It offers the possibility of using links from the *Censo Guía de Archivos* with other resources that are available on the Internet (Web pages, e-mails,

² Colombia, Argentina, Bolivia, Costa Rica, Chile, El Salvador, Guatemala, Nicaragua, Paraguay, Venezuela, Brazil, Puerto Rico, Peru, Cuba and Mexico.

videos related to the archives, finding aids for electronic formats, etc.).

The project also has a clear goal of standardization. The *Censo Guía de Archivos* is contributing to the standardization of archival descriptions and to the consolidation of international guidelines used amongst the community of archival professionals. Standardization has great importance for an instrument like this because the archival descriptions come from very diverse number of institutions and countries.

2. Information and Data Structure Model

For all the information that the electronic guide offers about the different *archival institutions*, the data structure model used by this project is an XML DTD called *EAG*. This XML is prepared by the Subdirección General de los Archivos Estatales of the Ministerio de Cultura in Spain.³ Currently it is in version Alfa 0.2. and regulates the market on archive centers, providing an electronic format for the storage, publication and exchange of archive center representation.

The descriptions for the *fonds and collections* managed in those archives are using a data structure according to ISAD(G) 2nd edition and to the DTD of XML *EAD* version 2002. We could search the documentation held at each center and see the classification scheme

³ The work group has been made up of the archivists Abelardo Santamaría, Francisco Javier Fito and Blanca Desantes. Abelardo Santamaría has been in charge of the XML encoding of the result of this group's work. The DTD created was verified by the development team of the Project "Censo Guía de los Archivos de España e Iberoamérica".

of the archives. All the information follows the standard ISAD(G)⁴ and is also available in XML format. The choice of this format is because it has become the international standard for electronic data exchange and has achieved broad diffusion. Also, EAG is very important because it allows to share and exchange information with other organizations that run projects like this.

Another important data component, linked to the previous ones, has been incorporated: description of *creators* that allows the searches by authority file. The description of creators follows the ISAAR (CPF) 2nd edition data structure. There is a complete list with all the results and three lists of persons, families and corporate bodies that make the access to the information easy. All this data is in XML format, too: the DTD XML *EAC* (Encoded Archival Context). The *authority file* is shared with the PARES project. There is an interconnection between these three components.

3. System Features

The *Censo Guía de Archivos* is divided into three access zones that are well differentiated: Internet, Extranet and Intranet.

Via *Internet*, the guide is accessible by any user and allows searches to be made of the archive centers and its fonds and/or collections. Searches can be either simple or advanced and interactive maps have been incorporated to allow users to search the archival resources by different geographical areas in Spain and Ibero

⁴ Identity Statement Area, Context Area, Content and Structure Area, Conditions of Access and Use Area, Allied materials Area, Notes Area and Description Control Area.

America. Likewise, pop-up windows have been used to provide the user with assisted searches. Combination searches can also be used which incorporate different classification criteria for the archives and allow an archive to be identified from diverse points of view. The queries can be viewed in different formats such as HTML and XML.

The *Extranet* is of restricted access for the archive systems, archive centers and accredited professionals for the maintenance of the information associated with the archive center. It allows information, which could be obsolete or enriched with the incorporation of new description units, to be added or modified. This involves the integration, via documental management, with the work flows of the Subdirección General de Archivos Estatales (SGAE) of Spain. Standardized templates have been designed for entering this data.

The *Intranet* offers exclusive access for SGAE users that coordinate the management of the *Censo Guía de Archivos*. The application allows expert searches that can be configured by the user and management of the temporary repository of information that allows the data supplied by the archive centers to be verified through the Extranet. Furthermore, the management of access statistics to the portal is possible (monthly number of accesses, access by countries etc.). Additionally, it allows the creation of reports, management of the assistant tables (countries, municipalities, types of archives) and data exchange through importing and exporting in different formats (XML, Access, etc.).

Throughout the year 2003, an on-line updating of the information contained on the platform had been implemented, which has involved a quantitative and qualitative leap in the process of up-

dating information. Functionality had been consolidated through the development of two pilot projects in the Autonomous Community of Murcia in Spain and with Mexico, respectively, in order to verify the proper functioning of the communication circuits and to be able to extend it to the rest of the Spanish and Ibero-American archives.

4. Why EAG DTD?

As professionals, we must answer questions of *what*, *who* and *where* on the Internet and in a structured and standardized way. We have standards to answer the question of *what* documentary resources we have custody of (ISAD-G; EAD DTD). Also from an archivist's point of view we can answer the question *who* is the corporate body, persons and family responsible for the creation of these records by providing contextual information (ISAAR-CPF; EAC DTD). And now we also have a standard to answer the question *where*: the ISDIAH (International Standard for Describing Institutions with Archival Holdings). The DTD of EAG, mentioned above, has been one of the basic documents for developing this recent standard (ISDIAH). ISDIAH and EAG quickly, clearly and adequately provide citizens with the practical information related to custody repositories which they need. The inclusion of 50,000 archives and the possibility of updating this information on-line from numerous countries and archival institutions on the platform of the *Censo Guía de Archivos* for Spain and Ibero America make an exchange format in XML necessary. All of these factors made us come up with the idea of this simple DTD in XML called EAG (Encoded Archival Guide), which pragmatically solved the exchange of information related to the archival institution.

Let me focus on important points: The EAG format is being used as a “fact” successfully and without any problems between Spain and Ibero America for the exchange of information related to the archives as document custodial centers. Now that the International Standard for Describing Institutions with Archival Holdings is approved, the DTD of EAG shall be modified and adapted to the new international standard. We must work with ISDIAH on establishing a mapping between EAG elements and attributes.

5. Encoded Archival Guide (EAG)/ Guía de Archivos Codificada

EAG is a simple XML DTD (currently in version Alfa 0.2), which regulates the market from general information on archival institutions, providing an electronic format for the storage, publication and exchange of archival institutions representation.

It is a DTD designed to operate as an XML DTD and a well-formed, valid XML document. It respects the basic rules for XML syntax and meets the restrictions expressed.

The EAG DTD has been prepared by a working group of the Subdirección General de los Archivos Estatales in October 2001/ July 2002. There are two essential technical documents:⁵ the EAG DTD and the Tag Library.

In the following, we will analyze in detail the elements and attributes of the EAG DTD and the mapping with the International Standard for Describing Institutions with Archival Holdings (ISDIAH), mak-

⁵ These technical documents are accessible to the public at the following address: <http://censoarchivos.mcu.es/CensoGuia/proyecto.htm>.

ing the correspondence with the elements of the standard. Due to lack of space, only a range of descriptive charts of the elements and attributes can be represented below.

5.1 Encoded Archival Guide `<eag>` `</eag>`:

This is the “wrapping” for an electronic description of the archival institution in accordance with the DTD of EAG. It is a digital reference instrument which contains the elements and attributes necessary to provide information about the archive, its history and formation, timetables and user services, etc. It is a mandatory element and unrepeatable. It can include an audience attribute which marks if the register can be consulted by the general public or has a restricted access.

EAG element	EAG attribute	Description
<code><eag></code>		The DTD of EAG contains the elements and attributes needed for the complete description of the archive centers
	<i>audience</i>	Can the archive be consulted by the general public or only internally, by personnel of the Subdirección General de Archivos Estatales
The <code><eag></code> element is the external element of our DTD and is made up of: 1 element <code><eagheader></code> (mandatory and unrepeatable) 1 element <code><archguide></code> (mandatory and unrepeatable)		

Fig. 2: Descriptive chart of `<eag>`

5.1.1 EAG Header `<eagheader>` `</eagheader>`:

This element contains all the information related to the creation of the description of the archival institution. It includes elements and

attributes about the unique identifier of the description as well as the maintenance history of this description: information about the different “events” performed on the description (creation, modification or deletion); the date of each “event” is included (standardized according to ISO-8601); the name of the person responsible for the creation or modification of the description (name, surnames, position or function and source used); likewise, a language declaration and the script employed in the archive description (ISO 15924; ISO 639-b).⁶

5.1.2 Archival guide <archguide> </archguide> :

This element contains all the data related to the archive which is being described: identification data which makes it unique with respect to the other archival institution and the descriptive information about the archive itself.

EAG element	Description
<archguide>	Description of the archive: names, location, data about access and people responsible for the archive, history, documentation, etc.

Fig. 3: Descriptive chart of the <archguide> element

⁶ Mapping with ISDIAH: 5.6.1 Institution description identifier; 5.6.2 Institution identifier; 5.6.3 Rules and/or conventions used; 5.6.4 Status, 5.6.5 Level of detail; 5.6.6 Dates of creation, revision or deletion; 5.6.7 Language(s) and script(s); 5.6.8 Sources; 5.6.9 Maintenance notes.

5.1.2.1 Identity Area <identity>:

This element groups all descriptive data related to the identification of the archive or, in other words, the information which allows us to unequivocally identify the archive we are describing from other similar centers.

It includes information related to the archive identifier (country code ISO 3166-1 and archive code) and also about the authorized form of the archive name, the parallel forms in case of multiple languages and the excluded forms of the name⁷ (see next page, Fig. 4).

5.1.2.2 Description Area <desc> </desc>:

This element includes the descriptive information of the archive repository which is being registered: details referring to its geographical location, contact details of the archival repository (address, Web page, telephone, fax, e-mail); details about access to the center (opening hours, holidays, free or restricted access); name of the director or responsible person at the center; typology of repository; institutional history of the repository; history of the building and its storage capacity; documentation referring to the archival repository (guides); its administrative structure; technical services it offers the public and administration as well as structure of its fond(s) and collection(s).

⁷ Mapping with ISDIAH: 5.1.1 Identifier; 5.1.2 Authorized form(s) of name; 5.1.3 Parallel form(s) of name; 5.1.4 Other form(s) of name.

EAG element	EAG attribute	Description
<identity>		Descriptive identification data about the archive
The element <identity> is made up of: 1 element <respositorid> (mandatory and unrepeatable) 1 element <autform> (mandatory and unrepeatable) 1 element <parform> (optional and repeatable) 1 element <nonpreform> (optional and repeatable)		
<respositorid>		Information which is used to unequivocally identify the archive, with the combination of the following attributes
	country-code	Country code in accordance with ISO standards
	repositoricode	Unique, internal, identification code of the archive, made up (in our case) by the concatenation of the INE (Instituto Nacional de Estadística) code of the municipality where it is located and by the acronym for the archive
<autform>		Authorized form of the archive name
<parform>		Parallel form(s) of the archive name
<nonpreform>		Excluded forms of the archive name

Fig. 4: Descriptive chart of the elements and attributes included in the <identity> element

This element, therefore, includes as many elements and attributes enabling the description of the center storing the documents which we might wish to consult. We will analyze them by sections (see next pages, Fig. 5).

In order to establish the archival institution typology, the description area element includes access points that constitute various

terms from controlled vocabularies. It is a basic element for search strategies in the platform *Censo Guía de Archivos*.⁸

EAG Element	Description
<desc>	General description of the Archive Repository
	<p>The element <desc> comprises:</p> <ul style="list-style-type: none"> 1 element <geogarea> (mandatory and non repeatable) 1 element <country> (mandatory and non repeatable) 1 element <firstdem> (optional and non repeatable) 1 element <secondem> (optional and non repeatable) 1 element <municipality> (mandatory and non repeatable) 1 element <localentity> (optional and non repeatable) 1 element <street> (optional and non repeatable) 1 element <postalcode> (optional and non repeatable) 0 or several <elements telephone> (optional and repeatable) 0 or several <elements fax> (optional and repeatable) 0 or several <elements email> (optional and repeatable) 1 element <webpage> (optional and non repeatable) 1 element <timetable> (optional and non repeatable) 1 element <access> (mandatory and non repeatable) 1 element <respositor> (optional and non repeatable) 1 element <repositorhist> (optional and non repeatable) 1 element <repositorfound> (optional and non repeatable) 1 element <repositor-sup> (optional and non repeatable) 0 or several elements <repositor-guides> (optional and repeatable) 0 or several elements <admin-hierarchy> (optional and repeatable) 1 element <building-info> (mandatory and non repeatable) 1 element <extent> (mandatory and non repeatable) 1 element <organization> (mandatory) 1 element <tech-services> (mandatory) 1 element <automation> (mandatory) 1 element <control-access> (optional) 1 element <notes> (optional)

Fig. 5: Descriptive chart of the elements included in the element <desc>

⁸ Mapping with ISDIAH: 5.1.5 Type of institution with archival holdings.

The first element gives geographical or location details of the archival institution. This is one of the key elements for search in the portal.⁹

EAG element	Description
<i><geogarea></i>	Geographical area in which the repository is located (Spain or Ibero America)
<i><country></i>	Country in which the repository is located
<i><firstdem></i>	Element that allocates the Spanish Autonomous Community in which the repository is located, or, for Ibero America or other regions, the first demarcation in which the repository is located
<i><secondem></i>	Second demarcation or province in which the repository is located
<i><municipality></i>	Municipality in which the archive is located
<i><localentity></i>	Local entity, below the level of municipality, in which the repository is located

Fig. 6a: Descriptive chart of the elements related to the geographic location of the archive in the *<desc>* element

⁹ Mapping with ISDIAH: 5.3.2 Geographical and cultural context.

The next area contains contact details of the repository.¹⁰

EAG element	EAG attribute	Description
< <i>street</i> >		Street and number at which the repository is located
< <i>postal-code</i> >		Post code
< <i>telephone</i> >		Contact telephone number(s) of the repository
< <i>fax</i> >		Fax contact number(s) of the repository
< <i>email</i> >		Contact e-mail address of the repository
	<i>href</i>	Virtual link of e-mail address
< <i>webpage</i> >		Repository Web page
	<i>href</i>	Virtual link of the page

Fig. 6b: Descriptive chart of the elements and attributes related to contact data included in the <desc> element

The description area also gives details of access to the archival repository (public opening hours, information on free or restricted access).¹¹

¹⁰ Mapping with ISDIAH: 5.2.1 Location and address(es); 5.2.2 Telephone, fax, e-mail.

¹¹ Mapping with ISDIAH: 5.4.1 Opening times; 5.4.2 Conditions and requirements for access and use.

EAG element	EAG attribute	Description
< <i>timetable</i> >		Public opening hours
The element < <i>timetable</i> > comprises: 1 element < <i>opening</i> > (optional and non repeatable) 1 element < <i>weekopen</i> > (optional and non repeatable) 1 element < <i>closing</i> > (optional and non repeatable)		
< <i>opening</i> >		Opening hours
< <i>weekopen</i> >		Number of hours per week that the repository is open to the public
The element < <i>weekopen</i> > comprises: 1 element < <i>num</i> > (mandatory and non repeatable)		
< <i>num</i> >		Numerical measurement fields
	<i>unit</i>	Measurement unit, in this case hours
< <i>closing</i> >		Time periods during which the repository is closed to the public
< <i>access</i> >		Free or restricted access to the repository
	<i>question</i>	No: restricted; yes: free access
If access is restricted, it will appear within the < <i>restaccess</i> > element, in which mentioned restrictions will be detailed		
< <i>restaccess</i> >		Detail of access restrictions

Fig. 7: Descriptive chart of the elements and attributes related to access data in the < *desc* > element

Further, the description area gives information referring to the Director or person responsible for the archival institution (surname, name, post or function, naming date). This area offers the possibility of preparing tags for correspondence.¹²

EAG element	EAG attribute	Description
<code><resprepositor></code> (optional and non repeatable)		Responsible person of the archive
The element <code><resprepositor></code> is made up of: 1 element <code><surnames></code> (mandatory and non repeatable) 1 element <code><name></code> (mandatory and non repeatable) 1 element <code><charge></code> (optional and non repeatable) 1 element <code><date></code> (optional and non repeatable)		
<code><surnames></code>		Surname(s) of the responsible person
<code><name></code>		First name of the responsible person
<code><charge></code>		Charge or position of the responsible person
<code><date></code>		Naming date of the responsible person
	<i>Normal</i>	Standardized date in line with standard ISO-8601

Fig. 8: Descriptive chart of the elements and attributes related to data about the director or responsible person of the archive included in the `<desc>` element

¹² Mapping with ISDIAH: 5.2.3 Contact persons.

Now are following elements which include the institutional history of the archive, its creation date, its administrative structure and guides published.¹³

Next the user finds elements which group together information on the archival building (general characteristics of the building, area and storage capacity, capacity of research room, facilities for the handicapped, etc.). This data can also be included in reports or statistics.¹⁴

EAG element	EAG attribute	Description
<buildinginfo>		Building housing the archive
<p>The element <buildinginfo> comprises:</p> <ul style="list-style-type: none"> 1 element <building> (optional and non repeatable) 1 element <repositorarea> (optional and non repeatable) 1 element <lengthshelf> (optional and non repeatable) 1 element <searchroom> (mandatory and non repeatable) 1 element <handicapped> (mandatory and non repeatable) 		
<building>		Information about the building via elements <p>
<repositorarea>		Storage area of the repository

¹³ Mapping with ISDIAH: 5.3.1 History of the institution with archival holdings; 5.3.3 Mandates/Sources of authority; 5.3.4 Administrative structure; 5.3.5 Records management and collecting policies; 5.3.7 Archival and other holdings; 5.3.8 Finding aids, guides and publications.

¹⁴ Mapping with ISDIAH: 5.3.6 Building(s); 5.4.3 Accessibility.

The element <i><repositorarea></i> comprises: 1 element <i><num></i> (mandatory)		
<i><num></i>		Numerical measurement fields
	<i>unit</i>	Measurement unit, in this case square metres
<i><lengthshelf></i>		Shelf length in metres of the repository
The element <i><lengthshelf></i> comprises: 1 element <i><num></i> (mandatory)		
<i><num></i>		Numerical measurement fields
	<i>unit</i>	Measurement unit, in this case linear metres
<i><searchroom></i>		Number of research stations in the search room
The element <i><searchroom></i> comprises: 1 element <i><num></i> (mandatory)		
<i><num></i>		Numerical measurement fields
	<i>unit</i>	Measurement unit, in this case site
<i><handicapped></i>		Provides the archive handicapped access or not
	<i>question</i>	No: it doesn't; yes: it does
<i><extent></i>		Archival volume in linear metres
The element <i><extent></i> comprises: 1 element <i><num></i> (mandatory)		
<i><num></i>		Numerical measurement fields
	<i>unit</i>	Measurement unit, in this case linear metre

Fig. 9: Descriptive chart of the elements and attributes related to building and repository data included in the *<desc>* element

The following elements provide information referring to the technical services offered by the archival institution: restoration, reproduction, auxiliary library and automation.¹⁵

¹⁵ Mapping with ISDIAH: 5.5.1 Research services; 5.5.2 Reproduction services; 5.5.3 Public areas.

Furthermore, the description area includes the classification table of archival repository fonds and collections. For each fond or collection, the ID code, level, title, date and volume or extension of the documentation is included. The area has a link with ISAD(G)/EAD.¹⁶

EAG element	EAG attribute	Description
<organization>		Description of the organisation of the repository fonds
The element <organization> comprises: 1 or several elements <descunit>		
<descunit>		Minimum information of a description unit from the organisation table
	<i>level</i>	Description level with controlled list: recordgrp (group of fonds), fonds (fond), collection (collection)
	<i>classcode</i>	Individual classification code
	<i>fathercode</i>	Father classification code of the element <class>
The element <descunit> comprises: 1 element<unitid> (mandatory and non repeatable) 1 element<unittitle> (mandatory and non repeatable) 1 element<date> (mandatory and non repeatable) 1 element<extent> (mandatory and non repeatable)		

¹⁶ Mapping with ISDIAH: Chapter 6. Relating descriptions of institutions with archival holdings.

<unitid>		Individual identifier of the description unit
	<i>href</i>	Virtual link to the XML document referring to the element (ISAD(G); EAD)
<unittitle>		Title of the description unit
<date>		Dates referring to the element (creation or accumulation dates)
	<i>Normal</i>	Standardized date in line with standard ISO-8601
<extent>		Volume or extension of the description unit
The element <extent> comprises: 1 element<num> (mandatory)		
<num>		Numerical measurement fields
	<i>unit</i>	Measurement unit, in this case linear metre

Fig. 10: Descriptive chart of the elements and attributes related to the classification chart for the fonds and collections in the <desc> element

6. Conclusions

Obviously, the mapping with ISDIAH is very superficial. It is necessary to introduce changes in EAG in order to adapt it perfectly to ISDIAH: to modify and add elements, to group these elements in the areas established in ISDIAH. And finally, to check attributes of the elements in order that the conformity with the aims and ISDIAH's rules could be complete.

It would be desirable to prepare the new EAG version within a working group which is not exclusively Spanish. We also feel that

EAG is a useful exchange format to the Archives Portal Europe (APE-net), as the information about archival repositories is included in the basic archival contents proposed for offering on the APEnet.

Austauschformate bei der Retrokonversion und ihr Nutzen für die deutschen Archive

von Sigrid Schieber

Seit dem Start der DFG-Förderung für die Retrokonversion archivi-scher Findmittel im Jahre 2008 hat die Verwendung von archivspezi-fischen Austauschformaten enorm an Bedeutung gewonnen. Dies kann zunächst einfach darauf zurückgeführt werden, dass die DFG (Deutsche Forschungsgemeinschaft) die Verwendung von bestimm-ten Austauschformaten als zwingende Voraussetzung für die Förde-rung der Retrokonversion definiert hat. Diese Vorgabe ist jedoch nicht willkürlich geschehen, es liegen ihr vielmehr zwei Prämissen zu Grunde: Durch die Verwendung von Austauschformaten wird

1. die Online-Stellung der retrokonvertierten Findmittel in übergrei-fenden Portalen (ArchivportalD, BAM-Portal etc.) und
2. die durch externe Dienstleister durchgeführte Retrokonversion vereinfacht und kostengünstiger.

Im Folgenden sollen die Erfahrungen mit der Anwendung von Aus-tauschformaten bei der Retrokonversion dargestellt und die Stichhal-tigkeit dieser Prämissen beurteilt werden.

1. Vorgegebene Austauschformate: SAFT und EAD

Als Zielformat der Findmitteldigitalisierung bietet sich XML (*Extended Markup Language*) an, das in den letzten Jahren einen Aufschwung als universelles Austauschformat zwischen Datenbanken erlebt hat. XML ist eine Auszeichnungssprache ähnlich wie HTML. Die logi-

schen Bestandteile von Informationen, die in einer Datenbanktabelle in unterschiedlichen Feldern stehen, werden hier durch sogenannte *tags* oder Elemente markiert bzw. ausgezeichnet. Der Titel einer Archivalie könnte zum Beispiel dadurch gekennzeichnet werden, dass er zwischen dem öffnenden Element `<Titel>` und dem schließenden Element `</Titel>` steht. Die durch spitze Klammern hervorgehobenen Elemente weisen dem Text, der sich zwischen ihnen befindet, eine Bedeutung zu – ähnlich wie ein öffnendes und ein schließendes Anführungszeichen den dazwischenliegenden Text als Zitat kennzeichnen. Die Elemente sowie deren Beziehung zueinander kann beliebig gestaltet werden, so dass sich jede komplexe Datenstruktur in XML abbilden lässt. Um als Austauschformat sinnvoll genutzt werden zu können, muss jedoch eine Übereinkunft über die benutzten Elemente zwischen den Interessenten getroffen werden, die XML als Austauschformat benutzen wollen. Dies ist möglich, indem man eine Grammatik für eine bestimmte XML-Ausprägung festlegt. In einer sogenannten DTD (*Document type definition*) oder einem XML-Schema wird festgelegt, welche Elemente in einem Dokument benutzt werden dürfen und in welcher Beziehung sie zueinander stehen. Für ein archivisches Austauschformat könnte man zum Beispiel definieren, dass die Verzeichnungseinheiten durch das Element `<Verzeichnungseinheit>` markiert werden, und dass dieses Element wiederum die Kindelemente `<Titel>`, `<Laufzeit>` und `<Signatur>` enthalten muss.

Im Archivwesen gibt es bereits zwei definierte XML-Austauschformate, SAFT-XML¹ und EAD-XML², die inzwischen beide als Austauschformate für die DFG-geförderte Retrokonversion zugelassen sind. Als umfassendes Austauschformat für die Retrokonversion von Findmitteln wurde SAFT-XML (*Standard-Austauschformat*) vom Landesarchiv Nordrhein-Westfalen im Rahmen einer DFG-Vorstudie entwickelt.³ Es zeichnet sich durch seine Orientierung an deutschen Verzeichnungstraditionen aus. SAFT-XML bildet die verschiedenen Objektarten bei den Verzeichnungseinheiten mit unterschiedlichen Mantelelementen (<Sachakte>, <Urkunde>, <Karte> etc.) ab. Allerdings ist die Objektauswahl begrenzt, z.B. fehlt das Element <Amtsbuch>, welches in vielen Archiven verwendet wird. Die Elemente zur Beschreibung der Verzeichnungseinheiten sind zwar sehr differenziert, aber es gibt auch bei SAFT das Element <Hilfsfeld>, um nicht vorgesehene Elemente abbilden zu können. Momentan wird SAFT bei der Retrokonversion und als Austauschformat für das Portal <<http://www.archive.nrw.de>> verwendet.

¹ Informationen zu SAFT-XML, die SAFT-DTD und eine tag library (interaktiv und als Word-Dokument) finden sich auf der Homepage der Koordinierungsstelle Retrokonversion unter <<http://www.archivschule.de/content/462.html>>. Alle im Beitrag genannten Webseiten wurden am 8.2.2010 aufgerufen.

² Informationen zu EAD inklusive einer deutschen Übersetzung der tag library finden sich auf der Homepage des Bundesarchivs unter <<http://www.daofind.de>>, eine interaktive tag library (englisch) unter <<http://www.loc.gov/ead/tglib/index.html>>.

³ Vgl. dazu zusammenfassend Ulrich Fischer/Wilfried Reininghaus, DFG-Vorstudie „Retrokonversion archivischer Findmittel“. Die wichtigsten Ergebnisse des Projekts, in: *Der Archivar* 59, 2006, S. 329-333, online abrufbar unter <http://www.archive.nrw.de/archivar/hefte/2006/Archivar_2006-4.pdf>.

Neben SAFT-XML gibt es vor allem den internationalen Standard EAD-XML, der in den USA entwickelt wurde und sich an angloamerikanischen Verzeichnungstraditionen orientiert. Er besitzt im Vergleich zu SAFT nicht ganz so viele Elemente zur Beschreibung der einzelnen Verzeichnungseinheiten. Außerdem unterscheidet EAD nicht zwischen Objektarten – allerdings kann die Information über die Objektart im Element <genreform> festgehalten werden. Jede Verzeichnungseinheit steht unter dem Mantelelement <c> (für *Component*) und wird mit dem gleichen Set an Kindelementen beschrieben. Im Vergleich zu SAFT bietet EAD etwas mehr Interpretationsspielraum bei der Verwendung. Derzeit wird EAD in Deutschland vor allem vom Bundesarchiv, vom Landesarchiv Baden-Württemberg und von den Hessischen Staatsarchiven als Austauschformat für die Datenlieferung an das BAM-Portal und an internationale Portale benutzt. Auf Basis der Erfahrungen, die die am BAM-Portal beteiligten Archive und das Bundesarchiv mit der Verwendung von EAD gemacht haben, wurden mittlerweile auch Übereinkünfte über die Anwendung dieses Standards im deutschen Archivwesen getroffen.⁴

2. Einsatzorte

Austauschformate kommen im Prozess der Retrokonversion an mehreren Stellen zum Einsatz. Wenn die Findmittel durch einen externen Dienstleister retrokonvertiert werden, muss dieser Dienstleister das

⁴ Vgl. hierzu den Bericht über ein Treffen von EAD-Anwendern im Bundesarchiv unter <<http://www.archivschule.de/forschung/retrokonversion-252/newsretro/arbeitsreffen-zur-abstimmung-eines-gemeinsamen-ead-austauschformats.htm>> und <<http://www.bundesarchiv.de/aktuelles/begegnungen/00201/index.html>>.

Ergebnis der Retrokonversion als SAFT- oder EAD-XML an das jeweilige Archiv liefern. Die DFG-geförderten Archive verpflichten sich außerdem, die Findmittel nicht nur in ihrem eigenen System bzw. Portal zu veröffentlichen, sondern sie – wiederum in einem der beiden Austauschformate – einem übergeordneten Portal wie dem Archivportal D oder dem BAM-Portal zur Verfügung zu stellen. Da die retrokonvertierten Findmittel in der Regel gemeinsam mit den originär digital erstellten Findmitteln eines Archivs im gleichen System verwaltet und gegebenenfalls noch weiter bearbeitet oder fortlaufend ergänzt werden sollen, müssen sie nach der Retrokonversion zunächst in die eigene Archivsoftware importiert und für eine Veröffentlichung anschließend wieder aus der Archivsoftware exportiert werden können. Bei einer DFG-geförderten Retrokonversion durch externe Dienstleister muss ein Archiv also über eine Austauschformat-Importschnittstelle und über eine Austauschformat-Exportschnittstelle verfügen. Da in vielen Archiven Archivsoftware im Einsatz ist, die noch nicht über derartige Schnittstellen verfügt, ist dies eines der Hauptprobleme bei der Antragstellung. Die DFG-Vorgaben werden natürlich auch erfüllt, wenn die in ein Austauschformat retrokonvertierten Findmittel nur in einem übergeordneten Portal online gestellt werden – zum Beispiel SAFT-Findmittel im nordrhein-westfälischen Portal <http://www.archive.nrw> oder EAD-Findmittel im hessischen Portal <http://www.hadis.hessen.de>. Für das einzelne Archiv kann dies aber nur als Übergangslösung dienen. Das Ziel sollte es sein, die retrokonvertierten Findmittel in absehbarer Zeit auch in die eigene Archivsoftware zu importieren, sobald diese zum Beispiel in einer neueren Version über geeignete Schnittstellen verfügt.

Die Konversion im eigenen Haus stellt sich unter dem Aspekt der Verwendung von Austauschformaten für die Archive etwas unkomplizierter dar. In diesem Fall wird das analoge Findmittel in der Regel händisch direkt in die eigene Archivsoftware übertragen. Anschließend muss es in einem Austauschformat aus der eigenen Archivsoftware exportiert werden können, um es in übergeordneten Portalen zu veröffentlichen. In diesem Fall benötigt die verwendete Archivsoftware also nur eine Austauschformat-Exportschnittstelle.

3. Anwendungsbeispiele

Beide Beispiele stammen aus der Pilotphase, so dass hier schon der Prozess vom Ausgangs- bis zum Endprodukt gezeigt werden kann. In beiden Fällen wurde EAD als Austauschformat benutzt.

Zu sehen ist in Abb. 1 ein für die Retrokonversion markiertes Findmittel des Sächsischen Staatsarchivs – Staatsarchiv Chemnitz (Findbuch 30049). Es handelt sich dabei um ein unkompliziertes Spaltenfindbuch mit wenigen Elementen.

Archivnummer	Aktenzeichen	Archivtitel	Laufzeit
583	XI,1,184 ^{II} /10	Umbezirkungen zwischen Johannegeorgenstadt und Jugel	1929-1935
584	XI,1,184/46	Umbezirkung zwischen dem Staatsforstrevier Lauter und der Gemeinde Bockau	1915-1918
585	XI,1,184/23	Umbezirkung von 6 Bahnwärterhäusern zwischen dem Staatsforstrevier Lauter und den Gemeinden Aue, Auerhammer, Bockau, Lauter	1922-1926
586	XI,1,184 ^{VI} /20	Umbezirkungen zwischen dem selbständigen Gutsbezirk Staatsforstrevier Lauter und der Gemeinde Lauter	1937-1939

- 49 -

Abb. 1: Analoges Findbuch 30049

Nach der Retrokonversion wird die oben markierte Verzeichnungseinheit in EAD-XML folgendermaßen dargestellt:

```
<c07 level=„file“>
  <did>
    <unittitle>Umbezirkungen zwischen Johanngeorgenstadt und
Jugel</unittitle>
    <unitid type=„Archivaliennummer“>583</unitid>
    <unitid type=„Aktenzeichen“>11,1, 184 II/10</unitid>
    <unitdate normal=„1929-01-01/1935-12-31“>1929-
1935</unitdate>
    <physcdesc>
      <genreform>Sachakte</genreform>
    </physcdesc>
    <abstract/>
  </did>
```

Abb. 2: EAD-Ausschnitt

Nach dem Import findet sich die betreffende Verzeichnungseinheit im Online-Findmittel des Sächsischen Staatsarchivs wieder (siehe nächste Seite, Abb. 3).

Im zweiten Beispiel ist ein Findmittel des Landesarchivs Baden-Württemberg zum Bestand B 63a des Hauptstaatsarchivs Stuttgart zu sehen. Dieses Findmittel ist deutlich komplexer, außerdem werden hier einige Aspekte deutlich, die unabhängig von der Verwendung eines Austauschformats geregelt und mit dem Dienstleister

abgestimmt werden müssen. So muss in diesem Fall zum Beispiel der Zusatz „Bü“ (für Büschel) vor die Signatur gesetzt werden, um die Eindeutigkeit der Signatur innerhalb des Bestandes zu gewährleisten. Das graphische Zeichen, das für die Quadrangelangabe verwendet wurde, kann so weder in das Austauschformat noch in die Archivsoftware übernommen werden. In diesem Fall wurde entschieden, das Symbol durch die Zeichenfolge „Qu.“ zu ersetzen.

30049 Amtshauptmannschaft Schwarzenberg
 Amtshauptmannschaft Schwarzenberg
 11. Gemeindeangelegenheiten
 11.01. Gemeindeverfassung

Nummer	Titel	Datum
1899		
659	Umbezirkungen zwischen den Gemeinden Raschau und Pöhl	1938 - 1939
604	Umbezirkungen zwischen der Gemeinde Tellerhäuser und dem selbständigen Gutsbezirk Oberwiesenthal	1929 - 1936
574	Umbezirkungen zwischen Elbenstock und Schönheide	1912
583	Umbezirkungen zwischen Johanngeorgenstadt und Jugel	1929 - 1935
589	Umbezirkungen zwischen Markersbach und dem selbständigen Gutsbezirk Staatsforstrevier Crottendorf	1937 - 1939
592	Umbezirkungen zwischen Obersachsenfeld und Schwarzenberg	1898 - 1899
606	Umbezirkungen zwischen Wittigsthal und Johanngeorgenstadt	1907 - 1910, 1926 - 1929
610	Umbezirkungen zwischen Zschorlau und Alberna	1917 - 1939
609	Umbezirkungen zwischen Zschorlau und Alberna	1914 - 1920
608	Umbezirkungen zwischen Zschorlau und Alberna	1909
751	Umfang des Ortsteils Haide im Gemeindebezirk Waschleithe und im Flurbezirk	

Abb. 3: Online-Findmittel 30049

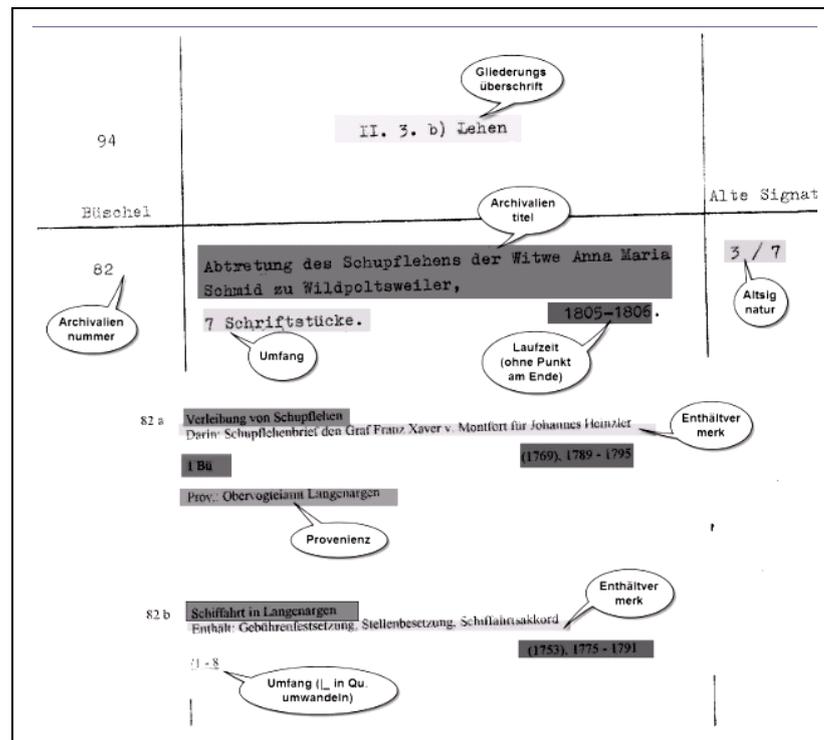


Abb. 4: LABW HSTAS B 63a

Im Austauschformat sieht die betreffende Verzeichnungseinheit nun so aus (siehe nächste Seite, Abb. 5).

Nach dem Import findet sich die betreffende Verzeichnungseinheit im Online-Findmittel des Landesarchivs Baden-Württemberg wieder (siehe Seite 231, Abb. 6).

```

<c06 level=„file>
  <did>
    <unittitle>Verleihung von Schupflehen</unittitle>
    <unitid type=„Archivaliennummer“>Bü 82a</unitid>
    <unitdate normal=„1769-01-01/1795-12-31“>(1769) 1789-
1795</unitdate>
    <physcdesc>
      <genreform>Sachakte</genreform>
      <extent>1 Bü</extent>
    </physcdesc>
    <abstract type=„Darin“>Schupflehenbrief den Graf Franz Xaver v. Mont-
fort für Johannes Heinzler</abstract>
    <origination>Obervogteiamt Langenargen</origination>
  </did>
</c06>

```

Abb. 5: EAD-Ausschnitt

Da das Landesarchiv Baden-Württemberg seine Findmittel zusätzlich im BAM-Portal veröffentlicht, kann die betreffende Verzeichnungseinheit auch bereits dort recherchiert werden (siehe Seite 232, Abb. 7).

4. Anpassungen für den Zweck der Retrokonversion

Für die Verwendung bei der Retrokonversion war es nötig, den Interpretationsspielraum von EAD einzuschränken und interne, d.h. deutschlandweite Festlegungen zu treffen. Dabei lehnte sich die Koordinierungsstelle Retrokonversion stark an das EAD-Profil des

LANDESARCHIV
BADEN-WÜRTTEMBERG
HAUPTSTAATSARCHIV STUTTGART

Bestand B 63 a ■ Strukturansicht DFG Impressum Home

Bestand B 63 a: Tettmang, Oberamt
II. Akten
II. 3. Cameraalia
II. 3. b) Lehen

Einträge 21 - 26 von 26 Alle

- I. Bände
- II. Akten
 - II. 1. Allgemeines
 - II. 2. Grenzregulierungen, -streitigkeiten und andere Angelegenheiten mit benachbarten Herrschaften
 - II. 3. Cameraalia
 - II. 3. a) Steuern, Abgaben, Gefälle
 - II. 3. b) Lehen
 - II. 3. c) Fronen und Leibeigenschaft
 - II. 3. d) Kauf, Tausch, Verkauf, Verpachtung
 - II. 3. e) Hochbau
 - II. 3. f) Straßen- und Wasserbau
 - II. 4. Forestalia
 - II. 5. Kirchen- und Schulsachen
 - II. 6. Stiftungen

B 63 a Bü 82
Abtretung des Schupflehens der Witwe Anna Maria Schmid zu Wildpoltzweiler
7 Schriftstücke 1805-1806
▶ Permalink

B 63 a Bü 82 a
Verleihung von Schupflehen
Darin: Schupflehenbrief den Graf Franz Xaver v. Montfort für Johannes Heinzler
1 Bü (1769) 1789-1795
▶ Permalink

B 63 a Bü 82 b
Schiffahrt in Longenargen
Enthält: Gebührenfestsetzung, Stellenbesetzung, Schiffsakkoord
Qu. 1-8 (1753) 1775-1791
▶ Permalink

Anfang Einführung Stichwortlisten Erweiterte Suche Hilfe

Beständeübersicht Übergreifende Suche Signatursuche Bestellkorb Lesezeichen

Abb. 6: Online-Findmittel B 63a

Bundesarchivs an,⁵ da dieses Profil für das zukünftige deutsche Archivportal verwendet werden soll. Die Zuordnung der logischen Elemente von Findmitteln zu SAFT- und EAD-Elementen wird in einem

⁵ Das EAD-Profil des Bundesarchivs ist online veröffentlicht unter http://www.bundesarchiv.de/imperia/md/content/archivportald/090407_ead_profil_findbuch_de.pdf

Mappingformular dargestellt, das von der Homepage der Koordinierungsstelle heruntergeladen werden kann.⁶

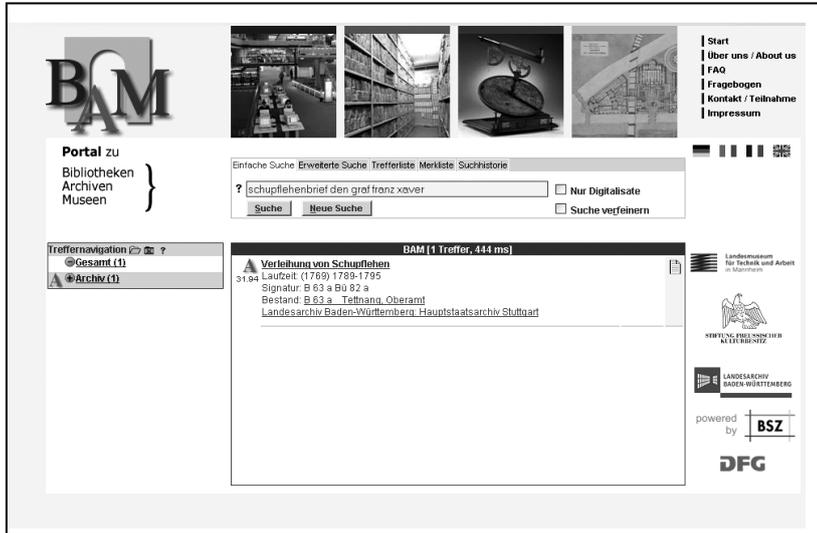


Abb. 7: BAM-Portal

So wurde etwa die Festlegung der Attributwerte zu <c> übernommen, die das Bundesarchiv getroffen hatte. Das Element <c> wird nach den EAD-Vorgaben als Mantelelement für alle Ebenen vom Bestand bis zur einzelnen Verzeichnungseinheit verwendet. Es kann dabei mit dem Attribut „level“ versehen werden, das wiederum ver-

⁶ Es handelt sich um Formular 3 bei den Unterlagen zur Antragstellung: <http://www.archivschule.de/forschung/retrokonversion-252/antragsstellung/unterlagen-zurantragstellung.html>

schiedene Attributwerte annehmen kann. Für die Retrokonversion wurde festgelegt, dass der Attributwert `<c level=„class“>` bei Klassifikationsebenen bzw. Gliederungsüberschriften verwendet werden soll, der Attributwert `<c level=„series“>` für Serien und der Attributwert `<c level=„file“>` für die Verzeichnungseinheiten auf der Ebene der einzelnen Archivalie.

Die angelsächsische Verzeichnungstradition kennt keinen Enthält-Vermerk. Es war also nötig, sich auf ein geeignetes Element für die deutschen Enthält-Vermerke zu einigen. In Frage kamen mehrere Elemente, im Wesentlichen `<scopecontent>` und `<abstract>`. Da sich das Bundesarchiv bereits auf `<abstract>` festgelegt hatte, wurde auch das EAD-Profil der Koordinierungsstelle Retrokonversion dahingehend angepasst.

Erst bei der Retrokonversion wurde deutlich, dass EAD die Objektart der Archivalien abbilden können muss. Dies ist bei anderen Verwendungszwecken nicht unbedingt nötig – bei der Weitergabe von Findmitteln zur übergreifenden Recherche an das BAM-Portal müssen zum Beispiel nur Kernelemente wie Titel, Laufzeit, Signatur sowie ein Link zur Verzeichnungseinheit in der ursprünglichen Umgebung enthalten und als solche markiert sein. Objektartspezifische Suchen – also beispielsweise Suchen nur nach Urkunden – sind im BAM-Portal gar nicht möglich, die Übermittlung dieser Information mithin unnötig. Die retrokonvertierten Findmittel sollten nun aber zunächst in die jeweilige Archivsoftware der geförderten Archive importiert werden. Diese sieht meist unterschiedliche Objektarten mit unterschiedlichen Elementen vor. Eine Urkundenverzeichnung muss beim Import automatisch als solche erkannt werden und darf nicht etwa als Sachakte

importiert werden. Um dies zu gewährleisten, wurde das EAD-Element <genreform>, das auch nach den EAD-Vorgaben die Art der Archivalie beschreiben soll, für die Angabe der Objektart verwendet.

Schließlich wurde das EAD-Element <odd> (*other descriptive data*) für Elemente verwendet, für die EAD kein spezifisches Element vorgesehen hat – analog zum Element <Hilfsfeld> in SAFT. Meist konnte jedoch ein spezielles Element gefunden werden. Auf <odd> musste bei den bisherigen Retrokonversionsprojekten nur selten zurückgegriffen werden, und zwar für das Element „Zeugen“ und das Element „Formalbeschreibung“ bei Urkundenverzeichnissen.

5. Fazit

Die DFG hatte aufgrund von zwei Vorannahmen die Verwendung eines der beiden Austauschformate für die Förderung von Retrokonversionsprojekten vorgegeben.

Die erste der beiden Annahmen, nämlich dass damit die Weitergabe an übergreifende Portale erleichtert wird, ist sicherlich unstrittig. Der Betrieb eines jeden übergreifenden Portals wird wesentlich erleichtert, wenn sich die Portalbetreiber nur auf die Integration von Daten in einem oder in wenigen Standardformaten einstellen müssen. Das ist selbst dann der Fall, wenn Standardaustauschformate, wie es immer wieder vorkommt, nicht vollständig einheitlich verwendet werden. Diese Vorgabe der DFG verursachte einen Zwang – oder positiver ausgedrückt einen Anreiz – zur Einrichtung von Austauschformat-Schnittstellen bei den Archiven und damit auch bei den Herstellern von Archivsoftware. Existiert eine solche Schnittstelle

erst einmal, dann kann das betreffende Archiv in der Folge natürlich nicht nur die retrokonvertierten Findmittel, sondern auch die originär digital erstellten Findmittel an andere Portale weitergeben.

Die zweite Annahme ging davon aus, dass die Retrokonversion durch externe Dienstleister mit Hilfe von Austauschformaten kostengünstiger und im Ablauf einfacher werden würde. Tatsächlich werden die Prozesse bei der Durchführung durch externe Dienstleister vereinfacht, vor allem, wenn ein Dienstleister mehrere Projekte abwickelt. Das Austauschformat dient als gemeinsame Bezugsgröße im Dialog zwischen Archiven und Dienstleistern, die sich nur einmal auf eine Datenstruktur einstellen müssen – je nach Komplexität der Findmittel entsteht aber dennoch bei jedem neuen Kunden ein mehr oder weniger großer Anpassungsbedarf an Besonderheiten.

Für die Archive ist eine Standardimportschnittstelle ebenfalls von Vorteil. Eine Importschnittstelle für SAFT- oder EAD-XML auf der Basis des überschaubaren Mappings der Koordinierungsstelle deckt je nach Komplexität des retrokonvertierten Findbuchs 90 bis 100 Prozent aller Importfälle ab. Es sind nur im Einzelfall noch geringe Anpassungen oder Ergänzungen der Schnittstelle nötig, zum Beispiel bei der Verwendung neuer Elemente aus SAFT oder EAD. Archiven, deren Archivsoftware noch nicht über Schnittstellen zu Austauschformaten verfügt, erschwert die Vorgabe zur Verwendung von Austauschformaten die Nutzung der Förderung allerdings. Sie müssen so nicht nur die Umwandlung ihrer Findmittel in ein Austauschformat für die Lieferung an übergeordnete Portale sicherstellen, sondern auch einen Datenimport ohne Standardschnittstelle bewerkstelligen. Dieser Aufwand kann derzeit nur vermieden werden, wenn die be-

troffenen Archive die Retrokonversion mit zusätzlichem Personal im eigenen Haus durchführen. Dass ein erhöhter Bedarf an Standardschnittstellen tatsächlich dazu führt, dass die Anbieter von Archivsoftware ihre Produkte entsprechend weiterentwickeln, ist in vielen Fällen leider noch nicht zu erkennen.

Datenaustausch in der Praxis: Das archivspartenübergreifende Portal „Archive in NRW“ nach dem Relaunch (2007)

von Peter Worm

Als die nordrhein-westfälischen Staatsarchive zusammen mit den Archivberatungsstellen der Landesteile Rheinland und Westfalen-Lippe gegen Mitte der 1990er Jahre die Konzeption eines archivspartenübergreifenden Internetportals begannen, steckte das Medium „Internet“ noch in den Kinderschuhen. Kaum ein deutsches Archiv hatte eine Homepage und die wenigsten Archivare bzw. Archivarinnen konnten sich vorstellen, welche Omnipräsenz diesem weltweiten Informationsnetz eine Dekade später zukommen würde. Im Rückblick muss man den von Wilfried Reininghaus und Frank Bischoff initiierten, spartenübergreifenden Ansatz fast als visionär bezeichnen.

Mit einem Schlag brachte das Portal 1998 insgesamt 420 Archive mit ihren Kontakt- und Beständeinformationen ins Netz und bildete dadurch über 80% der zu diesem Zeitpunkt überhaupt im Internet vertretenen deutschen Archive ab. 90% dieser Archive waren Kommunalarchive, über 175 aus dem Rheinland, über 220 aus Westfalen-Lippe. Dem Unternehmen kam dabei sehr zugute, dass 1994 für das Rheinland und 1996 für Westfalen ein gedrucktes „Handbuch der Kommunalarchive in NRW“ erschienen war, dessen Informationen für den Internetauftritt aufbereitet werden konnten und eine solide Basis für die weitere Arbeit lieferten.¹

¹ Peter Weber u.a. (Bearb.), Handbuch der Kommunalarchive in Nordrhein-Westfalen. Teil 1: Landesteil Nordrhein, Köln 1994 (Archivhefte 27); Alfred Bruns

Auf die näheren Umstände der Einrichtung und weiteren Entwicklung von Archive.NRW.de möchte ich hier nicht näher eingehen, da Mechthild Black-Veldtrup, Andreas Pilger und Martina Wiech darüber auf dem Westfälischen Archivtag 2009 in Detmold berichtet haben.²

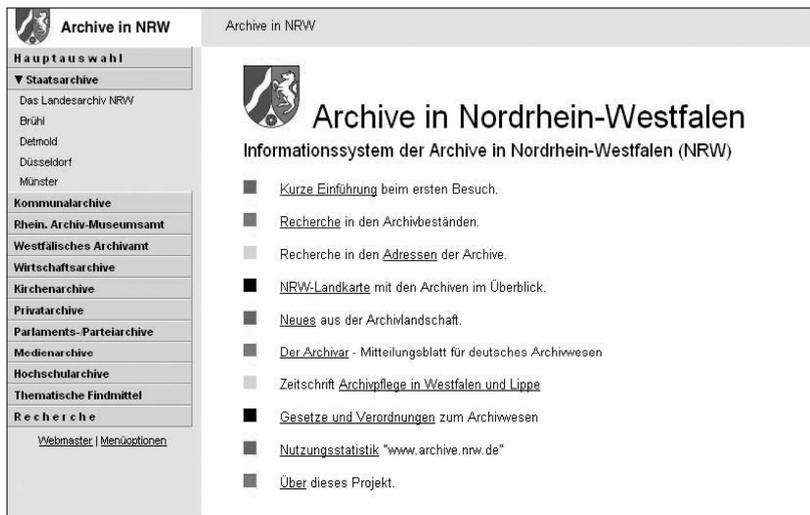


Abb. 1: Die Startseite des Internetportals „Archiv in NRW“ im ursprünglichen Web-Design (1998 – April 2007)

unter Mitarbeit der Kommunalarchive in Westfalen-Lippe, Handbuch der Kommunalarchive in Nordrhein-Westfalen. Teil 2: Landesteil Westfalen-Lippe, Münster 1996 (Westfälische Quellen und Archivpublikationen 21). Das Internetportal ist abrufbar unter <http://www.archive.nrw.de>. Alle im Beitrag genannten Webseiten wurden am 26.1.2010 besucht.

² Der Bericht mit umfassenden Literaturangaben über den Entstehungskontext des NRW Archivportals ist abgedruckt in: Mechthild Black-Veldtrup/Andreas

Die alte Version des Portals ermöglichte das Onlinestellen von Kontaktdaten, Beständeübersichten und aktuellen Nachrichten, doch konnten Online-Findbücher oder Archivalienabbildungen nicht eingebunden werden. Das framebasierte Layout entsprach nicht mehr den optischen und rechtlichen Anforderungen – Stichwort „Barrierefreiheit“ – als man sich 2004 zu einer Neukonzeption, einem sogenannten Relaunch, der Internetpräsenz entschloss.

Zwei Jahre später konnte ein Prototyp des neuen Portals auf dem westfälischen und rheinischen Archivtag vorgestellt werden. Erste Schulungen wurden für das Landesarchiv Nordrhein-Westfalen und sogenannte Multiplikatoren – u.a. die Archivämter – durchgeführt. Noch vor dem Startschuss des neuen Portals führte das LWL-Archivamt für Westfalen Ende 2006 und Anfang 2007 für interessierte Kolleginnen und Kollegen aus den Kommunalarchiven Fortbildungen im Redaktionssystem NPS durch. Als das neue Portal im Mai 2007 online gestellt wurde, enthielt es bereits neue und tiefere Informationen als das alte System. Zudem bot es ein frischeres, aus heutiger Sicht vielleicht ein wenig zu nüchternes Erscheinungsbild.

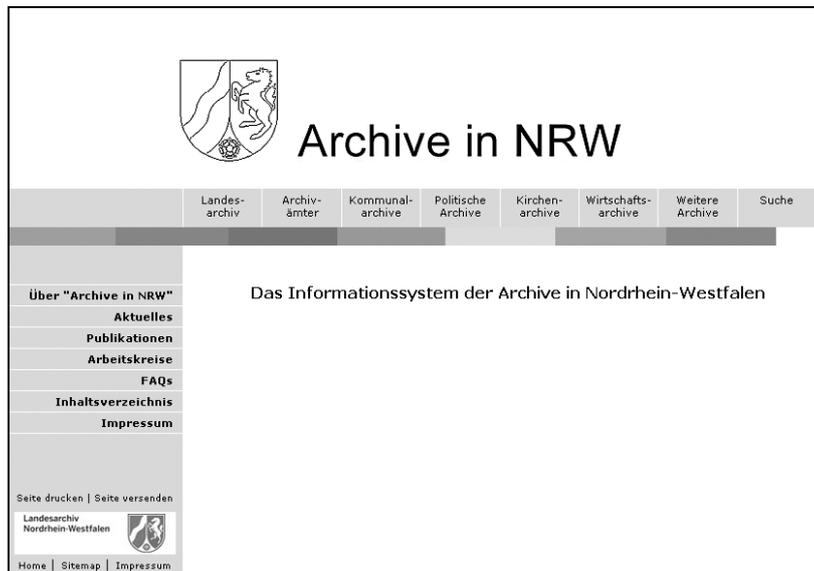


Abb. 2: Die Startseite des Internetportals „Archive in NRW“ <<http://www.archive.nrw.de/>> nach dem Relaunch im Mai 2007

Die Zugriffszahlen des Portals zeigen die hohe Akzeptanz bei den Internetnutzern: Das Portal transferierte im Jahr 2008 Daten im Umfang von knapp 500 GB, im Jahr 2009 waren es fast 630 GB. Die Anzahl der Besucher – gezählt wurden hierbei die unterschiedlichen IP-Adressen – lag im gleichen Zeitraum konstant bei 1,3 Millionen pro Jahr. Neben den wenigen, eher systematisch arbeitenden Benutzern, die den Recherche-Einstieg über Archive.NRW.de nehmen, kommen wohl die meisten Interessenten über die Trefferlisten der Suchmaschine Google auf Informationen aus dem Archivportal. Vielleicht betrübt es manchen Fachkollegen und manche Fachkollegin,

doch sollten die Archive m.E. froh sein, dass auf diese Weise archivfremde Nutzergruppen mit archivischen Quellen in Kontakt kommen. Positiv ist hervorzuheben, dass Archive.NRW.de eine hohe Glaubwürdigkeit im Google-Ranking besitzt, und die Treffer deshalb auch in den Ergebnislisten weit oben angezeigt werden. Hier haben die Archive bzw. das genannte Portal den Bibliotheken etwas voraus, deren Online-Kataloge i.d.R. keine Berücksichtigung beim Suchgiganten finden.

Wie verteilen sich nach heutigem Stand die Erschließungsinformationen rein mengenmäßig im Archivportal?³ Das Landesarchiv Nordrhein-Westfalen kann durch die Bemühungen der letzten Jahre auf dem Feld der Retrokonversion mehr als die Hälfte der insgesamt eingestellten Verzeichnungseinheiten beitragen. Dem folgen die rheinischen Kommunalarchive, interner Spitzenreiter ist hier das Historische Archiv der Stadt Köln mit ca. 175 Online-Findmitteln. Der Landesteil Westfalen steht an dritter Stelle. Aus dem LWL-Archivamt für Westfalen kommen noch rd. 30.000 Verzeichnungseinheiten aus den von dort betreuten Privatarchiven der Vereinigten Westfälischen Adelsarchive e.V. hinzu.⁴ Damit ist ein ungefährer Gleichstand mit den rheinischen Beständen erzielt.

³ Die folgenden Angaben beruhen auf dem Stand von Oktober 2009.

⁴ Die Vereinigten Westfälischen Adelsarchive e.V. nutzen ihre Seiten auf Archive in NRW auch als Hauptinternetpräsenz. Vereinsinformationen, Beständeübersicht und Findbücher sind erreichbar über den Link <http://www.adelsarchive.de> [Dez. 2009].

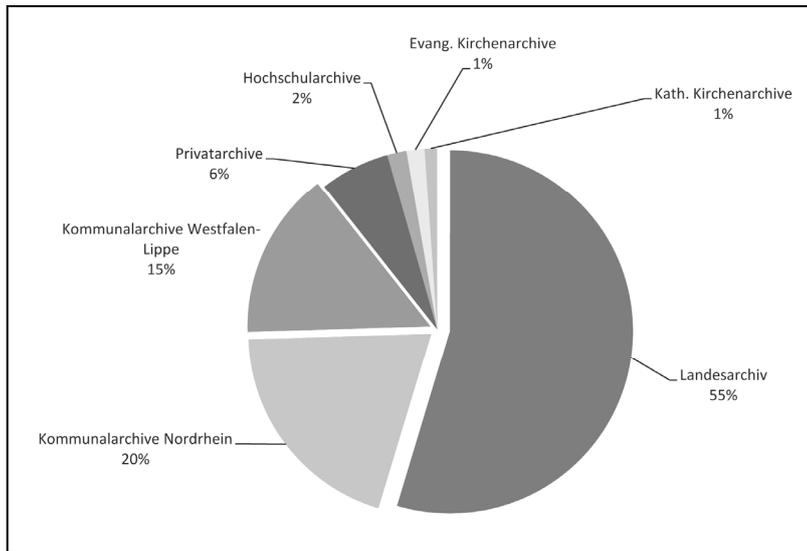


Abb. 3: Anzahl der Verzeichnungseinheiten pro Archivsparte (Stand Okt. 2009)

Die übrigen Sparten, d.h. Wirtschafts-, Kirchen- und Hochschularchive sowie die Parteien- und Parlamentsarchive haben bisher leider nur kleine Mengen an Erschließungsdaten online gestellt. Dieser Umstand liegt z.T. sicher an den institutionellen Rahmenbedingungen. Setzt man die knapp 250 westfälischen Kommunalarchive zu den 134 bisher eingestellten Findbüchern in Relation, bleibt im Schnitt jedoch auch nur ein halbes Findbuch pro Kommune übrig. Bei den Beständeübersichten haben seit Mai 2007 nur 11 bis 15% der Kommunalarchive Ergänzungen und Umstrukturierungen vorgenommen. Manche haben in der Vorgängerversion des Portals ihre Tektoniken fortgeschrieben und abgebildet, doch bei vielen größeren und kleineren Archiven findet man noch den Stand von Mitte der 1990er

Jahre. Alles in allem eine etwas ernüchternde Bilanz, die eine genauere Betrachtung verlangt.

Für die Frequenz der Findbucheinstellungen lässt sich nach einer regen Einführungsphase ein langes „Tal“ geringerer Aktivität ausmachen, aus dem wir uns aber in den letzten Monaten langsam herausbewegen. Die Gründe dafür liegen nicht zuletzt in der Problematik fehlender oder nicht gut funktionierender Schnittstellen zu den Archivprogrammen.

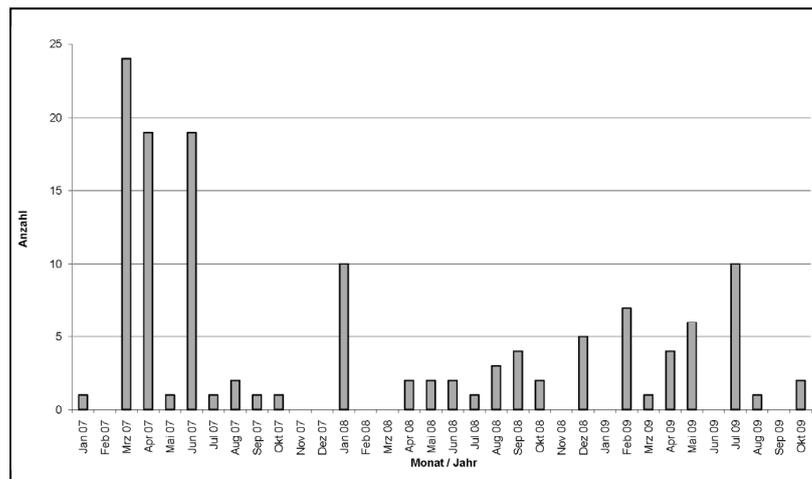


Abb. 4: Frequenz der neu eingestellten Online-Findbücher westfälischer Kommunalarchive

Betrachtet man die räumliche Verteilung von Online-Findbüchern und aktualisierten Beständeübersichten, ergibt sich ein erstaunliches Bild: Vorreiter sind die eher ländlich geprägten Gebiete v.a. des südlichen Münsterlands, des märkischen Sauerlands und Ostwestfalens. Die Großstadtarchive sind deutlich unterrepräsentiert.

Auf ihren eigenen Internetseiten bieten zumindest einige dieser Stadtarchive Beständeübersichten an, zwei sogar die Möglichkeit der Findbuchrecherche.⁵

Es ist somit klar, dass wir erst am Anfang einer Entwicklung stehen: Während mancher Archivar schon für das Web 2.0 und die Positionierung der Archive dort wirbt, sind viele Archive noch nicht im Web 1.0 angekommen.

Ein Grund wurde bereits angedeutet, er liegt im Zusammenwirken von Erschließungssoftware und Archivportal. In das Archivportal werden Beständeübersichten und Findbücher getrennt voneinander eingepflegt. Für die Beständeübersichten existieren drei Möglichkeiten:

1. Zum einen können Beständedaten, die nach der Struktur des alten Portals formatiert sind, hochgeladen werden. Diese Möglichkeit nutzt kein Archiv mehr, weshalb sie nicht näher betrachtet wird.
2. Die zweite Möglichkeit besteht im sogenannten SAFT (Standard-Austauschformat) aus dem Retrokonversionsprojekt. Diese Variante unterstützt das Programm V.E.R.A. des Landesarchivs NRW sowie das Programm ACTApro. Des Weiteren können valide XML-Beständeübersichten mit überschaubarem Aufwand aus

⁵ Das Historische Centrum Hagen stellt seine Findmittel unter <http://www.historisches-centrum.de/index.php?id=132> zur Verfügung. Das Gelsenkirchener Institut für Stadtgeschichte stellt seine Findbücher in Form von PDF-Dateien ins Internet: <http://www.institut-fuer-stadtgeschichte.de/Stadtarchiv/GZS.asp>.

Textdateien händisch erzeugt werden. Diese Dienstleistung bietet das LWL-Archivamt für Westfalen den Kommunalarchiven an.

3. Schließlich existiert eine zweite Schnittstelle, die Dateien nutzt, die mit dem Internet-Assistenten des Erschließungsprogramms AUGIAS erzeugt wurden. Nach anfänglichen Kompatibilitätsproblemen funktioniert diese Möglichkeit nun seit ca. Mitte 2009.

Findbuchdateien können über SAFT-konforme XML-Dateien aus V.E.R.A., ACTApro und MidosaXML nach Archive.NRW.de importiert werden. Seit Ende 2008 bietet Faust eine konfigurierbare Exportmöglichkeit, die SAFT abdeckt. Auch aus AUGIAS kann mittels des Findbuch-Assistenten und kleinen händischen Nachbearbeitungen SAFT-XML erzeugt werden. Schließlich existiert seit kurzem eine funktionierende Schnittstelle für AUGIAS-Exportdateien, die mit dem sogenannten Internetassistenten erzeugt wurden.⁶

⁶ Es handelt sich bei der erzeugten „*.upl-Datei“ nicht um durchgängig strukturiertes XML, sondern im Kern um eine RAR-komprimierte CSV-Datei. Diese ist für den Benutzer nicht unmittelbar einsehbar bzw. verständlich und wird serverseitig beim Import nach Archive.NRW.de ausgewertet.

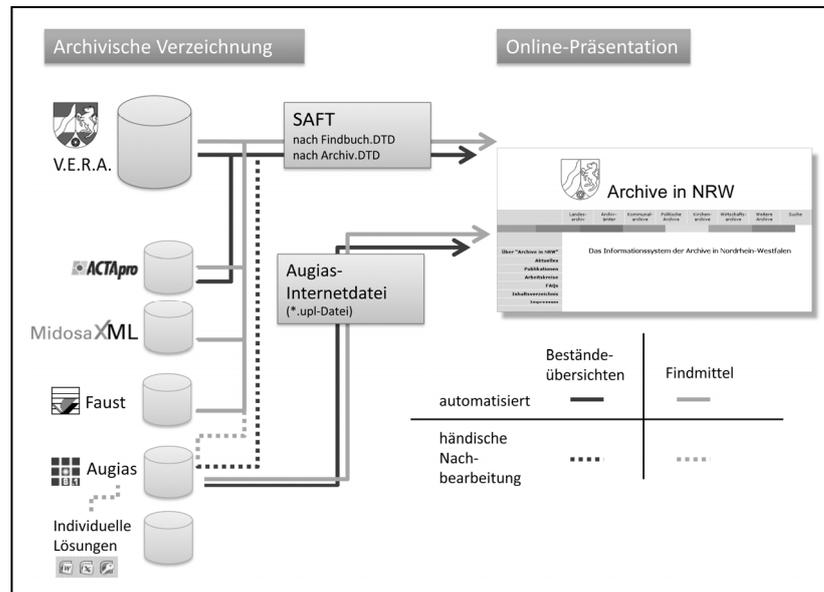


Abb. 6: Datenimporte von Beständeübersichten und Findbüchern in das Archivportal

Betrachtet man die ins Portal eingestellten Findbücher aus westfälischen Kommunalarchiven, so überwiegt derzeit klar der Anteil an Daten aus AUGIAS (> 80%). Ca. 2/3 der Archive haben den Weg über eine XML-Datei gewählt (via modifizierten Findbuchausdruck), das restliche Drittel nutzt die spezielle AUGIAS-Schnittstelle (via „AUGIAS-Internetassistenten“). Es folgen zehn Prozent an Erschließungsdaten, die in Excel- oder Accesstabellen erfasst wurden. Sie wurden zumeist über den Import der Tabellen nach AUGIAS ins Archivportal eingestellt. Diese Serviceleistung bietet das LWL-Archivamt besonders den kleinen Gemeindearchiven an, die über

keine Erschließungssoftware verfügen. Einzelne Findbücher wurden aus MidosaXML oder ACTApro erzeugt. Auch der SAFT-Export aus Faust funktioniert, wie ca. ein Prozent der Online-Findbücher belegen. Eine Im- und Exportschnittstelle zu EAD ist derzeit vom Landesarchiv NRW geplant.

Stefan Lang hat in seiner Transferarbeit über die Nutzeranforderungen an Internet-Auftritte von Kommunalarchiven deutlich gemacht, dass die Benutzer auf den Internetseiten der Archive mehr erwarten als nur Kontaktdaten und Öffnungszeiten.⁷ Aussagekräftige Beständeübersichten und Online-Findbücher werden in zunehmendem Maße eingefordert. Die Kommunalarchive haben in Nordrhein-Westfalen die Möglichkeit, kostenlos und mit überschaubarem Aufwand über das Internetportal auf ihre Bestände aufmerksam zu machen. Eines der Ziele des LWL-Archivamtes für Westfalen wird sein, noch stärker als in der Vergangenheit für das Medium Internet allgemein und das Archivportal im Besonderen zu werben. Als Minimum sollte eine aktuelle und ausführliche Beständeübersicht gelten. Die Voraussetzung dafür besteht in einer leicht bedienbaren und evtl. von Archivprogrammen unabhängigen Pflegemöglichkeit.

Ein profunder Internetauftritt macht die Arbeitsleistung gegenüber der eigenen Kommunalverwaltung und der Öffentlichkeit sicht-

⁷ Stefan Lang, Nutzeranforderungen und Informationsstandards. Internet-Auftritte von Kommunalarchiven am Beispiel Baden-Württembergs, Marburg 2009 (Archivschule Marburg, Transferarbeit), v.a. S. 46-47, online abrufbar unter: http://www.landesarchiv-bw.de/sixcms/media.php/120/49758/Lang_Transferarbeit.29487.pdf.

bar. Zudem holt er das Archiv aus der Ecke des „Verstaubten“ und der technischen Rückständigkeit heraus, die viele mit dieser Institution verbinden. Die bisherigen Erfahrungen der Kolleginnen und Kollegen, die stark auf Internetfindbücher setzen, zeigen, dass Benutzeranfragen sehr viel gezielter gestellt und somit leichter beantwortbar werden. Globale Anfragen zur Familiengeschichte o.ä. gehen zurück, es wird stärker nach einzelnen Archivalien oder deren Reproduktion gefragt. Es ergibt sich also für das einzelne Archiv ein arbeitsökonomischer Vorteil. Geografisch erweitert sich der Interessentenkreis: Es kommen verstärkt Anfragen aus dem gesamten deutschsprachigen Raum, aber auch aus dem Ausland. Es macht Sinn, die Zahl der Online-Nutzer als Kennziffer für den Rechenschaftsbericht des Archivs aufzunehmen. Der befürchtete Rückgang der persönlichen Benutzung ist übrigens bisher bei keinem im Netz engagierten Archiv zu verzeichnen.

Welche weiteren Ausbaustufen bieten sich? Die einfache Stichwortsuche wird durch das Portal aber auch durch die großen Suchmaschinen angeboten. Das Portal sollte deshalb neben der bereits existenten „Erweiterten Suche“ gezielt in Richtung intelligenter Suchmöglichkeiten ausgebaut werden. Eine phonetische Suche böte z.B. den Vorteil, dass bei archivischen Verzeichnissen, die sich stark am Buchstabenbestand der in den Quellen vorgefundenen Orts- und Personennamen orientieren, bessere Ergebnisse erzielt werden könnten. Durch assoziative Suchmöglichkeiten, wie eine sogenannte „Cloud-Suche“, würden dem Benutzer Eingrenzungen oder Erweiterungen zu einem eingegebenen Stichwort vorgeschlagen. Gerade archivunerfahrene Benutzer könnten dadurch größere Sucherfolge erzielen.

Das Landesarchiv NRW plant die umfangreiche Präsentation von digitalisiertem Archivgut auf Basis des METS-Standards. Diese Ausweitung des Portals wird erneut die Umsetzung von funktionierenden Schnittstellen aus den Archivprogrammen erfordern. Die Erfahrungen mit den Erschließungsdaten zeigen, dass bei den Herstellern von Archivsoftware mitunter „dicke Bretter“ zu bohren sind. Auch rechtliche und fachliche Vorbehalte gegen das Onlinestellen von Archivgut müssten hier im Vorfeld ausgeräumt werden.

Schließlich könnten die Trefferlisten, die Findbuchpräsentation und die Vorbestellfunktion für Archivalien noch in einzelnen Punkten verbessert werden. Die ein- bis zweimal im Jahr tagenden Power-User-Runden, an denen IT.NRW, das Landesarchiv NRW, die Archivämter und weitere interessierte Archivarinnen und Archivare teilnehmen, haben sich als gutes Gremium für die Bündelung und Abstimmung solcher Vorschläge bewährt. Ich bin sicher, dass ein fachlich und technisch gut aufgestelltes Archivportal mit regionalem Bezug auch in Zeiten eines gesamtdeutschen oder gar europäischen Portals seine Berechtigung und seinen Interessentenkreis behalten wird.

„Europeana“ – die europäische digitale Bibliothek und der Aufbau einer „Deutschen Digitalen Bibliothek“. Stand und Perspektiven

von Gerald Maier

1. Einleitung

Bei der Online-Präsentation von digitalisiertem Kulturgut kommt einer interdisziplinären Zusammenarbeit zwischen den verschiedenen Kulturgut-Sparten eine immer größere Bedeutung zu, die insbesondere auf der europäischen Ebene zunehmend propagiert wird, aber auch im Sinne eines verbesserten Dienstleistungsgedankens auf nationaler Ebene sinnvoll erscheint.

In der Europäischen Union erfolgt z.Zt. der Aufbau einer interdisziplinären „European Digital Library“, die den Namen „Europeana“ trägt.¹ Ziel ist es, den „eContent“ – d.h. das digitalisierte Kulturgut – aus allen 27 EU-Mitgliedsstaaten über ein Zugangsportale bereitzustellen. In Deutschland wird auf nationaler Ebene eine spartenübergreifende „Deutsche Digitale Bibliothek“ (DDB) in Zusammenarbeit von Bund, Ländern und Kommunen errichtet, die als eine Internet-Plattform mit Suchmaschine einen gemeinsamen Zugangspunkt zu digitalen und digitalisierten Beständen in deutschen Bibliotheken, Archiven, Museen sowie zu wissenschaftlichen Informationen schafft

¹ <<http://www.europeana.eu>>(Prototyp). Der Zugriff auf alle im Beitrag genannten Webseiten erfolgte am 8.2.2010.

und gleichzeitig auch die Funktion eines nationalen Aggregators² für die „Europeana“ übernehmen kann.

Im folgenden Beitrag werden im ersten Teil Aufbau und aktueller Stand der „Europeana“ beschrieben. Der zweite Teil beschäftigt sich mit der Errichtung der „Deutschen Digitalen Bibliothek“. Bei beiden Informationssystemen werden auch Perspektiven für die deutschen Archive aufgezeigt.

2. Die „Europäische Digitale Bibliothek – Europeana“

2.1 Ziele und Rahmenbedingungen

Der Aufbau einer interdisziplinären europäischen digitalen Bibliothek mit dem verbindenden Namen „Europeana“ resultiert aus den Zielen und Aktivitäten der europäischen Union zur Förderung der Informationsgesellschaft, die sich aktuell in der Initiative „i2010“ konkretisiert.³ Ein Ziel der Initiative ist auch die bereits in den „eEurope“-Aktionsplänen forcierte Digitalisierung des kulturellen Erbes.⁴ Ein wichtiger politischer Meilenstein waren dabei die Empfehlungen der EU-Kommission und die EU-Ratsentschließung „Digitalisierung und Online-Zugänglichkeit kulturellen Materials und dessen digitaler Be-

² Ein Aggregator ist ein Projekt oder eine Stelle, die Inhalte (eContent) kleinerer Einrichtungen sammelt, ggf. aufbereitet und dann geschlossen weiterreicht. Es gibt Aggregatoren, deren einziges Ziel die Bündelung von Inhalten zur Weitergabe an Dritte ist und solche, bei denen die Aggregation einen Nebeneffekt anderer Aufgaben darstellt.

³ Vgl. dazu <http://ec.europa.eu/information_society/eeurope/i2010/index_en.htm>.

⁴ Vgl. <http://ec.europa.eu/information_society/activities/digital_libraries/index_de.htm>.

wahrung“ vom August/November 2006.⁵ Hier wurde als vorrangiges Ziel im Bereich der Kulturgutdigitalisierung die Errichtung einer interdisziplinären „Europäischen Digitalen Bibliothek“ festgeschrieben. Begleitet und vorbereitet wird der politische Prozess durch Expertengremien der EU-Kommission wie die „Member States Expert Group“ (MSEG) und die „High Level Group“.⁶

Realisiert wird die „Europeana“ in Projekten mit dem Charakter eines thematischen Netzwerks. Ein erstes zweijähriges Projekt von Juli 2007 bis Januar 2009 mit etwa 90 Partnern unterschiedlicher Sparten aus verschiedenen europäischen Mitgliedsstaaten beschäftigte sich sowohl mit der organisatorischen, technischen und fachlichen Konzeption als auch mit der prototypischen Realisierung der „Europäischen Digitalen Bibliothek“. Eine erste Version der „Europeana“ mit rund zwei Millionen digitalen Objekten wurde im November 2008 der Öffentlichkeit präsentiert.

Seit Februar 2009 wird die Weiterentwicklung des Prototyps zu einer stabilen Anwendung und einem persistenten Service inklusive Geschäftsmodell im Rahmen eines als thematisches Netzwerk angelegten Nachfolgeprojekts „Europeana v1.0“ weiterentwickelt.⁷ Finan-

⁵ <http://ec.europa.eu/information_society/newsroom/cf/document.cfm?action=display&doc_id=161>.

⁶ Vgl. <http://ec.europa.eu/information_society/activities/digital_libraries/experts/mseg/index_en.htm> (MSEG) und <http://ec.europa.eu/information_society/activities/digital_libraries/experts/hleg/index_en.htm> (High Level Group).

⁷ <<http://version1.europeana.eu/web/europeana-project>>. Die Laufzeit des Projekts beträgt 30 Monate (Februar 2009-Juli 2011) mit einem Projektbudget von 6,2 Mio. EUR.

ziert wird das Projekt durch das eContentPlus-Förderprogramm der EU. An dem Netzwerk sind über 100 Partnerinstitutionen aus allen EU-Mitgliedstaaten beteiligt. Insgesamt wurden sechs Arbeitspakete definiert. Von archivischer Seite arbeitet das Landesarchiv Baden-Württemberg als Projektpartner im Arbeitspaket „Further Specification of Functionality and Interoperability aspects of Europeana“ mit, um zusammen mit den Vertretern des APEnet-Projekts⁸ die fachlichen Festlegungen im Hinblick auf die Präsentation von digitalisiertem Archivgut fortzuschreiben.⁹

Neben dem eigentlichen „Europeana“-Projekt gibt es ein Netzwerk mit über zehn Projekten – die sogenannte „Europeana-Group“.¹⁰ Die Projekte unterstützen die „Europeana“, indem sie als alleiniges oder zusätzliches Ziel neue Funktionalitäten und/oder neue Inhalte in die „Europeana“ einbringen. Die meisten dieser Projekte werden

⁸ Vgl. dazu <<http://www.apenet.eu>> und <<http://group.europeana.eu/web/guest/details-apenet/>>.

⁹ Weitere institutionelle Projektpartner aus Deutschland sind die Deutsche Nationalbibliothek, das Bibliotheksservice-Zentrum Baden-Württemberg, der Rundfunk Berlin-Brandenburg, das Leibniz-Institut für Sozialwissenschaften in Mannheim, die Universitäten Göttingen und Bielefeld, die Stiftung Preußischer Kulturbesitz und das Institut für Bibliotheks- und Informationswissenschaft (IBI) der Humboldt-Universität Berlin.

¹⁰ <<http://group.europeana.eu>>. Aktuelle Europeana-Group-Projekte sind: EuropeanConnect, EuropeanaLocal, European Film Gateway (EFG), EUScreen, PrestoPrime, Biodiversity Heritage Library Europe (BHLEurope), Access to cultural heritage networks across Europe (Athena), Archives Portal of Europe on the Internet (APEnet), EuropeanaTravel, MIMO (Musical Instruments Museum Online), JUDEICA (Jewish Urban Digital European Integrated Cultural Archive) und STERNA (Semantic Web-based Thematic European Reference Network Application).

ebenfalls von der EU-Kommission im Rahmen des eContentPlus- bzw. CIP-Programms unterstützt.¹¹ Die Projektbeteiligten sind in das „Europeana Thematic Network“ eingebunden, um voneinander zu profitieren und Redundanzen zu vermeiden. Zur „Europeana-Group“ gehört auch das oben genannte APEnet-Projekt, in dessen Kontext ein europäisches Archivportal aufgebaut werden soll und das im Grunde die archivische Sicht der „Europeana“ konzipiert.

2.2 Digitalisierte Inhalte – Aktueller Stand des „eContent“

In der Zwischenzeit kann in der „Europeana“ auf 4,6 Millionen digitalisierte Bücher, Karten, Fotografien, Filmclips und Zeitungen zugegriffen werden. Seit dem Launch des Prototyps im November 2008 hat sich die Anzahl an verfügbaren digitalisierten Objekten damit mehr als verdoppelt. Derzeit stellen mehr als 1.000 Kultureinrichtungen ihre Inhalte für „Europeana“ bereit und mehr als 150 Einrichtungen nehmen an ihrem Partnernetzwerk teil. Momentan liegt Deutschland mit der Anzahl seiner Beiträge auf Platz zwei hinter Frankreich.¹² Die Europäische Kommission hat in einem Strategiepapier vom Au-

¹¹ CIP – Competitiveness and Innovation Framework Programme (Rahmenprogramm für Wettbewerbsfähigkeit und Innovation 2007-2013) (<http://ec.europa.eu/cip/index_de.htm>): Das CIP-Program ist der Nachfolger des Ende 2008 abgeschlossenen eContentPlus-Programms. Erstmals ist nun auch die direkte Förderung für die Erstellung von eContent möglich. Ein erster Aufruf für Anträge endete im Juni 2009.

¹² 47% des eContent kommt von Institutionen aus Frankreich, 15,4% aus Deutschland. Alle weiteren EU-Staaten sind jeweils mit unter 10% vertreten. Größter deutscher eContent-Lieferant ist derzeit die SLUB Dresden mit über 530.000 digitalisierten Bildern.

gust 2009 erklärt,¹³ dass sie bis 2010 eine Erhöhung der Zahl der digitalisierten Objekte auf 10 Millionen anstrebt. Außerdem hat sie eine öffentliche Debatte über die künftigen Herausforderungen angestoßen, die sich im Zusammenhang mit der Digitalisierung von Büchern in Europa stellen, insbesondere die Möglichkeiten einer Partnerschaft zwischen öffentlichem und privatem Sektor sowie die Notwendigkeit einer Reform des allzu uneinheitlichen Urheberrechtsrahmens in Europa¹⁴ (siehe Abb. 1).

Von archivischer Seite ist die Bereitstellung von Inhalten für die „Europeana“ – wohl auch aufgrund der im Prototyp nur unstrukturierten Präsentation der Objekte – bis jetzt nur gering. Das Landesarchiv Baden-Württemberg hat z.Zt. 64.000 digitalisierte Objekte über das bisher für archivische Anforderungen unzureichende Austauschformat geliefert, die im bisherigen Prototyp auch recherchiert werden können.

Bislang werden nur Digitalisate in der „Europeana“ nachgewiesen. Es wird aber intensiv diskutiert und von archivischer Seite auch gefordert, dass künftig auch Metainformationen – also Erschließungsdaten – ohne digitale Objekte aufgenommen werden sollen (z.B. archivische Findbücher). Einziger deutscher „eContent“-Lieferant

¹³ <<http://europa.eu/rapid/pressReleasesAction.do?reference=IP/09/1257&format=HTML&aged=0&language=EN&guiLanguage=en>> KOM-Strategie-Papier vom 28.08.2009).

¹⁴ Zur Erörterung der Fragen hat die Kommission eine öffentliche Konsultation über die Zukunft von Europeana und die Digitalisierung von Büchern durchgeführt. Vgl. <http://ec.europa.eu/information_society/activities/digital_libraries/doc/communications/next_steps_2009/questions_de.pdf> .

aus dem Bereich der Archive ist bisher das Landesarchiv Baden-Württemberg, welches digitalisiertes Archivgut aus 24 Beständen mit mehr als 64.000 Digitalisaten bereitstellt. Darunter befinden sich illuminierte Urkunden, historische Ortsansichten, Fotografien, Plakat-sammlungen und Präsentationen von Patentakten mit Patentzeich-nungen.

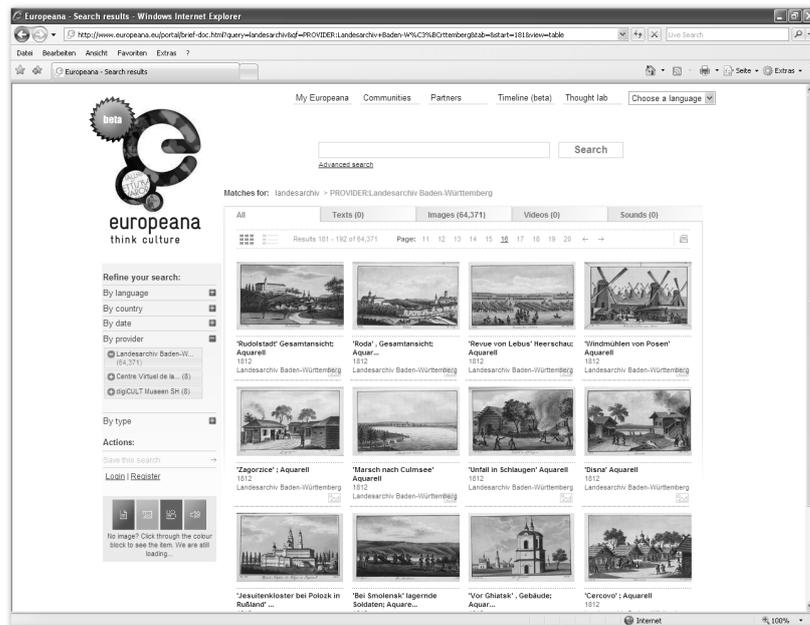


Abb. 1: „Europeana“ – Landesarchiv Baden-Württemberg, HStAS J56: Bilde-alben aus dem Nachlass Christian von Martens (1793-1882) aus dem Zeit-raum 1812-1865

2.3 Finanzierung und Organisation

Für die Aufbauphase der „Europeana“ (2009-2011) steht ein finanzielles Gesamtvolumen von 7,5 Millionen EUR (2,5 Millionen EUR jährlich) zur Verfügung. Dabei fließen 80% der Mittel aus dem eContentPlus-Programm der EU. Die restlichen 20% werden komplementär durch die Mitgliedsstaaten finanziert. Bis 2013 ist eine weitere Unterstützung von 9 Millionen EUR aus dem EU-Rahmenprogramm für Wettbewerbsfähigkeit und Innovation (CIP-Programm) vorgesehen. Die mittelfristige Finanzierung (ab 2013) ist noch nicht geklärt. Diskutiert werden verschiedene Möglichkeiten, u.a. öffentlich-private Partnerschaften (z.B. Sponsoring, Technologie-Partnerschaften) und/oder Beiträge der Mitgliedsstaaten (freiwillig oder nach einem Verteilungsschlüssel).

Verwaltet werden die Mittel durch die sogenannte EDL-Foundation.¹⁵ Dabei handelt es sich um eine spartenübergreifende Stiftung nach niederländischem Recht mit Sitz in Den Haag, die im November 2007 gegründet wurde. Sie ist Dachorganisation und Projektkoordinator für die eigentliche „Europeana“ und die „Europeana-Group“. Weitere Aufgaben der EDL-Foundation sind das Vertragsmanagement zwischen den beteiligten Partnern und die Schaffung finanzieller Rahmenbedingungen inklusive Finanzmittelverwaltung und -akquisition. Operativ bedient sich die EDL-Foundation eines „Executive Committee“ als Steuerungsgremium der „Europeana“-Geschäftsstelle, die an der Königlichen Bibliothek in Den Haag angesiedelt ist. Zu den Aufgaben der Geschäftsstelle gehören auch die

¹⁵ <<http://version1.europeana.eu/web/guest/edl-foundation>> .

zentrale Entwicklung und der operative Betrieb des „Europeana“-Portals.

2.4 Datenlieferung und Datenintegration

Die Datenlieferung an die „Europeana“ durch die Kultur- und Wissenschaftsinstitutionen erfolgte für die gegenwärtige Version des Portals im sogenannten ESE-XML-Format.¹⁶ Die ESE-XML-Dateien enthalten Hyperlinks zu Bild- bzw. Multimediadateien, die physisch auf den Servern der datenliefernden Institution bleiben. Nachdem die Daten in das zentrale „Europeana“-Informationssystem importiert worden sind, erfolgt eine automatisierte Vorschaubild-Generierung durch das zentrale System auf der Basis der in den Metadaten enthaltenen Hyperlinks. Bisher werden die Metadaten manuell an die „Europeana“ übermittelt, mittelfristig ist ein Harvesting über eine OAI-Schnittstelle vorgesehen. ESE ist zum jetzigen Zeitpunkt das Pflichtformat für alle Content-Provider. Es basiert auf Dublin Core und ist um eigene „Europeana“-Elemente erweitert worden. Die Ausgangsformate für ein Mapping in das ESE-Format bei den datenliefernden Einrichtungen sind v.a. DC (Dublin Core), METS/MODS (basiert auf MAB2) für digitalisierte Bücher,¹⁷ museumdat für digitalisierte Museumsobjekte und EAD (Encoded Archival Description) für digitalisiertes Archivgut. Das ESE-Format ist primär auf die Präsentation einzel-

¹⁶ ESE: „Europeana Semantic Elements“. Vgl. dazu die „Europeana Semantic Elements specifications v3.2.2“ unter http://version1.europeana.eu/c/document_library/get_file?uuid=c56f82a4-8191-42fa-9379-4d5ff8c4ff75&groupId=10602.

¹⁷ METS/MODS: „Metadata Encoding and Transmission Standard“ in Verbindung mit dem „Metadata Object Description Schema“.

ner Objekte ausgerichtet und für die kontextbezogene Präsentation archivischer Inhalte eher ungeeignet, da insbesondere Strukturinformationen nicht übernommen werden.



```

- <c level="file" id="DE_ArchLABW_1_100488">
- <did>
- <unitid>EA 99/002 B0 18</unitid>
- <unittitle>
Vorschläge von Dr. Ing. Carl H. Schweizer aus Stuttgart, Regierungsbaudirektor im Innenministerium, zu
einer Flagge und zu einem Wappen für Württemberg-Baden bzw. den Süweststaat
</unittitle>
- <physdesc>
<genreform>Sachakte
<dimensions>1 cm (2
</physdesc>
</did>
- <scopecontent>
- <list>
- <item>
Enthält 1 Entwurf
des Landeswappens
Herzogtums Schwab
</item>
</list>
</scopecontent>
- <otherfindaid>
<extref role="url" archiv
/olff/struktur.php?bestand
</otherfindaid>
</index>
- <indexentry>
- <indexentry>
<persname>
Schweizer, Carl H.
</persname>
</indexentry>
- <indexentry>
<geogname>Stuttgart
</indexentry>
</c>
<record>
<dc:identifier>DE_ArchLABW_1_100489</dc:identifier>
<dc:terms:isPartOf>Hauptstaatsarchiv/Archivdirektion Stuttgart:
Sachverständigenkommission Landesname, Landeswappen,
Landesfarben von Baden-Württemberg</dc:terms:isPartOf>
<dc:source>Landesarchiv Baden-Württemberg: Hauptstaatsarchiv Stuttgart</dc:source>
<dc:source>http://www.landearchiv-bw.de/hstas</dc:source>
<dc:relation>http://www.landearchiv-bw.de/plink/?f=1-1697</dc:relation>
<dc:identifier>EA 99/002 Nr. 18</dc:identifier>
<dc:description>Kontext: Hauptstaatsarchiv/Archivdirektion Stuttgart:
Sachverständigenkommission Landesname, Landeswappen, Landesfarben
von Baden-Württemberg >> 3. Entwürfe von Privaten, Vereinen >>
Vorschläge von Dr. Ing. Carl H. Schweizer aus Stuttgart,
Regierungsbaudirektor im Innenministerium, zu einer Flagge und zu
einem Wappen für Württemberg-Baden bzw. den Süweststaat</dc:description>
<dc:terms:temporal.xml:lang="de">Oktober 1949-Juli 1952</dc:terms:temporal>
<dc:title>Vorschläge von Dr. Ing. Carl H. Schweizer aus Stuttgart, Regierungsbaudirektor
im Innenministerium, zu einer Flagge und zu einem Wappen für Württemberg-Baden
bzw. den Süweststaat, Einer von zahlreichen Wappenentwürfen Dr. Schweizers mit
dem maßgeblichen Symbol der drei schwarzen Löwen und einer Wappenkrone</dc:title>
<europeana:unstored></europeana:unstored>
<europeana:isShownAt>http://www.landearchiv-bw.de/plink/?f=1-100489</europeana:isShownAt>
<europeana:object>http://www.landearchiv-bw.de/plink/?f=1-100489-1&ext=1</europeana:object>
<europeana:isShownBy>http://www.landearchiv-bw.de/plink/?f=1-100489-1</europeana:isShownBy>
<europeana:country>germany</europeana:country>
<europeana:provider>Landesarchiv Baden-Württemberg</europeana:provider>
<europeana:type>TEXT</europeana:type>
<europeana:language>de</europeana:language>
</record>

```

Abb. 2: „Europeana“ – Datenlieferung im ESE-Format. Beispiel-Datensatz des Landesarchivs Baden-Württemberg (EAD im Hintergrund, ESE im Vordergrund)

In den kommenden Versionen der „Europeana“, voraussichtlich ab April 2011, wird das ESE-Format durch ein neues Datenmodell abgelöst werden, das derzeit von den Projektpartnern entwickelt wird und ab März 2010 intern getestet werden soll. Es handelt sich

hierbei um das sogenannte „Europeana Data Model“ (EDM),¹⁸ das sich des Metadatenmodells und der Terminologie des W3C-Standards RDF¹⁹ bedient und sich zunehmend in Richtung „Semantisches Web“²⁰ entwickelt.

2.5 Nächste Schritte und Perspektiven

Das wesentliche Ziel des „Europeana v1.0“-Projekts ist die Anreicherung des Informationssystems mit weiteren Inhalten und der Ausbau der Funktionalitäten. Um die bis Ende 2010 angestrebte Bereitstellung von zehn Millionen digitalen Objekten zu erreichen, soll verstärkt auf nationale Aggregatoren zurückgegriffen werden und die Beteiligung insgesamt durch mehr europäische Mitgliedsstaaten erhöht werden.²¹ In der „Europeana Content Strategy“ vom August 2009 wird außerdem der Akquisition von digitalisiertem Archivgut eine hohe Priorität verliehen, da die archivische Sparte bisher noch stark unterrepräsentiert ist. Neben der Erhöhung des „eContent“ ist auch eine Weiterentwicklung der bestehenden Funktionen und Dienste vorgesehen, insbesondere in den Bereichen Suchfunktionen (v.a. semantische Suche), Schnittstellen und Community Features im

¹⁸ <https://version1.europeana.eu/c/document_library/get_file?p_l_id=10910&folderId=13001&name=DLFE-5676.pdf> .

¹⁹ RDF = Resource Description Framework, vgl. <<http://www.w3.org/RDF/>> .

²⁰ Beim Semantischen Web handelt es sich um eine Erweiterung des World Wide Web mit dem Ziel, Inhalte des Internets miteinander in Beziehung zu setzen und Informationen von Computern interpretierbar zu machen. Siehe auch <<http://www.w3.org/standards/semanticweb/>> .

²¹ Eine Analyse des eContents hat gezeigt, dass derzeit 70% von nur vier Content Providern (die Aggregatoren sind) und fast 80% aus vier Ländern stammen.

Sinne des Web 2.0. Für Juli 2010 und April 2011 ist die Onlinestellung neuer Versionen der „Europeana“-Plattform geplant.

Für eine Beteiligung der deutschen Kultur- und Wissenschaftseinrichtungen und speziell der Archive an der „Europeana“ ergeben sich dabei verschiedene Perspektiven. Die „Europeana“ bietet die Möglichkeit dafür, dass spartenspezifische Information international und interdisziplinär nachgewiesen werden kann. Ferner trägt sie zu einer erhöhten Wahrnehmung der einzelnen Einrichtungen als Kultur- und Wissenschaftseinrichtungen sowie als Informationsdienstleister bei, so dass höhere Nutzerzahlen für Kulturgut in institutionellen Online-Angeboten und vor Ort (z.B. mehr Lesesaalnutzer, mehr Museumsbesuche) zu erwarten sind. Schließlich eröffnet die „Europeana“ ebenso wie die geplante „Deutsche Digitale Bibliothek“ die Perspektive für eine stärkere interdisziplinäre Vernetzung mit anderen Kultur- und Wissenschaftseinrichtungen.

3. Die „Deutsche Digitale Bibliothek“ als Beitrag zur „Europeana“

3.1 Ziele und Rahmenbedingungen

Um das in den EU-Ratsentschlüssen vom 20. November 2008 verfolgte Ziel einer breiten Partizipation aller europäischen Kultureinrichtungen mit digitalisiertem Kulturgut an der „Europeana“ zu erreichen, müssen in Deutschland zunächst auf nationaler Ebene die Grundlagen dafür geschaffen werden. Eine wesentliche Voraussetzung ist der Aufbau eines nationalen Aggregators in Form einer spartenübergreifenden „Deutschen Digitalen Bibliothek“ (DDB), mit der ein nationaler, zentraler und interdisziplinärer Zugang zu Kulturgut und wissenschaftlicher Information geschaffen wird. Bei dem Begriff

„Deutsche Digitale Bibliothek“ handelt es sich um einen Arbeitsbegriff, der in Analogie zur „Europeana“ noch durch einen passenden Namen ersetzt werden muss, der dem spartenübergreifenden Anspruch gerecht wird.

Die Errichtung der DDB hat einen längeren Vorlauf. Anders als in zentralistisch organisierten Staaten wie Frankreich gibt es in Deutschland aufgrund der föderalen Strukturen eine Aufgabenverteilung zwischen Bund und Ländern. Wegen der Kulturhoheit der Länder ist also ein längerer politischer Abstimmungsprozess erforderlich, um die DDB auf den Weg zu bringen: Zunächst wurde seit Ende 2006 mit der „ad hoc-AG ‚Digitalisierung von Kulturgut‘“ eine Arbeitsgruppe der Kultusministerkonferenz bestehend aus Vertretern der Länder, des Bundes und der Kommunen ins Leben gerufen, welche die organisatorischen, rechtlichen und finanziellen Rahmenbedingungen vorbereitet hat (v.a. Bund-Länder-Eckpunktepapier, Verwaltungs- und Finanzabkommen zwischen Bund und Ländern).²² Unterstützt wird sie seit Juli 2007 durch die „Bund-Länder-Fachgruppe DDB“, die die fachlichen Grundlagen erarbeitet und fest schreibt. In der zweiten Jahreshälfte 2009 haben sich die Länder und der Bund politisch auf die Errichtung der DDB geeinigt,²³ so dass das als Organisationsform geplante „Kompetenznetzwerk DDB“ er-

²² <http://www.deutsche-digitale-bibliothek.de/pdf/gemeinsame_eckpunkte_finale_fassung_02122009.pdf>; <http://www.deutsche-digitale-bibliothek.de/pdf/verwaltungs_und_finanzabkommen_finale%20Fassung02122009.pdf>.

²³ Beschluss der Ministerpräsidentenkonferenz vom 30.10.2009 und Beschluss des Bundeskabinetts vom 02.12.2009. Eine Unterzeichnung des Abkommens durch die jeweiligen Fachministerien erfolgt voraussichtlich in der ersten Hälfte 2010.

richtet und ab 2011 auch finanziert werden kann. Ab Ende 2011 soll eine erste Fassung der DDB-Infrastruktur fertig sein, so dass dann sukzessive 30.000 Kultur- und Wissenschaftseinrichtungen ihre Inhalte einbringen können.

Im Rahmen der Öffentlichkeitsarbeit zur DDB hat die Bund-Länder-Fachgruppe eine Webseite eingerichtet, die Informationen und Dokumente über die Planung, den Aufbau und den aktuellen Stand der Realisierung enthält.²⁴ In der Namensfrage wurde nach intensiver Diskussion in der „Bund-Länder-Fachgruppe DDB“ bisher kein Vorschlag für geeignet befunden, so dass im Rahmen der Aufbaufinanzierung voraussichtlich ein externer Dienstleister damit beauftragt werden wird.

3.2 Finanzierung und Organisation

Der Aufbau der zentralen Infrastruktur wird mit „eGovernment-IT“-Mitteln in Höhe von 5 Millionen EUR aus dem Konjunkturpaket II des Bundes finanziert,²⁵ wohingegen der Dauerbetrieb und die fortwährende Arbeit des sogenannten Kompetenznetzwerks Deutsche Digitale Bibliothek mit einem jährlichem Budget von 2,6 Millionen EUR je zur Hälfte von Bund und Ländern getragen wird. Die Finanzierung der eigentlichen Digitalisierung muss durch die jeweiligen Unterhaltsträger der entsprechenden Einrichtungen erfolgen. Darüber hinaus sind Drittmittelgeber eine wichtige Finanzierungsquelle. Wichtigster Förderer für die Digitalisierung von Kulturgut mit dem Ziel einer wissen-

²⁴ <<http://www.deutsche-digitale-bibliothek.de>> oder <<http://www.ddb.de>>.

²⁵ Siehe dazu <<http://www.bundesregierung.de/Content/DE/Artikel/2009/01/2009-01-13-konjunkturpaket-2-materialien.html>>.

schaftlichen Nutzung ist seit vielen Jahren die Deutsche Forschungsgemeinschaft (DFG), die bisher jährlich ca. 13 Millionen EUR an Fördermitteln für diesen Bereich zur Verfügung stellt.²⁶

Als Organisationsform für die DDB ist durch das Verwaltungs- und Finanzabkommen die Errichtung eines „Kompetenznetzwerks Deutsche Digitale Bibliothek“ (KNW-DDB) geplant. Als Mitglieder des KNW-DDB sind Träger und Einrichtungen vorgesehen, die herausragende Fachkompetenz im Bereich der Digitalisierung von Kulturgut und wissenschaftlicher Information besitzen.²⁷ Die Aufgaben werden arbeitsteilig wahrgenommen. Dazu zählen der Aufbau und die Organisation der DDB inklusive der Datenlieferung an die „Europeana“, die Fortentwicklung technischer Werkzeuge für Datenlieferung und -haltung, die Festlegung von Standards zur Sicherstellung der Kompatibilität, die umfassende Bereitstellung von Informationen zur Digitalisierung und Bestandserhaltung für alle Kultur- und Wissenschaftseinrichtungen, die Öffentlichkeitsarbeit und schließlich die Entwicklung von Geschäftsmodellen unter Wahrung bestehender Urheber- und Verwertungsrechte. Diese Aufgaben werden innerhalb des Kom-

²⁶ <<http://www.dfg.de>> und insbesondere <http://www.dfg.de/foerderung/programme/infrastruktur/lis/lis_foerderbereiche_programme_aktionslinien/index.html#micro113676>.

²⁷ Als Gründungsmitglieder sind vorgesehen: Bayerische Staatsbibliothek, Bibliotheksservice-Zentrum Baden-Württemberg, Brandenburgisches Landesamt für Denkmalpflege, Bundesarchiv, Deutsches Filminstitut, Deutsche Nationalbibliothek, Landesarchiv Baden-Württemberg, Max-Planck-Institut für Wissenschaftsgeschichte, Stiftung Museum Hamburgische Geschichte (digiCULT), Niedersächsische Staats- und Universitätsbibliothek Göttingen, Sächsische Landesbibliothek – Staats- und Universitätsbibliothek Dresden, Stiftung Preußischer Kulturbesitz Berlin und das Kulturamt der Stadt Düsseldorf.

petenznetzwerks auf verschiedene Organe verteilt. Neben den eigentlichen Mitgliedern gibt es ein Kuratorium als politisches und rechtliches Aufsichtsgremium. Geleitet wird das Kompetenznetzwerk durch einen Vorstand, der durch eine Geschäftsstelle unterstützt wird, die bei der von Bund und Ländern gemeinsam finanzierten Stiftung Preußischer Kulturbesitz in Berlin eingerichtet wird. Durch die Geschäftsstelle werden auch die Haushaltsmittel des KNW-DDB verwaltet. Es ist vorgesehen, dass die Wahrnehmung der fachlichen Aufgaben einerseits von einem durch das Kompetenznetzwerk beauftragten Betreiber, andererseits durch einzelne Mitglieder gegen Kostenerstattung erfolgt. Zu den Pflichten des Betreibers gehören insbesondere der komplette technische Betrieb, aber auch die Weiterentwicklung der technischen Plattform. Das Kompetenznetzwerk wird darüber hinaus mit anderen Netzwerken wie dem Kompetenznetzwerk für Bibliotheken (KNB), der Allianz für Bestandserhaltung, dem NESTOR-Kompetenznetzwerk und dem Theseus-Projekt kooperieren.²⁸

Konzeption und Aufbau der technischen Infrastruktur erfolgen bis Ende 2011 durch das vom Bund beauftragte Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme (Fraunhofer IAIS) in

²⁸ <<http://www.langzeitarchivierung.de>>; <<http://www.theseus-programm.de>>.

Das „Theseus-Projekt“, finanziert mit einem Gesamtvolumen von 180 Mio. EUR (je zur Hälfte durch das Bundesministerium für Wirtschaft und Technologie und durch Industriepartner), soll die Entwicklung einer neuen internetbasierten Wissensinfrastruktur fördern. Dabei werden u.a. anwendungsorientierte Basistechnologien und technische Standards für eine semantische Suche im Internet erarbeitet und erprobt.

St. Augustin bei Bonn in der Eigenschaft als Generalunternehmer.²⁹ Begleitet wird der Auftrag zunächst durch die „Bund-Länder-Fachgruppe DDB“ und nach Unterzeichnung des Verwaltungs- und Finanzabkommens durch das KNW-DDB.

3.3 Fachkonzeption, „eContent“ und Funktionalitäten

Die fachliche und technische Realisierung der Portal-Plattform erfolgt auf der Basis eines Einführungsprojektes mit Anforderungsanalyse, Gesamtkonzeption und Pflichtenheften. Die „Bund-Länder-Fachgruppe DDB“ erstellte im Auftrag der „ad hoc-AG ‚Digitalisierung von Kulturgut‘“ bereits ein erstes Grobkonzept zum Aufbau und Betrieb einer DDB.³⁰ Wesentliche Themenbereiche des groben Fachkonzepts sind die digitalen Inhalte der DDB („eContent“), die technische Infrastruktur (u.a. Portalarchitektur, Schnittstellen), Standards und Metadaten (u.a. Formate für die Datenlieferung), Benutzeranforderungen und Präsentationsmodelle sowie die Verbindung zur „Europeana“. Weitere Vorarbeiten sind eine technische Machbarkeitsstudie,³¹ die durch einen IT-Experten erstellt worden ist, und eine durch eine Consulting-Firma erstellte Nutzerstudie. Momentan wird durch das Fraunhofer IAIS in Verbindung mit der „Bund-Länder-Fachgruppe DDB“ eine umfassende Anforderungsanalyse erstellt, die Grundlage für technische Pflichtenhefte sein wird. Die Anforderungsanalyse

²⁹ <<http://www.iais.fraunhofer.de/ddb.html>>.

³⁰ <http://www.deutsche-digitale-bibliothek.de/doc/fachkonzept_160208_ohne_organisation.doc>.

³¹ <http://www.deutsche-digitale-bibliothek.de/pdf/machbarkeitsstudie_20080723.pdf>.

thematisiert u.a. die Bereiche Informationsarchitektur, Recherche-funktionalitäten, Dienstplattform, Mehrwertdienste, Datenlieferung und -konsolidierung sowie den technischen Betrieb.

Was die digitalen Inhalte – also den sogenannten eContent – der DDB angeht, wird beabsichtigt, zunächst die bereits heterogen und dezentral vorhandenen digitalen Inhalte der deutschen Kultur- und Wissenschaftseinrichtungen zu integrieren. Strittig ist momentan noch, ob in einer ersten Phase auch reine Erschließungsinformationen ohne Digitalisate aufgenommen werden, was aus Sicht der Archive wünschenswert ist. Auf jeden Fall sollen schrittweise alle Formen von „eContent“, d.h. Images, Volltexte, 3D-Objekte, Video-/Audio-Dateien etc. im Kontext der Erschließung eingebunden werden. Eine bisher fehlende nationale Digitalisierungs- und Bereitstellungsstrategie kann im Rahmen der Arbeit des KNW-DDB entwickelt werden. Eine Aufgabe des Kompetenznetzwerks wird auch die Koordination zur Vermeidung von Dubletten bei nicht unikalener Überlieferung sein.

Grundsätzlich kann die DDB auch ohne die explizite Nennung eines Finanzvolumens für „eContent“ errichtet werden, da ja bereits große Mengen von „eContent“ vorliegen und mit der Errichtung der DDB grundsätzlich ein Anstoß für die vermehrte Digitalisierung gegeben wird, die je nach Kultur-/Wissenschaftseinrichtung bzw. ihrer Unterhaltsträger unterschiedlich ausfallen wird, ohne dass eine konkrete Zahl genannt werden muss.

Die technische Architektur der DDB wird nach einem ähnlichen Konzept wie bei „Europeana“ aufgebaut werden. Der eigentliche „eContent“ soll dabei bei den Kultur- und Wissenschaftseinrichtun-

gen verbleiben, wohingegen Erschließungsinformationen und sogenannte Derivate (z.B. Vorschaubilder von digitalen Abbildungen) zentral vorgehalten werden. Grundlage wird eine Mehrschichtenarchitektur sein, in deren Mittelpunkt sich ein „Metadastore“ in Form eines sogenannten Datawarehouse befinden wird. Er wird unterschiedliche Daten-Repositories für die verschiedenen Sparten berücksichtigen müssen. Auf der unteren Ebene werden ihm ein Aggregator und Harvester vorgelagert sein, der spartenspezifische Zielformate (z.B. EAD bzw. METS/EAD,³² museumdat, METS/MODS, MARC-XML) bedient. Auf der oberen Ebene ist ein freier Servicelayer sinnvoll, der verschiedene Dienste bedienen kann – wie das DDB-Portal mit übergreifender Suchmaschine oder die Schnittstelle für die Datenlieferung an die „Europeana“.

Als Recherchemöglichkeit ist zunächst in Analogie zu den großen Internet-Suchmaschinen ein Freitextsuchfeld („Google-Slot“) für einfache Suchanfragen geplant, die gegebenenfalls später in der Ergebnisübersicht verfeinert werden können. Gleichzeitig soll dieses einfache Freitextsuchfeld die Eingabe der von populären Suchmaschinen bekannten syntaktischen Suchoperatoren (z.B. Trunkierung, Verknüpfung von Suchbegriffen) unterstützen, so dass auch kom-

³² Für den archivischen Bereich ist in der technischen Machbarkeitsstudie als Datenlieferungsformat für Erschließungs- und Strukturinformationen ein definiertes EAD-XML-Format vorgesehen. Ein solches Format wird bereits im DFG-Projekt „Ausbau des Netzwerks SED-/FDGB-Archivgut“ (<<http://www.archivgut-online.de>>), im Rahmen des DFG-Projekts „Retrokonversion archivischer Findmittel“ (<<http://www.archivschule.de/retrokonversion>>) und für die Datenlieferung an das BAM-Portal (<<http://www.bam-portal.de>>) verwendet. Zum DFG-Projekt Retrokonversion vgl. auch den Beitrag von Frank M. Bischoff in diesem Band.

plexere Suchanfragen formuliert werden können. Für konkrete bzw. zielgerichtete Suchanfragen stellt das Portal ergänzend zur Standard-suche eine erweiterte Suche zur Verfügung. Diese enthält zusätzliche Filterelemente und Freitextsuchfelder, mit denen eine Suchanfrage nach definierten Metadatenfeldern (u.a. Person, Ort, Sparte, Institution, Zeitraum, Schlagworte und Objekt- bzw. Medientyp) der zu durchsuchenden Objekte gefiltert werden kann. Um diese Metadatenfelder für alle beteiligten Sparten anbieten zu können, ist die nachträgliche Bearbeitung der von den einzelnen Einrichtungen gelieferten Erschließungsinformationen mittels Werkzeugen notwendig, die eine semantische Analyse und Anreicherung ermöglichen. Neben der Volltextsuche ist auch navigierende Recherche („Browsing“) z.B. über einen Strukturbaum und gegebenenfalls Klassifikationen geplant.

3.4 Zeitplan und Perspektiven

Der aktuelle Zeitplan (Stand: Februar 2010) für die Realisierung des Zugangsportals und der technischen Infrastruktur sieht vor, dass noch in der ersten Jahreshälfte 2010 die Anforderungsanalyse, technische Konzeption und Erstellung von Pflichtenheften abgeschlossen werden und mit der Entwicklung der Portalsoftware und der technischen Infrastruktur begonnen wird. Aufgrund der restriktiven Auflagen für die Verwendung von Mitteln aus dem „Konjunkturpaket II“ der Bundesregierung ist geplant, dass bereits im Sommer 2011 ein erster Probetrieb starten und Ende 2011 der Launch einer ersten Version des zentralen Zugangsportals sowie die Inbetriebnahme der zentralen Infrastruktur der DDB erfolgen kann.

Die DDB wird dem Internet-Nutzer mehrere Vorzüge gegenüber dem Projekt „Google Books“ bieten: Dort sind z.Zt. in erster Linie englischsprachige Bücher digitalisiert. Durch ihren spartenübergreifenden Ansatz wird die DDB nicht nur digitalisierte Bibliotheksbestände in deutscher Sprache anbieten, sondern z.B. auch digitalisiertes Archivgut, digitalisierte Museumsobjekte, digitalisierte Fotobestände oder audiovisuelles Material. Anders als bei „Google Books“ ist ein umfassendes Rechtemanagement vorgesehen, das z.B. die Gewährleistung der Urheber- oder Besitzrechte für die jeweiligen Rechteinhaber sicherstellt.

Für die deutschen Kultur- und Wissenschaftseinrichtungen ergeben sich mit der Errichtung der DDB verschiedene Perspektiven. Mit der DDB wird ein überregionaler, nationaler Zugang zu digitalisiertem Kulturgut bzw. wissenschaftlicher Information geschaffen und dies sowohl in einem interdisziplinären als auch spartenspezifischen Kontext. Verbunden ist damit auch – soweit von den teilnehmenden Einrichtungen gewünscht – die automatische Partizipation an der „Europeana“ über den nationalen Aggregator DDB. Darüber hinaus bietet die DDB auch eine Plattform für die Realisierung, den dauerhaften Betrieb und die Weiterentwicklung spartenspezifischer Portale wie z.B. ein deutsches Archivportal in Form einer spartenspezifischen Sicht innerhalb der DDB-Architektur. Schließlich kann durch die mit der DDB verbundene Organisation des KNW-DDB eine dauerhafte Unterstützung der Kultur- und Wissenschaftseinrichtungen in Fragen der Digitalisierung und Online-Präsentation (u.a. Workflow, Standards, Bereitstellung von Werkzeugen, ggf. Hosting) und der Datenlieferung erfolgen.

4. Zusammenfassung und Ausblick

Zusammengefasst können sowohl die „Europeana“ als auch die „Deutsche Digitale Bibliothek“ neue Perspektiven für die Nutzung digitalisierter und digitaler Bestände oder Sammlungen in europäischen bzw. deutschen Bibliotheken, Archiven, Museen und anderen Sammlungsbereichen (u.a. Filmarchive, Denkmalpflege) sowie zu deren Erschließungsinformationen schaffen. Die DDB kann gleichzeitig als vorstrukturierter Zulieferer zur „Europeana“ dienen und so eine umfassende Teilnahme deutscher Kultureinrichtungen an der „Europeana“ ermöglichen. Allerdings ist die Entwicklung und Verwaltung beider Informationssysteme ein sehr komplexer Prozess. Es treffen unterschiedliche fachliche Traditionen und Anforderungen, divergente Zielsetzungen und Benutzeranforderungen sowie verschiedene Entwicklungsstrategien und eine verteilte Datenhaltung aufeinander. In Deutschland kann für den Aufbau einer DDB auf ein breites Spektrum an Erfahrungen und Projektergebnissen im Bereich der Digitalisierung von Kulturgut zurückgegriffen werden. Eine Reihe von Kultur- und Wissenschaftseinrichtungen hat, oft mit finanzieller Unterstützung durch die DFG, bereits umfangreiche praktische Erfahrungen gewonnen.³³

Neben den bereits aufgezeigten Perspektiven gibt es auch noch offene Fragen, die einer Klärung bedürfen. Dazu gehört das Verhältnis der institutionellen Online-Informationssysteme zur DDB und zur „Europeana“. Insbesondere im archivischen Bereich stellt sich die

³³ Vgl. dazu das Verzeichnis der Bund-Länder-Arbeitsgruppe EUBAM unter <<http://www.eubam.de>> und <<http://www.kulturerbe-digital.de>>.

Frage nach der Abgrenzung der Präsentationsebene für archivische Informationen (Tektonik/Online-Beständeübersichten, Online-Findbücher, Digitalisate). Unklar ist auch noch, wie der kleinste gemeinsame Nenner für das spartenübergreifende Zielformat aussehen wird und in welcher Form semantische Verknüpfungen möglich sind, wenn keine Erschließung mit „Normdaten“ erfolgt. Geklärt werden muss auch, auf welcher Ebene „Mehrwertdienste“ (z.B. Online-Bestellsystem für Archivalienbestellung oder digitale Reproduktionenlieferung) angeboten werden können. Schließlich stellt sich die Frage, wie Hostingangebote für Einrichtungen geschaffen werden können, die selbst kein eigenes Online-Informationssystem besitzen. Dies gilt insbesondere auch für kleinere Archive im kommunalen Bereich.

Trotz der noch ungeklärten Fragen ist für die deutschen Archive eine Partizipation an der DDB und damit auch an der „Europeana“ aufgrund des bereits vorliegenden, die deutschen Verhältnisse berücksichtigenden EAD-XML-Austauschformats für digitale Findmittel sinnvoll und zeitnah realisierbar. Dabei sollte von archivischer Seite der Aufbauprozess der „Europeana“ und der DDB aktiv mitgestaltet werden. Durch die intensive Förderung der Retrokonversion analoger Findmittel im Rahmen der DFG-Förderung können die Archive ihren Rückstand gegenüber den Bibliotheken aufholen, so dass in absehbarer Zeit eine kritische Masse an archivischen Strukturinformationen und Titelaufnahmen im Netz verfügbar sein wird und im Rahmen einer interdisziplinären Recherche mit Informationen anderer Kulturinstitutionen wie Bibliotheken und Museen verknüpft werden kann. Die in deutschen Archiven vorhandenen Filmkopien der Sicherungs- und Schutzverfilmung bieten darüber hinaus eine reichhaltige und gute Ausgangslage für eine massenhafte rationelle Digitalisierung

von Mikrofilmen, die dann im Kontext digitaler Erschließungsinformationen auch in der DDB und „Europeana“ bereitgestellt werden können.

The Digitization of Archives in France: Projects and Perspectives

by Claire Sibille-de Grimoüard

Digitization in French archives began in the 1990's: the first projects took place at the National Archives in 1993 and in 1994 in the territorial archives (the pioneers were the Departmental Archives of Mayenne in West of France, and the Departmental Archives of Aveyron in South-West of France).

Initially, digitization was considered as an alternative format to the original material and as the best way to answer the frequent requests of users about the most popular documents: parish and civil status registers, of course, but also the cadastre, local newspapers and graphical records such as postcards, etc. At that time, users could only consult the digital images on local computer networks in reading rooms.

However, issues of digitization have expanded considerably because of the improvement of computer performance and the Internet development. The increase of digitization projects has been rising steadily, while digitized collections are increasingly accessible on the Internet.

1. Overview of digitization projects and nature of the concerned documents

To present an overview of the digitization of French-written heritage, the Ministry of Culture and Communication created in 2001 a na-

tional collective catalogue of digitized cultural content, the "Patrimoine numérique" (Digital Heritage Catalogue), which is accessible online.¹ This catalogue is a navigating tool of digitized funds. The aims are to enhance the digitized collections of France and to facilitate the access to digitized resources. The collections listed are the result of initiatives of the government, local authorities, associations and foundations.

The Digital Heritage Catalogue currently lists some 567 French cultural institutions involved in over 1,500 digitization operations. The statistics from this catalogue give a good overview of the dynamism of the archival network.

In late 2008, there were 678 references published by the archival institutions since the creation of the catalogue (including 408 references from the departmental archives), on a total of 1,469 references, i.e. 46%. The libraries had recorded 382 references (26%), museums, 133 (9%) and cultural heritage services, 153 (10%). Besides the fact that the archives were the first, they have also created the largest number of new records in 2008: 102 new references, as opposed to 38 for libraries, 18 for museums, and 1 for cultural heritage services.

In late 2009, 1,578 collections are listed, including:

- 728 references from archival institutions (448 departmental archives);
- 406 references from libraries;

¹ See: <<http://www.numerique.culture.fr>> (all the mentioned websites have been accessed on 27/11/2009).

- 153 references from cultural heritage services;
- 149 references from museums;
- 88 references from audiovisual institutions.

A record represents a project, but it should be noted that the importance of a project can be very variable. For example, the digitization of a small photo collection does not obviously generate the same amount of data as digitization of the civil status registers of a department does.

The digitization projects carried out by archives have focused on:

- various graphic documents (photo collections or albums, postcards, prints, posters, etc.); this category includes certainly the most important part of the projects;
- cadastral plans;
- civil status (parish and civil status registers, decennial tables);
- other types of records or documents frequently consulted (population census, notarial archives, registers of military recruitment, municipal registers of deliberations);
- local newspapers.

This distribution should not obscure the fact that digitization of a series of records generally represents many documents and a very high budget in comparison to other types of projects: two to three million pages for parish and civil status registers for a department, for amounts that can approach or exceed one million Euros.

The Departmental Archives of Var are very representative of the digitization projects carried out by territorial archives. The website of this archival institution provides a specific access to digitized docu-

ments (cadastral maps and records, decennial tables, civil status, population census, medieval notarial records). The user can also consult finding aids and visit virtual exhibitions.²

Digitization projects are very different. Besides these usual types of documents, other types of material have emerged through the last few years, for example acts of the administration of Avignon from 1790 until 1940 (municipal archives of Avignon)³ or inventories of churches established when the law of separation of church and state was applied or notebooks of teachers during the First World War (departmental archives of Charente). Several projects also concerned sound recordings such as oral testimonials about the beginnings of the Social Welfare Administration, employees of the National Institute for Agronomic Research,⁴ proceedings of territorial assemblies, local radio emissions, etc. As an extension of the Program "Saving the oral heritage of the Sarthe department", the digitization of oral heritage of this department includes audio records such as poetry, stories and testimonials about some aspects of history of the department of Sarthe (Resistance, May 1968) or the economic life of the department (testimonials on former activities: pottery, wheelwright, car industry, etc.).⁵ Among the digitized documents, we should point out the audio dictionary of the Sarthe language. Written and recorded in 1987 by its author, Gaston Cheverreau (1909-2003),

² See: <<http://www.archives.var.fr>>.

³ See: <<http://recherche.archives.avignon.fr>>.

⁴ See: <<https://www2.nancy.inra.fr/collectif/archorales>>.

⁵ See: <<http://www.fonds-sonores-archives.sarthe.com>>.

“My dialect of Hommedaire” is a unique audio dictionary of the language of Saint-Mars d’Outillé. Digitized in 2008, it contains more than a thousand words and is a unique oral testimony directly produced by a native speaker, giving almost an entire dialect dictionary.

Also note that archival finding aids are increasingly digitized, especially if they are old and rare and have some importance from a historiographical point of view. The aim is to generate an electronic structured document from a printed or typed document, which will therefore be searchable by computer, usually in XML-format according to the EAD-DTD. Thus, the National Archives of Overseas and several departmental archives have published on the Web an XML/EAD-holdings guide which can provide access to more detailed finding aids. They may also be structured in XML/EAD. Currently, the National Archives are developing their archival information system for the new center of Pierrefitte-sur-Seine and are initiating an intense campaign of digitization/retrospective conversion into EAD of their archival finding aids.

The number of pages and images that have been digitized since January 1, 2009 has been estimated: from the Cultural Heritage Catalogue and from annual reports sent to the Archives of France.

Total number of pages digitized or being digitized	133,443,421
thereof civil status registers	120,385,656
Total number of images digitized or being digitized	1,340,900

Tab 1: Number of digitized pages and images

2. Online-access to digitized materials

Online publishing of digitized record groups is the major evolution since early 2000. This is also encouraged by the Ministry of Culture through the National Annual Digitization Plan, which obliges the recipients to upload the digitized resources in the year of their creation.

Like operations of digitization, the archives upload records related to civil status, which are often requested by genealogists, and required frequently for cadastre as well as graphical documents. Concerning graphical documents, digital platforms have been developed: the Ulysse base of the National Archives of Overseas, "History through images" in collaboration with the Réunion des Musées nationaux and the French museums or Archim, developed by the National Archives.

At the beginning of 2009, 64 archival institutions uploaded archival material: 47 departmental archives, 14 municipal archives and three centres of the National Archives (Paris, Fontainebleau and Archives of Overseas at Aix-en-Provence).

Regarding civil status registers, Web publication has grown increasingly since 2006 (see Tab. 2):

2003: Mayenne, Savoie, Vendée	3
2004: Yvelines	1
2005: Alpes-Maritimes, Côtes d'Armor, Sarthe, Tarn, Vienne	5
2006: Ardèche, Aube, Bouches-du-Rhône, Côte-d'Or, Eure, Eure-et-Loir, Loire-Atlantique, Lozère, Haute-Marne, Val-de-Marne	10
2007: Ain, Hautes-Alpes, Corrèze, Gironde, Oise, Orne, Morbihan, Var	8
2008: Hérault, Drôme, Cantal, Charente-Maritime, Maine-et-Loire, Meurthe-et-Moselle, Nord, Saône-et-Loire, Seine-et-Marne, Deux-Sèvres, Alpes-de-Haute-Provence	11
Total	38

Tab. 2: Departmental archives that have uploaded their civil status records (registers and tables)

Total number of digitized pages online or partially online	90,537,988
National Archives	1,873,247
Departmental Archives	87,856,602
Municipal Archives	808,139
Total number of digitized images online	438,572
National Archives	11,987
Departmental Archives	406,784
Municipal Archives	19,801

Tab. 3: Number of resources available online at the beginning of 2009

Number of digitized and uploaded record groups	164
Parish and civil status registers, decennial tables	56
Census lists	6
Military rolls	4
Notarial archives	3
Cadastre	27
Photo collections and postcards	18
Maps and plans	13
Posters collections	6

Tab. 4: Different types of archival material available online

The most frequent ways of access are the following:

- creation of specialized databases, providing access to a specific type of material (image banks, cadastre, marital status, etc.). This is available through a pre-established list (for example: list of cities for the cadastre), a search form or a map of the territory.
- or integration in encoded finding aids of the digital images corresponding to the document(s) described. The users are provided with the digitized documents, their description and archival higher levels, which enable them to situate archival information in its production context.

An example of a specialized database is ARCHIM⁶, which offers digital images of valuable historical documents, pictures and photographs and

⁶ See: <<http://www.culture.gouv.fr/documentation/archim/accueil.html>>.

thereby preserves the extensive holdings of France's National Archives' historical department. Historians of French history from the medieval period onward will find the site invaluable. Highlights include: medieval manuscripts and especially illuminated manuscripts (unfortunately there is no zoom function), casts of medieval seals, eye-catching atlases from the Ancien Régime, French Revolutionary documents, virtual albums of 19th century art purchased from Parisian salons (organized chronologically), images from Second World War internment camps, and other World War II documents dealing with French Jews and occupation. Not all documents are online, but there are descriptions for some 3,740 items. Of those, users can browse over 3,700 images through a search engine. There is also an online archival catalogue.

Another example is the Ulysses database allowing access to digitalized images of documents belonging to the repository's icon and posters archive.⁷ Started in 2002, this database aims to progressively make available spare photographs, or albums, postcards, posters, drawings and engravings, maps and layouts.

The National Archives of Overseas at Aix-en-Provence (ANOM, in its French acronym) are in charge of the preservation of archives of the French colonial expansion. This set of an exceptional richness for the history of the first and second French colonial empires comes both from public archives (Secretariats of State and Ministries that have ruled the colonies from the 17th century until half of the 20th century, general governments, etc.) and private ones, received as donation, legacy, gift or purchased.

⁷ See: <<http://caom.archivesnationales.culture.gouv.fr/sdx/ulyse/index>>.

Both a simplified descriptive notice and a detailed one give all the necessary information in order to understand the documents. The documents' titles or caption are originals either from the authors or the colonial administration. In order to round out one's research, the user can go further and have a look at the funds and detailed inventory, being progressively put online, which enables a scientific approach of each document held in the Ulysses database.

The database will grow richer at the pace of the digitalization campaigns and data entering. In 2005, images of some photographs (Indochina, Sub-Saharan Africa, Madagascar) coming from the collection of the Economic Agency of Overseas France, maps and layouts of North America (Colonies' Fortifications Deposit), and the whole collection of posters, are open to consultation.

Working in partnership with the Ministry of Education, the Digital platform "L'Histoire par l'image" aims at promoting an innovative and complex approach of history from 1789 to 1939 through interpretation of visual archives. It required to set up a selection committee with historians, museums mediators, educators, etc., and an editorial team working continuously on the content of the website.⁸

Another example is provided by the Departmental Archives of Val-de-Marne.⁹ EAD finding aids are displayed according to the classification schema of the archives, based on the status of records creators and on their activities. Several finding aids include links to digital images of documents.

⁸ See: <<http://www.histoire-image.org>>.

⁹ See: <<http://www.archives94.fr>>.

3. Recent trends: sharing and new uses

Digitization is more than an operation of reproduction and communication and it has become a priority of the archives, framed by technical standards and increasingly regulated in terms of access and reuse of digital cultural resources.

It also results from real strategic goals. With the national plan for digitalization, the Ministry of Culture, that financially supports projects of cultural services, encourages cultural themes but also rules: required web publication of digital resources, free access to documents, use of tools based on Web 2.0, dissemination on platforms for sharing content for general audience, use of open and shared technical standards for harmonizing digital resources and federate them around shared thematic portals (data interoperability). The Ministry encourages networking of resources of its services and agencies as well as the complementarity of the digitization programs.

Furthermore, the Ministry of Culture and Communication has developed many guides about digitization of cultural heritage. Topics range from the advantages and disadvantages of digital cameras to recommendations for different types of storage. More detailed documents about ensuring the quality of digital cultural heritage are also available. The Ministry also links specifically to the eEurope Minerva guides and has translated them into French.

a) Thematic portals

The issue is to use the flexibility that is made possible by digitization and online publication. Indeed, geographical distance is no longer a constraint today. It is possible to integrate digital resources from various institutions and distant in space around a theme or a shared project.

Archival portals have been developed, including the portal New France - New horizons, where digital resources are distributed from Canadian and French archives.¹⁰ Another example is the Genealogy Gateway led by the Archives of France and the Ministry of Culture. It integrates several genealogical databases.¹¹ The portal now offers a single access point to various digitized genealogical resources, and it is ultimately intended to be as comprehensive as possible. Parish registers and civil status records are of course commonly available, but new and valuable additional resources should be digitized and made available: naturalization documents, records of Legion of Honor, military rolls or population censuses will be valuable resources to supplement the knowledge of a family, of a city, of a region, of a society at a special period, etc.

The Archives of France also participate in projects of international portals, including Europeana, the European digital library. Currently, the Archives of France are actively participating in the development of the European archival APEnet portal which will enable users to access digital and archival resources from the archival institutions of the partner countries. Therefore, the interoperability of digital resources has become a primary concern in these projects. The harvesting of records using the OAI-PMH for the collections web site was from this point of view an important advancement for the archival network.

¹⁰ See: <<http://www.archivescanadafrance.org/francais/accueil.html>>.

¹¹ See: <<http://www.culture.fr>>.

b) Educational use

Archival resources also attract young audiences, for example:

- “Suburbs Buddies”: animation flash game developed by the Municipal Archives of Saint-Denis. The principle is to tell the story of the suburbs from 1905 to today through archival documents.¹²
- “The Mystery of Beautiful Cordeliere”: directed by the departmental archives of Aube, the principle of this game is to solve enigmas by exploring records. It is a particularly successful adaptation.¹³

The Ministry of Culture and the Archives of France wish to encourage the development of such initiatives. The services begin to provide users with training online modules, including paleography (Departmental Archives of Indre-et-Loire): players can transcribe digitized documents.¹⁴ The next step is to capitalize what users can provide for enrichment of online resources.

c) Users and Web 2.0

To facilitate access of users to cultural heritage, projects inspired by the philosophy of the Web 2.0 have been developed, the user being considered both as a consumer and as a creator of resources: the archives begin to capitalize on this philosophy and created a few software appli-

¹² See: <<http://www.copainsdebanlieue.com/gene.html>> .

¹³ See: <<http://www.archives-aube.com>> .

¹⁴ See: <http://archives.cg37.fr/Apprendre/INITIATION_A_LA_PALEOGRAPHIE-ACC.html> .

cations enabling readers and volunteers to enrich the digital content (indexing or collaborative annotation). This practice exists in the departmental archives of the Mayenne,¹⁵ Yvelines,¹⁶ and recently Ain¹⁷. The Municipal Archives of Rennes have also adopted this method.¹⁸

The principle is the following: a means of access to the digital image of a document is to retrieve it by recording in the search interface keywords such as personal or geographical names. For example, images of civil status records will be retrieved by the user once the user has recorded criteria such as date range, name of the town of birth, type of act (marriage, baptism, burial), etc.

The images of civil status records can then be very finely indexed, for instance, the name of the person mentioned in the act can be recorded: this work was done by the National Archives of Overseas for the civil status of Algeria. But it is very painstaking work that agents do not always have time to do. Therefore, we can offer volunteer users to participate in this work. They are assigned a set of images, and for each image they will record index terms once they have decoded the document.

However, such a solution requires the services to develop a software interface specifically adapted and rigorous control procedures about the quality of the indexation made by the users. Some digitization companies or software publishers provide such a module of shared

¹⁵ See: <<http://www.lamayenne.fr>>.

¹⁶ See: <<http://www.yvelines.fr/archives/home.html>>.

¹⁷ See: <<http://www.ain.fr>>.

¹⁸ See: <<http://www.archives.rennes.fr>>.

indexation. This can eventually enrich the databases related to digitization.

The departmental archives of Yvelines have just launched an application of this kind for the digitized newspapers of the former department of Seine-et-Oise, and it is now available on the portal "The Kyosque".¹⁹ Users can volunteer to index newspapers (events mentioned in headlines for example).

Other "innovative cultural practices" are being developed (insert RSS feeds, sharing content on digital platforms such as Daily Motion, Flickr, etc.). But this assumes that the archival institution is attentive to its website and especially that it has the means to make it live. Some softwares provide archival features to easily link the data entered in the production modules and the website that will broadcast the same data. But for the moment the website management is mostly entrusted to partners, for instance the territorial organization on which the archival institution or a company depends (including digitization companies).

In conclusion, public interest and demand for digitization of archival materials are very important. This has led to the introduction of archival standards and the development of new services and software tools (direct encoding of the description, annotation modules, management of website, digital images viewer etc.).

¹⁹ See: <<http://img-dad.cg78.fr/path/pac/LeKyosque/internet.html>>.

Digitalisierung und Onlinestellung von Archivgut im Bundesarchiv: Ziele, Verfahren und Werkzeuge

von Angelika Menne-Haritz

1. Warum Digitalisierung?

Digitale Formen spielen in den Archiven seit einigen Jahren eine zunehmend größere Rolle. Dabei sind verschiedene Aspekte zu unterscheiden, die Auswirkungen auf die Methoden der Bearbeitung und der Erhaltung haben. Zum einen geht es um die Archivierung genuin digitaler Aufzeichnungen aus Verwaltungsprozessen. Diese digitalen Aufzeichnungen sind unmittelbar von der digitalen Technik abhängig, bei deren Nutzung sie entstanden sind.¹ Ihre Erhaltung und Zugänglichkeit über längere Zeiträume hinweg ist ständig bedroht durch die kontinuierliche Weiterentwicklung der Software und Dateiformate sowie durch Systemwechsel und Speicherkosten. Es ist nicht unwahrscheinlich, dass heute gängige und als sicher eingeschätzte Formate bald schon durch noch bessere Formate abgelöst werden. Ebenso werden zukünftige Systeme leistungsfähiger sein. Zusätzlich wird die zu beobachtende Reduktion der absoluten Speicherkosten durch wachsende Dateigrößen und Datenumfänge kompensiert. Viele Probleme sind bisher ungelöst und die Erfahrungen

¹ Zur archivischen Behandlung des Themas mit vielen weiterführenden Literaturverweisen vgl. Thekla Kluttig, 10 Jahre Arbeitskreis zur Archivierung von Unterlagen aus digitalen Systemen. Bilanz und Ausblick, in: Der Archivar 60, 2007, S. 51-53.

der kommenden Jahre werden noch einiges zur Strategieentwicklung beitragen können und müssen.

Von den Anforderungen der Arbeit mit den genuin digitalen Aufzeichnungen unterscheidet sich der Umgang mit Digitalisaten, weil sie digitale Sekundärformen analoger Vorlagen sind. Dabei können die digitalen Sekundärformen sowohl als bildliche Reproduktionen wie als digitalisierte Texte bereitstehen. Das Bundesarchiv hat seit langem die Edition der Kabinettsprotokolle der Bundesregierung von 1949 bis 1964, deren erste Bände zunächst ausschließlich gedruckt publiziert wurden, mit dem vollständigen wissenschaftlichen Apparat zusätzlich als Textdigitalisierung ins Internet gestellt² und in gleicher Form die Edition der Akten der Reichkanzlei aus der Weimarer Republik³ zugänglich gemacht. Solche Textdigitalisierungen sind als Nebenprodukte wissenschaftlicher, gedruckter Editionen enorm nützlich, weil sie umfangreiche Möglichkeiten der Auswertung geben. Sie sind durchsuchbar und bieten alle erforderlichen Informationen zur Einordnung der Aussagen bis hin zu den biographischen Angaben der erwähnten Personen.

Inzwischen denkt man beim Begriff Digitalisate eher an Schriftgutbestände, die bildlich digitalisiert wurden, um für eine Einsichtnahme im Internet angeboten zu werden. Popularisiert wurde diese Form durch die großen Projekte zur Digitalisierung von Büchern

² Direkt erreichbar über <http://www.bundesarchiv.de> oder <http://www.bundesarchiv.de/cocoon/barch/0000/index.html>. Alle im Beitrag genannten Webseiten wurden am 25.2.2010 aufgerufen.

³ <http://www.bundesarchiv.de/aktenreichskanzlei/1919-1933/0000/index.html>.

durch Google für seine Buchsuche, bei der bildliche Reproduktionen angezeigt und die konvertierten Texte im Hintergrund für die Suche verwendet werden. Die seit einiger Zeit erhältlichen Lesegeräte für digitalisierte Bücher nutzen dagegen die Textdigitalisierung für eine Anzeige, z.T. in einem offenen Format,⁴ das die Anpassung der Schriftgröße und des Seitenformats an unterschiedliche Geräte für das mobile Lesen erlaubt. Inzwischen bietet Google ebenfalls die Texte zahlreicher Bücher in diesem Format an.

Digitalisierung im Archiv steht weiterhin vor allem für die bildliche Reproduktion. Um diese Form soll es im Folgenden gehen. Sie macht nicht nur den Text, sondern auch das Layout sowie Bearbeitungsspuren sichtbar und stellt damit einen großen Teil der unter dem Begriff des intrinsischen Werts⁵ zusammengefassten äußerlichen Merkmale der Schriftstücke für die Interpretation bereit. Die Eigenschaft von bildlichen Digitalisaten als Kopien von analogen Vorlagen unterscheidet sie grundsätzlich von den genuin digitalen Unterlagen. So sind einerseits die Entstehungsformate der digitalen Abbildungen nicht von einer Software abhängig und können bei der Herstellung selbst bestimmt werden. Ihre Wirkung entfalten bildliche Digitalisate vor allem als komfortabel nutzbare Kopien. Sie sind jedoch denkbar schlecht als Erhaltungsform geeignet. Nur in integrier-

⁴ Genutzt wird das ePub-Format, eine XML-Kodierung, die vom International Digital Publishing Forum dokumentiert und gepflegt wird (<<http://www.openbook.org>>).

⁵ Vgl. Angelika Menne-Haritz/Nils Brübach, Der intrinsische Wert von Archiv- und Bibliotheksgut. Kriterienkatalog zur bildlichen und textlichen Konversion bei der Bestandserhaltung. Ergebnisse eines DFG-Projektes, Marburg 1997 (Veröffentlichungen der Archivschule Marburg 26).

ten Strategien zur Bestandserhaltung, die analoge und digitale Formen kombinieren, können sie als Medium zum Schutz anderer Formen einen Platz einnehmen.⁶

Der Einsatz von Digitalisaten im Archiv ist nicht unumstritten, widerspricht er doch auf den ersten Blick durchaus den Wünschen von Nutzergruppen, besonders aus der wissenschaftlichen Forschung. Sie sind daran interessiert, das zu entdecken, was noch niemand weiß, um es selbst als neue Erkenntnis in den Forschungsprozess einbringen zu können. Doch auch in häufig genutzten Unterlagen konnte es bisher bereits passieren, dass eine neue Fragestellung noch unbekannte Zusammenhänge erkennen ließ. Da zudem kaum abzusehen ist, dass zu einem bestimmten Zeitpunkt das Archivgut komplett digitalisiert sein könnte, und weil in überschaubarer Zukunft im Verhältnis zum Gesamtbestand nur sehr ausschnittshafte Mengen digitaler Reproduktionen im Internet angeboten werden können, bleibt immer noch viel Neues vor Ort zu entdecken.⁷

Das Bundesarchiv bietet zum Zeitpunkt der Erstellung dieses Textes rund 600.000 Digitalisate von 14 Beständen im Internet an und steht vor der Herausforderung, in den kommenden Jahren deut-

⁶ Vgl. Anna Haberditzl/Martin Luchterhand, Positionspapier der ARK „Digitalisierung von Archivgut im Kontext der Bestandserhaltung“, in: *Der Archivar* 61, 2008, S. 395-399 (mit Abdruck des Papiers), dazu auch Wolfgang-Valentin Ikas, Vom "prodesse et delectare" der Digitalisierung – Sieben gute Gründe für die Überführung analoger Altbestandsmaterialien in Nullen und Einsen, in: *Zeitschrift für Bibliothekswesen und Bibliographie* 57, 2010, S. 5-14.

⁷ Zu den Dimensionen vgl. Hartmut Weber, Attraktiv weil digital – Archivgut auf dem Weg ins Internet, in: *Mitteilungen aus dem Bundesarchiv* 13, 2005, Heft 2, S. 4.

lich höhere Mengen ins Internet zu stellen, die entweder bereits vorliegen oder zu erwarten sind. Wie es dazu kam und welche Strategien das Bundesarchiv einsetzt, soll im Folgenden mit einigen Schlaglichtern dargestellt werden.

1.1 Die Ziele der Digitalisierungsaktivitäten

Digitalisate sind eine neue Reproduktionsform, mit der zunächst die Zugänge zu den digitalisierten Akten selbst verbessert und neue Komfortfunktionen für die Nutzung angeboten werden können, die bei analogen Reproduktionen auf Mikrofilm nicht vorhanden sind. Als Nutzungsalternative zum Mikrofilm machen sie den Benutzern den Verzicht auf die Vorlagen bei der Auswertung attraktiver und erlauben dadurch einen besseren Schutz der Unterlagen durch geringere Frequenz der Aushebung.

Digitalisate im Internet können auch ungeübten Benutzern beispielhaft zeigen, was sonst oft beim ersten Besuch im Archiv erklärt werden muss, nämlich, dass es sich beim Archivgut um Schreiben aus Kommunikationsprozessen handelt, die zwischen den beteiligten Personen oder Stellen ausgetauscht wurden und dass diese Zusammenhänge bei einer Auswertung einbezogen werden sollten, um die Ergebnisse plausibel zu machen. Es ist immer wieder zu beobachten, dass Digitalisate neugierig auf diese Zusammenhänge machen. Sie können durchaus auch zufällige Leser in Sachen hinziehen, nach denen sie nicht gefragt haben und von denen sie vorher auch gar nichts wussten. Digitalisiertes Archivgut macht neugierig auf seine Kontexte, weil es Fragen aufwirft und offen ist für Interpretationen. Es fordert selbst zum weiteren Recherchieren auf.

Ein weiterer Nutzen liegt darin, dass Digitalisate die Kooperationen zwischen Archiven beim Aufbau gemeinsamer Zugangspunkte und Portale und bei der Beteiligung an digitalen Bibliotheken unterstützen. Von den neuen Möglichkeiten der gemeinsamen Nutzung von Erschließungen einschließlich der Digitalisate befördert, hat die Diskussion um die gemeinsame Präsentation von spartenübergreifenden Zugängen zu Bibliotheken und Archivgut sowie Museumsobjekten eine neue Dynamik bekommen und wird besonders in den Projekten zum Aufbau digitaler Bibliotheken auf deutscher und europäischer Ebene mit großer Erwartung geführt.⁸

Schließlich ist die Digitalisierung ein Weg dazu, die Kompetenz der Archive auch auf diesem in Politik und Öffentlichkeit stark diskutierten Gebiet zu demonstrieren. Archive beherrschen die Verfahren der Bereitstellung für eine ertragreiche Nutzung von Archivgut auch in digitaler Form. Archive können mit der Digitalisierung Ergebnisse vorweisen, die für sich selbst sprechen, und die umfangreiche Erläuterungen von Vorbehalten gegenüber fachfremden Vorstellungen erübrigen oder zumindest illustrieren können.

1.2 Neue technische Rahmenbedingungen

Die inzwischen verfügbaren technischen Rahmenbedingungen erleichtern die Bereitstellung digitaler Reproduktionen im Internet. Die Bandbreiten, die für eine Übertragung von Bildern zur Verfügung stehen, reichen völlig aus für eine unterbrechungsfreie Beschäftigung

⁸ Als Beispiele für bisherige Portale mit übergreifendem Zugriff siehe <<http://www.bam-portal.de>> und <<http://www.sondera.kb.se>>. Vgl. Sigrid Schieber,

mit den digitalen Abbildungen. Außerdem stehen internationale Fachstandards zur Verfügung, die eine Anschlussfähigkeit der eigenen Arbeiten durch Bereitstellung der Daten auch außerhalb eigener Präsentationen gewährleisten. Mit der Formulierung dieser Fachstandards in der Form strukturierter Beschreibungen von Dokumenten wie den Findbüchern oder Beständeübersichten erlauben sie es,⁹ Daten und Präsentation von einander zu trennen und ermöglichen eine dezentrale Haltung der Daten bei ihren Eigentümern mit gleichzeitiger Bereitstellung für zentrale Zugangspunkte. Ein Grund ist die technische Trennung von Daten und Anwendungen ähnlich wie in der Textverarbeitung. Es werden vollständige Dateien für komplette Findbücher erstellt, die mit anderer Software, so wie eine Word-Datei mit Open Office, geöffnet und bearbeitet werden können. Gleichzeitig ist jedoch bei Nutzung der Standards im Gegensatz zu Texten aus der Textverarbeitung oder den daraus erstellten PDF-Dateien die gezielte Adressierung von Strukturteilen der Dateien möglich, so dass beides, eine Volltextsuche und eine strukturbasierte Navigation miteinander kombiniert werden kann. Die Eigenschaften, die aus der Softwareunabhängigkeit sowie der direkten Lesbarkeit und der Bearbeitungsmöglichkeit mit Hilfe verschiedener Werkzeuge resultieren, machen diese Datenformate offen für eine große Vielfalt

Das neue BAM-Portal. EAD als Austauschformat, in: *Der Archivar* 61, 2008, S. 41-44.

⁹ Es geht dabei vorrangig um die XML-formulierten Standards EAD (Encoded Archival Description) für Findbücher und Beständeübersichten sowie EAC (Encoded Archival Context) für die Beschreibung der Herkunftsstellen und EAG (Encoded Archival Guide) für die Informationen zu den Einrichtungen, die Daten bereitstellen.

an inhaltlichen Ausprägungen, soweit sie die formalen Strukturbeschreibungen respektieren. Deshalb kommen sie der Arbeit mit Archivgut entgegen. Kommt es doch hier auf die jeweils besonderen Eigenschaften eines Bestandes oder seiner Teile an, wie ein spezielles Schriftstück oder einen Vermerk, die, wenn sie im Archivgut neu entdeckt werden, weil bisher niemand daran dachte, danach zu suchen, zum Schlussstein eines Thesengebäudes werden können.

2. Wie digitalisieren?

Das Bundesarchiv hat sein Konzept im Jahre 2006 in einer Digitalisierungsstrategie formuliert und auf seiner Webseite veröffentlicht, um die Grundsätze, die bei der Digitalisierung verfolgt werden auch nach außen transparent zu machen.¹⁰ Darin enthalten ist die Zielvorstellung, die Zugänglichkeit von Archivgut unabhängig von Ort und Zeit weiter zu verbessern und es vor Gefährdungen bei der Benutzung zu schützen, den Nutzungskomfort zu erhöhen und dabei wirtschaftliche Verfahren auf der Basis abgesicherter Standards zu entwickeln und zu nutzen.

Um die Nachhaltigkeit der eingesetzten Maßnahmen zu erreichen, wurde festgelegt, dass weiterhin der Mikrofilm die Form der

¹⁰ Vgl. <<http://www.bundesarchiv.de/aktuelles/fachinformationen/00046/index.html>> (wird abgelöst durch <<http://www.bundesarchiv.de/fachinformationen/01435/index.html>>). Weitere Digitalisierungsstrategien finden sich etwa auf der Webseite des Landesarchivs Baden-Württemberg (<<http://www.landearchiv-bw.de/web/43033>>); vgl. auch die Empfehlung „Digitalisierung von archivischem Sammlungsgut“ sowie die „Eckpunkte für Verträge über die Digitalisierung durch Dritte“ der BKK (<<http://www.bundskonferenz-kommunalarchive.de/empfehlungen.html>>).

Sicherung bleibt und die Digitalisierung vom Mikrofilm aus vorgenommen wird. Die Obsoleszenz der Computertechnik, die aus ihrem stetig wirkenden, immanenten Modernisierungszwang herrührt, wirft bei der Erhaltung genuin digitaler Aufzeichnungen große Probleme auf. Bei digitalen Reproduktionen können die daraus resultierenden Probleme durch kombinierte Strategien, die analoge und digitale Formen einsetzen, umgangen werden. Im Fall einer Veralterung der heutigen Techniken der Speicherung oder Übertragung ist eine spätere Neudigitalisierung vom Mikrofilm wirtschaftlicher als die Verwaltung und Migration der Digitalisate auf Vorrat über längere Zeiträume hinweg, vor allem, wenn sie etwa als Masterkopie in hoher Auflösung aufbewahrt werden sollen. Bei dieser integrierten Bestandserhaltungsstrategie ist das Authentikum mit dem höchsten Schutzbedarf auf Grund des vollen Informationsreichtum die Vorlage im Magazin, der Mikrofilm bleibt die alterungsbeständige Sicherungskopie und das Digitalisat erhält den neuen Platz als komfortable Nutzungsform, deren Bereitstellung die beiden anderen Formen vor Abnutzungserscheinungen schützen kann.

2.1 Auswahlkriterien für Digitalisierungsmaßnahmen

Auf Grund der engen Abhängigkeit einzelner Stücke, Seiten oder Dokumente vom Bestandszusammenhang, der gleichzeitig die Zusammenhänge der Entstehung demonstriert, hat es sich herausgestellt, dass es sinnvoll ist, ganze Bestände zu digitalisieren. So kann die Erklärungskraft der Gesamtheit auch bei der Lektüre einzelner Seiten zum Tragen kommen. Eine eventuell erreichbare Einsparung

an Digitalisierungskosten wird zudem mehr als kompensiert durch den erforderlichen Aufwand für die Auswahl und die physische Separierung der Vorlagen.

Bei der Festlegung von Beständen für eine Digitalisierung würden idealerweise zunächst Referenzbestände in Betracht kommen, die als Ergänzung zu den Findmitteln im Internet den Zugang zu weiteren Beständen eröffnen. Dazu gehören Bestände zentraler Stellen, wie etwa das Politbüro der SED, in denen bei der Recherche nach Unterlagen in bestimmten Angelegenheiten Hinweise für Verbindungen zu anderen Stellen und deren Beständen gefunden werden können. Unter dem Gesichtspunkt der Schonung häufig genutzter Bestände haben auch sie eine höhere Priorität, wenngleich sich die Nutzungsfrequenz mit einer Verlagerung des Fokus historischer Debatten oder mit besseren Informationen über andere Bestände verändern kann.

Da bisher keine zusätzlichen Haushaltsmittel für Digitalisierungsmaßnahmen bereitstanden, wurden Drittmittel oder Finanzierungen von Kooperationspartnern genutzt, die an bestimmten Beständen interessiert waren. Dazu gehört etwa das umfangreiche und mehrjährige Projekt mit dem Verlag K.G. Saur für die Mikrofiche-Edition „Partei und Staat in der DDR“. Der Verlag eröffnete eine neue Vertriebsform für die zuvor schon beim Bundesarchiv zu erwerbenden Mikrofiches der Protokolle des Politbüros der SED. Dafür erbat sich Saur das Recht zur Digitalisierung der Filme unter Einräumung eines Rechts zur Nutzung der digitalen Reproduktionen durch das Bundesarchiv in dessen eigenem Internetangebot nach einer Karenzzeit von drei Jahren. Dadurch können nun die Protokolle des Politbüros sowie die Unterlagen aus den Büros der Generalsekretäre

der SED sukzessive in das Internetangebot aufgenommen werden. Zwei weitere Bestände von Büros des ZK der SED wurden als Beispiele im <daofind>-Projekt mit Mitteln der Andrew-W.-Mellon-Foundation digitalisiert.¹¹ Die Digitaliste zu den 230 Akten des Bestands Einsatzstab Reichsleiter Rosenberg (NS 30) wurden im Rahmen der Provenienzforschung erstellt. Die Protokolle des Nationalen Verteidigungsrates wurden mit dem Militärgeschichtlichen Forschungsamt zusammen erschlossen sowie digitalisiert und stehen sowohl innerhalb von ARGUS wie auf der dortigen Projektseite bereit.¹²

2.2 Technische Parameter

Die technischen Parameter sind ebenfalls in der Digitalisierungsstrategie beschrieben. Sie sind dem Zweck der Digitalisierung angepasst und sichern die bestmögliche Nutzbarkeit bei gleichzeitiger Beachtung der Wirtschaftlichkeit der Maßnahmen. So wird einheitlich das Format PNG (Portable Networks Graphic) für die Speicherung und die Anzeige der Digitalisate verwendet, um die Notwendigkeit der Konversion für die Internetnutzung auszuschalten und eine verlustfreie Komprimierung einzusetzen. Zudem verschärft PNG im Gegensatz etwa zu JPEG die Kontraste und fördert so die Lesbarkeit der Schriften und die Erkennbarkeit von Randnotizen oder Anmerkungen.¹³ Die

¹¹ <<http://www.daofind.de>>.

¹² <<http://www.nationaler-verteidigungsrat.de>>.

¹³ Orientierungsrahmen für diese Festlegungen waren die Ausführungen von Gerald Maier, Qualität, Bearbeitung und Präsentation digitaler Bilder, in: Hartmut Weber/Gerald Maier (Hrsg.), Digitale Archive und Bibliotheken. Neue Zugangs-

Digitalisate werden in einer auf den üblichen Computerbildschirmen gut darstellbaren und von Druckern nutzbaren, relativ niedrigen Auflösung in Graustufen formatiert und gespeichert.

2.3 Arbeitsabläufe

Seit Beginn der Digitalisierungsmaßnahmen ist es ein Prinzip, dass alle damit verbundenen Arbeitsschritte, soweit es eben geht, am archivischen Facharbeitsplatz stattfinden. So behalten die Referate die volle Verantwortung für ihre Bestände und deren Bereitstellung unabhängig von der Nutzungsform. Das betrifft alle Arbeitsschritte von der Vorbereitung der Bestände ähnlich wie für die Mikroverfilmung, über die Kontrolle der gelieferten Digitalisate, ihre Speicherung in strukturierten Verzeichnissen, ihre Integration in die Internetpräsentation bis zum Hochladen auf den Server.

Die Integration der Digitalisate in die Internetpräsentation geschieht über ihre Einbindung in die dort bereits vorhandenen Online-Findbücher.¹⁴ Dazu wird für jede Verzeichnungseinheit ein digitales archivistisches Objekt gebildet, das aus den zugehörigen Digitalisaten besteht. Repräsentiert, zusammengehalten und intern strukturiert wird es mithilfe einer Datei im METS-Format (Metadata Encoding and Transmission Standard).¹⁵ Die METS-Datei mit der Signatur der

möglichkeiten und Nutzungsqualitäten, Stuttgart 2000, S. 129-178, v.a. S. 149-152.

¹⁴ Das entspricht den Empfehlungen zur Weiterentwicklung der Präsentationen von Erschließungsinformationen der ARK (Archivreferentenkonferenz von Bund und Ländern) vom 30.4.2009 (<<http://www.ark.bundesarchiv.de>>).

¹⁵ Zur Dokumentation des Standards vgl. <<http://www.loc.gov/standards/mets>>. Weitere Informationen zur Verwendung finden sich unter <

Verzeichnungseinheit als Dateiname nimmt die Adressen aller zugehörigen Bilder in der Reihenfolge, in der sie in ihrem Verzeichnis liegen, auf. Dieser Vorgang läuft weitgehend automatisch ab.

Danach kann das digitale Objekt intern strukturiert werden, indem einzelnen Reproduktionen bestimmte Eigenschaften, etwa als Eingang in neuer Sache, als Titelseite eines mehrseitigen Dokuments oder als Bild oder Karte zugewiesen und eine Beschriftung hinzugefügt wird. Dabei kann gleichzeitig geprüft werden, ob einzelne Seiten personenbezogene Informationen enthalten, die nur auf Grundlage eines Benutzungsantrags zugänglich sind. Solche Abbildungen werden im gleichen Zuge durch Seiten mit dem Hinweis auf die Möglichkeit einer Bereitstellung vor Ort nach Ausfüllen eines Benutzungsantrags ersetzt.

Die markierten Abbildungen werden in der Internet-Präsentation auf einer Seite untereinander zusammengestellt. Sie können wie Lesezeichen als Einstiegspunkte in die Akte genutzt werden und zeigen wie ein Inhaltsverzeichnis deren innere Struktur. Diese Seite bietet eine Orientierungsansicht, in der die Bilder je nach Kennzeichnung mit dem oberen Drittel oder auf ein Viertel verkleinert angezeigt werden. Mit dem Klick auf ein Bild öffnet sich die Anzeige der Einzelansicht zum Durchblättern oder gezielten Aufruf einzelner Seiten. Von hier kann zurück zur Orientierungsansicht gewechselt werden, die sich wieder an der Stelle öffnet, an der man sie verlassen hat.

online.de>, der Webseite zum DFG-Projekt für den Ausbau des Netzwerks SED-/FDGB-Archivgut. Zu den technischen Aspekten der Integration von Digitalisaten in Online-Findbücher unter Verwendung von METS vgl. auch den Beitrag von Paul Bantzer in diesem Band.

Für die Erstellung und Bearbeitung der METS-Datei wird das Werkzeug MINDEX eingesetzt. Es ist eine Weiterentwicklung von MEX, dem Midosa-Editor für XML-Standards, der für die Nutzung durch die Staatsarchive der neuen Bundesländer zur Aufbereitung ihrer Daten für das Verbundfindmittel im Netzwerk SED-/FDGB-Archivgut zusätzlich mit Konversionsskripten für die Umwandlung von lokalen Erschließungsangaben in ein zentrales Profil von EAD (Encoded Archival Description) ausgestattet wurde.¹⁶

Der METS-Editor ist das Werkzeug, mit dem die Präsentation der Digitalisate für das Internet vorbereitet wird. Damit werden alle Angaben erfasst und in je eine Datei für jede Akte zusammengeführt. Am archivischen Arbeitsplatz zeigt MINDEX die Digitalisate in einem speziellen Fenster an und erlaubt gleichzeitig die Erfassung der entsprechenden Angaben in der METS-Datei, ohne dass dabei Kenntnisse von XML oder der Elemente bzw. Attribute von METS erforderlich sind. Die METS-Dateien können einzeln oder als Block zusammen mit dem Findbuch, zu dem sie gehören, nach HTML ausgeneriert werden. Die HTML-Präsentation kann vor dem Hochladen in die Rechercheplattform ARGUS (ARchivGUtSuche) auch ohne Internetanbindung auf dem eigenen Computer überprüft werden. Nach dem Hoch-

¹⁶ MEX war im Zuge der Entwicklung von Verfahren und Werkzeugen für die Bereitstellung von Digitalisaten im Internet im Rahmen des 2005 bis 2007 mit einer Förderung durch die Andrew W. Mellon-Stiftung, New-York durchgeführten <daofind>-Projektes entstanden. Zugang zur Projektwebseite über <<http://www.archivgut-online.de>> oder direkt unter <<http://www.daofind.de>>. Seit Beendigung des Projektes steht MEX als Open-Source-Software auf SourceForge (<<http://sourceforge.net/projects/mextoolset>>) zum Download bereit. Der heute

laden in ARGUS wird sie mit der Beständeübersicht verknüpft und steht zur Recherche bereit.

MIDEX kann zusätzlichen zu den digitalen Abbildungen die Texte der Aktenseiten als Transkriptionen einbinden und sie für eine Volltextsuche in den Digitalisaten nutzen. Falls Adressen von Textdateien in einer METS-Datei vorhanden sind, stellt ARGUS am Kopf der jeweiligen Präsentation des digitalen Objektes ein Suchfeld bereit. Wird ein dort eingegebener Begriff in den Transkriptionen gefunden, öffnet die Präsentation die Abbildung der ersten zugehörige Seite und markiert in der anklickbaren Liste der Seitenzahlen am Kopf des Browserfensters die weiteren Seiten mit diesem Begriff. Die Volltextsuche in Digitalisaten wird aktuell in der ARGUS-Präsentation für einen kleineren Bestand angeboten und soll es erlauben, ihre Nützlichkeit auszuprobieren. Auf Grund eines doch eingeschränkten Nutzens im Verhältnis zum erforderlichen Aufwand besteht derzeit keine Absicht, die Funktion auf andere Bestände auszuweiten. Allerdings eignet sich das vorhandene Angebot als Demonstration, wie eine solche Funktion auch in größerem Umfang realisierbar wäre.

2.4 Erste Erfahrungen

Seit 2008 werden mit dem Werkzeug MEX sowie seinem Nachfolger MIDEX Digitalisate in Online-Findbücher integriert und in ARGUS bereitgestellt. Mit den Ende 2009 bereitgestellten Mengen liegen nun einige Erfahrungen im Umgang mit den Digitalisaten und mit den Arbeitsabläufen im Massenverfahren vor.

in MIDEX genutzte METS-Editor ist immer noch weitgehend identisch mit dem in MEX eingebundenen Modul.

Die Arbeit mit dem Werkzeug erfüllt die Erwartungen. Damit ist es möglich, dass archivische Fachkräfte ihre Beständekenntnis einsetzen, um Digitalisate besser nutzbar zu machen. Es erlaubt zudem eine Arbeitsteilung zwischen der Anlegung der METS-Dateien mit der automatisierten Integration der Dateiadressen sowie der anschließenden Auswahl der Seiten für die Orientierungsansichten. Mehrere umfangreiche Bestände sind auf diese Art bearbeitet worden und stehen zur Nutzung bereit.

Als kritischer Faktor im gesamten Arbeitsablauf der Digitalisierung hat sich die Verwaltung der großen Mengen an Digitalisaten, die aus den Kooperationsprojekten kamen, vor ihrer Bearbeitung herausgestellt. Es wurde klar, dass sie möglichst zügig in Online-Findbücher eingebunden werden müssen, um nicht Gefahr zu laufen, für den Zugriff verloren zu gehen. Die Findbücher sind auch intern die am besten geeigneten Ordnungsmittel und Zugriffsmechanismen über längere Zeiträume. Über sie behalten die digitalen Abbildungen ihre Identität und bleiben adressierbar. Bewährt hat es sich dabei, als Dateinamen die Signaturen der Verzeichnungseinheiten zusammen mit einer internen Durchnummerierung zu verwenden.

Diese Erfahrungen bedeuten auch für neue Digitalisierungsprojekte, dass die dabei zu bewältigenden Quantitäten vorrangig durch die Kapazitäten für die Einbindung in die Online-Findbücher gesteuert werden. Die nachfolgende Bearbeitungsgeschwindigkeit bestimmt die vorausgehende Digitalisierungsgeschwindigkeit. Beim Versuch, diese Abhängigkeit umzukehren, steigt das Risiko für die Digitalisate, längerfristig nicht mehr identifizierbar zu sein und damit aus dem Zugriff zu geraten.

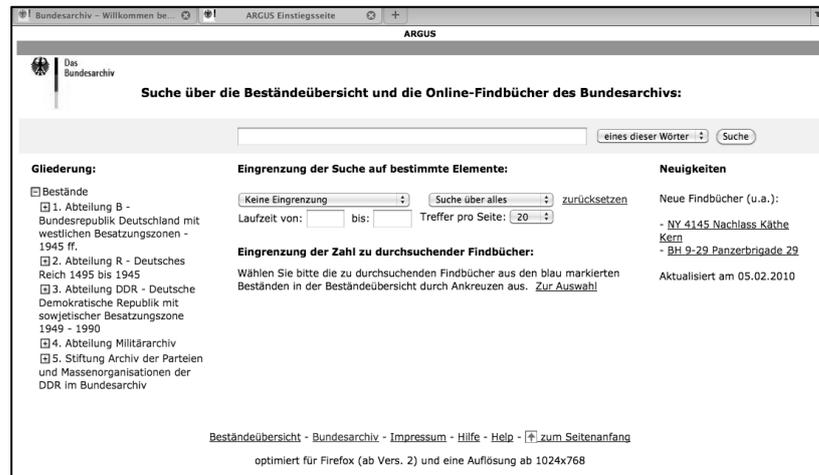


Abb. 1: Einstiegsseite von ARGUS

Ein zweiter kritischer Faktor sind spezielle Grundkenntnisse von Fotobearbeitungssoftware bei einigen der Mitarbeiter, um Auswirkungen von Formatkonvertierungen, Stapelverarbeitungen und Bilderverwaltung vorhersehen und erklären zu können sowie Leistungsbeschreibungen für eine Vergabe der Digitalisierung als Dienstleistung detailliert genug formulieren zu können. Diese Kenntnisse sind jedoch andere, als man sie für die Bearbeitung von digitalen Fotos benötigt und sind relativ eng eingrenzbar. Um Probleme bei einer eventuell nötig werdenden Überarbeitung der Dateien zu reduzieren, können die Ursprungsdigitalisate, von denen aus ein angefordertes Format eventuell heruntergerechnet wird, mit übernommen und für eine begrenzte Zeit während der Qualitätskontrolle gespeichert werden. Die Löschung überflüssiger Kopien erfordert dann jedoch eine

Auf der Einstiegsseite von ARGUS wird eine Volltextsuche über alle Erschließungsangaben sowie ein strukturierter Einstieg angeboten. Über den Strukturbaum auf der linken Seite können in der Beständeübersicht die Kurzbeschreibungen aller Bestände angesehen und bereits vorhandene Online-Findbücher aufgeschlagen werden. Mithilfe des Strukturbaums können außerdem in der Beständeübersicht ausgewählte Findbücher markiert werden, auf die eine nachfolgende Suche begrenzt werden soll. Suchergebnisse werden in einer strukturierten Resultatliste angezeigt. Die Reihenfolge der Treffer entspricht der Reihenfolge der Bestände in der übergeordneten Struktur sowie der Gliederung der jeweiligen Findbücher. Diese Anordnung ersetzt ein automatisches Ranking, das in diesem Kontext keine sinnvolle Aussagekraft besäße.

In der Resultatliste werden jeweils die untersten Gliederungsüberschriften mit der Zahl der zutreffenden Titel von Verzeichnungseinheiten in dieser Gliederungsgruppe angezeigt. Mit dem Plus-Zeichen kann die vollständige Gliederung nach oben und unten sowie die Liste der gefundenen Titel in dieser Gruppe angezeigt werden. Der Klick auf einen gefundenen Titel öffnet als Detailansicht das jeweilige Findbuch wie ein elektronisches Buch an der Stelle der Verzeichnungseinheit, in deren Erschließungsangaben der gesuchte Begriff gefunden wurde, und zeigt den gefundenen Titel in seinem Umfeld an.

Unterhalb des Titels erscheint bei Vorhandensein einer METS-Datei, die mit dem Findbuch verknüpft ist, ein Link zu den digitalen Reproduktionen mit der Beschriftung „Akte einsehen“. Der Text der

Beschriftung kann in der Datei des Findbuchs bearbeitet werden. Die Art der Präsentation wird mit der METS-Datei gesteuert.

Das digitale Objekt wird mit der Orientierungsansicht geöffnet, von der aus mit einem Klick auf jedes Bild zur Einzelansicht gewechselt werden kann. Die Darstellungsgröße entspricht den bei der Bearbeitung der METS-Datei eingestellten Vorgaben.

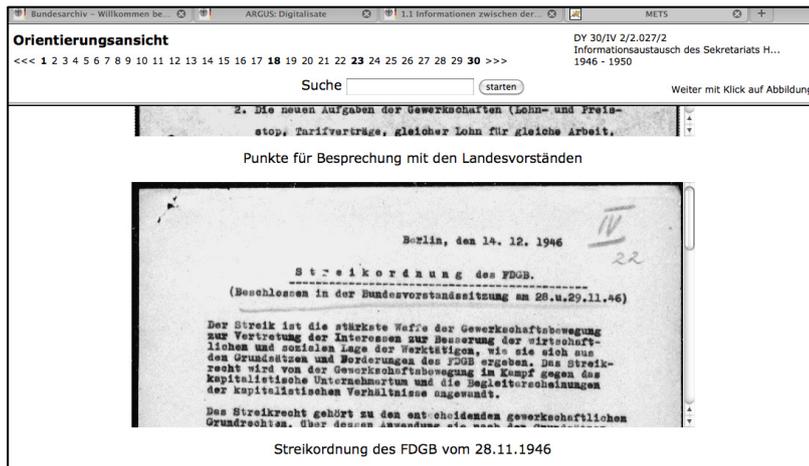


Abb. 3: Abbildung von Seiten mit dem oberen Drittel

METS erweist sich dabei als ein universelles Werkzeug zur Zusammenfassung der zusammengehörenden Digitalisate als digitale archivische Objekte und zu ihrer gleichzeitigen inneren Strukturierung für die Präsentation. Dabei werden nur geringe Teile des Standards genutzt, der auch weitere Angaben zur Verwaltung der Daten, zu ihrem rechtlichen Status oder zur Herstellung und Technik aufnehmen kann.

Die METS-Dateien können bei ihrer Bearbeitung in MIDEX solche Angaben bereits enthalten oder später mit ihnen ergänzt werden, wenn gewünscht.

4. Strategische Einordnung

Digitalisierung von Archivgut hat durchaus strategischen Stellenwert. Sie zeigt einerseits die Geltung und den Nutzen archivischer Grundsätze auch im Internet. Die Präsentation der digitalen Reproduktionen innerhalb der Kontexte ihrer Entstehung ist der Schlüssel zu ihrer Nutzbarkeit und beugt Missverständnissen über ihre Bedeutung vor. Die archivische Digitalisierung antwortet andererseits auf massive politische Forderungen, die von dem inzwischen verbreiteten Grundsatz ausgehen, dass was nicht im Netz ist, nicht in der Welt ist. Die Einführung der Buchsuche bei Google hat Bibliotheken und Politik aufgeschreckt und Vorstellungen von der Verdrängung bekannter und vielfach genutzter Medien durch neue Formen erneute Nahrung gegeben.

Bekannte und lange eingeübte Techniken, die nur deshalb bisher nicht formuliert wurden, weil sie selbstverständlich sind, sind damit noch lange nicht obsolet. Dazu gehören etwa die Techniken der Informationsaufbereitung mit Hilfe der Form von Büchern. So hat etwa David Gelernter in seinem Manifest über die fundamentalen Änderungen des Denkens durch das Internet festgestellt, dass gerade die Instrumentalität eines Buches nicht verzichtbar ist: „Whenever we

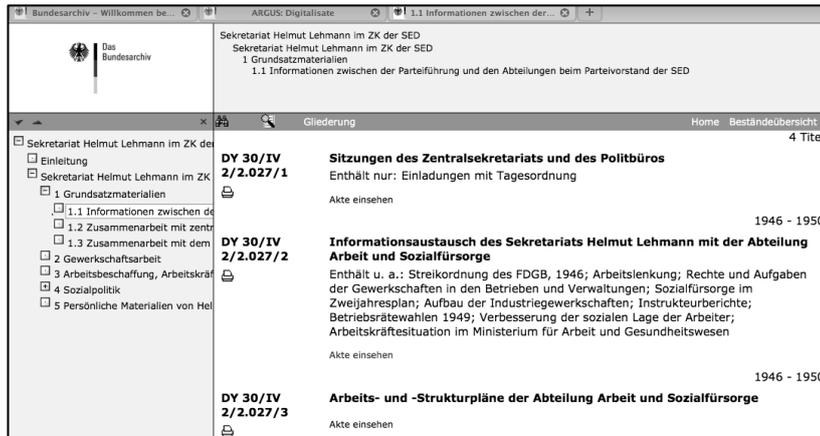


Abb. 4: Titel im Findbuch mit Links zu den zugehörigen digitalen archivischen Objekten

are organizing information, the book is too powerful an idea to do without it in some form or other.”¹⁷

Auch die Idee von den umfassenden Wissensinhalten, die mit Hilfe des Internet als grenzenlose Gesamtheit bereitgestellt würde, ist weit entfernt von einer Realisierung. Die Vision der Aufhebung aller gescannten Bücher durch das eine umfassende Buch des Internets, deren Mächtigkeit Jaron Lanier angesichts der Digitalisierungskampagnen Googles befürchtet, könnte nach seiner Einschätzung wahr werden, wenn genau diese traditionellen Techniken der Infor-

¹⁷ David Gelernter, *The second Coming – A Manifesto*, 1999, <www.edge.org/erd_culture/gelernter/gelernter_index.html>, Punkt 5.

mationsorganisation nicht mehr greifen können, weil keine Ecken und Kanten, keine Grenzen zwischen den Objekten mehr da sind, wenn man also nicht mehr weiß, woher ein Fragment eines Textes stammt, wer eine Notiz aufnahm oder einen Kommentar geschrieben hat.¹⁸ Die Idee der digitalen Bibliothek, die ähnlich, wie es der Mythos der Bibliothek von Alexandria beschreibt, das Wissen umfassend speichern und bereitstellen soll, läuft Gefahr, dem Anspruch zu erliegen, das eine Buch zu konstruieren, das alle anderen aufhebt. Doch mit der Aufhebung der Grenzen, also der Formen, in denen sich Wissen und Fragen artikulieren, wird das Wissen selbst unverstündlich und dadurch unbrauchbar. Ebenso sind die Grenzen zwischen den Beständen, zwischen den Zuständigkeitsbereichen einzelner Archiveinrichtungen wie die Strukturen innerhalb der Findbücher formale Grenzen, die der Verstehbarkeit des Archivguts dienen.

Die Besonderheit von Archivgut ist es, dass es dazu verhilft, sich selbst neues Wissen, das noch niemand kennt, zu erarbeiten. Dazu sind auch solche formalen Merkmale wie die Grenzen zwischen Aufzeichnungen und den Kommentaren oder Vermerken dazu, die eindeutige Autorenschaft von Texten sowie der Platz einer einzelnen Äußerung innerhalb einer Kommunikationskette unverzichtbar. Ziel archivischer Fachkompetenz ist es, diese äußeren Rahmenbedingungen der im Archivgut nachvollziehend zu beobachtenden Kommunikationen präsent und nutzbar zu machen. Deshalb bilden die Strukturen, innerhalb derer Erschließungsinformationen präsentiert werden, die Entstehungszusammenhänge ab. Sie zeigen keine Wissenszu-

¹⁸ Jaron Lanier, *You are not a gadget – A Manifesto*, New York 2010, S. 46-47.

sammenhänge, die dem Archivgut erst nachträglich übergestülpt würden. Sie geben den Benutzern des Archivguts Hilfestellungen wie Brücken oder Geländer, damit sie ihren eigenen Weg bahnen und eigene Erkenntnisse finden. Die Auswertung generiert neues Wissen in den Köpfen der Benutzer, das diese dann publizieren und in den wissenschaftlichen oder den öffentlichen Kommunikationsprozess einbringen können.

Eine strategische Entscheidung ist es ebenfalls, die Arbeitsabläufe bei der Digitalisierung so anzulegen, dass sie der archivischen Fachkompetenz bestmögliche Entscheidungs- und Gestaltungsspielräume eröffnen. Die dezentrale Datenhaltung und die Aufbereitung der Digitalisate durch archivische Fachkräfte integriert ihre Kompetenz, während für die technischen Aufgaben, wie die Bereithaltung ausreichender Speicher und die Sicherung der Daten die Kompetenz von IT-Fachkräften benötigt wird. Diese Arbeitsteilung sichert die Professionalität bei der Nutzung der Möglichkeiten der Digitalisierung wie auch beim Umgang mit ihren Risiken und Gefahren. Die Digitalisierung wird damit zu einer archivischen Fachaufgabe auf allen Ebenen archivischer Tätigkeiten, sei es als Sachbearbeitung, sei es als Leitungstätigkeit, ähnlich wie es Hartmut Weber für die Entwicklung von der Archivtechnik als früherem Unterrichtsfach der Archivausbildung zur Bestandserhaltung als Fach- und Führungskompetenz beschrieben hat.¹⁹ Die Digitalisierung ist dabei Teil eines Bestandserhaltungskonzepts, dass die Erhaltung des Archivguts als Voraussetzung

¹⁹ Vgl. Hartmut Weber, Bestandserhaltung als Fach- und Führungsaufgabe, in: Hartmut Weber (Hrsg.), Bestandserhaltung in Archiven und Bibliotheken, Stuttgart 1992, S. 135-155.

für die Nutzung sieht und nicht als ein ihr widersprechendes Ziel. Deshalb ist Digitalisierung nicht ein technischer Vorgang, sondern Bestandteil der fachlichen Konzepte, die eine Bereitstellung des Archivguts zur Nutzung als ihr oberstes Ziel anstreben und sie mit dem höchsten Schutz des Archivguts vor Gefahren gerade auch durch die Nutzung verbinden. Nur was für eine Nutzung aufbereitet wurde und bereitsteht, kann erhalten werden und nur was in seiner Qualität und Aussagekraft erhalten wird, kann für jeden Nutzungszweck und jede Fragestellung in unveränderter Weise unabhängig vom Nutzungszeitpunkt bereitgehalten werden, der durchaus gestern stattgefunden haben mag und morgen, vielleicht auch in ferner Zukunft erneut eintreten kann.

5. Zusammenfassung und Ausblick

Mit dem Einstieg in die Digitalisierung zum aktuellen Zeitpunkt kann der Rückenwind der öffentlichen Aufmerksamkeit genutzt werden und es ergibt sich die Chance, Erfahrungen zu sammeln und zu lernen. Bei der Digitalisierung ganzer Bestände werden schnell große Mengen an Digitalisaten erreicht, so dass der routinierte und professionelle Umgang mit den Daten erforderlich ist. Dabei entstehen neue Anforderungen an archivarische Fachkräfte. Doch ist auch bei dieser Arbeit eine größere Arbeitsteilung und Spezialisierung sowie die Zusammenarbeit in Gruppen möglich und erforderlich.

Voraussetzung für den Einstieg in die Digitalisierung sind im Internet bereitstehende Erschließungsinformationen etwa in der Form von Online-Findbüchern, in die Digitalisate eingebunden werden können. Damit können sie auch in Kooperationsprojekte wie das Netzwerk SED-/FDGB-Archivgut und seine Weiterentwicklungen einge-

bracht werden. Auch das Projekt für ein europäisches Archivportal (APEnet) nutzt die hier verwendeten archivischen Standards. Das Portal wird auf dieser Basis als Publikationsmöglichkeit zur Verfügung stehen. Die Konzepte für den Aufbau gemeinsamer digitaler Bibliotheken sind zum jetzigen Zeitpunkt noch nicht weit genug entwickelt, um die Schnittstellen für die Integration von Archivgut zu erkennen, die jedoch deutlich formuliertes Ziel bleibt.

Bisher gibt es kaum Rückmeldungen über den Nutzen der Digitalisate für Besucher der Internetangebote des Bundesarchivs trotz deutlich ansteigender Nutzerzahlen für die Internetrecherche in Findmitteln. Mit der Online-Beständeübersicht und den damit verknüpften 1.700 Online-Findbüchern zu Beständen des Bundesarchivs in ARGUS²⁰ ist offenbar eine kritische Masse erreicht, die eine Recherche nützlich werden lässt, sei es zur Vorbereitung eines Archivbesuchs, sei es zur Vorbestellung weiterer Unterlagen vor einer erneuten Archivreise. Viele neue Nutzer, die zuvor noch nicht in Archiven gearbeitet haben, kommen inzwischen über die Online-Findbücher zum Bundesarchiv.

Unklar ist noch, ob es dieselben Nutzer sind, die in den Findbüchern recherchieren und die an Digitalisaten interessiert wären oder ob mit den digitalen Abbildungen andere, neue Nutzungsfälle bedient werden können. Eine übliche Archivbenutzung benötigt mehrere Bestände, unabhängig davon, in welcher Form sie vorliegen. Wäre es sinnvoll, einen besonderen Zugang anzubieten, der speziell zu den digitalisierten Beständen führt? Oder werden damit nur die an der

²⁰ <<http://www.argus.bundesarchiv.de>>.

Digitalisierung als solcher interessierten Besucher angesprochen? Wo fängt die kritische Masse bei Digitalisaten an oder kann sie in diesem Fall kaum definiert werden, da jeweils nur Spezialinteressen angesprochen werden?

Ein Ergebnis kann aber in jedem Fall bereits jetzt festgehalten werden: Archivgut kann durchaus digital im Internet präsentiert und genutzt werden, wenn es innerhalb der Erschließungsinformationen und ihrer Strukturen präsentiert wird. Diese Bedingung wird genauso auch für die Integration von Archivgut in spartenübergreifende Zugriffspunkte gelten und es ist nicht möglich, aber auch nicht nötig, sie wegen angeblich technisch bedingter Sachzwänge zu opfern.

Realisierung der technischen Integration von Digitalisaten unter der Verwendung des Austauschstandards METS

von Paul Bantzer

1. Einleitung – Akten im digitalen Archiv

Die Motivation zur Digitalisierung von Archivgut liegt in zwei Bereichen: Zum einen ist die elektronische Bereitstellung sinnvoll, wenn Akten voraussichtlich häufig angefordert werden und eine hohe Sichtbarkeit des Archivs gewährleistet werden soll. Ein frühes und für diese Gattung typisches Beispiel sind die Erich Honecker betreffenden Unterlagen in Mielkes „Rotem Koffer“, die das Bundesarchiv mit MidosaXML erschlossen hat. Der zweite Bereich dient dem Erhalt der Archivalien: Ist die Ausleihe aus konservatorischen Gründen nicht mehr möglich, so hilft die virtuelle Bereitstellung der Akten. Liegt ausgewähltes Archivgut in digitalisierter Form vor, dann ist es technisch sofort benutzbar. Zur ersten Beurteilung der Relevanz einer Akte blättert ein Benutzer in der Papierakte. Er verschafft sich mit dem „Daumenkino“ einen schnellen Einstieg. Eine elektronische Nachbildung dieser Nutzung stellt sich nicht dar. Die inhaltliche Erschließung der Akte ist mit der Erstellung der Verzeichnungseinheit im Findbuch abgeschlossen, eine tiefer gehende inhaltliche Erschließung wäre wirtschaftlich nicht machbar. Der im Bundesarchiv realisierte Weg geht dahin, ausgewählte Seiten als Einstieg in die Akte hervorzuheben und in einer Übersichtsseite zusammenzustellen.

Je nach Benutzer und Stand der Freigabe können automatisch abgeleitete Freigaben realisiert werden:

- Exponate (Beispiel „Roter Koffer“) können ohne Bedenken im Internet bereitgestellt werden. Sind die Unterlagen von überragendem Interesse, so bietet sich die Bereitstellung des Volltextes der Akte an, um darin suchen zu können.
- Urheberrechtliche Gründe, die eine Benutzung im Lesesaal erlauben, während die Publikation im Internet nicht möglich ist, liegen im Fall externer Gutachten zur Druckgenehmigung im Bestand Ministerium für Kultur (DR1) des Bundesarchivs vor.
- Personenkarteien (Beispiele sind die Heeresbeförderungskartei und die Gefangenenkartei im Bundesarchiv) werden zur Beantwortung von Anfragen intern genutzt. Die Benutzung im Lesesaal beschränkt sich auf besonders berechnigte Benutzer.



Abb. 1: Aus der Pressemitteilung vom 13.2. 2003, „Mielkes „Roter Koffer“ ist jetzt im Bundesarchiv zugänglich.“

Die Präsentation von Image-Digitalisaten aus Archivgut verwendet Metadaten über die Bilddaten, die aus den Filepointern stammen, über die Art der Hervorhebungen für den Aufbau einer Übersichtsseite sowie über die Verknüpfung zwischen unterschiedlichen Ausprägungen wie Image und codiertem Text.

1.1 Was ist METS und was kann es?

Die Verknüpfung der Images einer elektronischen Akte kann als Multipage in TIFF, als PDF oder mittels METS erfolgen. METS steht für „Metadata Encoding and Transmission Standard“ und ist ein „Abstract Model“ zur Beschreibung darin enthaltener Objekte, das von der amerikanischen Digital Library Federation initiiert wurde und als Standard bei der Library of Congress gehostet wird. In allen drei hier vorgestellten Projekten wird METS im OAIS Referenzmodell als Liefereinheit (SIP Submission Information Package), als Archivierungseinheit (AIP) sowie als Bereitstellungseinheit (DIP) verwendet.

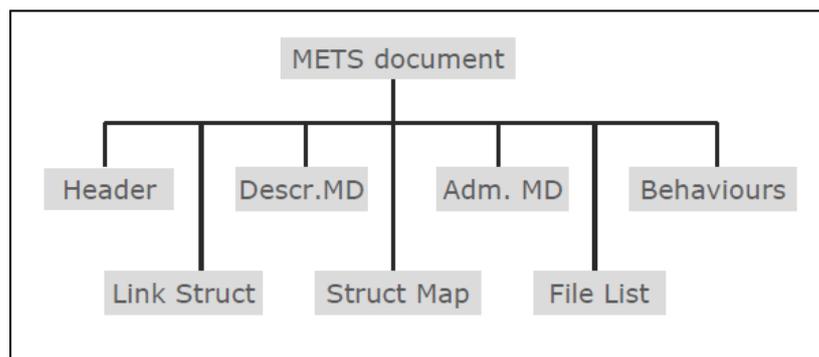


Abb. 2: METS für die Akteneinsicht

METS gliedert sich in sieben Hauptabschnitte, von denen in der archivischen Anwendung die beiden Abschnitte „Struct Map“ und „File List“ verwendet werden. Während die „File List“ eine Abbildung der Directory Struktur der Images darstellt, definiert die „Struct Map“ die Abfolge der Images innerhalb der Akte.

METS ist auf hohe Flexibilität hin ausgelegt. Dazu trägt insbesondere bei, dass das Container-Format weitere Dokumentations-Schemata aufnehmen kann. Die gewonnene Flexibilität ist jedoch ausgesprochen hinderlich, wenn Open Source Komponenten wie der SAXON XSLT Interpreter in der Community Edition nur eingeschränkt verwendbar sind.

In der Konsequenz kommt bei den vorgestellten Projekten die Akteneinsicht ohne weitere Schemata aus. Die Personenkarteien im Projekt PERSEUS betten die Namen im EAC-Format ein. Der VERA Bilder Viewer verwendet, nach der Vorgabe des DFG METS Viewers für erschließende Metadaten, das MODS-Format für die beschreibenden Metadaten.

2. Akteneinsicht – Navigieren und Entdecken

Die Online-Findbücher des Bundesarchivs ermöglichen ein Navigieren im Kontext für ein Entdecken der Zusammenhänge, z.B. wie Entscheidungen entstanden sind. Die Akteneinsicht unterhalb der Ebene der Verzeichnung stellt eine wirkungsvolle Ergänzung dar.

Die Auswahl von Beständen, für die eine elektronische Akteneinsicht realisiert wird, erfolgt unter zwei Gesichtspunkten: Einmal kommt sie für herausragende Bestände in Betracht, die ein hohes Interesse erwarten lassen. Aber auch die Erhaltung hat bei der Aus-

wahl ein hohes Gewicht, da das einmalige Scannen das Original weit weniger belastet als eine wiederholte Ausleihe.

Ein weiterer Gesichtspunkt zur Bestandserhaltung ist, dass die Benutzer eine Akteneinsicht in verfilmte Bestände am Lesegerät eher ablehnen und den Einblick in die Papierunterlage bevorzugen. Der Scan vom Mikrofilm ist überaus preiswert geworden. In der Konsequenz werden verfilmte Bestände, die häufiger benutzt werden, mit niedrigerer Schwelle auch als Scan bereitgestellt. Die elektronische Akteneinsicht wird an Bedeutung gewinnen, je stärker die Benutzer auf den Komfort der sofortigen Verfügbarkeit drängen. Die ersten Findbücher mit Akteneinsicht entstanden im Jahr 2004 für Bestände der Stiftung Archiv der Parteien und Massenorganisationen der DDR (SAPMO). In der Tektonik des Bundesarchivs finden sie sich unter:

Beständeübersicht des Bundesarchivs

Bestände

5. Stiftung Archiv der Parteien und Massenorganisationen der DDR im Bundesarchiv

5.1. Archivbestände

5.1.1. Deutsche Demokratische Republik und SBZ

5.1.1.2. Parteien

5.1.1.2.5. Sozialistische Einheitspartei Deutschlands

5.1.1.2.5.3. Abteilungen, Arbeitsgruppen, Kommissionen beim Politbüro

Die Scans werden für die Publikation im Web mit einer Auflösung von 75 dpi mit Graustufen vom Scan-Dienstleister bezogen.

In den Verzeichnungseinheiten findet sich ein Link „Akteneinsicht“, der zu einer Übersichtsseite der Akte führt. Sie ermöglicht dem Benutzer einen schnellen Einstieg (siehe Abb. 3). Die Anwendung ist so konzipiert, dass einzelne Seiten oder Dokumente durch einfaches Blättern und Springen direkt erreicht werden (siehe Seite 326, Abb. 4). Transkribierte Texte werden in der METS-Datei eingebettet und sind über eine Suchtechnologie auffindbar (siehe Seite 326, Abb. 5). Die Nutzung im Intra- oder Internet erfolgt über einen Webserver. Die komplette Darstellung und Einsehbarkeit einer Akte ist auch offline möglich (siehe Seite 327, Abb. 6).

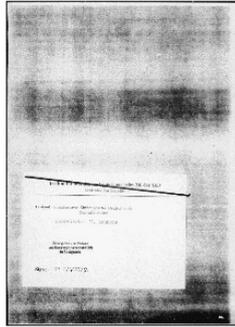
3. PERSEUS: Die Online Bereitstellung von Personenkarteien unter Verwendung von EAC

PERSEUS dient dazu, digitalisierte und sortierte Personenkarteien komfortabel und schnell am PC zu recherchieren. Sie besteht aus zwei Teilen:

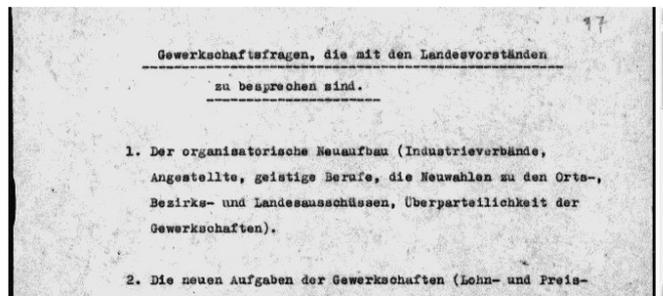
Die Webanwendung dient ausschließlich der Suche. Eine einfache Suchfunktion dient der ersten Positionierung in den Karteien. Von diesem Einstieg aus kann in den Karteien in den Stufen 1, 3, 10 und 25 vorwärts und rückwärts geblättert werden.

Im Web Client wird die Vorderseite in einem Format dargestellt, das den Abgleich von Ordnungskriterien mit dem Rechercheauftrag zulässt. Die Rückseite ist verkleinert dargestellt, um einen ersten Hinweis auf weitere auswertbare Informationen zu liefern. Mit einem Klick auf die Karte wechseln Vorder- und Rückseite in der Größe.

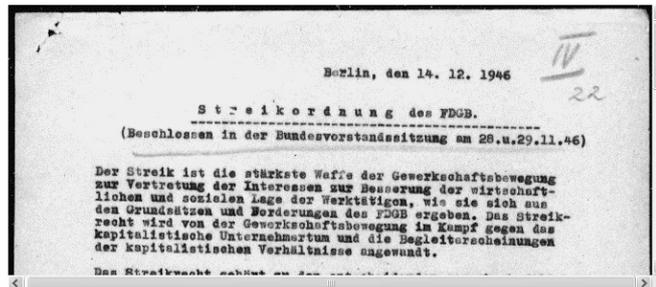
Eine Druckfunktion, sowie ein Aktenkorb zur Bündelung von Suchergebnissen runden die Anwendung ab (siehe Seite 328, Abb. 7).



Aktendeckel



Punkte für Besprechung mit den Landesvorständen



Streikordnung des FDGB vom 28.11.1946

Abb. 3: Scans mit einer Auflösung von 75 dpi

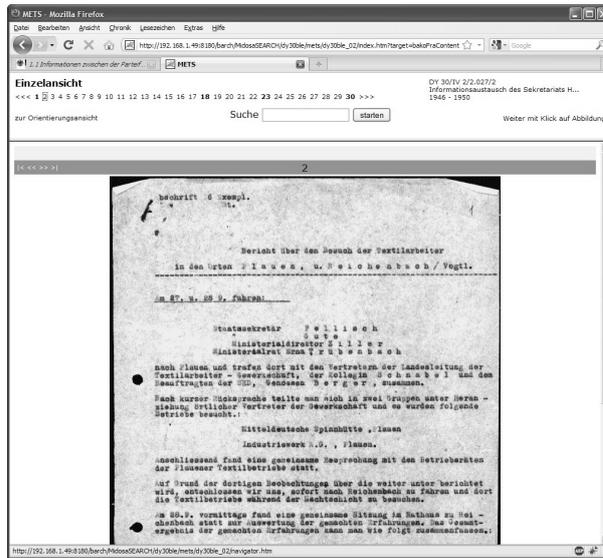


Abb. 4: Darstellung zum Blättern in der Akte

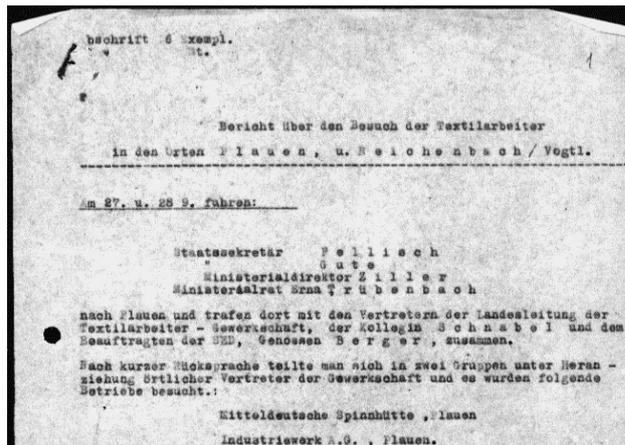


Abb. 5: Scan aus der Akte

```
<file GUID="59eddf1c-3050-409c-9d18-d1e4202d236d"
  ID="FID_12_01_0002" MIMETYPE="text/plain " USE="transkr">
  <Locat GUID="aabe33e0-69ad-488f-8d91-f9bc03afdaa1"
    LOCTYPE="URL" xlink:href="texte/DY30_01_IV_2_2_027_2/01_0002.txt"/>
  <FContent GUID="83208097-6931-4888-80f6-f2b4529d0bde">
    <xmlData GUID="97d127e3-c24c-45b8-b2a4-699ece91a962">
      <mex:content GUID="a57d5591-a5aa-444e-bae3-d372208dba38">&lt;sw
Abschrift 6 Exempl.
&lt;unl/&gt; St.

Bericht über den Besuch der Textilarbeiter in den Orten Plauen, u. Reichenbach / Vogtl.

Am 27. u. 28. 9. führen:

Staatssekretär Fellisch
Staatssekretär Gute
Ministerialdirektor Ziller
Ministerialrat Erna Trübenbach

nach Plauen und trafen dort mit den Vertretern der Landesleitung der Textilarbeiter - Ge-
werkschaft, der Kollegin Schnabel und dem Beauftragten der SED, Genossen Berger, zu-
sammen.

Nach kurzer Rücksprache teilte man sich in zwei Gruppen unter Heranziehung örtlicher Ver-
treter der Gewerkschaft und es wurden folgende Betriebe besucht:

Mitteldeutsche Spinnhütte, Plauen
Industriewerk A. G., Plauen

Anschließend fand eine gemeinsame Besprechung mit den Betriebsräten der Plauener Tex-
tilbetriebe statt.

Auf Grund der dortigen Beobachtungen über die weiter unter berichtet wird, entschlossen
uns, sofort nach Reichenbach zu fahren und dort die Textilbetriebe während der Nacht-
schicht zu besuchen.
```

Abb. 6: Transkription des Scans aus Abb. 5 in XML Text

Die Arbeitsanwendung ist als Erweiterung des MEX-Editors reali-
siert, mit der die METS-Dateien bearbeitet werden. Sie bietet neben
der Suchfunktion eine einfache und schnelle Möglichkeit der Nachin-
dexierung und kann zum Anbringen von Bemerkungen genutzt wer-
den (siehe Seite 328, Abb. 8).

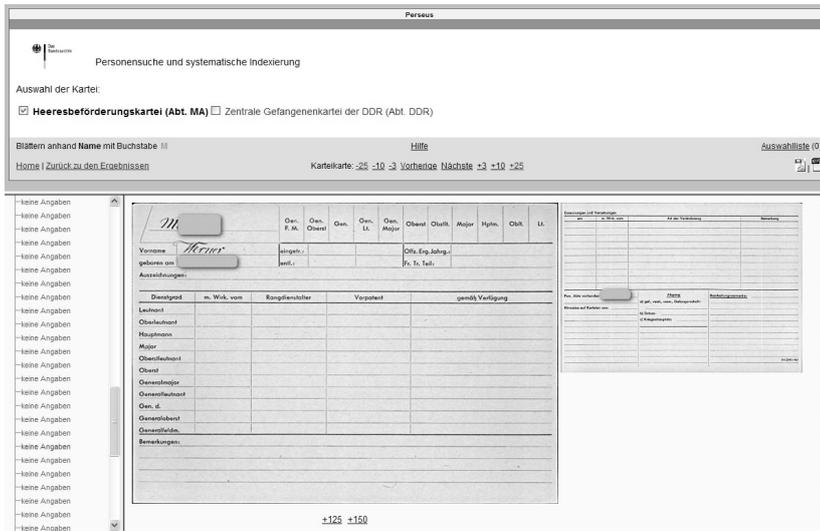


Abb. 7: Personenkartei WEB Client Anzeige der Scans

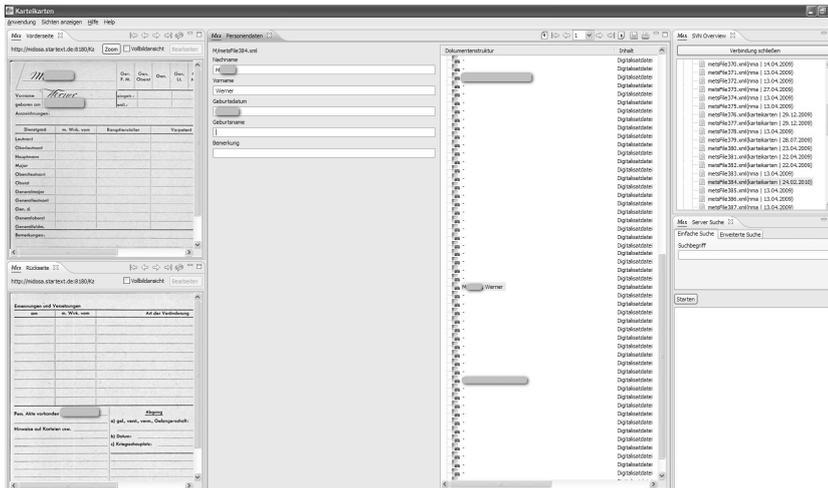


Abb. 8: Personenkartei Eclipse Client

Die Scans der Karten werden auf einem Server im Dateisystem in der Gliederung abgelegt, wie sie vom Digitalisierungsunternehmen geliefert werden. Zu den Scans wird für jeweils 50 Karten eine METS-Datei generiert.

```
<fileSec>
  <fileGrp ID="ID_0" USE="navigation">
    <file ID="FID_0_previous" MIMETYPE="application/xml" USE="navigation">
      <FLocat LOCTYPE="URL" xlink:href="M/metsFile383.xml"/>
    </file>
  </fileGrp>
  <fileGrp ID="ID_1" USE="vorderseite">
    <file ID="FID_1_M_00075439" MIMETYPE="image/png" USE="vorderseite">
      <FLocat LOCTYPE="URL" xlink:href="75dpi-Graustufen/M/M_00075439.png"/>
      <FContent>
        <xmlData>
          [...]
        </xmlData>
      </FContent>
    </file>
    <file ID="FID_1_M_00075441" MIMETYPE="image/png" USE="vorderseite">
      <FLocat LOCTYPE="URL" xlink:href="75dpi-Graustufen/M/M_00075441.png"/>
    </file>
    <file ID="FID_1_M_00075443" MIMETYPE="image/png" USE="vorderseite">
      <FLocat LOCTYPE="URL" xlink:href="75dpi-Graustufen/M/M_00075443.png"/>
    </file>
    <file ID="FID_1_M_00075445" MIMETYPE="image/png" USE="vorderseite">
      <FLocat LOCTYPE="URL" xlink:href="75dpi-Graustufen/M/M_00075445.png"/>
    </file>
  </fileGrp>
</fileSec>
```

Abb. 9: Personenkartei METS fileSec in XML Text (1)

Die Dateien sind untereinander vorwärts und rückwärts verpointert, im Beispiel über die fileGrp mit USE = "NAVIGATION".

Für die Suche in der Personenkartei wird in jedes METS-File die erste Karteikarte als Einstiegspunkt verwendet. Name, Vorname und Geburtsdatum werden im EAC-Format erfasst und in die fileSec als XML Insel im EAC-Format eingebettet (siehe Seite 330, Abb. 10).

```

<xmlData>
  <eac:eac type="persname"
    xmlns:eac="http://jefferson.village.virginia.edu/eac"
    xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
    xsi:schemaLocation="http://jefferson.village.virginia.edu/eac eac.xsd">
    <eac:eachheader countryencoding="iso3166-1"
      dateencoding="iso8601"
      langencoding="iso639-2b"
      ownerencoding="iso15511"
      scriptencoding="iso15924" status="edited">
      <eac:eacid countrycode="DE"/>
      <eac:mainhist>
        <eac:mainevent maintype="create">
          <eac:name/>
          <eac:maindate/>
        </eac:mainevent>
      </eac:mainhist>
    </eac:eachheader>
    <eac:condesc>
      <eac:identity>
        <eac:pershead>
          <eac:part type="surnames">M [redacted] /eac:part>
          <eac:part type="birthname"/>
          <eac:part type="forenames">Werner</eac:part>
          <eac:existdate
            calendar="gregorian"
            era="ce" scope="begin" type="birth"/>
        </eac:pershead>
      </eac:identity>
    </eac:condesc>
  </eac:eac>
</xmlData>

```

Abb. 10: Personenkartei METS fileSec in XML Text (2)

Vorder- und Rückseite werden in der structMap zusammengeführt, indem die IDs der fptr in div-Abschnitten gegliedert werden (siehe Abb. 11).

3.1 Technologie

Die Technologie der PERSEUS-Anwendung kommt ohne Datenbanken aus und nutzt das Dateisystem des Servers, um die elektronischen Karteien bereitzustellen.

```
<structMap>
  <div>
    <div ID="ID_SGFprevious" TYPE="navigation">
      <fptr FILEID="FID_0_previous"/>
    </div>
    <div ID="ID_SGFM_00075439" TYPE="standard">
      <fptr FILEID="FID_1_M_00075439"/>
      <fptr FILEID="FID_2_M_00075440"/>
    </div>
    <div ID="ID_SGFM_00075441" TYPE="standard">
      <fptr FILEID="FID_1_M_00075441"/>
      <fptr FILEID="FID_2_M_00075442"/>
    </div>
    <div ID="ID_SGFM_00075443" TYPE="standard">
      <fptr FILEID="FID_1_M_00075443"/>
      <fptr FILEID="FID_2_M_00075444"/>
    </div>
    <div ID="ID_SGFM_00075445" TYPE="standard">
      <fptr FILEID="FID_1_M_00075445"/>
      <fptr FILEID="FID_2_M_00075446"/>
    </div>
  </div>
```

Abb. 11: Personenkartei METS structMap in XML Text

Die Scans von Vorder- und Rückseite werden für jeweils 50 Karten in einem METS File gekapselt. Die resultierende XML-Datei ist für den MEX-Editor über SVN mit Password-Schutz sowie HTTPS chiffriertem Protokoll geschützt zugreifbar. Die Scans liegen im Dateisystem und können über einen Web Service vom Eclipse- und Web Client gelesen werden.

Die Webanwendung ist mit dem Framework XTF der California Digital Library realisiert. Der Indexer liest die METS-Dateien mit den eingebetteten EAC-Namenseinträgen und indexiert sie mit der Such-

maschine Lucene. XTF dient auch der Eclipse-Anwendung als SRU-Server, um die Suche in den Namen zu ermöglichen. Vom Zugriff her ist die XTF-Anwendung über User und Password geschützt.

4. Bilder und MODS – Beispiel V.E.R.A.

Die Anwendung entstand im Auftrag der Landesarchivverwaltung NRW (LAV), um Digitalisate in die mit V.E.R.A. erstellten Findmittel zu integrieren. V.E.R.A. ist das Verwaltungs-, Erschließungs- und Recherchesystem für Archive. V.E.R.A. beinhaltet alle vom Archivar benötigten Funktionen für die Bestände- und Archivalienerschließung, die Archivaliendokumentation und die Archivgutverwaltung. Entwickelt wurde V.E.R.A. für die Landesarchivverwaltung NRW.

Die Digitalisate zu einer Akte werden im METS-Format erfasst und mit der Verzeichnungseinheit in der Datenbank verknüpft. Die Zuordnung erfolgt über einen Pfad zwischen der Verzeichnungseinheit der Akte und der METS-Datei des Digitalisats.

Die Entwicklung zielt darauf, in den Konventionen des DFG-Viewers Scans im Jpeg Format verfügbar zu machen und die Möglichkeiten der gekachelten Darstellung mit Zoomify bereitzustellen.

Die Auslegung des DFG-Viewers erwartet eine XML-Schnittstelle zum Austausch von Meta- und Strukturdaten im METS/MODS-Format. Mit Hilfe dieser Daten wird die Anzeige eines Digitalisats angereichert und gegebenenfalls um zusätzliche Funktionen, z.B. Zoomify, ergänzt.

MODS wird analog zu EAC in das METS-Dokument eingebettet und enthält eher die technischen Metadaten zu den Digitalisaten – Bestandsbeschreibung, Angaben zur Herkunft etc. (siehe Abb. 12).

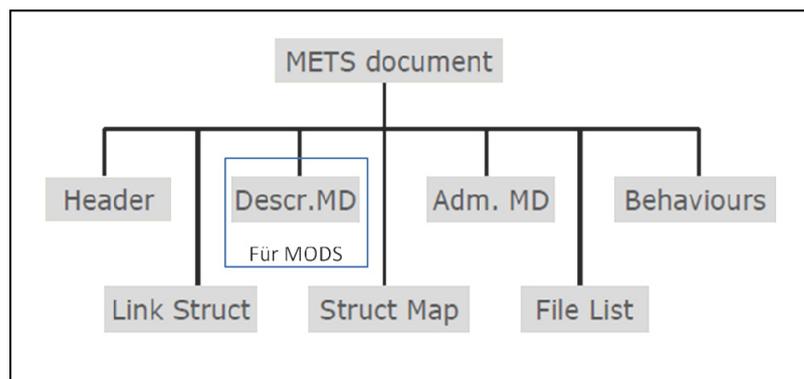


Abb. 12: METS für den Bildviewer

Eingebettet in die METS-Datei sind MODS-Angaben mit weiteren vornehmlich technischen Details zum Digitalisat. Die MODS-Angaben sind innerhalb von METS in die Descriptive Metadata Section (Desc.MD) eingebettet. Sowohl METS- als auch MODS-Angaben werden im BilderViewer dargestellt (siehe Seite 334, Abb. 13).

Das Beispiel illustriert die hohe Flexibilität, mit dem xmlData Tags weitere XML-Schemen in das Austauschformat integrieren können. In der Anwendung des LAV wird lediglich der mods:title verwendet, um eine Suche innerhalb des Findmittels zu ermöglichen.

Die Jpeg Bilder werden in der Struktur des Dateisystems in der METS File List eingetragen. Dabei erfolgt die Anbindung von Zoomify als Verknüpfung des Dienstes über die Namenskonvention der fileGrp (siehe Seite 334, Abb. 14).

```

<mets xmlns="http://www.loc.gov/METS/" xmlns:xlink="http://www.
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi
  mets.xsd http://www.w3.org/1999/xlink xlink.xsd">
  <dmdSec ID="DMD_00">
    <mdWrap MDTYPE="MODS" MIMETYPE="text/xml">
      <xmlData>
        <mods:mods xmlns:mods="http://www.loc.gov/mods/v3" vers
          xsi:schemaLocation="http://www.loc.gov/mods/v3 http://w
            <mods:identifiertype="VERA dockey">Vz      00001572M
            <mods:titleInfo>
              <mods:title>Ein siebenjähriger Junge ist in Berlin
            </mods:titleInfo>
            <mods:relatedItem>
              <mods:identifiert>Best      00000115MST2</mods:identif
            </mods:relatedItem>
            </mods:mods>
          </xmlData>
        </mdWrap>
      </dmdSec>

```

Abb. 13: Personenkartei METS dmdSec in XML Text

```

<fileSec>
  <fileGrp USE="ZOOMIFY">
    <file ID="0001_vonu_01_I_ZFY">
      <Flocat LOCTYPE="URL" xlink:href="00001/0001_vonu_01_I"></Flocat>
    </file>
  </fileGrp>
  <fileGrp USE="DEFAULT">
    <file ID="00001_MSC_IV_00001_001_jpg_DEF" MIMETYPE="image/jpeg" SIZE="21
      <Flocat LOCTYPE="URL" xlink:href="Store0/arbeitsdateien/Staatsarchiv_1
    </file>
    <file ID="00001_MSC_IV_00001_002_jpg_DEF" MIMETYPE="image/jpeg" SIZE="31
      <Flocat LOCTYPE="URL" xlink:href="Store0/arbeitsdateien/Staatsarchiv_1
    </file>
    <file ID="00001_MSC_IV_00001_003_jpg_DEF" MIMETYPE="image/jpeg" SIZE="11
      <Flocat LOCTYPE="URL" xlink:href="Store0/arbeitsdateien/Staatsarchiv_1
    </file>
    <file ID="00001_MSC_IV_00001_004_jpg_DEF" MIMETYPE="image/jpeg" SIZE="41
      <Flocat LOCTYPE="URL" xlink:href="Store0/arbeitsdateien/Staatsarchiv_1
    </file>

```

Abb. 14: Bilder Viewer fileSec in XML Text

Die Darstellung als Folge von Bildern erfolgt wiederum in der structMap. Die mit div tags gegliederte Struktur enthält die Reihenfolge der Abbildungen sowie die Verweise auf die Abbildungen über fptr Tags.

```
<structMap TYPE="LOGICAL">
  <div DMDID="DMD_00" ID="01700_ff_LOG" TYPE="File"/>
</structMap>
<structMap TYPE="PHYSICAL">
  <div TYPE="physSequence">
    <div ID="01700_ff_Bilderslg_01700_01" ORDER="1" TYPE="page">
      <fptr FILEID="01700_ff_Bilderslg_01700_01_jpg_DEF"/>
      <fptr FILEID="01700_ff_Bilderslg_01700_01_jpg_THB"/>
    </div>
    <div ID="01700_ff_Bilderslg_01700_02" ORDER="2" TYPE="page">
      <fptr FILEID="01700_ff_Bilderslg_01700_02_jpg_DEF"/>
      <fptr FILEID="01700_ff_Bilderslg_01700_02_jpg_THB"/>
    </div>
    <div ID="01700_ff_Bilderslg_01701_01" ORDER="3" TYPE="page">
      <fptr FILEID="01700_ff_Bilderslg_01701_01_jpg_DEF"/>
      <fptr FILEID="01700_ff_Bilderslg_01701_01_jpg_THB"/>
    </div>
  </div>
</structMap>
```

Abb. 15: Bilder Viewer structMap in XML Text

Um die METS/MODS-Dateien zu erstellen und zu pflegen, wurde im Projekt ein Plugin für den MEX-Editor erstellt, der die Bilder im Dateisystem darstellt und aus diesen Strukturen die fileSec erzeugen kann. Weiter dient die Anwendung dazu, die Abbildungen über eine tabellarische Sicht mit den Verzeichnungseinheiten des Findmittels zu verknüpfen:

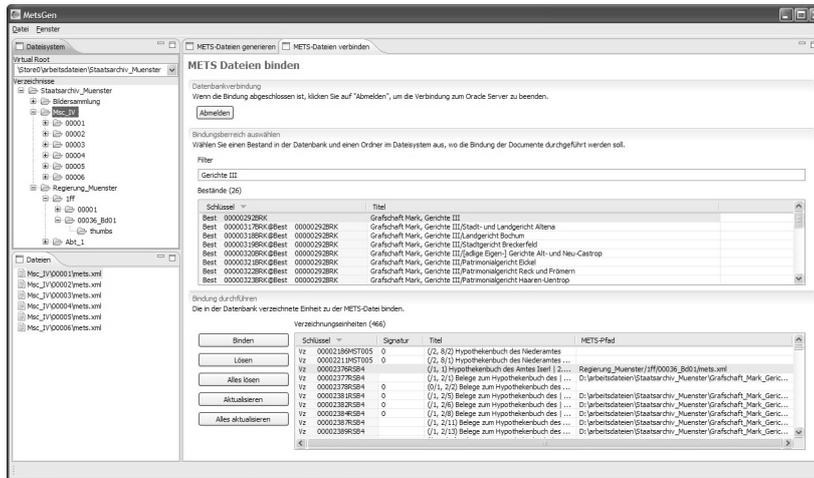


Abb. 16: Bilder Viewer Eclipse Client

Im gewählten Format der Veröffentlichung sind die Details der Bedienungselemente kaum zu erkennen. Für den interessierten Leser stehen die Screenshots über die Web-Seite zum Kolloquium zum Download bereit.

Die Benutzer finden eine Webanwendung vor, die die in METS/MODS bereitgestellten Digitalisate in den Konventionen des DFG-Viewers darstellt. Neben der Darstellung des Digitalisats ermöglicht er die Vorschau auf weitere Bilder (siehe Abb. 17):

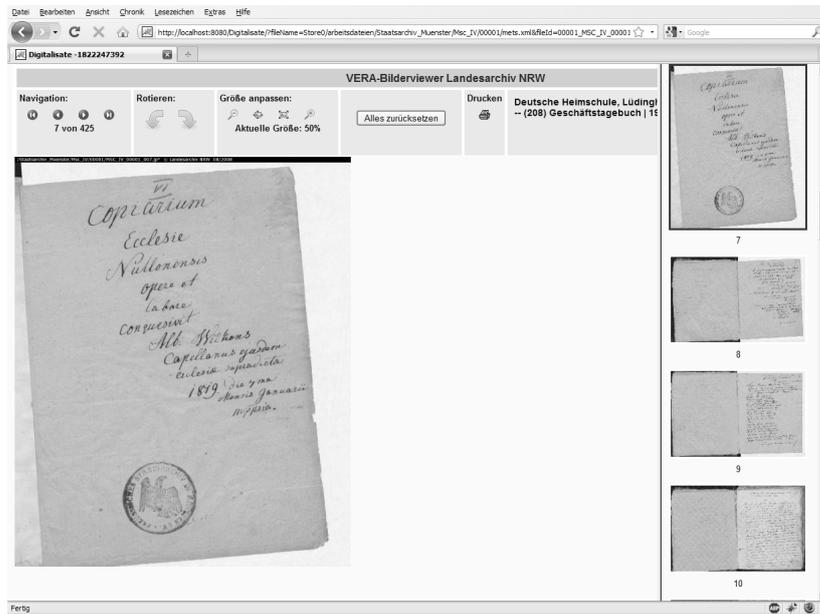


Abb. 17: Bilder Viewer Web Client Anzeige

Archivgutdigitalisierung im Rahmen der Bestandserhaltung. Die Praxis im Landesarchiv Nordrhein-Westfalen

von Johannes Kistenich

In der aktuellen, zumeist stark auf zeitgemäßen Nutzungskomfort für die Kund(inn)en sowie eine bessere Online-Präsenz der Archive fokussierten Fachdiskussion um die Bereitstellung von Digitalisaten im Lesesaal und in Archivportalen gerät gelegentlich der Stellenwert von Digitalisierung zur Erzeugung von Schutzmedien als Element moderner Bestandserhaltungskonzepte für Archive aus dem Blick. Dabei verbindet die Archivgutdigitalisierung geradezu kongenial heutige Kundenerwartungen mit den Anliegen des Schutzes der Originale und führt damit zwei Bereiche zusammen, nämlich Nutzung und Erhaltung, die mit Blick auf die Originale in einem scheinbar unauflösbaren Interessenkonflikt zueinander stehen.

Denn die Benutzung des Originals führt mit der Zeit unvermeidlich zu Schäden, Substanz- und Informationsverlusten, selbst wenn bei Kundinnen und Kunden wie im Übrigen auch bei Mitarbeiter(inne)n im Archiv das Problembewusstsein für den sachgemäßen, sorgfältigen Umgang mit Archivgut bei Aushebung, Transport, Auswertung, Anfertigung von Reprografien, Verwahrung in der Rücklage und Reponierung geschaffen ist und vorhandene Hilfsmittel von Buchkeilen als Unterlagen für Amtsbücher bis hin zu ummantelten Bleibandschnüren zur Beschwerung gefaltet aufbewahrter Pergamenturkunden im Sinne der Schadensprophylaxe bestmöglich eingehalten werden. Die als Archivgut übernommenen Unterlagen sachgemäß zu verwahren, zu sichern, zu erhalten und instand zu setzen,

zählt zwar unstreitig zu den gesetzlich verankerten Kernaufgaben und geradezu konstitutiven Merkmalen von Archiven, doch ist Bestandserhaltung kein Selbstzweck. Zweck und Ziel der Bestandserhaltung – wie der Archivarbeit insgesamt – ist es vielmehr, Archivgut dauerhaft für die Nutzung zur Verfügung zu stellen, oder, um es mit der im angelsächsischen Sprachraum eingängigen Formulierung auf den Punkt zu bringen: *preservation for access*.

Zugang zu Archivgut zu eröffnen und dies für die Nutzung bereit zu stellen, bedeutet nun aber nicht grundsätzlich Benutzung der Originale. Hier bilden Schutzmedien ein zentrales Instrument, um den Interessenkonflikt zwischen Nutzung und Erhalt ein Stück weit aufzulösen. Dies gilt v.a. dann, wenn es um den Informationsgehalt, eine rein inhaltliche Auswertung, also um die Text- oder Bildinformation einer Quelle geht, was getrost als Regelfall der Benutzung angesehen werden darf. Neben der Nutzung im engeren Sinne, also der Auswertung im Lesesaal, entfalten Schutzmedien auch einen Schutz der Originale bei der Benutzerreprografie, bei der Erstellung von Vorlagen für Veröffentlichungs- oder Ausstellungszwecke, erlauben sie doch die nahezu uneingeschränkte Herstellung von Reproduktionen in vergleichbarer Qualität, ohne erneuten Rückgriff auf das Archivgut selbst.

Fragestellungen die sich über den Informationsgehalt hinaus auch auf die Materialität, also die Substanz des Archivguts erstrecken, z.B. auf Einbandtechniken, Schreib- und Beschreibstoffe, auf Herkunftsnachweise von Dokumenten anhand von Wasserzeichen oder der Bearbeitungsart von Pergamentoberflächen, bleiben selbstredend auf die Nutzung der Originale angewiesen. Selbst wenn im Einzelnen

Techniken zur Verfügung stehen, auch solche Informationen z.B. mit Einsatz von Durchleuchttechnik oder Streiflicht bei der Digitalisierung erkennbar zu machen, so wird man diese Optionen bei Standarddigitalisierungsprojekten nicht durchgängig zum Einsatz bringen. Die berechtigten, gleichsam auf den intrinsischen Wert bezogenen Forschungen, entziehen sich in der Praxis daher grundsätzlich einer Diskussion um die Nutzung von Ersatzmedien.¹

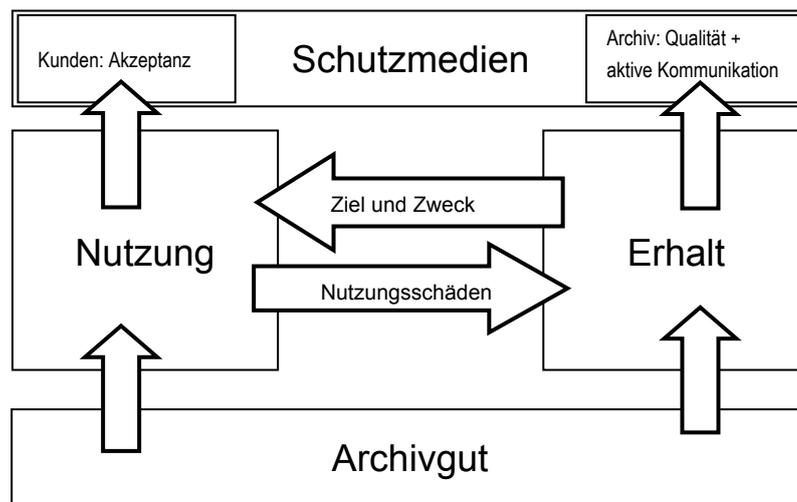


Abb. 1: Bedeutung der Schutzmedien im Spannungsfeld von Nutzung und Erhalt des Archivguts

¹ Vgl. zum intrinsischen Wert: Angelika Menne-Haritz/Nils Brübach, Der intrinsische Wert von Bibliotheks- und Archivgut. Kriterienkatalog zur bildlichen und textlichen Konversion bei der Bestandserhaltung. Ergebnisse eines DFG-Projekts, Marburg 1997 (Veröffentlichungen der Archivschule Marburg 26).

Beim Thema Schutzmedien sind zwei Aspekte von zentraler Bedeutung, die wiederum eng miteinander zusammenhängen:

1. Von Seiten der Archive müssen wir an unsere Kundinnen und Kunden aktiv kommunizieren, warum es im Hinblick auf den Originalerhalt weitgehend unikalere Überlieferung in unseren Archiven notwendig ist, einzelne Bestände im Regelfall anhand von Schutzmedien statt im Original mit seiner unstrittig ganz eigenen Aura des Authentischen, dem „goût des archives“, zu nutzen.
2. Dies wird uns umso besser gelingen, je mehr das Schutzmedium bei unseren Kunden akzeptiert ist, sprich je höher die Qualität und der Nutzungskomfort der Reprografie sind. Die Erstellung eines Schutzmediums, das nicht akzeptiert wird, sodass doch wieder auf die Originale zurückgegriffen werden muss, ist letztlich nicht nur nutzlos, sondern auch unwirtschaftlich.

Heute stehen zwei Varianten der Schutzreprografie zur Verfügung: Mikroverfilmung und Digitalisierung. Das 2008 veröffentlichte gemeinsame Positionspapier der ARK-Fachausschüsse „Bestandserhaltung“ sowie „Sicherung und Nutzung durch bildgebende Verfahren – Fototechnik“ zur Digitalisierung von Archivgut im Kontext der Bestandserhaltung hat sehr ausgewogen die Vor- und Nachteile beider Varianten gegenübergestellt.² In der seit Jahren lebhaft geführten Diskussion um Mikrofilm und/oder Digitalisat bleibt es wichtig fest-

² Positionspapier der ARK „Digitalisierung von Archivgut im Kontext der Bestandserhaltung“, in: Der Archivar 61, 2008, S. 395-398.

zuhalten, dass beide Formen sinnvoll miteinander zu verzahnen sind: Mithilfe von Mikrofilmscannern können Mikrofilme nach derzeitigem technischen Standard schneller und damit im Vergleich zum Aufsichtsscan wirtschaftlicher digitalisiert werden, wenn man einmal von Einzugsscannern absieht, die sich nach unseren Erfahrungen gerade unter bestandserhalterischen Gesichtspunkten nur sehr begrenzt im archivischen Bereich für Digitalisierungsprojekte einsetzen lassen. Soweit Duplikatfilme aus der Sicherungsverfilmung oder andere qualitativ geeignete Schutzfilme vorhanden sind, geht auch das Landesarchiv Nordrhein-Westfalen in erster Linie den Weg der Digitalisierung vom Film. Einzelne Vorbehalte der Kundinnen und Kunden gegenüber dem Mikrofilm, werden schon alleine durch die digitale Bereitstellung selbst behoben, z.B. der schnellere und gezieltere Zugriff auf einzelne Bilder. Es bleiben aber angesichts der bislang erst punktuell eingesetzten Farbmikroverfilmung Einschränkungen des Nutzungskomforts gerade dann bestehen, wenn beispielsweise zur Interpretation der Quelle die Farbinformationen wesentlich sind, etwa bei Fragen der genetischen Aktenkunde. Tendenziell wird sich bei Digitalisaten vom schwarz/weiß-Mikrofilm häufiger die Notwendigkeit des Rückgriffs auf die Originale ergeben, als bei einer Farbproduktion, fernab der Nutzungsgewohnheiten in der „schönen, bunten“ virtuellen Welt.

Daher gehen andere Projekte den umgekehrten Weg, ausgehend von der Direkt-(Aufsichts-)Digitalisierung in Verbindung mit einer Ausbelichtung auf Mikrofilm, insbesondere vor dem Hintergrund offener Fragen bei der Langzeitsicherung digitaler Unterlagen. Zu den „Leuchtturmprojekten“ zählen etwa die Digitalisierungs- und Ausbelichtungsprojekte bei Beständen der Herzogin Anna Amalia

Bibliothek in Weimar oder auch beim Edith-Stein-Archiv der Kölner Karmelitinnen, bei denen übrigens jeweils Farbmikrofilm zum Einsatz gekommen ist.³

Die Vorstellung, dass man zumindest bei größeren Archivverwaltungen den kompletten Originalbestand in überschaubarer Zeit mittels Schutzmedien, egal ob Mikrofilm oder Digitalisat, vor Nutzungsschäden bewahren könne, ist aus heutiger Sicht illusorisch und angesichts der sehr unterschiedlichen Nutzungsfrequenz von Beständen letztlich auch unter bestandserhalterischen Gesichtspunkten gar nicht zwingend erforderlich. Zur Einordnung mögen einige wenige Daten und Fakten genügen: Die im Rahmen der Bundesauftragsverwaltung durchgeführte Sicherungsverfilmung staatlichen Archivguts in der Verfilmungsstelle des Landesarchivs Nordrhein-Westfalen wird seit Jahren mit sechs Kameras und (teils nicht in Vollzeit besetzten) Stellen durchgeführt.⁴ Die Verfilmungsleistung lag zuletzt bei rund 1,6 Mio. Aufnahmen pro Jahr, entsprechend etwa 250 lfd. m Archivgut. Verglichen mit der für die staatlichen Archive in Nordrhein-Westfalen auf 2,2 lfd. km pro Jahr gedeckelten Übernahmemenge

³ Zum Projekt mit Beständen der Herzogin Anna Amalia Bibliothek und dem ARCHE-Pilotprojekt zur Ausbelichtung von Farbdigitalisaten vgl. <<http://www.klassik-stiftung.de/einrichtungen/herzogin-anna-amalia-bibliothek/projekte/sicherungsverfilmung.html>>; <<http://www.landesarxiv-bw.de/web/49137>> Alle im Beitrag genannten Websites wurden am 28.1.2010 besucht.

⁴ Zum Stand der Sicherungsverfilmung im Landesarchiv vgl. Landesarchiv Nordrhein-Westfalen, Grundsätze der Bestandserhaltung – Technisches Zentrum, Düsseldorf 2009 (Veröffentlichungen des Landesarchivs Nordrhein-Westfalen 19), S. 33-35.

(1%-Quote), wird man mit Sicherheit davon ausgehen können, dass die Kapazitäten der Sicherungsverfilmung nicht einmal reichen, um mit dem aktuellen Zuwachs an Archivgut staatlicher Provenienz der Wertstufe 1 (15% der Akten und Amtsbücher aus der Zeit nach 1800) schrittzuhalten. Gemessen am heutigen Gesamtbestand von ca. 160 lfd. km Archivgut im Landesarchiv Nordrhein-Westfalen entspricht die Verfilmungsleistung seit Beginn der statistischen Erhebung im Jahre 1961 bis heute durch die Verfilmungsstelle bei der staatlichen Archivverwaltung Nordrhein-Westfalens mit rund 77 Mio. Aufnahmen etwa sieben bis zehn Prozent.

Mit Digitalisierungsprojekten startete die staatliche Archivverwaltung in Nordrhein-Westfalen gegen Ende der 1990er Jahre im Personenstandsarchiv Brühl, wobei wiederum insbesondere die fehlende Akzeptanz der Kunden im Umgang mit den dort als Schutzmedien eingesetzten Mikrofiches und eine z.T. unzureichende Qualität derselben wesentliche Ausgangspunkte darstellten.⁵ Das Landesarchiv führt Digitalisierungsprojekte seit einigen Jahren sowohl als „In-

⁵ Zur Startphase des Projekts im Personenstandsarchiv Brühl vgl. Wolf-Rüdiger Schleidgen, Pixel contra Mikrofiche. Erfahrungen mit neuen Formen der Nutzung von Archivgut im Nordrhein-Westfälischen Personenstandsarchiv Rheinland, in: Ders./Verena Kinle (Hrsg.), Zwischen Tradition und Innovation. Strategien für die Lösung archivischer Aufgaben am Beginn des 21. Jahrhunderts, Siegburg 2002 (Veröffentlichungen der staatlichen Archive des Landes Nordrhein-Westfalen Reihe E Nr. 8), S. 247-271. Vgl. ferner: Marcus Stumpf, Bestandserhaltungskonzepte für das Landesarchiv Nordrhein-Westfalen, in: Arbeitsblätter des Arbeitskreises Nordrhein-Westfälischer Papierrestauratoren 11, 2007, S. 7-13. Zum aktuellen Stand: Landesarchiv Nordrhein-Westfalen, Grundsätze der Bestandserhaltung – Technisches Zentrum (wie Anm. 3), S. 37-39.

house-Lösungen“ als auch in Kooperation mit Dienstleistern durch. Für die in Eigenregie durchgeführten Digitalisierungen verfügt es über mehrere Aufsichtsscanner (DIN A2 bis DIN A0) zur Direktdigitalisierung, über einen Wolfenbütteler Buchspiegel, ausgestattet mit einer Mittelformatdigitalkamera, für Objekte, bei denen auch durch Einsatz von Buchwippen an den Aufsichtsscannern Einbandschäden drohen, und über derzeit einen Mikrofilmscanner sowie einen Einzugsscanner.⁶

Fragt man nach dem Stand der Schutzdigitalisierung im Herbst 2009, zehn Jahre nach Beginn der ersten Digitalisierungsprojekte und knapp vier Jahre nach Inbetriebnahme des Technischen Zentrums, so werden derzeit rund 53 TB digitale Master (Bilddateien) im Format JPEG2000 aufbewahrt. Seit 2004, dem Zeitpunkt der Einführung einer einheitlichen statistischen Erfassung im Landesarchiv und gleichzeitig in dessen Gründungsjahr, wurden rund 118.000 Archivguteinheiten mit 5,8 Mio. Bilddateien durch Dienstleister, im Technischen Zentrum (seit 2006) und in den Werkstätten der Archivarabteilungen erzeugt. Der Hinweis auf die am ehesten vertraute Größenordnung „Archivguteinheit“ macht deutlich, dass der Anteil digitalisierten Archivguts mit rund 0,5% gemessen am Gesamtbestand des Landesarchivs auch hier ungeachtet erheblicher Mittel bei Investition und Wartung der Scantechnik, Personalkosten für Digita-

⁶ Zu Fragen der Konzeption und Praxis von Digitalisierungsprojekten vgl. jetzt Marcus Stumpf, Grundlagen, Planung und Durchführung von Digitalisierungsprojekten, in: Ders./Katharina Tiemann (Hrsg.), Kommunalarchive und Internet, Münster 2009 (Texte und Untersuchungen zur Archivpflege 22), S. 111-132.

lisierung bei der Inhouse-Lösung, jährliche Vergaben an Dienstleister im Umfang von zuletzt etwa 500.000 Digitalisaten, Qualitätskontrolle, Datenaufbereitung und Speicherung durchaus „überschaubar“ bleibt.

Umso drängender stellt sich die Frage nach den Kriterien für die Priorisierung von Digitalisierungsprojekten. Die Formulierung dieser Kriterien erfolgt im Landesarchiv NRW bislang überwiegend unter dem Gesichtspunkt, Nutzungsinteressen und Bestandserhaltung gleichermaßen zu berücksichtigen. Im Wesentlichen werden Bestände digitalisiert,

- die viel benutzt sind,
- die aufgrund spezifischer Lagerungsformen wie z.B. Großformate besonderen Schadensrisiken beim Transport im Archiv von der Aushebung bis zur Reponierung ausgesetzt sind,
- deren Erhaltungszustand eine Vorlage im Original nicht mehr erlaubt oder
- bei denen sich aufgrund von Schadensbildern mithilfe der Digitalisierung verbesserte Auswertungsmöglichkeiten ergeben.

Als Beispiel für stark frequentierte Bestände ist zunächst die Digitalisierung von Kirchenbüchern und Zivilstandsregistern aus der Personenstandsüberlieferung zu nennen, die auch chronologisch am Anfang der Digitalisierungsaktivitäten des Landesarchivs standen. Digitalisate dieser Projekte werden angereichert durch Kontextinformation (Einführungstexte und -tabellen) in Verbindung mit dem Verlag

Patrimonium Transcriptum als Edition Brühl bzw. Detmold auf CDs bzw. DVDs veröffentlicht und vermarktet.⁷

Nach dem Einsturz des Historischen Archivs der Stadt Köln hat das Landesarchiv NRW für 2009 bis 2011 einen Schwerpunkt auf die Digitalisierung der im Landesarchiv aufbewahrten Zweitüberlieferung der Kölner Personenstandsunterlagen gelegt. 2009 wurden bei einem Dienstleister die bislang noch nicht digitalisierten knapp 400.000 Seiten aus der Zeitschicht bis 1875 gescannt, strukturiert abgelegt und zugleich auf Mikrofilm ausbelichtet. In diesem konkreten Fall erweist sich die zusätzliche Sicherung auf Mikrofilm durchaus als fachlich geboten, weil die Benutzung der Erstüberlieferung, die wenige Tage vor dem Einsturz an das Historische Archiv übergeben worden war, auf Jahre hinaus nicht oder nur eingeschränkt möglich sein wird. Außerdem ist zumindest davon auszugehen, dass die Erstüberlieferung in einzelnen Fällen als verloren gelten muss, sodass ohnehin nach dem Personenstandsgesetz dann die entsprechenden Zweitregister an die Stelle der Erstregister treten werden und eine neue Zweitüberlieferung geschaffen werden muss. Das Landesarchiv NRW begreift diese Digitalisierungsprojekte auch als ein Element seiner „Kölnhilfe“: Kopien der Digitalisate werden nach

⁷ Edition Brühl. Kirchenbücher Vol. 110 ff., Bonn 2006 ff. (Vol. 1-109 = Veröffentlichungen der staatlichen Archive des Landes Nordrhein-Westfalen, Reihe H Nr. 1); Edition Detmold. Kirchenbuchduplikate und Zivilstandsregister, Bonn 2006 ff. (Veröffentlichungen des Landesarchivs Nordrhein-Westfalen 3). Beide CD- bzw. DVD-Reihen werden laufend fortgesetzt.

Abprache mit dem Historischen Archiv auch für die Benutzung im dortigen (provisorischen) Stadtarchiv zur Verfügung stehen.

Ein anderes Beispiel bietet der Bestand Geheime Staatspolizei - Staatspolizei(leit)stelle Düsseldorf („Gestapo Düsseldorf“, LAV NRW Abt. R, RW 58) der Abteilung Rheinland des Landesarchivs mit über 71.000 Archivguteinheiten in 1.700 Archivkartons (rund 190 lfd. m), der derzeit ausgehend von den Duplikatfilmen aus der Sicherungsverfilmung digitalisiert wird.⁸ Die hohe Benutzungsfrequenz, auch die enorme Nachfrage nach Benutzerreprografien zu diesem Bestand seitens verschiedener Forschungseinrichtungen, und damit die zunehmenden Nutzungsschäden, gepaart mit der zeittypisch problematischen Papierqualität der 1930er/40er Jahre rückt den Bestand in der Prioritätensetzung der Bestandserhaltungsmaßnahmen unter mehreren Gesichtspunkten weit nach oben. Gleiches gilt für die aktuell laufende Planung zur Digitalisierung der rund 1,2 Mio. Entnazifizierungsakten, die in der Abteilung Rheinland des Landesarchivs aufbewahrt werden.

⁸ Zu diesem Bestand vgl. das gedruckte Auswahlinventar von Peter Dohms, Flugschriften in Gestapoakten. Nachweis und Analyse der Flugschriften in den Gestapoakten des Hauptstaatsarchivs Düsseldorf, Siegburg 1977 (Veröffentlichungen der staatlichen Archive des Landes Nordrhein-Westfalen, Reihe C Nr. 3). Zur Bedeutung des Bestands vgl. z.B. Holger Berschel, Bürokratie und Terror. Das Judenreferat der Gestapo Düsseldorf 1935-1945, Essen 2001 (Düsseldorfer Schriften zur Neueren Landesgeschichte und zur Geschichte Nordrhein-Westfalens 58), S. 14f.

Die Digitalisierung solcher für die Zeitgeschichtsforschung zentraler Quellenbestände wirft die Frage auf, ob und wie solche Projekte zukünftig als drittmittelfinanzierte Kooperationsprojekte mit Forschungseinrichtungen gemeinsam geplant und durchgeführt werden können. Dabei werden wir davon ausgehen dürfen, dass die Interessen von Forschungsseite und Archiv an solchen Digitalisierungsprojekten durchaus nicht kongenial sind, bis hin zu Fragen von Formaten und Auflösung, die ja nicht zuletzt auch Kostenfragen sind. Umso wichtiger wird es sein, von archivischer Seite – anders als bei Erschließungsinformationen und Austauschformaten – bislang archivübergreifend noch weitgehend fehlende fachliche Standards für Scanprozesse, Strukturdatenerfassung, Dateiformate etc. im Hinblick auf die digitale Langzeitsicherung zu formulieren und auf deren Einhaltung bzw. Umsetzung zu achten. Das verwahrende Archiv bleibt für den gesamten Prozess von der Auswahl der Dienstleister bis hin zur Qualitätskontrolle in seiner Verantwortung.

Einen weiteren Schwerpunkt hat das Landesarchiv NRW bei seinen laufenden Digitalisierungsprojekten auf Bestände gelegt, die formatbedingt besonderen Schadensrisiken beim Transport im Archiv ausgesetzt sind, also bei Aushebung, Transport vom Magazin zum Arbeitsplatz, bei der Benutzerreprografie und bei der Reponierung. Hierzu zählen beispielsweise Großformate: Karten, Pläne, Urkunden und Plakate. Dabei handelt es sich übrigens auch um Bestände, bei denen es überdurchschnittlich häufig zu „Fehlbestellungen“ kommt, wenn z.B. für die Vorbereitung archivpädagogischer Veranstaltungen oder bei der Suche nach Abbildungen für Publikationen ganze Serien solcher Großformate – oft nur flüchtig – auf geeignete Objekte hin durchgesehen werden.

Gerade „Großformate“ öffnen auch den Blick für die Frage des richtigen Zeitpunkts einer Digitalisierung im Ablauf von Bestandserhaltungsmaßnahmen. Im Landesarchiv Nordrhein-Westfalen werden z.B. die laufenden Großprojekte zur Restaurierung und Umbettung der Urkunden aus den Abteilungen Rheinland und Westfalen mit der Erstellung von Schutzdigitalisaten verknüpft, wobei die Digitalisierung (von Vor- und Rückseiten der Urkunden und Siegel) sinnvollerweise im Stadium des Planlegens der Urkunden erfolgt, bevor die Urkunden in der Regel anschließend für die Lagerung in Archivsammelmappen und Archivboxen montiert und – soweit erforderlich – gefaltet werden. Die Frage, an welcher Stelle des Workflows Schutzmedien sinnvollerweise erstellt werden, ist selbstverständlich nicht auf den Sonderfall „Großformate“ beschränkt. Auch die bereits erwähnten Entnazifizierungsakten bedürfen auch einer konservatorischen Behandlung (einschließlich Entsäuerung). In diesem Fall liegt folgender Arbeitsablauf nahe, um im Zustand optimaler Lesbarkeit und ohne Beeinträchtigung durch die gerade bei Mengenverfahren letztlich nie ganz auszuschließenden „Nebenwirkungen“ bei der Entsäuerung wie z.B. ausblutende Tinten, das Schutzmedium zu erstellen:

1. konservatorische Arbeiten/Vorbereitung für die Digitalisierung und Entsäuerung wie Glätten, Entfernung von Schmutz, Schließen von Rissen, Folierung usw.
2. Digitalisierung (einschließlich Strukturdatenerfassung, Qualitätskontrolle, Erstellung von Master- und Arbeitsdateien bis hin zur Bereitstellung der Digitalisate im Lesesaal)
3. Entsäuerung, Verpackung/Umbettung
4. Ausbelichtung auf (s/w-)Mikrofilm als Sicherungsform

Digitalisiert werden des Weiteren im Landesarchiv NRW diejenigen Bestände, deren Erhaltungszustand eine Vorlage im Original nicht mehr erlaubt, etwa wegen stark abgebauter, fragiler Papiere z.B. durch Säurezerfall oder Schäden durch Mikroorganismen. Hier verfolgt das Landesarchiv den Weg: Originalerhalt durch die erforderlichen konservatorischen Maßnahmen, um das Archivgut sicher in geeigneten Magazinen zu lagern und nur noch für spezielle, ausschließlich am Original zu klärende Fragestellung vorzulegen, in Verbindung mit hochwertigen, akzeptierten Digitalisaten als Schutzmedien.

Dabei geht es um eine wichtige Stellschraube im Bestanderhaltungsmanagement, denn die konservatorisch-restauratorische Bearbeitung eines Bestandes für die Vorlage im Original als Regelnutzungsform ist im Normalfall weitaus zeitintensiver und ressourcen-trächtiger, als eine Bearbeitung, die die (einmalige) Reproduktion und die Originalnutzung nur noch im Ausnahmefall im Blick hat. Vergleichbar der Diskussion um flache Erschließung geht es also auch hier angesichts stets begrenzter Ressourcen um das Prinzip: „Nicht bei Allem alles, was möglich ist, sondern für möglichst viel Alles, was nötig ist“. Grundsätzlich auf einem hohen Standard der Restaurierung (Vollrestaurierung von Einzelobjekten) für alle in den Werkstätten behandelten Archivalien zu beharren, hieße in letzter Konsequenz, offenen Auges dem fortschreitenden Verfall anderer dringend behandlungsbedürftiger Stücke und ganzer Bestände untätig zuzusehen bzw. in Kauf zu nehmen, dass andere Objekte und Bestände auf

Jahr(zehnt)e hinaus unbenutzbar bleiben. Die Entscheidung über Prioritäten bei Bestandserhaltungsmaßnahmen und Behandlungstiefe ist insoweit immer auch ein Stück „zweite Bewertung“.⁹ Um hier nicht missverstanden zu werden: Es wird immer auch „Leuchtturm-Objekte“, Zimelien geben, bei denen eine Vollrestaurierung die angemessene Behandlungsart ist, und wo sich z.B. über „Patenschaften“ für einzelne Archivalien großartige Einzelobjektrestaurierungen realisieren lassen.

Ein Beispiel, bei dem wir das Verfahren weg von der Vollrestaurierung hin zu einer konservatorisch-restauratorische Basisbehandlung zur Vorbereitung einer anschließenden Digitalisierung umstellen, sind die so genannten „Kahnakten“. Es handelt sich dabei um jene ca. 20.000 Archivalien (etwa 3 Mio. Blatt), die 1945 beim Transport in ein Ausweichquartier mit der MS Main 68 durch Bombenabwürfe auf den Hafen Hannover-Linden teils in Brand gesetzt wurden, mit dem Schiff gesunken sind und erst Monate später wieder gehoben werden konnten.¹⁰ Bislang wurden die teils stark fragmentierten Akten nach der Lösung der Verblockungen und Trockenreinigung grundsätzlich einer aufwändigen Nassbehandlung unterzogen (Anfa-

⁹ Mario Glauert, Die zweite Bewertung. Prioritäten in der Bestandserhaltung, in: Für die Zukunft sichern! Bestandserhaltung analoger und digitaler Unterlagen. 78. Deutscher Archivtag 2008 in Erfurt (Tagungsdokumentation zum Deutschen Archivtag 13), Fulda 2009, S. 49-60.

¹⁰ Zu Hintergrund- und Restaurierungsgeschichte der Kahnakten bereitet der Verfasser eine Publikation vor, die 2010 in der Reihe der Veröffentlichungen des Landesarchivs Nordrhein-Westfalen erscheint.

serung, Stabilisierung mit Restaurierungsseidenpapier usw.), wobei durch die Nassbehandlung eine geringfügige weitere Verschlechterung der Lesbarkeit gelegentlich unvermeidlich ist. Zukünftig werden nur noch „Zimelien“ unter den „Kahnakten“ einer Vollrestaurierung zur Nutzung und Präsentation im Original unterzogen. Im Übrigen bleibt es auch hier überwiegend bei konservatorischen Arbeiten.

Derzeit wird das Vergabeverfahren für die Entwicklung und Teststellungen mit einem Hochleistungsscanner vorbereitet, der unter Einsatz von IR-, UV-, Streiflicht usw. beim Scanprozess selber einerseits sowie mittels Nachbereitung des digitalen Bildes andererseits zugleich eine spürbare Verbesserung der Lesbarkeit bringen soll.

Digitalisierungsprojekte durchzuführen heißt selbstverständlich nicht nur scannen. Eine Digitalisierungsstrategie, die die Langzeitsicherung nicht mitbedenkt, wäre letztlich hochgradig unwirtschaftlich (s. Abb. 2, nächste Seite).

Die Mikrofilme werden mit einer physikalischen Auflösung von 150 dpi (bezogen auf die Vorlage) gescannt und als JPEG-Dateien mit Komprimierungsstufe 100% gespeichert. Diese Images dienen als Master für die anschließenden Konvertierungsprozesse und werden abschließend auf LTO-Band gesichert. Sowohl die Auflösung wie die Formatwahl wurden im Rahmen von Tests ermittelt, bei denen sich zeigte, dass aufgrund der Vorlagenqualität weder eine höhere Auflösung noch die Speicherung in einem unkomprimierten Format (TIFF) zu besseren Ergebnissen führte.

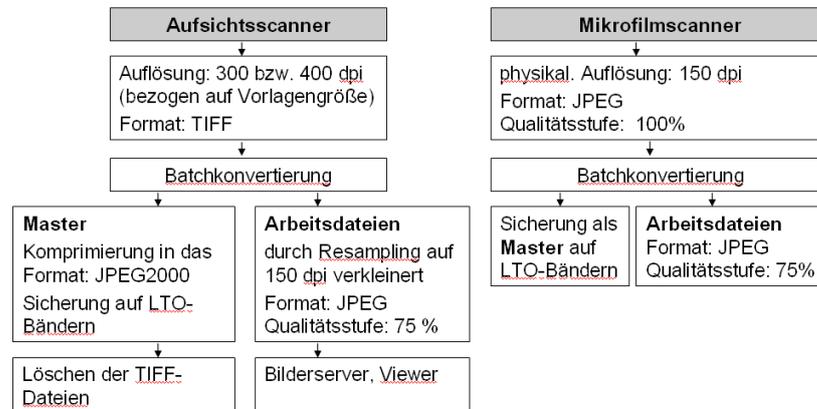


Abb. 2: Standardworkflow für die als Inhouse-Projekte durchgeführten Digitalisierungsmaßnahmen im Landesarchiv NRW

Bei der Aufsichtsdigitalisierung wird das Archivgut bei einer Auflösung von 300 dpi bei Akten (bis zum Format A1) bzw. 400 dpi bei Großformaten (A0) sowie Urkunden bezogen auf die Vorlagengröße gescannt und in unkomprimierten TIFF-Dateien abgelegt. Diese Rohdaten werden zunächst lokal auf der Festplatte gesichert. Durch Batchkonvertierung werden aus den Rohbilddateien einerseits onlinefähige, durch Resampling auf 150 dpi verkleinerte "Arbeitsdateien" im JPEG-Format in der Qualitätsstufe 75% erzeugt und auf dem Bilderserver abgelegt.¹¹ Von dort können die Digitalisate in den Lese-

¹¹ Die Batchverarbeitung im Umfeld der Konversion von Digitalisaten basiert auf einem einfachen Programm, mit dem in einer definierten Abfolge von Schritten Grafikdateien automatisiert in andere Formate umgewandelt werden. Resampling

sälen über den web-basierten Viewer geladen werden. Gleichzeitig werden aus den Rohbilddateien im Rahmen der Batchkonvertierung verlustfrei komprimierte Masterdateien im Format JPEG2000 erstellt, damit auf die Hälfte der Datenmenge im Vergleich zu unkomprimierten TIFF-Dateien reduziert, auf LTO-Magnetbändern gesichert und im Medienmagazin eingelagert. Die Rohdaten (TIFF-Dateien) werden gelöscht.

Die Verwendung des JPEG2000-Formats als verlustfreies Masterformat ist anders als im angelsächsischen Bereich im deutschsprachigen Raum eher untypisch, obwohl es sich, im Gegensatz zu TIFF, um einen ISO-Standard handelt.¹² Das Landesarchiv NRW hat sich insbesondere mit Blick auf den geringeren Speicherbedarf (30-50% Einsparung gegenüber unkomprimiertem TIFF!) bei gleicher Qualität für diesen Weg entschieden.

Bei der Erstellung der Digitalisate und der Batchkonvertierung werden zugleich die Metadaten erfasst bzw. generiert. Bei der Schutzdigitalisierung im Landesarchiv sind es schwerpunktmäßig technische Metadaten (Dateiname, -größe, Erstellungsdatum, Auflösung, Informationen zu Scanner und Software) neben einigen organi-

nennt man einen Vorgang, bei dem eine Grafikdatei mittels eines Programms durch interne Neuberechnung der Bildpunkte in ihrer Größe und/oder Farbtiefe verändert wird. Die unterschiedlichen Berechnungsmethoden versuchen, die zwangsläufig auftretenden Qualitätsverluste zu verringern.

¹² Paolo Buonora/Franco Liberati, A Format for Digital Preservation of Images. A Study on JPEG2000 File Robustness, in: D-Lib Magazine 14, 2008, 7/8 <<http://www.dlib.org/dlib/july08/buonora/07buonora.html>> .

satorischen Informationen wie der Abteilung, des Bestandsnamens und der Archivsignatur. Die Metadaten werden derzeit im XML-Format in einer gesonderten Datei abgelegt und durch das Metadaten-Extraktions-Tool der National Library of New Zealand (NLNZ), einer free-ware, ausgelesen. Die XML-Dateien werden gemeinsam mit den digitalen Masterdateien in einem Metadaten-Repository abgelegt. Teile der in XML vorliegenden Informationen werden zusätzlich in eine Datenbank übertragen, um Auswertungen zielgerichtet durchführen zu können.

Mit einer kürzlich entwickelten Erweiterung der Software V.E.R.A. können die Benutzer(innen) bald im Lesesaal nicht nur in elektronischen Findmitteln recherchieren sondern sich auch in sehr komfortabler Weise die dazu ggf. vorhandenen digitalen Archivalienabbildungen anzeigen lassen.¹³ Der neue web-basierte Viewer bietet den Benutzer(inne)n komfortable Anzeigoptionen, wie etwa Navigations-, Zoomify- und Drehfunktionen. Im Hinblick auf die Online-Stellung von Archivalienabbildungen wird mit einem der nächsten Entwicklungsschritte des Internetportals *www.archive.nrw.de* die Voraussetzung dafür geschaffen, die Präsentation von Archivalienabbildungen unter Nutzung des internationalen Standards METS¹⁴

¹³ Johannes Kistenich/Martina Wiech, Auf dem Weg zum elektronischen Landesarchiv?, in: Archive im digitalen Zeitalter. Tagungsdokumentation zum 79. Deutschen Archivtag in Regensburg (im Druck).

¹⁴ Zu METS vgl. die aktuelle Dokumentation auf der Website der Library of Congress unter <http://www.loc.gov/standards/mets> und die Materialien auf der Website der <daofind>-Projekte des Bundesarchivs unter <http://www.daofind.de>.

an die aktuellen technischen Möglichkeiten anzupassen und damit eine Plattform für die Verknüpfung und Nutzung der Digitalisate mit Metadaten und Erschließungsinformation, also die Verbindung von Informationsgehalt des Archivguts mit Informationen über das Archivgut, in anderen Portalzusammenhängen zu schaffen, z.B. im geplanten Archivportal D¹⁵ oder auch spartenübergreifend in der Europeana¹⁶, der Deutschen Digitalen Bibliothek¹⁷ oder in einem in Planung befindlichen Digitalen Archiv Nordrhein-Westfalen. Digitalisierung ist letztlich nur sinnvoll, wenn auch die Tektonik (Beständeübersicht) und Erschließungsinformationen digital vorliegen.

Mit der Bereitstellung von Digitalisaten analogen Archivguts sind viele Archive auf einem ambitionierten Weg, selbst zu Produzenten digitaler Unterlagen zu werden. Deren Authentizität muss in besonderer Weise gewährleistet werden: Wie am Beispiel der verblassten „Kahnakten“ erwähnt, kann wegen des vorhandenen Originals sogar bewusst gerade eine „nicht authentische“ Abweichung durch Scan-technik und Bildbearbeitung instrumentalisiert werden, um eine bessere Auswertbarkeit zu erreichen. Der standardmäßige Einsatz beispielsweise von Maßstab, Farbkarte und Grauskala bei der Digitali-

¹⁵ Vgl. dazu die Projektwebsite: <<http://www.archivgut-online.de>>.

¹⁶ <<http://www.europeana.eu>>. Vgl. dazu auch den Beitrag von Gerald Maier in diesem Band.

¹⁷ Gerald Maier, Der Aufbau einer "Deutschen Digitalen Bibliothek" und der "European Digital Library – Europeana": Ausgangslage und Perspektiven für die deutschen Archive, in: Der Archivar 61, 2008, S. 399-401.

sierung muss in diesen Fällen neben technischen Metadaten eine Referenzierung auf den Originalzustand gewährleisten.

Für die Langzeitverfügbarkeit von Digitalisaten stellen sich in mancher Hinsicht vergleichbare Herausforderungen, wie sie auch im Kontext von originär digitalen Unterlagen, sogenannten digital-born-documents, diskutiert werden. Freilich mit dem wesentlichen Unterschied, dass wir bei digitalisiertem Archivgut als Schutz- und Nutzungsmedium immer noch auf das weiterhin bestehende analoge Original zurückgreifen können, für dessen Erhalt langjährig bewährte Techniken zur Verfügung stehen und auf dessen Erhalt sich das Hauptaugenmerk richten muss.

Wollen wir nicht in einer zunehmend digitalen Welt aus der Perspektive unserer Kund(inn)en und auch aus der Sicht unserer Registraturbildner, für deren archivwürdige born-digitals wir verantwortlich zeichnen, zum „toten Archiv“ werden, bedarf es funktionierender Lösungen für die Langzeitverfügbarkeit von Digitalisaten. Hierbei kann und sollte der OAIS-Standard¹⁸ als Referenzmodell fungieren. Im Unterschied allerdings zu der Vielfalt an Dokumentenmanage-

¹⁸ Vgl. hierzu: Reference Model for an Open Archival Information System (OAIS), <<http://public.ccsds.org/publications/archive/650x0b1.pdf>>. In deutscher Zusammenfassung Nils Brübach, OAIS – Das „Open Archival Information System“. Ein Referenzmodell zur Organisation und Abwicklung der Archivierung digitaler Unterlagen <http://www.archiv.sachsen.de/download/pp_bruebach.pdf>. Vgl. zum Stellenwert des OAIS-Modells für die Frage der Langzeitverfügbarkeit von Digitalisaten und darüber hinaus zur Praxis im Landesarchiv NRW die 2009 bei der Archivschule Marburg eingereichte Transferarbeit von Danny Weber über „Langzeitarchivierung von Schutzdigitalisaten“.

mentssystemen, Fachverfahren, Datenbanken, geobasierten oder Web-Anwendungen bei den von uns betreuten Behörden, handelt es sich bei den von uns selbst angelegten Digitalisaten i.d.R. um übersichtlich strukturierte Daten, die meist bereits in Standardformaten vorliegen und mit zahlreichen Metadaten angereichert sind. Deshalb sollte die kontextualisierte Sicherung und Zugänglichkeit von Bilddateien eine durchaus handhabbare Herausforderung stellen. Die Archive müssen dennoch von Beginn eines jeden Digitalisierungsprojekts an darauf achten, die Standards selbstverantwortlich so zu definieren, dass die bestmöglichen Voraussetzungen für eine Langzeitverfügbarkeit geschaffen werden. Denn eines gilt es zu vermeiden: Dass die Schutzdigitalisierung als Maßnahme der Bestandserhaltung analogen Archivguts zu unlösbaren Problemen oder unwirtschaftlichen Lösungen beim Erhalt der digitalen Schutzmedien führt, damit am Schluss nicht der Aufwand für die Erhaltung authentischer Schutzmedien zulasten des Authentischen selbst, also der Mittel für die Bestandserhaltung der analogen Originale, erfolgt.

Das virtuelle Urkundenarchiv „Monasterium“ – bayerische Urkunden im internationalen Kontext¹

von Joachim Kemper

1. Einleitung: Digitale Urkundenpräsentationen

In diesem Beitrag steht eine spezielle Präsentationsform von Archivalien und historischen Quellen im Mittelpunkt: digitale Urkundenpräsentationen bzw. Urkundenportale. Solche Präsentationen sind einerseits eine Aufgabe der Archive, aber auch der Anteil der mediävistischen Grundlagenforschung und der Historischen Hilfswissenschaften (Stichwort „Digitale Diplomatie“) an der Materie ist nicht zu unterschätzen. Dies sollte insbesondere im Hinblick auf mögliche Kooperationen und Synergieeffekte bedacht werden.

Im Zentrum des vorliegenden Textes steht ein „bayerischer“ Blick auf das derzeit größte Urkundenportal <<http://www.monasterium.net>>. Mehrere Projekte der staatlichen Archive Bayerns zur Urkundendigitalisierung sollen beschrieben werden, ebenso

¹ Geringfügig überarbeitete und mit Anmerkungen versehene Fassung des Vortrags am 2.12.2009 (Titel: „Das internationale virtuelle Urkundenarchiv Monasterium“). Der Vortrag basierte in wesentlichen Teilen auf den Ausführungen des Verfassers auf der Fortbildungsveranstaltung „Digitale Urkundenpräsentationen im Vergleich“ (79. Deutscher Archivtag Regensburg, 22.9.2009). Allen Teilnehmern dieses Workshops sei an dieser Stelle nochmals für die rege Diskussion gedankt; ebenso ein herzlicher Dank an die Mitorganisatorin Dr. Maria Magdalena Rückert (Landesarchiv Baden-Württemberg) sowie an die beiden Gastreferenten Redmer Alma (Drents Archief, Niederlande) und Mag. Thomas Just (Haus-, Hof- und Staatsarchiv Wien).

aber auch der gegenwärtige Stand des gesamten Urkundenportals. Abschließend erfolgt ein kleiner Blick in die Zukunft.

Die Urkunden des Mittelalters und der Frühen Neuzeit sind – dies muss nicht eigens betont werden – eine der herausragenden Quellengruppen zur Erforschung der deutschen und europäischen Geschichte. Urkunden stellen insbesondere für das Mittelalter eine zentrale Quellengruppe dar. Sie enthalten inhaltlich gesehen ein breites Spektrum von Rechtsgeschäften aller Art (Verträge, Verkäufe, Privilegien, Belehnungen – um nur ein paar Beispiele zu nennen). Sie liefern weiterhin zentrale Informationen über politische Aktivitäten, über das Recht, die Verfassung, das kirchliche Leben oder können als Quelle zur Wirtschaftsgeschichte dienen. Die Diplomatik als Teil der Historischen Hilfswissenschaften beschäftigt sich als Spezialwissenschaft mit den Urkunden, wobei in den letzten Jahren digitale Arbeitsmethoden und Editionsformen immer mehr an Bedeutung gewonnen haben.

Was die Verwendung für Onlinepräsentationen angeht, eignen sich Urkunden im Vergleich zu den in den Archiven verwahrten Akten und Amtsbüchern ausgesprochen gut für Digitalisierungsprojekte, da es sich um Einzelstücke und nicht um vielblättrige Stücke handelt. Für die Präsentation werden lediglich Images von Vorder- und Rückseite sowie eventuell der Siegel benötigt. Die Vorderseite enthält in der Regel die relevanten Informationen, den Urkundentext. Hinzu kommen Metadaten, bekanntlich im besten Fall ein Vollregist

oder abgestufte Varianten („Archivregest“,² Kurzregest). Nach Meinung des Verfassers ist ein Kurzregest völlig ausreichend – namentlich, wenn dieses mit der Abbildung der Urkunde online zur Verfügung steht – und es ist zeitökonomisch effektiv.³

Mit welchen Zahlen hat man es zu tun, wie groß ist die Menge der in den deutschen Archiven überlieferten Urkunden? Allgemein gültige Zahlen liegen praktisch nicht vor. Für Deutschland kann man grob geschätzt von vielleicht 3 bis 4 Millionen Stücken ausgehen. Dieser Annäherungswert beruht auf einer Hochrechnung, die auf den Zahlen in den staatlichen Archiven in Bayern, Baden-Württemberg und Rheinland-Pfalz basiert. In den staatlichen Archiven Bayerns zum Beispiel werden allein knapp 600.000 Urkunden vom frühen Mittelalter bis zum Beginn des 19. Jahrhunderts verwahrt. Die angenommene Zahl von 3 bis 4 Millionen Urkunden ist einerseits – nimmt man die bibliothekarische, den Umgang mit Massendigitalisierungsvorhaben gewohnte Perspektive ein – keine „unbezwingbare“ Menge. Dennoch wird derzeit kaum ein Archivar der Digitalisierung eines erheblichen Teils oder gar des Gesamtbestands der deutschen Ur-

² Besonders in Archiven findet vielfach das gegenüber dem Vollregest leicht gekürzte, jedoch selten normativ geregelte sog. „Archivregest“ Verwendung. Generell sind die Überlegungen hinsichtlich einer angemessen-zeitgemäßen Anfertigung von Regesten in den letzten Jahren wieder stärker geworden; dies gilt insbesondere für Kürzungsmöglichkeiten.

³ Vgl. Joachim Kemper, Archivische Urkundenregesten im Kontext der Digitalisierung, in: Tehnici in vsebinski problemi klasicnega in elektronskega arhiviranja (Technical and field related problems of traditional and electronic archiving), Maribor 2009, S. 473-482; Ders., Neue Richtlinien der staatlichen Archive Bayerns für die Erstellung von Urkundenregesten, in: Archivalische Zeitschrift 91, 2009 (im Druck).

kunden Priorität einräumen. Die deutschen Archive arbeiten gerade mit Nachdruck und unterstützt von der Deutschen Forschungsgemeinschaft (DFG) daran, die Ebene unterhalb der Digitalisate – die Findmittel – für das Internet aufzubereiten. Bezüglich der Digitalisierung von Archivgut kann als *communis opinio* gelten, dass überhaupt nur ein kleiner Prozentsatz des Archivguts realistisch online gestellt werden kann. Dass es sich bei den in den Archiven verwahrten Urkundenbeständen zweifellos oft um die wertvollsten Archivalien des jeweiligen Hauses handelt und diese nur einen geringen Prozentsatz des gesamten Archivgutes bilden, sei an dieser Stelle ausdrücklich erwähnt.

Im Folgenden werden mehrere Beispiele für digitale Urkundenpräsentationen aus dem In- und Ausland kurz vorgestellt. Es wird sich dabei lediglich um eine (kleine) Auswahl handeln. Bei den Beispielen erfolgt – sieht man von den *Regesta Imperii* ab – eine Beschränkung auf die Verbindung von Online-Regesten mit digitalen Abbildungen der Urkunden. Die „einfache“ Onlinepublikation von Regesten, d.h. ohne Abbildungen der Urkunden, ist noch weiter verbreitet, denkt man an die im Internet präsentierten Regestenrepertorien von Archiven oder an jüngere Regestenpublikationen aus der Forschung.⁴

Die *Regesta Imperii* zählen bis heute als Inventar aller urkundlichen Quellen der deutschen Kaiser und Könige von den Karolingern bis zum Beginn des 16. Jahrhunderts zu den großen mittelalterlichen

⁴ Vgl. insgesamt den guten, freilich v.a. den archivischen Bereich noch unzureichend berücksichtigenden Überblick auf der „VL Geschichte“ (Historische Hilfswissenschaften, Urkundeneditionen, Regesten etc.): <<http://www.vl-ghw.uni-muenchen.de/chartularia.html>> [11.12.2009].

Quellenwerken. Zugleich sind sie für den deutschen Sprachraum ein einschlägiges Referenzunternehmen im Bereich der Regestenerstellung. Wie andere Editionsprojekte auch, sahen sich die Regesta Imperii schon bald mit dem Problem der Quantität der spätmittelalterlichen Überlieferung konfrontiert: existiert aus dem Früh- und Hochmittelalter bekanntlich eine zumeist überschaubare Anzahl an Herrscherurkunden, so nimmt ihre Zahl im 14. und 15. Jahrhundert enorm zu. Die Beurkundungs- bzw. Korrespondenzaktivität des letzten bearbeiteten Herrschers, Maximilian I. (1486/1493-1519), kann allein auf über 100.000 Urkunden und Briefe geschätzt werden. Aufgrund dieser Problematik sind die Regesta Imperii bereits vor knapp drei Jahrzehnten z.T. zu einer sukzessiv-bandweisen Publikation der Regesten übergegangen. Beispielsweise erscheinen die Regesten Kaiser Friedrichs III. (1440-1493) geordnet nach regionalen Gesichtspunkten oder gar archivweise. Die Publikation erfolgt nicht erst nach vielen Jahrzehnten der Sammlung und Aufarbeitung, sondern in regelmäßigen Abständen.⁵ Vergleichsweise früh haben die Regesta Imperii auch damit begonnen, ihre bislang publizierten Regestenbände zu digitalisieren. Zunächst wurde 1998 eine CD-R mit den ersten zehn Bänden der Regesten Friedrichs III. veröffentlicht, dann erfolgte gemeinsam mit der Bayerischen Staatsbibliothek die Digitalisierung des gesamten Altbestandes samt Bereitstellung im Internet („Regesta Imperii Online“)⁶. Sämtliche Bände der Regesta Imperii

⁵ Seit dem Jahr 1982 sind auf diese Weise bereits weit über 20, teils sehr umfangreiche „Hefte“ der Regesten Friedrichs III. erschienen, dazu etliche Sonderbände bzw. „Supplemente“ sowie eine CD-ROM.

⁶ <<http://www.regesta-imperii.de>> [11.12.2009].

sind jetzt übergreifend durchsuchbar. Noch nicht abschließend bearbeitete Regestenbände, Ergänzungen und Arbeitsdatenbanken werden als „work in progress“ auf der Homepage vorab schon zur Verfügung gestellt. Die Verlinkung der Regesten mit anderen digitalen Präsentationen ist geplant bzw. bereits erfolgt, z.B. mit der „Online-Edition der Urkunden der Reichsabtei Fulda“. Bezüglich des Portals „Monasterium“ ist ähnliches vorgesehen, ergänzt um eine tiefer gehende Kooperation mit einem geplanten Projekt der staatlichen Archive Bayerns. Derzeit kann man über die Internetseite der Regesta Imperii in ca. 150.000 Urkunden recherchieren.

Zu den digitalen Pilotprojekten im deutschen Archivbereich zählt insbesondere das *digitale Archiv von Duderstadt*.⁷ Im Stadtarchiv Duderstadt (Niedersachsen) waren zwischen 1996 und 1999 in Zusammenarbeit mit dem Max-Planck-Institut für Geschichte die älteren Archivalien (bis 1650) digitalisiert und erschlossen worden. Über 77.000 Archivalienseiten sind seitdem im Internet frei zugänglich. Darunter befinden sich auch sämtliche, durch kurze Regesten erschlossene Urkunden (ca. 1.000 Stücke).

Zu den älteren Projekten zählen ebenfalls die *Fontes civitatis Ratisponensis* (FCR-Online)⁸. Die FCR-Online sind ein Kooperationsprojekt des Stadtarchivs Regensburg mit der Karl-Franzens-Universität Graz (Institut für Geschichte, Abteilung für Historische Fachinformatik und Dokumentation). Erklärtes Ziel ist insbesondere, Editionen wichtiger Regensburger Archivalien und Amtsbücher online zur Ver-

⁷ <<http://www.archive.geschichte.mpg.de/duderstadt>> [12.12.2009].

⁸ <<http://www.fcr-online.com>> [12.12.2009].

fügung zu stellen (z.B. Bürgeraufnahmebücher, Rechnungsbücher und Urkunden der Stadt Regensburg sowie das sogenannte „Schwarze Stadtbuch“). Ein weiterer Schwerpunkt ist das „Virtuelle Archiv“, das unabhängig vom Lagerort alle reichsstädtisch-regensburgischen Archivalien bis zum Beginn des 19. Jahrhunderts nachweisen soll. Es enthält derzeit nach eigenen Angaben über 1.100 Archivalien und annähernd 17.000 Images.

Mittlerweile hat auch das altherwürdige *Marburger Lichtbildarchiv älterer Originalurkunden bis 1250* (LBA) als Schnittstelle zwischen hilfswissenschaftlichen und archivischen Vorhaben über ein DFG-Projekt den Weg ins Internetzeitalter gefunden: die Urkundendatenbank sowie z.T. auch schon die digitalisierten Fotobestände des Instituts stehen online zur Verfügung.⁹ Eine Verknüpfung mit anderen Urkundenpräsentationen dürfte aufgrund des umfassenden Sammelgebiets des LBA (mit ca. 15.000 Urkundenfotos aus Deutschland und der Schweiz) mehr als sinnvoll und für alle Seiten gewinnbringend sein. Dies gilt auch für die eventuell nötige Mithilfe bei der Recherche nach heutigen, geänderten Lagerorten bzw. Signaturen der Urkunden (z.B. gehäuft in Bayern aufgrund von Beständeberichtigungen).

Ein weiteres, aktuell laufendes Projekt kommt ebenfalls aus Marburg: im Rahmen eines DFG-Projekts werden im Staatsarchiv Marburg derzeit die *Urkunden der Reichsabtei Fulda* ausführlich erschlossen und online samt Digitalisaten der Urkunden zur Verfügung

⁹ <<http://www.uni-marburg.de/fb06/mag/lba/df>> [12.12.2009].

gestellt.¹⁰ Die sukzessive Verfügbarmachung über das Hessische Archiv-, Dokumentations- und Informationssystem HADIS erfolgt bereits.¹¹ Bisher sind die Urkunden bis 1500 registriert und z.T. mit den Online-Regesten der Regesta Imperii verknüpft.

Von den Online-Projekten des Landesarchivs Baden-Württemberg sollen an dieser Stelle das *Württembergische Urkundenbuch Online* (WUB-Online) sowie die Internetversion des altwürttembergischen urkundlichen Kernbestands *Württembergische Regesten* genannt werden.¹² Während es sich beim WUB-Online um eine verbesserte und ergänzte Version des alten „Württembergischen Urkundenbuchs“ (1849/1913) handelt, bieten die zeitlich anschließenden „Regesten“ zusätzlich die Images mehrerer Tausend Urkunden. Beide Angebote zusammen stellen eine hervorragende Arbeitsplattform für die südwestdeutsche Landesgeschichte dar.

Zum Abschluss ein kleiner Blick in das benachbarte Ausland. Auch unabhängig von „Monasterium“ (bzw. den daran beteiligten Staaten) zeigt sich, dass teils erhebliche Anstrengungen unternommen werden, um das urkundliche Kulturerbe per Internet zugänglich zu machen. Ein eindrucksvolles niederländisches Projekt ist das *digitale Urkundenbuch der Provinzen Groningen und Drente*: zu beiden Regionen steht jetzt bis zum „Stichjahr“ 1600 ein umfangreicher Bestand von ca. 34.000 Urkunden und anderen Schriftstücken on-

¹⁰ <http://www.kulturerbe-digital.de/de/projekte/9_38_31303438.php> [12.12.2009].

¹¹ <<http://www.hadis.hessen.de>> [12.1.2010].

¹² <<http://www.landesarchiv-bw.de/web/4940>> [12.12.2009].

line zur Verfügung.¹³ An italienischen Beispielen seien ein größeres Projekt aus dem Trentino (*Pergamene online*) und die Urkundenpräsentation des Staatsarchivs Florenz genannt.¹⁴

2. Digitale Urkundenprojekte der staatlichen Archive Bayerns

Die staatlichen Archive Bayerns haben die Stufe der Onlinefindmittel vor wenigen Jahren zunächst „übersprungen“:¹⁵ seit 2005 wurden verstärkt Überlegungen angestellt, Urkunden und andere ausgewählte Archivalien in Kooperation mit anderen Einrichtungen oder Portalen zu digitalisieren und online bereitzustellen. Für größere Digitalisierungsaktionen waren keine eigenen Finanzmittel vorhanden und auch heute läuft ein erheblicher Teil der Projekte über Drittmittel.

Gemeinsam mit der Bayerischen Staatsbibliothek wurden 2006 und 2007 bedeutende *Freisinger Amtsbücher* des Mittelalters digitalisiert.¹⁶ Sie stehen seitdem im Rahmen der „Bayerischen Landesbibliothek Online“ im Netz zur Verfügung und bilden dort einen zentralen Zimelienbestand.

¹³ <<http://www.cartago.nl>> [12.12.2009].

¹⁴ <http://www.trentinocultura.net/catalogo/cat_fondi_arch/pergamene/cat_pergamene_DE_h.asp; <http://www.archiviodistato.firenze.it/nuovosito/index.php?id=369>> [12.12.2009].

¹⁵ Seit 2008 wird online unter <<http://www.gda.bayern.de>> eine immer größere Menge von Findmitteln im PDF-Format angeboten; diese sollen in Bälde durch eine umfassende datenbankgestützte Recherchemöglichkeit ersetzt werden.

¹⁶ <<http://www.bayerische-landesbibliothek-online.de/hauptstaatsarchiv>> [12.12.2009].

Als wichtiger noch hat sich das Konzept für eine umfassendere Digitalisierung der Urkunden des Bayerischen Hauptstaatsarchivs erwiesen. Das Bayerische Hauptstaatsarchiv ist bekanntlich mit insgesamt ca. 280.000 Urkunden eines der großen Urkundenarchive Deutschlands. Die Bedeutung der grenzüberschreitenden Überlieferung gegenüber Österreich ist besonders groß, weswegen der Weg nicht weit war, gemeinsam mit dem Partner „Monasterium“ Digitalisierungsprojekte in Angriff zu nehmen. „Monasterium“ war zunächst als Plattform vorgesehen, daneben aber auch als Dienstleister für die Digitalisierung und die technische Umsetzung. Der letzte Punkt liegt mittlerweile zu einem Teil schon in den Händen der staatlichen Archive (was weiter ausgebaut werden wird). Natürlich sollen die Urkunden bald genauso über die Homepage der staatlichen Archive abrufbar sein, wie über das virtuelle Urkundenarchiv „Monasterium“.

Zur Chronologie: nach einer ersten Vereinbarung mit „Monasterium“ über die (kostenfreie) Digitalisierung einer überschaubaren Zahl von Urkundenbeständen wurde bei der DFG im Rahmen des Programms „Wissenschaftliche Literaturversorgungs- und Informationssysteme“ (LIS) ein relativ ambitionierter und umfassender Antrag mit dem Titel „Aufbau eines internetbasierten Portals für größere Bestände von digitalisierten Urkunden des süddeutschen Raumes“ gestellt. Das Projekt (Kurztitel: „Urkundenportal“) wurde mit der Laufzeit vom 1. Januar 2008 bis 31. Dezember 2009 bzw. 31. März 2010 genehmigt. Ziel des Projektes war es, ein bayerisches Urkundenportal innerhalb von „Monasterium“ aufzubauen, ohne freilich dessen internationalen Gesamtkontext zu verlassen. Die DFG hatte Sachmittel für die Digitalisierung von Urkunden und die technische Entwicklung bereitgestellt, ebenso Geldmittel für zwei Mitarbeiter

sowie zwei studentische Hilfskräfte. Zur Erreichung der Projektziele wurden 2008 in knapp sechs Monaten 25.000 Urkunden geistlicher Provenienz aus dem südostbayerischen Raum digitalisiert (zusammen mit den bereits zuvor digitalisierten Stücken waren es ca. 30.000 Urkunden). Im Rahmen der technischen Arbeiten wurde nicht nur das bestehende Portal „Monasterium“ modifiziert und den Projektbedürfnissen angepasst, sondern auch ein „Länderportal Bayern“ entworfen, das jetzt neben weiteren Länderportalen steht (Ober- und Niederösterreich, seit neuestem auch Italien). Das Länderportal enthält einführende Texte und Übersichten zu den Projekten in Bayern sowie zur Finanzierung. Mittels entsprechender Logos und Schriftzüge („Das virtuelle Urkundenarchiv Bayerns“) ist deutlich zu erkennen, wer hinter dem Projekt steht. Eine komplette Liste der noch zu bearbeitenden und bereits online stehenden Bestände ist gegeben, ebenso selbstverständlich der Zugang über eine Recherchefunktion. Natürlich kann man auch auf dem „üblichen“ Weg auf die bayerischen Bestände innerhalb des Gesamtportals zugreifen.¹⁷

Der Stand der Urkundenerschließung im Bayerischen Hauptstaatsarchiv ist nicht besonders gut. Für einen Großteil der Bestände liegen lediglich kombinierte Aussteller- und Sieglerlisten vor. Zwar wurden natürlich für das Projekt „Urkundenportal“ auch Bestände ausgewählt, bei denen hand- oder maschinenschriftliche Findmittel vorhanden waren oder bereits in der archivischen Fachdatenbank (FAUST) Regesten vorlagen. Dennoch kann der Anteil der fast nicht

¹⁷ <http://www.monasterium.net/ieekq/de/portal_bayern00_de.php; <http://www.monasterium-bayern.net>> [12.12.2009].

erschlossenen Stücke auf knapp 20.000 (von 30.000) Urkunden geschätzt werden. Von der bislang üblichen Verzeichnung mittels Voll- oder Archivregesten musste natürlich abgesehen werden. Ziel war es, für den größten Teil der Urkunden während der Projektlaufzeit Kurzregesten oder wenigstens Betreffe bzw. Stichworte anzufertigen. Gemeinsam mit den vorhandenen „Metadaten“ (wie Datierung, Angaben zu Aussteller und Siegler) sowie den Images ist dem Benutzer geholfen und ein erheblicher Mehrwert erzielt. Spätere Verbesserungen der Erschließung sind ohne weiteres machbar und auch von „außen“, d.h. über das Internet, sind Bearbeitungen möglich.

Innerhalb des Projekts wurden zunächst Bestände bearbeitet, zu denen bereits brauchbare Erschließungsinformationen vorlagen (z.B. neuere maschinenschriftliche oder datenbankmäßig aufbereitete Regesten). Die Onlinestellung konnte dementsprechend relativ rasch erfolgen, was wiederum für die Öffentlichkeitswahrnehmung des Projekts positiv war. Die Priorität lag von Anfang an darauf, das Projekt auch in der Region bekannt zu machen: sei es durch kleinere oder größere Veranstaltungen („Präsentationen“), aber ebenso durch eine gedruckte Projektbroschüre.¹⁸ Entsprechend trafen und treffen immer wieder Rückmeldungen, aber auch konkrete Korrekturvorschläge zu den Beständen bzw. zu einzelnen Urkunden ein. Selbst die Finanzierung von Digitalisierungsarbeiten an einem Klosterbestand (Kloster Rohr, Niederbayern) konnte jetzt über „Mund zu Mund-Propaganda“ und das Engagement der entsprechenden Gemeinde erreicht werden.

¹⁸ Zusätzlich als Download zu erhalten unter http://www.monasterium.net/ieekq/de/doc/Folder_Bayern.pdf [12.12.2009].

Durch zwei weitere, derzeit laufende Projekte wird ebenfalls die Digitalisierung und Onlinestellung von Urkunden vorangetrieben (die Präsentation erfolgt über „Monasterium“). Die Projekte sollen im Folgenden knapp vorgestellt werden.¹⁹

Gemeinsam mit der Bayerischen Staatsbibliothek betreibt das Bayerische Hauptstaatsarchiv seit dem Herbst 2008 ein DFG-Projekt zum Thema „Schriftlichkeit in süddeutschen Frauenklöstern“ innerhalb der Förderlinie „Archive und Bibliotheken im Verbund mit der Forschung“. Ziel ist es, die Bibliotheken und Archive ausgewählter bayerischer Frauenklöster virtuell wiederherzustellen. Die am Vorhaben beteiligte Forschergruppe (Leitung: Prof. Dr. Eva Schlotheuber, Münster) ist dem Ansatz der vergleichenden Ordensgeschichte verpflichtet. Das Projekt endet bezüglich der Erschließungsarbeiten mit dem Beginn der Frühen Neuzeit. Dies ist aus archivischer Sicht durchaus problematisch, denn es können nur die Urkunden und Amtsbücher der Klöster bis zum Stichjahr 1550 bearbeitet werden. Immerhin werden die Urkundenbestände im Internet mit sämtlichen Images präsentiert. Ausgewählte Amtsbücher werden die urkundliche Überlieferung ergänzen.

Durch Mittel aus dem EU-Projekt „Charters network“ werden im Bayerischen Hauptstaatsarchiv seit Dezember 2009 die Urkunden des Hochstifts und Domkapitels Freising digitalisiert. Der für die altbayerische Geschichte bedeutende, teilweise bereits durch gute Regesten erschlossene Bestand umfasst ca. 8.000 Urkunden und wird die bislang digitalisierten geistlichen Urkundenbestände gut

¹⁹ <<http://www.gda.bayern.de/projekte/index.php>> [12.12.2009].

ergänzen. Die Adaptierung und Ergänzung der Regesten wird derzeit durch einen Mitarbeiter des Bayerischen Hauptstaatsarchivs vorgenommen. Das Gesamtprojekt ist ein Gemeinschaftsvorhaben von Archiven aus Tschechien, der Slowakei, Slowenien und Ungarn – Deutschland ist über das Bayerische Hauptstaatsarchiv beteiligt, Österreich durch das Österreichische Staatsarchiv sowie das Diözesanarchiv St. Pölten und das „International center for archival research“ (ICARUS, Wien).

3. Das virtuelle Urkundenarchiv „Monasterium“²⁰

Der seit etwas mehr als eineinhalb Jahren existierende Verein ICARUS kann als Nachfolger des Monasterium-Konsortiums bezeichnet werden, in dem sich die an „Monasterium“ Beteiligten regelmäßig getroffen haben. ICARUS gehören derzeit Mitglieder aus zehn Staaten an. Der Verein verfolgt insgesamt mehrere Projekte, die nicht alle mit dem sicher bedeutendsten Standbein „Monasterium“ zu tun haben (z.B. das Kirchenbuchprojekt „Matricula“). Aus Sicht des Verfassers ist aber explizit hervorzuheben, dass „Monasterium“ jetzt deutlich institutionalisierter als noch vor einigen Jahren ist: für alle relevanten Bereiche bzw. Arbeitsgebiete (Metadaten, Digitalisierung, Technik, Wissenschaft/Projekte, Administration) sind jetzt feste Ansprechpartner vorhanden – das Diözesanarchiv St. Pölten ist nur mehr ein (freilich wichtiger) Teil des Ganzen. Sämtliche Staaten sind im ICARUS-Vorstand durch Vizepräsidenten vertreten.²¹

²⁰ <<http://www.monasterium.net>> [12.1.2010].

²¹ <<http://www.icar-us.eu>> [12.12.2009].

Das Angebot des Portals „Monasterium“ umfasst derzeit einen Urkundenbestand von über 120.000 Urkunden aus neun Staaten (Deutschland/Bayern, Italien, Österreich, Schweiz, Serbien, Slowakei, Slowenien, Tschechien und Ungarn; in Bälde vermutlich auch: Kroatien und Spanien). Die Urkunden sind natürlich kostenfrei online zugänglich. Zu den zentralen Funktionalitäten des Portals zählen einerseits der Bereich der Ansicht und Recherche, andererseits auch zunehmend ein Tool für eine „kollaborative“ Bearbeitung der Urkunden. Der Benutzer kann in den Beständen übergreifend navigieren und recherchieren. Die Urkundenbestände folgen in ihrer Strukturierung weitestgehend dem Archivprinzip. Der Zugriff auf das Einzelstück kann demnach über die Tektonik (Staat-Region/Bundesland-Archiv-Bestand) oder aber über Suchmasken erfolgen. Der Suchmodus umfasst eine „Einfache Suche“ und eine „Erweiterte Suche“, bei der bis zu vier Suchparameter miteinander kombiniert werden können. Die Darstellung eines Einzelobjekts erfolgt in Wort und Bild (Image der Urkunde). Im oberen Bildschirmteil wird die digitale Abbildung gezeigt (Vorder- und Rückseite), welche stark vergrößert werden kann. Im unteren Bildschirmteil werden die vorhandenen Metadaten angezeigt. Deren Form kann sehr unterschiedlich sein: von knappsten Angaben bis hin zu modernen Kurz- und Vollregesten bzw. sogar Transkriptionen finden sich viele Varianten – das Portal ist in dieser Hinsicht sehr offen (und muss dies auch sein). Das „kollaborative“ Tool ermöglicht zusätzlich die Bearbeitung und Ergänzung des bestehenden Datenbestandes. Die Qualitätssicherung erfolgt über ein Moderatorensystem. Jeder Bearbeiter muss sich bei seiner erstmaligen Anmeldung einen Moderator wählen (z.B. des die Urkunde verwahrenden Archivs), welcher nach Beendigung der Arbeit

das Ergebnis überprüft. Erst danach erfolgt die endgültige Freischaltung der Neuerung. Als gut geeignet hat sich das Tool übrigens auch in der akademischen Lehre erwiesen.

Aus Deutschland sind bislang nur bayerische Archive beteiligt, darunter das Bayerische Hauptstaatsarchiv mit derzeit etwas weniger als 20.000 Urkunden. Das Hauptstaatsarchiv ist somit der größte inhaltliche Beiträger („content provider“) zum Urkundenarchiv. Die Urkunden stammen derzeit weitgehend aus dem Projekt „Urkundenportal“. Zum Sommer 2010 sollen aus dem dann abgeschlossenen Projekt knapp 30.000 Urkunden im Netz sein, bis zum Jahr 2011 soll sich diese Zahl auf ca. 40.000 Stücke erhöhen. Von anderen bayerischen Staatsarchiven sind bislang lediglich Urkundenregesten online verfügbar, doch sollen sukzessive auch ergänzend Abbildungen online gestellt werden. Insbesondere das Staatsarchiv Würzburg ist mit mehreren Tausend neu erstellten Urkundenregesten von Klöstern und Stiften seines Zuständigkeitsbereiches am Portal beteiligt.

4. Ausblick

Bevor ein abschließender Ausblick gegeben wird, sollen noch aus der Projektpraxis heraus einige, zweifellos symptomatische „Nebengeräusche“ und „Nebenprodukte“ bei der Arbeit an einer Online-Urkundenpräsentation umrissen werden: zu Beginn wurde die Tätigkeit der Projektmitarbeiter oft fehlgedeutet. Man vermutete, diese würden in klassischer Manier die Urkunden durch Vollregesten erschließen, um sie dann ausgedruckt lediglich im Repertorienzimmer bereitzustellen. Die Digitalisierung wurde als Mittel begriffen, um (vor dem Hintergrund von Gebühreneinnahmen) einen bequemen Zugriff auf die Digitalisate zu erhalten. Kurz gesagt: Die Mitarbeiter

hatten anfangs einen schweren Stand in der Erklärung, was sie eigentlich und warum sie dies machen. Trotz der Beschränkung auf Kurzregesten wurde und wird der Erschließungsstand erheblich verbessert – und ist sowohl im Internet als auch selbstverständlich im Repertorienzimmer sichtbar. Aus der Projektarbeit heraus sind neue Regestenrichtlinien entstanden, die für die staatlichen Archive Bayerns verbindlich gemacht worden sind. Sie sehen Kurzregesten als Standard vor, Vollregesten oder auch sog. „Archivregesten“ sind nur ausnahmsweise gestattet.²²

Für die zukünftige Etablierung einer als umfassend konzipierten Digitalisierung der deutschen Urkunden benötigt man eine belastbare Anzahl erfolgreich durchgeführter Projekte, um aufzeigen zu können, was, in welcher Zeit sowie mit welchem Personal- und Kostenaufwand machbar ist. Benötigt wird ein überregional bekanntes, bereits bestehendes Portal, auf das exemplarisch zu verweisen ist und in das die eigenen Bestände einzubringen sind: „Monasterium“ ist mit mehr als 120.000 Urkunden das mit weitem Abstand größte Internet-Urkundenarchiv. Mit „Monasterium“ können ohne weiteres andere Angebote oder digitale Präsentationen verknüpft werden, ohne dass diese ihre Eigenständigkeit verlieren. Schließlich wird auch ein „Masterplan“ nötig sein, ein Plan für ein zukünftiges virtuelles Urkundenetzwerk. Hier ist vielleicht in naher Zukunft ein Projekt zu erwarten, in dem einerseits prototypische Workflows entwickelt werden, die die sukzessive, langfristige, verteilte und kooperative Digitalisierung

²² <http://www.gda.bayern.de/download/erschliessung/richtlinien_regesten.pdf> [12.12.2009].

eines erheblichen Teils des deutschen Urkundenbestandes organisierbar machen. Andererseits soll ein fortschrittliches technisches Instrumentarium geschaffen werden, das als Online-Arbeitsumgebung dient. Diese soll durch die historische Forschung im Projekt getestet werden.

Doch ist dies noch Zukunftsmusik, auch wenn die Aussicht, in ein paar Jahren oder Jahrzehnten einen beachtlichen Teil der in den deutschen Archiven verwahrten Urkunden im Internet greifbar zu machen, so unrealistisch nicht ist. Vor allem ist es eine reizvolle Aussicht.

Autorenverzeichnis

Paul Bantzer
Startext GmbH
Kennedyallee 2
D – 53175 Bonn

Dr. Frank M. Bischoff
Landesarchiv Nordrhein-
Westfalen, Abteilung Rheinland
Mauerstraße 55
D – 40476 Düsseldorf

Dr. Nils Brübach
Sächsisches Staatsarchiv
Marienallee 12
D – 01099 Dresden

Blanca Desantes
Subdirección General de los Ar-
chivos Estatales de España
Plaza del rey 1
E – 28004 Madrid

Dr. Ulrich Fischer
Histor. Archiv der Stadt Köln
Willy-Brandt-Platz 2
D – 50679 Köln

Michael Fox
Minnesota Historical Society
345 Kellogg Blvd West
USA – Saint Paul, MN 55102

Dr. Jeanette Godau
Wirtschaftsarchiv Baden-
Württemberg
Schloss Hohenheim
D – 70593 Stuttgart

Jan Jäckel
Koordinierungsstelle
Retrokonversion
Bismarckstraße 32
D – 35037 Marburg

Dr. Joachim Kemper
Staatsarchiv München
Schönfeldstraße 3
D – 80539 München

Dr. Johannes Kistenich
Landesarchiv Nordrhein-
Westfalen, Abteilung Westfalen
An den Speichern 11
D – 48157 Münster

Dr. Gerald Maier
Landesarchiv Baden-
Württemberg
Eugenstraße 7
D – 70182 Stuttgart

Prof. Dr. Angelika Menne-Haritz
Bundesarchiv
Finckensteinallee 63
D – 12205 Berlin

Dr. Sigrid Schieber
Hessisches Hauptstaatsarchiv
Mosbacher Straße 55
D – 65187 Wiesbaden

Karin Schmidgall
Deutsches Literaturarchiv
Marbach
Postfach 1162
D – 71666 Marbach am Neckar

Claire Sibille-de Grimoüard
Direction Générale des Patri-
moines
56 rue des Francs-Bourgeois
F – 75141 Paris Cedex 03

Stefano Vitali
Soprintendenza Archivistica per
l'Emilia-Romagna
Strada Maggiore 51
I – 40125 Bologna

Katrin Wenzel
BStU-Außenstelle Erfurt
Petersberg, Haus 19
D – 99084 Erfurt

Dr. Peter Worm
LWL-Archivamt für Westfalen
Jahnstraße 26
D – 48147 Münster

Veröffentlichungen der Archivschule Marburg

Stand: September 2010

- Nr. 1: DÜLFER, K.; KORN, H.-E.: Gebräuchliche Abkürzungen des 16.-20. Jh.**, bearb. von K. Uhde, Nachdr. d. 9., überarb. Aufl., 2009. 93 S.
ISBN 978-3-923833-76-4 € 10,00
- Nr. 2: DÜLFER, K.; KORN, H.-E.: Schrifttafeln zur dt. Paläographie des 16.-20. Jh.**, bearb. von K. Uhde, 12. Aufl., 2007. 184 S, 50 Tafeln.
ISBN 978-3-923833-79-5 € 28,60
- Nr. 3: PAPRITZ, J.: Die Kartentitelaufnahme im Archiv**, 6. unveränd. Aufl., 1998. 87 S. und ein Maßstabschlüssel.
ISBN 978-3-923833-02-3 € 6,10
- Nr. 4: PAPRITZ, J.: Die archivi-sche Titelaufnahme bei Sachakten**, 6. erw. Aufl., 1997. 144 S.
ISBN 978-3-923833-56-6 € 10,40
- Nr. 7: DEMANDT, K. E.: Laterculus notarum.** Lateinisch-deutsche Interpretationshilfen für spätmittelalterliche und frühneuzeitliche Archivalien, 8. unveränd. Aufl., 2006. 332 S.
ISBN 978-3-923833-06-1 € 24,80
- Nr. 14: ECKHARDT, W. A. (Hrsg.): Wissenschaftliche Archivarsausbildung in Europa**, Marburger Vorträge, 1989. 163 S.
ISBN 978-3-923833-25-2 € 12,80
- Nr. 17: SCHELLENBERG, Th.: Die Bewertung modernen Verwaltungsschriftguts**, übers. und hrsg. von A. Menne-Haritz, 1990. 113 S.
ISBN 978-3-923833-32-0 € 6,10
- Nr. 18: POLLEY, R. (Hrsg.): Archivgesetzgebung in Deutschland.** Beiträge eines Symposions, 1991. 322 S.
ISBN 978-3-923833-15-3 € 16,40
- Nr. 19: DOLLAR, C. M.: Die Auswirkungen der Informationstechnologien auf archivische Prinzipien und Methoden**, übers. und hrsg. von A. Menne-Haritz, 1992. 160 S.
ISBN 978-3-923833-82-5 € 8,20

Nr. 20: MENNE-HARITZ, A.: Schlüsselbegriffe der Archivterminologie. Lehrmaterialien für das Fach Archivwissenschaft, Nachdr. d. 3., durchges. Aufl., 2006. 126 S.
ISBN 978-3-923833-09-2 € 16,80

Nr. 21: WETTMANN, A. (Hrsg.): Bilanz und Perspektiven archivischer Bewertung. Beiträge eines Archivwissenschaftlichen Kolloquiums, 1994. 285 S.
ISBN 978-3-923833-91-7 € 14,40

Nr. 23: BRÜNING, R.; HEGEWALDT, W.; BRÜBACH, N. (Übers. & Hrsg.): Internationale Grundsätze für die archivische Verzeichnung, Durchgesehener Nachdruck der 2., überarb. Ausgabe, 2006. 60 S.
ISBN 978-3-923833-07-8 € 10,40

Nr. 24: STEIN, W. H.: Französisches Verwaltungsschriftgut in Deutschland. Die Departementalverwaltung in der Zeit der Französischen Revolution und des Empire, 1996. 154 S., 5 Karten und 48 Tafeln.
ISBN 978-3-923833-51-1 € 18,40

Nr. 25: MENNE-HARITZ, A.: Akten, Vorgänge und elektronische Bürosysteme. Mit Handreichungen für die Beratung von Behörden, 1996. 121 S.
ISBN 978-3-923833-50-4 € 10,40

Nr. 26: MENNE-HARITZ, A.; BRÜBACH, N.: Der intrinsische Wert von Archiv- und Bibliotheksgut. Kriterienkatalog zur bildlichen und textlichen Konversion bei der Bestandserhaltung. Ergebnisse eines DFG-Projekts, 1997. 99 S.
ISBN 978-3-923833-52-8 € 8,20

Nr. 27: UHDE, K. (Hrsg.): Qualitätssicherung und Rationalisierungspotentiale in der Archivarbeit. Beiträge des 2. Archivwissenschaftlichen Kolloquiums der Archivschule Marburg, 1997. 246 S.
ISBN 978-3-923833-53-5 € 20,40

Nr. 30: MENNE-HARITZ, A. (Hrsg.): Archivische Erschließung: Methodische Aspekte einer Fachkompetenz. Beiträge des 3. Archivwissenschaftlichen Kolloquiums der Archivschule Marburg, 1999. 265 S.
ISBN 978-3-923833-58-0 € 20,40

Nr. 31: METZING, A. (Hrsg.): Digitale Archive – Ein neues Paradigma? Beiträge des 4. Archivwissenschaftlichen Kolloquiums der Archivschule Marburg, 2000. 325 S.
ISBN 978-3-923833-62-7 € 20,40

Nr. 32: ECKELMANN, S. (u.a.): Neue Medien im Archiv: Onlinezugang und elektronische Unterlagen. Bericht

Über eine Studienreise nach Nordamerika, 10.-21. Mai 1999, 2000. 85 S.
ISBN 978-3-923833-64-1 € 10,40

Nr. 33: BRÜBACH, N. (Hrsg.): Der Zugang zu Verwaltungsinformationen. Transparenz als archivische Dienstleistung. Beiträge des 5. Archivwissenschaftl. Kolloquiums der Archivschule Marburg, 2000. 247 S. (vergriffen, PDF-Version online)
ISBN 978-3-923833-65-8 € 20,40

Nr. 34: BÜTTNER, S.; KRETZSCHMAR, R.; STAHLSCHEIDT, R.: Der archivische Umgang mit großen Fallaktenserien. Bericht der Arbeitsgruppe „Archivierung großer Fallaktenserien der Archivreferentenkonferenz des Bundes und der Länder. 2001, X, 89 S.
ISBN 978-3-923833-67-2 € 8,20

Nr. 35: MENNE-HARITZ, A. (Hrsg.): Online-Findbücher, Suchmaschinen und Portale. Beiträge des 6. Archivwissenschaftlichen Kolloquiums der Archivschule Marburg, 2002, 238 S.
ISBN 978-3-923833-68-9 € 20,40

Nr. 36: BRÜBACH, N. (Hrsg.): Archivierung und Zugang. Transferarbeiten des 34. wissenschaftlichen Kurses der Archivschule Marburg, 2002, 327 S. (vergriffen, PDF-Version online)
ISBN 978-3-923833-69-6 € 20,40

Nr. 37: BRÜBACH, N.; MURK, K.: Zur Lage der Universitätsarchive in Deutschland. Beiträge eines Symposiums, 2003, 99 S.
ISBN 978-3-923833-70-2 € 10,00

Nr. 38: POLLEY, R. (Hrsg.): Archivgesetzgebung in Deutschland. Ungeklärte Rechtsfragen und neue Herausforderungen. Beiträge des 7. Archivwissenschaftlichen Kolloquiums der Archivschule Marburg, 2003, 191 S.
ISBN 978-3-923833-72-6 € 20,40

Nr. 39: UNGER, S. (Hrsg.): Archive und ihre Nutzer – Archive als moderne Dienstleister. Beiträge des 8. Archivwissenschaftl. Kolloquiums der Archivschule Marburg, 2004, 261 S. (vergriffen, PDF-Version online)
ISBN 978-3-923833-75-7 € 20,40

Nr. 40: BISCHOFF, F.M.; HOFMAN, H.; ROSS, S. (Hrsg.): Metadata in Preservation. Selected papers from an ERPANET seminar at the Archives School Marburg, 3.-5. September 2003, 2004, 259 S., 8 Tafeln.
ISBN 978-3-923833-77-1 € 23,80

Nr. 41: UNGER, S. (Hrsg.): Archivarbeit zwischen Theorie und Praxis. Ausgewählte Transferarbeiten des 35. und 36. Wissenschaftlichen Kurses an der Archivschule Marburg, 2004, 458 S.
ISBN 978-3-923833-78-8 € 28,60

Nr. 42: BISCHOFF, F. M.; KRETZSCHMAR, R. (Hrsg.): Neue Perspektiven archivischer Bewertung. Beiträge zu einem Workshop an der Archivschule Marburg am 15. November 2004, 2005, 218 S.
ISBN 978-3-923833-01-6 € 23,80

Nr. 43: UHDE, K. (Hrsg.): Berufsbild im Wandel – Aktuelle Herausforderungen für die archivische Ausbildung und Fortbildung. Beiträge zum 9. Archivwissenschaftlichen Kolloquium der Archivschule Marburg, 2005, 286 S.
ISBN 978-3-923833-04-7 € 24,80

Nr. 44: LUTZ, A. (Hrsg.): Neue Konzepte für die archivische Praxis. Ausgewählte Transferarbeiten des 37. und 38. wissenschaftlichen Kurses an der Archivschule Marburg, 2006, 399 S.
ISBN 978-3-923833-08-5 € 28,60

Nr. 45: BISCHOFF, F. M.; HONIGMANN, P. (Hrsg.): Jüdisches Archivwesen. Beiträge zum Kolloquium aus Anlass des 100. Jahrestags der Gründung des Gesamtarchivs der deutschen Juden, zugleich 10. Archivwissenschaftliches Kolloquium der Archivschule Marburg, 2007, 430 S.
ISBN 978-3-923833-10-8 € 28,60

Nr. 46: BISCHOFF, F. M. (Hrsg.): Aktuelle Anforderungen an archivische Erschließung und Findmittel. Beiträge zum 11. Archivwissenschaftlichen Kolloquium der Archivschule Marburg, 2007, 338 S. (vergriffen, PDF-Version online)
ISBN 978-3-923833-11-5 € 26,80

Nr. 47: HIRSCH, V. (Hrsg.): Archivarbeit – die Kunst des Machbaren. Ausgewählte Transferarbeiten des 39. und 40. wissenschaftlichen Kurses der Archivschule Marburg, 2008, 276 S.
ISBN 978-3-923833-34-4 € 29,80

Nr. 48: UHDE, K. (Hrsg.): Quellenarbeit und Schriftgutverwaltung – Historische Hilfswissenschaften im Kontext archivischer Aufgaben. Beiträge zum 12. Archivwissenschaftlichen Kolloquium der Archivschule Marburg, 2009, 224 S.
ISBN 978-3-923833-35-1 € 19,80

Nr. 49: LUTZ, A. (Hrsg.): Zwischen analog und digital – Schriftgutverwaltung als Herausforderung für Archive. Beiträge zum 13. Archivwissenschaftlichen Kolloquium der Archivschule Marburg, 2009, 292 S.
ISBN 978-3-923833-36-8 € 29,80

Nr. 50: POLLEY, R. (Hrsg.): Anbiertung von Unterlagen öffentlicher Stellen an die Archive: Rechtslagen, Probleme, Lösungswege. Beiträge zu einem Workshop an der Archivschule Marburg, 27. November 2008 (in Vorbereitung zum Druck)
ISBN 978-3-923833-37-5

Nr. 51 WENZEL, K.; JÄCKEL, J. (Hrsg): Retrokonversion, Austauschformate und Archivgutdigitalisierung. Beiträge zum Kolloquium aus Anlass des 60jährigen Bestehens der Archivschule Marburg, zugleich 14. Archivwissenschaftliches Kolloquium der Archivschule Marburg, 2010, 385 S.
ISBN 978-3-923833-38-2 € 27,80

Nr. 52: HIRSCH, V. (Hrsg.): Ausgewählte Transferarbeiten des 41. und 42. wissenschaftlichen Kurses an der Archivschule Marburg, 2010 (in Vorbereitung zum Druck)
ISBN 978-3- 923833-39-9

Midosa-XML:

XML-basiertes Werkzeug zur Erfassung und redaktionellen Bearbeitung von Findbüchern. Hrsg. von der Archivschule Marburg, Handbuch Angelika Menne-Haritz, Bundesarchiv.
€ 250,00

Tel. +49 - 06421 - 1 69 71-0
Fax: +49 - 06421 - 1 69 71-10
Mail: archivschule@staff.uni-marburg.de
Internet: www.archivschule.de

