

Veröffentlichungen der Archivschule Marburg

Hochschule für Archivwissenschaft

Nr. 60

Irmgard Christa Becker, Stephanie Oertel (Hrsg.)

Digitalisierung im Archiv –

Neue Wege der Bereitstellung des Archivguts

Beiträge zum 18. Archivwissenschaftlichen Kolloquium der
Archivschule Marburg

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 978-3-923833-47-4

HESSEN



Die Archivschule Marburg ist eine Einrichtung des Landes Hessen.

Gedruckt auf alterungsbeständigem Papier ∞

©Archivschule Marburg 2015

Bismarckstraße 32, D 35037 Marburg

Konzeption der Umschlaggestaltung: Designbüro Glas, Seeheim-
Jugenheim

Druck: msi - media serve international GmbH, Marburger Strasse 92,
35043 Marburg

Inhalt

Vorwort

von Irmgard Christa Becker	11
----------------------------------	----

Priorisierung ist von den Zielen der Digitalisierung abhängig! Ergebnisse des Workshops am 6. Mai 2013 und Vorstellung der Entscheidungsmatrix

von Irmgard Christa Becker und Stephanie Oertel	13
---	----

Digitalisierung eines ganzen Archivs?

Strategien zur Auswahl, Priorisierung und Durchführung von Massendigitalisierungen

von Andreas Berger	23
--------------------------	----

Digisam – towards a Coordinated Digital Cultural Heritage in Sweden

by Rolf Källman	33
-----------------------	----

Reconceptualising access: sustainability in the digital future

by Barbara Reed	49
-----------------------	----

Analoge Bildarchive auf dem Weg ins digitale Zeitalter – Chancen und Herausforderungen für die Bereitstellung und Benutzung bildhafter Materialien	
von Elke Bauer	61
The View from Outside: Amerikanische Erwartungen an das digitale Angebot deutscher Archive	
von Rainer Hering	75
Die automatisierte Volltexterkennung historischer Handschriften als gemeinsame Aufgabe von Archiven, Geistes- und Computerwissenschaft. Das Modell einer zentralen Transkriptionsplattform als virtuelle Forschungsumgebung	
von Guenter Muehlberger	87
Wenn ich mir was wünschen dürfte. Wunsch(t)raum Archiv für NutzerInnen im digitalen Zeitalter	
von Sylvia Necker	117
Nutzen und Gefahren des digitalen Dokuments. Erfahrungsbericht aus Sicht der Unternehmensgeschichte	
von Andrea H. Schneider	135

Instrumente, Skizzen, Briefe...	
das besondere Handwerkszeug der Wissenschaftsgeschichte?	
von Susan Splinter	151
 HerBalt, Hereditas Baltica –	
„Virtueller Lesesaal für baltisches Archivgut“.	
Ein länderverbindendes Digitalisierungsprojekt	
von Peter Wörster	169
 Digitalisierung und die DDB: Standards und Perspektiven	
von Reinhard Altenhöner	181
 Scanner, Standards, Strategien – Digitalisierung im	
Digitalisierungszentrum der Niedersächsischen	
Staats- und Universitätsbibliothek Göttingen	
von Martin Liebetruth	213
 Mit langem Atem zu gutem Ende	
– Erfahrungen bei „museum-digital“ –	
von Stefan Rohde-Enslin	233
 Die Podiumsdiskussion des 18. Archivwissenschaftlichen	
Kolloquiums – eine Zusammenfassung	
von Stephanie Oertel	265
 Autorenverzeichnis	275

Vorwort

von Irmgard Christa Becker

Digitalisate gibt es in deutschen Archiven seit Mitte der 1990er Jahre. Der erste vollständig digitalisierte Archivbestand wurde im Stadtarchiv Duderstadt bereitgestellt. Trotz vieler weiterer Digitalisierungsprojekte gibt es aber keine nationale Digitalisierungsstrategie für Archivgut wie in anderen europäischen Ländern. Die Etablierung einer solchen Strategie setzt voraus, dass die Ziele und Zielgruppen, die Prioritäten, die Standards, die Abläufe und die Werkzeuge einer solchen Digitalisierungskampagne geklärt sind.

In einem von der DFG geförderten Pilotprojekt sind sieben Kooperationspartner aus dem deutschen Archivwesen, die Landesarchive Baden-Württemberg und Nordrhein-Westfalen, die Generaldirektion der staatlichen Archive Bayerns, das Sächsische Staatsarchiv, das LWL Archivamt für Westfalen und das Institut für Stadtgeschichte – Stadtarchiv Mannheim damit befasst, diese Punkte zu klären, um die Digitalisierung des Archivguts in Deutschland voranzutreiben.

Im Rahmen dieses Projekts hat die Archivschule Marburg in Abstimmung mit den Kooperationspartnern das 18. Archivwissenschaftliche Kolloquium mit dem Titel Digitalisierung im Archiv - Neue Wege der Bereitstellung des Archivguts durchgeführt. Das Kolloquium hatte das Ziel, mit Vertretern historischer Fachdisziplinen und verwandter Kultureinrichtungen wie Bibliotheken und Museen Fragen der Digitalisierung zu diskutieren. Die Historiker berichteten, welche Informationen

sie im Internet erwarten und wie sie sich den Zugang zu Archivgut vorstellen. Bibliothekare und Museumsmitarbeiter stellten ihre Vorgehensweisen bei der Bereitstellung digitaler Informationen und Kulturgüter vor. Mit einem Blick über die Grenzen – nach Schweden und Australien wurde deutlich, wie unterschiedlich die Strategien sein können.

Die Ergebnisse des Kolloquiums sind im vorliegenden Band zusammengefasst. Sie werden in den Abschlussbericht des DFG-Projekts einbezogen.

Im Rahmen des Kolloquiums wurden auch die Ergebnisse eines Workshops präsentiert, bei dem im Mai 2014 Archivare aus verschiedenen Archivsparten Fragestellungen und Kriterien der Priorisierung bei der Digitalisierung diskutiert hatten. Dazu gehörte die Erkenntnis, dass die Priorisierung in einem formalisierten Verfahren mit Hilfe einer Matrix erfolgen soll. Der Entwurf der Matrix wurde beim Kolloquium präsentiert. Informationen zur aktuellen Fassung sind im vorliegenden Band zu finden.

In den Band wurde darüber hinaus ein Beitrag von Guenter Mühlberger, Universität Innsbruck, aufgenommen, in dem er den Einsatz von Handschriftenerkennungssoftware für die Recherche in Digitalisaten mittelalterlicher Handschriften vorstellt.

Die Archive stehen erst am Anfang der Digitalisierung des Archivguts. Sie werden sich vielen Fragen zur Recherechtfreundlichkeit durch OCR-Texterkennung, zur Bereitstellung geeigneter Nutzungsumgebungen für ihre Digitalisate und zum Zugang zu Archivalien, die aus rechtlichen Gründen nicht oder nur eingeschränkt nutzbar sind, stellen müssen. Mit dem vorliegenden Band wollen wir dazu einen Beitrag leisten.

Priorisierung ist von den Zielen der Digitalisierung abhängig!

Ergebnisse des Workshops am 6. Mai 2013 und Vorstellung der Entscheidungsmatrix

von Irmgard Christa Becker und Stephanie Oertel

Im Rahmen des DFG-Projekts Produktivpilot ‚Digitalisierung von archivalischen Quellen‘ hat die Archivschule Marburg unter anderem den Arbeitsschwerpunkt, Priorisierungsmerkmale zu entwickeln. Erste Überlegungen dazu sind im Projektantrag formuliert, vor allem die Aufforderung, bei der Entwicklung mit der Forschung zusammenzuarbeiten, um deren Bedarf angemessen berücksichtigen zu können. Wir haben dazu ein zweistufiges Vorgehen gewählt. In einem Workshop am 6. Mai 2014 haben die Projektpartner mit der archivischen Fachwelt erste Überlegungen zur Erarbeitung von Priorisierungskriterien angestellt. In der zweiten Stufe haben wir die Ergebnisse des Workshops und den zwischenzeitlich erarbeiteten Entwurf einer Priorisierungsmatrix beim 18. Archivwissenschaftlichen Kolloquium für die Forschung zur Diskussion gestellt.

Der Workshop Priorisierung bei der Digitalisierung des Archivguts

Nach der Vorstellung des DFG-Projekts durch Frank Bischoff, Landesarchiv Nordrhein-Westfalen, erläuterte Niklaus Bütikofer, Universitätsarchiv Bern, in einem Impulsreferat Konzepte und Methoden der Priorisierung. Er hob hervor, dass zuerst in einer Strategie die Ziele

der Digitalisierung erarbeitet werden müssen. Hauseigene Ziele müssen dabei lokal und regional abgestimmt werden. Eine solche Strategie kann auch in nationale und internationale Strategien eingebettet sein. Anhand der internationalen Instrumente SOUDAAM (Source-Oriented User-Driven Asset-Aware Model)¹ und NINCH-Guide (National Initiative for a Networked Cultural Heritage)² zeigte er wie Ziele und Priorisierungskriterien entwickelt werden können. Ziele basieren meistens auf Fragenkatalogen. Aus den Zielen werden die Priorisierungskriterien abgeleitet. Die Priorisierung kann entweder mit Nutzwertanalysen (SOUDAAM) oder mit Entscheidungsbäumen (NINCH-Guide) durchgeführt werden. Im Ergebnis machte Bütikofer deutlich, dass eine Priorisierung immer von den Zielen der Digitalisierung abhängig ist.

Auf dieser Grundlage diskutierten die drei Arbeitsgruppen, deren Themen im Vorfeld festgelegt worden waren. In der AG 1: Nutzerinteressen und Qualität der Nutzung kristallisierte sich der Mehrwert für die Nutzung durch die Digitalisierung als wichtigstes Kriterium heraus. Als weitere Kriterien wurden unter anderem die Nachfrage nach einem Bestand sowie der Nutzungskomfort diskutiert. Die Arbeitsgruppe bestätigte die zentrale Aussage des Impulsvortrags:

¹ Zu SOUDAAM: Ian Anderson, Archival digitisation: breaking out of the strong box, in: Alistair Tough, Michal Moss (Hrsg.), Record Keeping in a Hybrid Environment. Managing the creation, use, preservation and disposal of unpublished information objects in context, Oxford 2006, S. 203-226.

² Zu NINCH-Guide: <<http://www.ninch.org/programs/practice/>> [07.04.2014].

Priorisierungskriterien und ihre Gewichtung sind in hohem Maße von den Zielen der Digitalisierung abhängig. Die AG 2: Archivgutbezogene Kriterien beschäftigte sich vorwiegend mit Kriterien, die zur Bestimmung der Bedeutung eines Bestandes aus archivarischer Sicht geeignet sind. Dazu zählen unter anderem die Fragen, ob ein Bestand ein Referenzbestand mit vielen Bezügen zu anderen Beständen ist, ob Informationswert und –dichte des Bestandes besonders groß ist, ob ein Bestand singulär ist oder überregionale Bedeutung hat oder ob ein Bestand besonders häufig genutzt wird. Bei einzelnen Archivalien kann der Memorialwert oder ihre Eigenschaft als Zimelie relevant sein. Die von AG 2 erarbeiteten Kriterien müssten aber immer mit denen der AG 1 abgeglichen werden und könnten dann zur Bestimmung der Bedeutung eines Bestandes herangezogen werden. Die AG 3: Rahmenbedingungen der Digitalisierung diskutierte drei Themenkomplexe: Technik, Bestandserhaltung und Rechte. Die technischen Fragen bezogen sich auf die Durchführung der Digitalisierung wie Vorlagenschonung, auf die Erschließung wie eine Verknüpfung zur Archivsoftware und auf die Bereitstellung im Internet. Im Ergebnis sind technische Fragen mit besonders vielen k.o. Kriterien belegt. Bei Fragen der Bestandserhaltung wurden besonders viele Priorisierungskriterien genannt, z.B. ob und wie die Digitalisierungsfähigkeit hergestellt werden kann. Rechtsfragen beziehen sich auf fremde urheberrechtliche Verwertungsrechte und archivrechtliche Schutzfristen, hier sind mehr Priorisierungs- als k.o.-Kriterien zu verzeichnen. Wie oben formuliert hat besonders AG 1 die zentrale Aussage des Impulsvortrags bestätigt. Auch in den beiden anderen Arbeitsgruppen war sie letztlich die Quintessenz der Diskussionen.

Die Ergebnisse des Workshops haben wir im Teilprojekt der Archivschule weiter verarbeitet und eine Entscheidungsmatrix entwickelt.

Entwicklung einer Entscheidungsmatrix

Bei der Erarbeitung von Priorisierungsmerkmalen genügt es nicht, diese einfach aneinander zu reihen. Wenn man mit ihnen wirklich arbeiten will, müssen sie zueinander in Beziehung gesetzt und gewichtet werden. Ziel ist eine Entscheidungsgrundlage für die Planung der Digitalisierungsmaßnahmen und die Visualisierung des damit verbundenen Aufwands in einem Archiv. Schon während des Workshops hat Frank Bischof dazu eine Matrixstruktur vorgeschlagen. Diese Idee haben wir gerne aufgegriffen und den ersten Entwurf einer Entscheidungsmatrix erstellt. Wir haben dafür das Rad nicht neu erfunden, sondern eine Priorisierungsmatrix des Historischen Archivs der Stadt Köln als Grundlage genommen.³ Auf den ersten Blick erscheinen die beiden Matrizen ähnlich, die Schwerpunkte sind aber unterschiedlich gesetzt. Das Ziel der Kölner Matrix ist die Bestandserhaltung, deshalb sind dazu die meisten Kriterien enthalten.

³ Franz-Josef Verscharen, Gisela Fleckenstein, Andreas Berger, Was restaurieren wir zuerst? Priorisierungsmatrix für die Restaurierung und Zusammenführung beim Wiederaufbau des Historischen Archivs der Stadt Köln, in: *Archivar* 64,1, 2011, S. 29-32.

Darüber hinaus wird nach der Wertigkeit des Bestands und nach der Benutzung gefragt.

Die Entscheidungsmatrix, die wir für die Digitalisierung im Archiv vorschlagen, soll dabei helfen Digitalisierungsprojekte effektiv in Abhängigkeit von der jeweiligen Ausgangssituation, den Ressourcen und der Nachhaltigkeit zu planen. Die archivischen Ziele, die sich auf die Priorisierung der zu digitalisierenden Bestände auswirken, beziehen sich auf die Sichtbarkeit und Zugänglichkeit der Bestände sowie die Schonung der Originale und die Bereitstellung gefährdeter Bestände in Abhängigkeit von den Rechtsvorschriften und der technischen Ausstattung. Die Matrix erfasst hierfür ein umfangreiches Set an Merkmalen⁴ und Kriterien.⁵ Sie bildet einen Ablauf ab, mit dem zuerst Ziele definiert, dann Priorisierungskriterien gewichtet und schließlich Werte zu den Archivbeständen zugeordnet werden können. Die am Ende errechneten Werte geben eine Rangfolge der Bestände wieder, die die oben erwähnte Entscheidungsgrundlage für Digitalisierungsmaßnahmen sein kann. Darüber hinaus sollen Archive mit Hilfe der Entscheidungsmatrix einen Anhaltspunkt erhalten,

⁴ Merkmale: Bedeutung, Zugänglichkeit, Bestandserhaltung, Rechte und Ausstattung.

⁵ Kriterien: Überlieferungsdichte / historische Relevanz / Nutzungshäufigkeit / Nutzungspotential, Findmittel / Erschließungstiefe, Erhaltungszustand / physische Gegebenheit, urheberrechtliche Verwertungsrechte / eingeschränkte Verwertungsrechte / Schutzfristen, Digitalisierungsinfrastruktur / Aufbewahrungsinfrastruktur / Restaurierungsabteilung und Nutzungsinfrastruktur.

ob ihre Digitalisierungsmaßnahmen förderfähig im Sinne der DFG sind. Das übergeordnete Ziel unserer Matrix ist es, herauszufinden, welche Bestände des abgebildeten Gesamtbestandes unter geringsten Digitalisierungsaufwand die festgelegten Ziele effektiv realisieren.

Die Matrix ist als Anwendung in Excel 2010⁶ in drei Tabellenblätter eingeteilt, die miteinander in Beziehung stehen. Ein viertes Tabellenblatt enthält inhaltliche Erläuterungen zu den drei vorangegangenen Tabellenblättern. Auf dem ersten Tabellenblatt, das für den ersten Schritt der Priorisierung steht, befindet sich auch die Anleitung zum Ausfüllen der Matrix. Dort legt der Nutzer die Ziele fest und gewichtet damit die Merkmale⁷ der Priorisierung untereinander. Das Ergebnis ergibt immer 100 %, den einzelnen Merkmalen können wegen ihrer Abhängigkeit voneinander Gewichtungswerte zwischen 4 % und 60 % zugewiesen werden. Das Ergebnis wird durch eine Berechnungsformel im zweiten Tabellenblatt abgebildet. In diesem zweiten Blatt legt der Nutzer im zweiten Schritt fest, wie er die Priorisierungskriterien⁸ innerhalb der Merkmale gewichten möchte. In Schritt 3, auf dem dritten Tabellenblatt, wählt der Nutzer mit Hilfe von Drop-down-Menüs für die Bestände, die priorisiert werden sollen, die Werte zu den Kriterien aus. Hierbei werden die Werte

⁶ Ältere Versionen von Excel unterstützen nicht alle Funktionen der Auswertung.

⁷ Vgl. Anm. 4.

⁸ Vgl. Anm. 5.

unterschieden in Angaben, die nicht in die Berechnung einbezogen werden, bspw. allgemeine Angaben⁹ und Einschätzungen der Kriterien, die mit hoch / mittel / gering und ja / nein sowie schlecht / mittel / gut bewertet werden. Die Zahlen, die diesen Bewertungen zugrunde liegen, werden addiert und in der letzten Spalte der Tabelle als Summe angezeigt. Der höchste Wert wird dabei dem Bestand zugewiesen, der mit dem geringsten Digitalisierungsaufwand die gesetzten Ziele realisiert.

Auf dem ersten Tabellenblatt haben wir für häufige Digitalisierungsziele¹⁰ der Archive und zwei weitere entscheidungsrelevante Punkte¹¹ Fragen formuliert, mit deren Beantwortung die dazu gehörenden Merkmale untereinander gewichtet werden können. Die Fragen erstrecken sich auf die Themen: Wahrnehmung der Bestände, Recherchemöglichkeiten in den Beständen, Schonung der Originale / Bereitstellung von gefährdeten Originalen, Einschränkung der Zugänglichkeit durch Rechtsvorschriften und Ausstattung des Archivs in Hinblick auf die Digitalisierung.

⁹ Allgemeine Angaben: Archivname, Signatur, Name und Laufzeit des Bestandes, Umfang und VZE.

¹⁰ Digitalisierungsziele: Steigerung der Bedeutung und Zugänglichkeit der Bestände unter Berücksichtigung der Bestandserhaltung.

¹¹ Die Digitalisierungssziele werden unter Berücksichtigung der Rechtsvorschriften und der technischen Ausstattung festgelegt.

Den Digitalisierungszielen und den beiden weiteren entscheidungsrelevanten Punkten haben wir Priorisierungsmerkmale zugeordnet, die mit Priorisierungskriterien hinterlegt sind. Diese Priorisierungskriterien sind im zweiten und dritten Tabellenblatt zu finden. Im zweiten Tabellenblatt können sie zueinander gewichtet werden, im dritten Tabellenblatt werden die Werte zugeordnet. Beim Merkmal Rechte tritt eine Schwierigkeit auf: Urheberrechtliche Verwertungsrechte und archivrechtliche Schutzfristen sind zwar keine Ausschlusskriterien für eine Digitalisierung im Allgemeinen, die DFG fördert die Digitalisierung solcher Bestände aber nicht.¹² Deshalb haben wir hier zwei parallele Wertungen eingebaut. In der einen sind die Kriterien urheberrechtliche Verwertungsrechte und archivrechtliche Schutzfristen Priorisierungskriterien, in der zweiten Wertung sind sie k.o.-Kriterien. Die Nutzer der Matrix können damit einschätzen, ob eine Förderung durch die DFG in Frage kommt oder nicht. Das Merkmal technische Ausstattung liefert Anhaltspunkte zum finanziellen Aufwand bei der Digitalisierung: Je weniger Infrastruktur im eigenen Haus vorhanden ist, desto mehr Finanzmittel müssen eingesetzt werden.

¹² Urheberrechtliche Verwertungsrechte sind ein Digitalisierungshindernis, wenn der digitalisierte Bestand im Internet bereitgestellt werden soll und die urheberrechtlichen Verwertungsrechte nicht geklärt sind. Wenn eine Einverständniserklärung des Rechteinhabers vorliegt, dass die Digitalisate dauerhaft im Internet präsentiert werden dürfen, können DFG-Mittel für eine Digitalisierung beantragt werden.

1. Fazit

Der Erarbeitungsstand der Entscheidungsmatrix ist ein Zwischenergebnis, das nach dem Kolloquium in der Praxis erprobt wurde und im Laufe des Projekts überarbeitet wird.

Wir haben bisher die drei wichtigsten Digitalisierungsziele¹³ mit einem bestimmten Set an Priorisierungskriterien¹⁴ berücksichtigt. Welche weiteren Ziele und Merkmale aufgenommen werden müssen, wird der weitere Verlauf des Projekts zeigen. Denkbar ist zum Beispiel, beim Merkmal Zugänglichkeit noch ‚Texterkennung möglich‘ als Priorisierungskriterium aufzunehmen oder weiter gehende Kriterien einzubauen, wie ‚Daten auslesbar‘. Das berührt aber einen Aufgabenbereich, in dem die Grenze zwischen der Bereitstellung von Archivgut durch Archivare und der Datenerhebung für die Forschung durch Historiker verschwimmt. Deshalb ist das ein Punkt, der im Projekt und in der Fachwelt diskutiert werden muss.

Grundsätzlich ist die Frage zu stellen, ob in unserer Matrix die Ziele berücksichtigt sind, die aus der Sicht der Nutzer relevant sind. Man könnte auch ‚Kooperation‘ als Ziel einfügen. Denkbar wäre auch die flexible Möglichkeit zur Anreicherung von weiteren Zielen in der Matrix.

¹³ Vgl. Anm. 10.

¹⁴ Vgl. Anm. 5.

Es bleiben jedenfalls noch einige offene Fragen, die im Projektverlauf weiter ausgelotet werden sollen. Nach Abschluss des Projekts wird die Entscheidungsmatrix über die Archivschulhomepage¹⁵ der Öffentlichkeit zur Verfügung gestellt. Weitere Angaben zum Projekt und den beteiligten Archiven finden Sie auf der Projekthomepage¹⁶ und dem Projektflyer¹⁷.

¹⁵ Archivschulhomepage: <<http://archivschule.de/>> [07.04.2014].

¹⁶ Projekthomepage: <<http://archivschule.de/DE/forschung/retrokonversion/>> [07.04.2014].

¹⁷ Projektflyer:
<http://archivschule.de/uploads/Forschung/Digitalisierung/Publikationen/Flyer_Pilotprojekt_Digitalisierung_von_archivalischen_Quellen_Druckversion.pdf> [07.04.2014].

Digitalisierung eines ganzen Archivs?

Strategien zur Auswahl, Priorisierung und Durchführung von Massendigitalisierungen

von Andreas Berger

Der folgende Beitrag ist ein Werkstattbericht über die im Historischen Archiv der Stadt Köln eingesetzte Strategie in Bezug auf die Massendigitalisierung von Archivgut, das durch den Archiveinsturz vom 03.03.2009 betroffen ist. Daneben ist die Strategie in Köln auch auf Neuübernahmen übertragen worden. So ergibt sich auch eine Übertragbarkeit auf andere Archive oder wenigstens die Möglichkeit Elemente oder Anregungen zu übernehmen.¹ Dabei liegt der Fokus auf den Entscheidungsprozessen, die im Vorfeld der eigentlichen Digitalisierung ablaufen müssen. Der Digitalisierungsprozess, die Verarbeitung der Digitalisate und die Speicherung werden in der Darstellung nicht berücksichtigt.

Schon zu Beginn kann die Frage, die im Titel gestellt wird, eindeutig beantwortet werden: Auch in Köln wird *kein* ganzes Archiv digitalisiert. Das ist durch die reine Masse des Archivguts in absehbarer Zeit nicht mit den vorhandenen Mitteln zu bewältigen. Die Speicherkosten für die dadurch entstandenen Daten wären immens.

¹ Da es sich um einen Erfahrungsbericht handelt, wird zum größten Teil auf Literaturverweise verzichtet. Anders als bei der Abfassung dieses Beitrags sind für die Erstellung der Strategie Publikationen wie auch persönliche Kontakte mit Kolleginnen und Kollegen, die in diesem Bereich Erfahrungen gesammelt haben, berücksichtigt worden.

Zudem ist die komplette Digitalisierung aller Archivalien eines Archivs unnötig, da die Nutzungsfrequenz vieler Bestände (noch) sehr niedrig ist und gerade die jüngere städtische Überlieferung und Teile der Deposita und Nachlässe aus rechtlichen Gründen noch nicht nutzbar sind. Hier müssen andere Strategien, wie z.B. „Digitalisierung on Demand“ greifen. Das Ziel in Köln ist vielmehr, den ganzen Bestand (wieder) nutzbar zu machen. Entweder soll eine Nutzung im Original stattfinden oder, wenn das aus konservatorischen Gründen nicht oder noch nicht möglich ist, als digitale Reprographie möglich sein. Das stellt eine Kölner Besonderheit dar, die im ersten Teil des Beitrags im Mittelpunkt stehen wird. Im zweiten Teil können die dabei gemachten Erfahrungen und eingesetzten Abläufe auch auf Neuübernahmen und damit auf andere Archive übertragen werden.

Alle Entscheidungen, die einer Digitalisierung zugrunde liegen, lassen sich auf vier Fragen reduzieren. 1) Kann das Archivale digitalisiert werden? Das ist die Frage nach dem physischen Zustand des Archivales und den rechtlichen Rahmenbedingungen, die einer Veröffentlichung oder schon der Digitalisierung an sich im Weg stehen können. 2) Soll das Archivale, der Bestand oder der Teilbestand digitalisiert werden? Damit wird die archivische Entscheidung getroffen, ob ein Digitalisat des Archivales etc. notwendig oder sinnvoll ist. 3) Wie soll digitalisiert werden? Hier stehen die Parameter der Digitalisierung im Mittelpunkt, die durch das Ziel der Digitalisierung bestimmt werden. 4) In welcher Reihenfolge soll digitalisiert werden? Diese vier Fragen können theoretisch beantwortet werden und dadurch eine Strategie festgelegt werden, die dann die praktischen Abläufe steuert. Dadurch werden klare nicht von Einzelstücken abhängige Regeln geschaffen und feste Strukturen vorgegeben. Wir

haben uns gegen dieses deduktive Vorgehen bei der Erstellung einer Strategie entschieden, auch wenn alle Parameter, die Zahlengerüste und die Ziele bekannt oder annähernd geschätzt waren. Der Grund ist, dass wir aus der praktischen Arbeit heraus eigene Erfahrungen sammeln wollten, dann gesammelt haben und auf die Erfahrungen möglichst vieler Kolleginnen und Kollegen zurückgegriffen haben. Das ist der besonderen Situation geschuldet, die durch den Einsturz hervorgerufen worden ist und durch die in vielen Fällen neue nicht vorhersehbare Umstände auch in den Entscheidungsprozessen auftreten. Auf der Grundlage von Einzelentscheidungen, die wir immer in größeren Gruppen mit allen Beteiligten (Restauratoren, Mitarbeitenden der Digitalisierung und Archivaren) diskutiert haben, wurden diese verallgemeinert. Aus den Entscheidungen bei der Behandlung einer überschaubaren Zahl von Einzelstücken ist also eine Auswahlmatrix erzeugt worden, die nun auf alle Stücke angewendet wird.

Ebenso problematisch, wie die Erstellung der Kriterien ist es, die zu treffenden Entscheidungen tatsächlich rationell durchzuführen und dadurch nicht noch zusätzliche Arbeitsschritte zu erzeugen. Zuerst sind die vorhandenen Prozesse betrachtet worden und innerhalb derer werden die Entscheidungen für die Digitalisierung getroffen, ohne dass neue Prozesse geschaffen werden. Dabei werden die vorhandenen Kompetenzen der Mitarbeiter genutzt, sodass nur jeder das entscheidet, was er fachlich auch entscheiden kann. Ein Restaurator trifft in der Regel keine inhaltlichen Entscheidungen und der Archivar in der Regel keine bestandserhalterischen. Daraus werden klare Entscheidungsstrukturen etabliert, die möglichst in Prozessbeschreibungen münden und dadurch das Handeln normieren und nachvollziehbar machen. Die Entscheidungen müssen klar definiert

werden und die Kriterien transparent und eindeutig festgelegt werden. Um keine Unklarheiten über die Kompetenzen entstehen zu lassen, werden diese auch eindeutig festgelegt.² Grob gesagt: die Fachkraft der Bestandserhaltung trifft die Entscheidungen über die Digitalisierbarkeit, Archivare treffen die Entscheidung, ob, in welcher Reihenfolge und zu welchem Zweck digitalisiert wird und die Fachkraft der Digitalisierung trifft die technischen Entscheidungen.

Der Ablauf in Köln kann dies verdeutlichen: Zuerst wird über die Digitalisierungsfähigkeit in zwei Stufen entschieden. Vor der konservatorischen Behandlung der Stücke wird für jedes einzelne Archivale beim Ausfüllen eines Laufzettels, der die konservatorischen Prozesse steuert, durch eine restauratorische Fachkraft entschieden, ob es nach der Trockenreinigung im Original genutzt werden kann (Kat. A), es nicht im Original nutzbar aber digitalisierbar ist (Kat. B) oder die Schäden so stark sind, dass eine Nutzung vor weitergehenden restauratorischen Maßnahmen nicht möglich ist (Kat. C). Diese Kategorisierung hat natürlich später auch Einfluss auf die archivische Entscheidung darüber, ob digitalisiert werden soll. So muss ein Archivale der Kat. B entweder für eine Nutzung digitalisiert werden oder ggf. aufwändig restauriert werden, um eine Nutzung im Original zu ermöglichen. Nach der Konservierung durch Restaurierungshelfer und während der Qualitätssicherung durch einen Restaurator wird

² Bei Problemfällen oder bei sich unterscheidenden Beurteilungen ist es immer nötig, gemeinsam Entscheidungen zu treffen. Soll z.B. eine von der Bestandserhaltung für die Nutzung gesperrte Handschrift ausgestellt werden, müssen gemeinsam Maßnahmen beschlossen werden, diese ausstellungsfähig zu machen oder eine Alternative gesucht werden.

festgelegt, wie das Archivale digitalisiert werden kann, ohne dass dadurch Schäden am Original entstehen. Dabei bedient sich der Restaurator einer Buchstaben-Zahlen-Kombination, die auf dem beiliegenden Laufzettel notiert wird. Für die Festlegung dieser Bedingungen existieren Richtlinien, die jedoch immer der Einschätzung der Fachkraft in Bezug auf das Einzelstück bedürfen. Der Code wird dann beim Digitalisieren interpretiert und umgesetzt.³

Während der endgültigen Identifizierung oder der Erschließung des Archivales wird vom Archivar festgelegt ob es digitalisiert werden soll und mit welchem Ziel digitalisiert werden soll. Zwei Ziele der Digitalisierung sind möglich. Das häufigste Ziel ist die digitale Nutzung am Bildschirm. Dazu sind in der Regel keine besonders hoch aufgelösten Digitalisate und keine verlustfreien Speicherungen nötig. Das Ziel kann auch eine qualitativ hochwertige Reproduktion sein. Folgende Kriterien haben dabei Einfluss auf die Entscheidung: Ist das Archivale im Original nutzbar? In dem Fall muss nicht unbedingt digitalisiert werden. Ist eine häufige Nutzung zu erwarten? Dann sollte unbedingt eine Digitalisierung erfolgen. Werden alle relevanten Inhalte im Digitalisat sichtbar sein? Nur dann ist eine Digitalisierung überhaupt sinnvoll. Ist das Stück sehr fragil oder z.B. durch sein Format nur mit Aufwand nutzbar? Ein schlecht nutzbares Archivale muss digitalisiert werden. Liegen lange Schutz- und Sperrfristen vor?

³ Es wird der mögliche Öffnungswinkel angegeben (z.B. D100 = maximaler Öffnungswinkel 100 Grad). Es werden die Hilfsmittel angegeben, die verwendet werden dürfen oder müssen (z.B. F9 = Fixieren der Seiten mit Bleischlangen). Viele weitere Angaben werden bei Bedarf gemacht, bis hin zur Festlegung, dass beim Scannvorgang ein Restaurator anwesend sein muss.

Eine Digitalisierung macht dann zum jetzigen Zeitpunkt keinen Sinn. Liegt eine hohe inhaltliche Bedeutung oder ein hoher materieller Wert vor? Dann sollte vordringlich digitalisiert werden. Werden Reproduktionen häufig angefragt oder deren Anfrage erwartet? So ist eine Digitalisierung dringend geboten. Zuletzt spielt ein nur für Köln relevantes Kriterium eine Rolle: Kann das Archivale identifiziert werden? Wenn nicht oder nicht vollständig muss es digitalisiert werden, um eine Identifizierung am Digitalisat – gegebenenfalls über das Internet – zu ermöglichen. Diese Entscheidungen sind am wenigsten standardisierbar und bei der Anwendung der Kriterien kommt es auf die – nur zu Teilen objektivierbare – Beurteilung des Archivars an. Hierbei handelt es sich immer um Entscheidungen auf Ebene des einzelnen Archivales. Das ist eine Besonderheit für Köln und durch die Arbeits- und Entscheidungsprozesse zur Einsturzbewältigung bedingt.

Es wird in der Community diskutiert in welcher Qualität digitalisiert werden soll und ob immer ein hochaufgelöstes Master gespeichert werden soll. In Köln haben wir aufgrund des benötigten Speicherplatzes entschieden, nicht immer Masterdigitalisate zu erstellen bzw. diese nicht zu speichern, sondern direkt im Nutzungsformat und der nötigen Auflösung zu speichern. In welchen Fällen ein hochaufgelöster Scan erstellt wird, ist genau festgelegt. Die Vorlagenart und der Inhalt bestimmen dabei die Auflösung und das Speicherformat. So werden z.B. durchgehend illuminierte Handschriften hochaufgelöst und unkomprimiert gespeichert. Diese Speicherung erfolgt immer zusätzlich zum Benutzungsdigitalisat in geringerer Auflösung in einem komprimierten Format. Dabei wird auf die Verwendung von langzeitstabilen und/oder konvertierbaren Formaten geachtet. Auch eine Speicherung einzelner Seiten in hoher Auflösung ist vorgesehen,

wenn diese Einzelseiten bestimmte Kriterien erfüllen. Diesen Vorgang leiten die Mitarbeitenden an den Scannern automatisch anhand von vorgegebenen Kriterien ein. So wird ein hochaufgelöster Scan für eine Seite ausgelöst, wenn z.B. eine Miniatur in einer Handschrift vorliegt. Es gibt jedoch auch die Möglichkeit für den Archivar festzulegen, dass ein Archivale oder einzelne Teile in hoher Auflösung und unkomprimiert gespeichert werden, wenn die formalen Kriterien nicht erfüllt sind. Das wird angewendet, wenn z.B. eine wertvolle aber groß und deutlich geschriebene Handschrift ohne Miniaturen vorliegt, für deren Reproduktion häufige Anfragen zu erwarten sind. Die Dokumentation auch dieser Entscheidung erfolgt auf einem Laufzettel über den dann die weiteren Schritte gesteuert werden. In Zukunft wird die komplette Dokumentation und Steuerung der Prozesse über eine Software erfolgen, die im Moment als neues Modul der bisher eingesetzten Software konzipiert wird.

Damit sind alle anfangs gestellten Fragen zur Digitalisierung eines Einzelstückes mit einer Ausnahme beantwortet. Einzig die Reihenfolge wurde nicht festgelegt. Das hat den Grund, dass zurzeit die vom Einsturz betroffenen Archivalien bearbeitet werden, wie sie in den Asylarchiven vorgefunden werden. Durch die Verunordnung ist eine sinnvolle inhaltliche Steuerung im Moment nicht möglich, da alle Archivalien gerade den Prozess der Trockenreinigung möglichst zeitnah durchlaufen müssen, um fortschreitende Schäden zu vermeiden.

Genutzt werden für die Entscheidungen ausschließlich schon vorhandene Prozesse, in denen die jeweils zuständigen Fachkräfte neben anderen Entscheidungen auch die zur Digitalisierung treffen. Die Einrichtung neuer Prozesse ist also nicht nötig. Es muss vielmehr

nur eine Betrachtung der vorhandenen Prozesse stattfinden und die zusätzlich durchzuführenden Entscheidungen mussten an den entsprechenden Stellen eingegliedert werden. Hilfreich ist es dabei die einzelnen Prozesse, Workflows und Arbeitsabläufe genau zu kennen, textlich wie bildlich festzuhalten und regelmäßig zu evaluieren. Die praktische Umsetzung der Entscheidungen durch die Mitarbeitenden an den Scannern und bei der weiteren Verarbeitung der Digitalisate folgt vorher festgelegten Regeln, die durch das zu scannende Archivgut und den beiliegenden Laufzettel bestimmt werden. Dieser Prozess läuft weitestgehend automatisch ohne die Einbindung von Fachkräften in die operativen Abläufe.

Neben der Behandlung der vom Einsturz betroffenen Archivalien werden in Köln auch die Digitalisierungsentscheidungen für Neuübernahmen anhand leicht modifizierter Abläufe getroffen. Dabei wurde dieser Prozess erst nach der vollständigen Etablierung der oben beschriebenen Abläufe und den ersten sicheren Erfahrungen mit diesen Abläufen in einem gemeinsamen Team erarbeitet. Die Abläufe wurden von Mitarbeitern der Digitalisierung, der Bestandserhaltung und des Magazindienstes festgelegt, da alle diese Arbeitsbereiche betroffen sind. Auch für die Behandlung von Neuübernahmen wurde ein Laufzettel eingeführt, der von der Akzessionierung, über die Anlieferung, die Erstkontrolle durch die Bestandserhaltung, ggf. den bestandserhaltenden Maßnahmen, die Erschließung, die Umbettung, die Folierung und die Digitalisierung bis zur Einlagerung die Abläufe steuert. Auch hier ist langfristig eine Steuerung über eine Software vorgesehen. Bisher ist in Kernbereichen eine Softwareunterstützung über eine Magazinverwaltungssoftware vorhanden.

Die Entscheidungen zur Digitalisierung werden hier in fünf Schritten, die in die immer an Neuübernahmen ablaufenden Prozessen integriert wurden, getroffen. Während des Übernahmevorgangs bei der Anlieferung im Archiv wird eine Erstkontrolle durch die Bestandserhaltung durchgeführt. Dabei wird neben der Beurteilung, ob bestandserhalterische Maßnahmen nötig sind, auch festgelegt, ob eine Nutzung im Original, als Digitalisat oder nur nach der Durchführung bestandserhalterischer Maßnahmen möglich ist (Kat. A, B, C). Ebenso wird nach demselben Muster, wie oben beschrieben, festgelegt, wie digitalisiert werden kann und welche Besonderheiten bei der Digitalisierung zu beachten sind. Diese Festlegungen werden zu diesem Zeitpunkt immer getroffen. Das gilt auch für Archivgut, das noch längeren Sperrfristen unterliegt.

Die Festlegung der Auflösung und des Speicherformats wird analog zu den vom Einsturz betroffenen Archivalien durch die Archivalienart und den Inhalt bedingt. Das kann sich auf das ganze Archival, aber auch abweichend auf Einzelseiten beziehen. Während der Erschließung kann der Archivar jedoch auch davon abweichende Angaben machen. Im Gegensatz zu den vom Einsturz betroffenen Archivalien, wird hier eine Priorisierung beständeweise vorgenommen, die sich an den Kategorien orientiert, die in der Bundessicherungsverfilmung eingesetzt werden. Der Archivar priorisiert auf der Bestandsebene während oder nach der Erschließung. Die Digitalisierung von Einzelstücken kann jedoch on Demand, d.h. bei einer Anfrage durch Benutzer etc. geschehen.

Die Digitalisierungsstrategie im Historischen Archiv der Stadt Köln wurde durch die besonderen Umstände geprägt und folgt dem

eindeutigen Ziel der schnellstmöglichen Benutzbarmachung der vom Einsturz betroffenen Archivalien. Die eingesetzten Steuerungsinstrumente, die durch die Betrachtung und Verallgemeinerung von Einzelfällen gewonnen wurden, sind zum überwiegenden Teil auch auf Neuübernahmen und nicht beschädigte Altbestände übertragbar. Im letzten Fall müssen die Entscheidungsprozesse an andere Arbeiten angegliedert werden. Ggf. müssen die Entscheidungen auch anhand von Beständeübersichten oder des Archivkatasters getroffen werden. Grundlegend sind aber immer die genaue Kenntnis der Prozesse, der Kompetenzenverteilung im Archiv und die Möglichkeit die getroffenen Entscheidungen über Laufzettel oder eine Software festzuhalten und später nutzen zu können.

Wenn diese Voraussetzungen gegeben sind und die Ziele einer Digitalisierung klar sind, muss jedes Archiv für sich selbst Entscheidungsprozesse festlegen und eine Strategie erstellen. Diese sollten sich aber unbedingt an den archivübergreifend verwendeten Vorschlägen, bzw. Vorgaben zur Digitalisierung, wie den Praxisregeln der DFG oder dem an der Archivschule Marburg koordinierten Projekt „Digitalisierung archivischer Quellen“⁴ orientieren, da alle Archive unter den gleichen oder stark ähnlichen Bedingungen handeln und nur durch ein abgestimmtes Vorgehen in diesem Bereich mittelfristig Erfolge für das ganze Archivwesen erreicht werden können.

⁴ Vgl.: <http://www.dfg.de/formulare/12_151/> [28.02.2014];
<http://www.archivschule.de/DE/forschung/digitalisierung>> [28.02.2014].

Digisam – towards a Coordinated Digital Cultural Heritage in Sweden

by Rolf Källman

Introduction

In 2011 the Swedish Government established a national strategy for digitisation, digital access and digital preservation. The aim of the strategy is to regulate the work for the government agencies and institutions that collect, preserve and make cultural heritage information available for the end users. The strategy covers the whole range of issues on digitisation, ranging from selection criteria for digitisation of collections and holdings to preservation and usability. The strategy has a strong focus on use, re-use and creativity, and the necessary prerequisites for that to happen.

The Swedish strategy for digitisation of cultural heritage information and material follows closely the European framework and is strongly connected to Europe 2020, the European Council conclusions, Commission recommendations, National policies and agendas. There are in total 23 central Swedish authorities and institutions, responsible for different perspectives of the cultural heritage, which are targeted by the strategy.

Digisam, a secretariat for national coordination of digitisation, digital preservation and digital access to cultural heritage, is established by the Swedish Government as a part of the national strategy, in order to coordinate the activities taken place within the task of the strategy. Digisam is assigned the role to be a node in

competence and knowledge, to develop proposals of a common infrastructure including cost effective solutions for preservation and to suggest distribution of roles and responsibilities. Digisams mission is focused on the 23 state authorities and institutions as a core target group, also pointed out as collaborative partners in the strategy.

Digisams main task is to coordinate the work with the national digital strategy within the timeframe of 2011-2015 and to:

- Be a node for knowledge and competence.
- Present recommendations for coordinated digital information management of collections and holdings.
- Develop proposals for cost-effective long-term digital preservation of collections and holdings.
- Define roles and responsibilities for the work on aggregation, access and preservation of digital cultural heritage information.

Key words in Digisams work are: infrastructure, interoperability, data harmonisation, competence, legal frameworks, recommendations, regulations and cost effective workflows.

Infrastructure as a prerequisite for success

The overall goal set up in the national strategy is to increase, the amount, use and preservation of accessible digital heritage information. It is also specified that this shall be done within the existing budget frames. In order to accomplish this, there is a need for a stable and sustainable infrastructure. Infrastructure is here defined in a broad sense as a technical platform for several layers of

information, with a range of accessible or implemented tools and services. When working with the complex task of modeling and proposing an infrastructure for digital cultural heritage information the Swedish national research infrastructure SUNET (Swedish University Computer Network) and the Royal Institute of Technology (KTH), must be held forward as important partners.

SUNET provides an e-infrastructure, to which most of the national cultural heritage authorities and institutions are connected. A deepened collaboration between Digisam and SUNET could lead to seamless connections and access to services for storage and preservation.

The Government and the Public Authorities

As a background to Digisams task to coordinate the national agencies and institutions, some short facts concerning the Swedish public administration must be noted. The authorities are the Government's most important instrument in carrying out its policies. The Swedish Government decides upon the preconditions for the individual authorities' operations via annual appropriations directives, ordinances and special assignments. The authorities' work is followed-up through annual reports and evaluation. This is a quite substantial scope for steering the operations of government authorities, but at the same time the Government has not the mandate to intervene in the authorities' decisions in specific matters relating to the application of the law or the due exercise of their authority.

The digitisation process as a balanced score card

During modeling Digisams complex field of task, we developed a model at a more conceptual level. As a general concept and overall covering terms we have defined the following words: Steer, produce, use, preserve and follow-up. Later on we decided that follow up is a part of steer, which left us with four general terms.

There is absolutely nothing surprising or radically new with these four terms; terms that could have been picked out from any common project model. What turned out to be very useful was, for example during interviews with the cultural heritage agencies and institutions, the use of the terms as components of a kind of balanced score card. Our simple thesis is that for the management of a successful digitisation process, in a long term sustainable way, there is a need for balancing the activities according to the terms steer, produce, use and preserve. The interviews held with the authorities and institutions show that none of them manages their digitisation processes in a balanced way, and that most of them have an obvious focus on production and use.

Plans for digitisation

In the national strategy it is also decided that all 23 governmental authorities and institutions covered by the strategy shall deliver plans for digitising and making their archives, collections and libraries accessible and usable. The plans shall include criteria for how they prioritise when they digitise their collections and holdings, the quantity of information they plan to digitise and how they aim to make the digital information accessible and used.

From infrastructure to an evolving interstructure

With the complex task that is assigned to Digisam, it soon became obvious that we needed to decide which of our different tasks that should be established as a baseline to which all other tasks could relate. We chose infrastructure as such a baseline, with the definition as mentioned above, in infrastructure that embrace all the relevant digitisation processes concerning selection, conversion, harvesting, management, use, storage and preservation.

When modeling the existing infrastructure, used by the national heritage agencies and institutions, it became clear that at present there are not one, but several different infrastructures. These infrastructures are interconnected in some parts, and there is not a single one that has all necessary elements for the entire digitisation process. Regarding the fact that there is no allocated state funding for building an entirely new infrastructure for the Swedish digital cultural heritage, the conclusion was drawn that we must build on the existing structures and aim for an evolving interstructure. Digisams role will then be, together with the heritage agencies and institutions, to identify tools and services that support interoperability and data harmonisation and that interconnect the different infrastructures. In fact, we believe this is a more efficient, flexible and sustainable way to address the infrastructural issues. The technical development is running fast which is an argument for open, scalable and flexible solutions with services and tools that can be implemented and removed as resource-economic as possible. Such an approach could also be expressed in terms of design for change, change by increment and change by encapsulation.



Fig. 1: Modeling infrastructure at KTH (R. Källmann, CC-BY)

If the work is successful, we will be able to establish an evolving interstructure that supports communication and collaboration with maximised possibilities for emerging positive synergetic supra structures.

A breakthrough for GLAM

We now stand at a point in time where it has become possible to reach the goals for a seamless flow and use of information that is stored in the archives, libraries, museums and the institutions that keep record of the immobile heritage. For each day there is a

growing amount of cultural heritage information that can be accessed via portals and search engines, without that the users have to know at which institute the original files are kept. Developed interoperability and harmonisation of data are necessary prerequisites for this development, and there is still a lot to do in this field. In order to be able to use the often scattered information in a meaningful context the terminologies have to be sorted out and modeled. With harmonisation of different ontologies, the quality and cross domain use of heritage information can be radically increased.

With harmonised data models as a fundament, the rapidly evolving semantic web can truly elevate the use of heritage information to new levels.

An issue high on the heritage agencies' and institutions' wishing list, when Digisam started, was a national coordination and assignation of roles concerning authority files. Some of the categories of authorities could probably be assigned to a specific agency or institution, but it will obviously be difficult to find a managing partner for all of them. In the light of the ongoing development of the semantic web, Linked Open Data seems to be the most fruitful way to go.

Under 2013 Digisam managed a project called DISKA (Digital Semantic Cultural Heritage Authorities), funded by VINNOVA, which is an authority with the task to develop Sweden's innovation capacity¹. During the project, an inventory of digital and analogue registers and authority files at the national heritage institutions was carried out. More than 120 such registers and files were identified. In the

¹ VINNOVA: <<http://vinnova.se/en>> [11.03.2014].

next phase of the project, researchers at KTH (Royal Institute of Technology) transformed a selection of the registers and files to RDF-format and published them as Linkable Open Data at KTH².

At the moment there is an ongoing discussion on how to bring the results from the project further in a broad open collaboration, with the aim to increase the amount of authority data in a cost effective way, and to go from linkable to Linked Open Data.



Fig. 2: Constructive discussions during a DISKA-seminar
(R. Källmann, CC-BY)

A bottom-up perspective has been an important focus in the DISKA-project. The goal of the project was to process authority files from the individual institutions perspective, to open up for user participation and to hold forward the importance of keeping track of who

² KTH web platform: <<http://kth.se/en>> [11.03.2014].

stands for which concepts and thereby also model the relationship if the reports are different, not to say this is right and this is wrong. In this way, the project has promoted a method of non-authorial authority files.

The values is in the use

The increasing amount of published Linkable Open Data also emphasizes that the information is valuable at the point when used. Until used, information is just building blocks waiting for someone to put them together to meaningful and contextual stories. It is therefore of outmost importance that the digital heritage information is clear and machine-readable licensed, as open and free for use as possible.

The user perspective is natural for GLAM institutions. They have always built their collections and holdings with future use in mind and with a focus on knowledge building and dissemination.

With the publishing of heritage data as Linked Open Data follows a huge potential for knowledge building and creative use from comparing and integration with data from other sectors. After all, heritage information is a partial set of the total amount of societal information and with a width and breadth that could be made useful in almost all other contexts.

There are many actors who can creat applications and make creative use of the digital heritage information and in the future most of them is to be found outside the heritage institutions. It is important to remember though, is that institutions databases are and will be the basis for all use. The value of well managed databases, with continuously enhanced quality of the data cannot be overestimated.

Digital cultural heritage for learning

Another area that Digisam operates on behalf of cultural heritage institutions is cooperation with KTH concerning technology enhanced learning (TEL). KTH has a long tradition of conducting research in this area and now wants to strengthen its position in association with cultural institutions. The reason for this is that KTH find digital heritage information being complex, having interoperational challenges and to a wide amount being free and open which makes it interesting to bring into the “laboratory”. The research field of TEL is naturally associated with the use and development of semantic tools and services.

In the rapidly developing web based information society, the tasks run horizontal and vertical at the same time and there are no specific limits between natural sciences, humanities and social sciences. If we want to achieve common as well as domain specific goals, collaboration is crucial. This is particularly true when it comes to heritage and learning. The responsibility for this area is in Sweden today divided between the Ministry of Culture and the Ministry of Education.

Cost Effective Workflows

To obtain the positive effects of open data, we need to reach a critical mass of information. This means that we have to promote mass digitisation. Mass digitisation is the only way through which we can increase the amount of digitised heritage information fast enough. There are of course collections and holdings that can't be

transported to a digitisation factory, but the option to mass digitise must always be considered and if proven not applicable the method used instead must be as effective as possible. We must not forget that also a broad public support through crowd sourcing can be extremely efficient, with a broad user engagement as a positive side effect.

With the equation we are set to solve in Sweden, to increase the amount of digitised collections and holdings, and to increase the data being preserved, used and re-used, there are really no other options at hand.

To be successful in the huge task of digitising our common heritage we need to collaborate and split tasks and responsibilities. The most cost effective way can not be that all institutions acquire their own high quality scanner, allocate resources for handling, calibration and set up stable work flows, resulting in production of digital files to a substantially higher cost. There is a lot to gain by sharing the common task and by distributing the responsibility.

European collaboration

The European heritage portal, Europeana, which is now moving in the direction of becoming a platform, has since the establishment connected to a European Council conclusion in 2008 been a driving force and a catalyst for change and collaboration between the member states. It is no exaggeration to say that digital heritage would be further from the agenda than what is the case today. There are and have been many projects, funded by the European Commission, that are related to Europeana and that have resulted in harvested

meta data, creative use and the development of services and tools to the benefit of all.

An important development is also the growing interest for collaboration between digital cultural heritage and the infrastructures for research information. The major research infrastructures are today at the forefront in the work to benefit from economies of scale in large interconnected digital infrastructures as a means to increase access for research and access to shared services. In Sweden, a participation in the EU project DC-NET led to discussion over cooperation between the national research and education network, SUNET and heritage institutions represented by the National Archives and Digisam. In ongoing EU project DCH RP (Digital Cultural Heritage, Roadmap for Preservation³) in which the National Archives and Digisam participate, positive experiences from DC-NET are brought forward with further collaboration between cultural heritage, e-Infrastructure, GRID networks and infrastructures for research information. From a Swedish perspective we intend to implement the road map for preservation, the major outcome from the project, on a national level. DCH RP project will evaluate opportunities to apply preservation services available through e-Infrastructures and GRID network, partly by using data sets from a number of Digisams participating authorities and institutions.

It's reasonable to assume that, from the research communities' point of view, there is a lot to gain from infrastructural cooperation. Beyond collections and holdings, The GLAM-institutions also hold

³ DCHRP: <<http://www.dch-rp.eu>> [11.03.2014].

research information. Accumulated, this information is a significant resource for research but also for broad public use and re-use where access and use of information can be controlled by permissions and licensing. A common infrastructure would reasonably be considerably more cost effective than a few parallel ones and probably lead to more research.

In collaboration between research infrastructures information and digital heritage could also services and tools developed for each area largely be commonly used. At European level, there is currently an initiative to create a common and interdisciplinary data infrastructure for research data using different services facilitating the handling of large data sets. The EUDAT project's mission is to create a common research infrastructure with general services (currently divided into four areas: "Find - Safe - Stage - Share") that can be implemented in different ways. Within the DCH - RP project, some of these services relevant for digital preservation of cultural heritage materials will be tested.

The role of the archives in the field of GLAM

The archives' role, competences, long tradition management, storage and preservation of the analogue and digital information, make them a perfect GLAM partner providing an infrastructure for information. When culture heritage collections and holdings are being digitised the result is digital master files and meta data. That is per definition archival material and should be treated as such. Digital heritage files don't differ from files from other sectors and much could be gained by applying the same regulations and recommendations on other data sets in which public information is used.



Fig 3: The value is in the use
(NASA Goddard Photo and Video, CC-BY)

Storage and preservation

Regarding issues of storage and preservation, the archives play an important role. Within Digisams task to present solutions for cost-effective storage and preservation of digitised collections and holdings, we are currently performing a pre-study together with seven heritage agencies and institutions.

Interviews with the institutions have so far given several interesting results. Several of the institutions don't see storage as a huge problem today since the amount of digitised cultural heritage information is still rather manageable. However, they anticipate a rapid growth of information, especially of Audio Visual and 3D-information, and they see a problem in allocating budget means to finance the managing of the growing amount of data.

Most of the institutions also manage their data within the collection management systems and have no solutions for long term preservation.

Since the type of databases and collection management systems used within archives, libraries and museums differ, the information is handled with different formats for the description of metadata, due to how the single system is constructed and what formats and standards it can handle. The same types of objects, for example maps, can be treated in different ways in the archives, museums and libraries.

The pre-study clearly showed a wish for common central solutions for storage and preservation. A need for a coordinated support, and to some extent regulations, for the use and implementation of standards was also clearly asked for. The conclusion could be made that the heritage institutions ask for more centralised solutions, coordination and steering regulations.

Archives in today's landscape of information and social media

Archive as a conceptual term is mentioned in many more ways today than before. For example, you often hear people refer to their use of different social media platforms as an archive of personal memories. Such kinds of archives are, beyond doubt, an unstable solution and the risk is extremely high for loss of valuable local history individual memories. But it is an interesting perspective that archiving of information is a topic for discussion. This development opens up for various opportunities of collaboration between the archives and the end users. There is a lot of contextual information that could be harvested and preserved in the archives and the users

could make significant contributions in different crowd sourcing projects.

Scalable solutions

Even though the goals for the national strategy and Digisams mission are focused on the governmental authorities and institutions, an important goal is that the results of the work and the proposed infrastructure will be scalable and usable for all heritage institutions. We believe that here is another important role to be played by the archives. There are seven regional archives that are part of the Swedish National Archive. They are all important actors in the regional GLAM networks and can as such be connectors to the national infrastructure and regional nodes for knowledge, development and support within the field of digitisation.

A critical mass of information

In order to reach the goals for the Swedish national strategy for the digitisation of our common cultural heritage we have to radically increase the amount of information being digitised preserved, used and re-used. When it comes to digitising the archival holdings it is often expressed that it is an impossible task to digitise each document. This may be true, but we know very little about the future technical and logistical solutions, and in a world where information must be digital to be relevant, we need a mind shift. If you instead say that the goal is to digitise everything, without taking into account whether it is possible or not, you start looking for new approaches and solutions. So there is really no other way than to say: Digitise it all!

Reconceptualising access: sustainability in the digital future

by Barbara Reed

Introduction

There is no dispute to be had with the proposition that the digital is the future and digitisation is a key strategy in opening up accessibility to our venerable paper legacy. As archivists we face particular challenges in this space. Every one of our individual repositories collections are unique. Every one has been valued and supported by a funding community, either implicitly supported by social participation, or increasingly explicitly in community archives. Our holdings run to kilometres of unique paper resources. Our funders are seeing the digital world and the power of making these resources available to the public, without necessarily appreciating the specific challenges that our profession face in making the paper material accessible. So, for me, the key challenge archival institutions (and archivists) face is to ensure we approach the digitisation agenda in a strategic way.

My contribution is based around 5 main points:

1. Public (and funders) assumptions
2. Priorities for digitisation
3. Archival descriptive practice
4. Be clever – use 3rd party arrangements strategically
5. Let it go – but what then is the role of archives

Each of these areas could support several papers, and I have to be brief, so this represents a summary of my views and outlines what I see as some of the core challenges that face our profession.

Public (and funders) assumptions

We will all have encountered simplistic reductionist comments, almost as throw-away lines: 'oh, just digitise it'. In previous technology lives the same cry was common for microfilm, so we should have experience at dealing with it. But my observation is that funders assume that quick digitisation is achievable and should be our key priority. This is reinforced by calls from campaigns such as the 'Yes, We Scan'¹ project in the US. On the one hand, they ask very good questions about priorities. On the other hand they feed popular assumptions that everything is going to be digital. Business cases containing costings involved in digitising the archival record of archives in Europe have been conducted. In 2010 a study by the Collections Council estimated the cost of digitising Europe's archives: approximately the same cost as the research and development budget for the Joint Strike Fighter programme.² Clearly the figures are extrapolated from assumptions but they do provide a scale of expenditure which is not impossible (although perhaps unlikely).

¹ Yes We Scan <<https://yeswescan.org/>> [05.03.2014].

² Nick Poole, 'The Cost of Digitising Europe's Cultural Heritage' The Collections Trust, November 2010, available at <http://ec.europa.eu/information_society/activities/digital_libraries/doc/refgroup/annexes/digiti_report.pdf> [05.03.2014].

Not that archives and archivists oppose digitisation, but we certainly do have a richer appreciation of some of the complexities. Our realities include tight budgets and hundreds of kilometres of paper accumulated over time. We have specific challenges, which include, but are not limited to:

- Pages of a document rendered as a single object are not the same as that document in a network of related documents (that is, records not document objects).
- Some documents are not able to be found/interpreted without significant reworking of the way their original (paper based) systems supported them through contemporary webs of control records– and then the question becomes whether we are accurately representing the record, and whether that matters in the digital age.
- Many of our records pre-date automation. Handwriting is now seen as a barrier to interpretation.
- Technology for converting handwriting to digital bits and bytes is always improving, but currently Optical Character Recognition (OCR) technology needs to be supplemented by human quality checking to ensure the right meaning is conveyed.

We are not alone in facing these challenges, but archival institutions do need to accept that these are challenges of particular importance to them. They create quite legitimate problems for archival institutions when faced with simplistic reductions that digitisation means access.

In fact most attention to digitisation in Australia is not about digitising the existing holdings of our archival institutions, but rather focussing on eliminating the paper barriers to digital working in current business environments. An example is the National Archives of Australia's current campaign 'Digital Transition' which establishes the goal that the majority of Australian Government records will be created, stored and managed digitally and, where possible, incoming paper records will be scanned so that new paper files are not created'.³ Similarly standards for digitisation, some of which I have written,⁴ within Australasia focus less on the existing paper records in archives, and much more on enabling digital working in current business offices.

Wherever we face the challenge of digitisation, some clear trends are observable. Where previously archivists could provide hand-crafted responses to preferred clients, such as academic researchers, now digitisation and the internet age clearly shift the focus to anyone and everyone. In that dramatically democratised environment, what in our archival practice scales to the challenge and what will need to be recreated for it? Is this a distraction from our core professional responsibilities or an amazing opportunity to engage with a whole new audience for our resources?

³ National Archives of Australia 'Digital Transition Policy' <<http://www.naa.gov.au/records-management/digital-transition-policy/>> [05.03.2014].

⁴ For example, Archives New Zealand, Digitisation Standard No 6 (now replaced by ISO TR 13028) and National Archives of Australia, Digitisation Guide, 2010.

Priorities for digitisation

The emphasis of the digital environment of the past 5 years has been leading towards a 'personalised' future. The records that contain details of individuals are clearly those in greatest demand as can be seen by ballooning usage of genealogical resources and the success of releasing digitised census material where the records were so popular the best planning underestimated user demand on release.

In the homogenised environment of the internet, archives relation to location, events and people have something unique to offer. This distinguishes us from many of our library colleagues who are struggling with roles in the world of ubiquitous 'google' resources. Our holdings are very specific to communities. We should use this as a major strength when assessing our role in the digital future. An emphasis on people, places and events connect what we have in archives to the emerging world of the semantic web. Following this advantage in determining digitisation priorities may present huge strategic advantage into the future.

However, if we reconceptualise our thinking to prioritise people in our records, then we must take seriously our responsibilities in relation to personal information. Digitising recent records, or enabling access to recent records (such as war service records this century) lift the veil of what has been called 'practical obscurity'. Practical obscurity is the principle that private information in public records is effectively protected from disclosure as the result of practical barriers to access, including access include travel to view the

record, the passage of time, and the limits of indexing.⁵ Digitised images on the internet lose this protection and the consequences must be considered for personal information.

Privileging personal information is also a challenge to our traditional appraisal techniques which have typically privileged the records of policy or high officials, rather than the mass of 'case files' recording the detail of individual citizens interaction with the state. How much of this was technological – that is, the considerations of volume, redundancy and space available – that drove those decisions? Many of those arguments are invalidated in the relative abundance of capacity in the digital present and future. Just what should we be taking into our archival holdings?

These changes will clearly create challenges for current government programs and archival thinking alike. If people are the centre, how do we establish and maintain trust of the 'data subjects' while enabling the creative reuse of these records for multiple future purposes. This links us very closely to current debates and controversies of how to document, manage and enable access to personal and private information. One of our critical roles becomes as trusted agents – that ensure personal information is protected. Can archives be trusted? How do we position ourselves (or perhaps maintain our already nascent position) as institutions of trust in the digital world.

⁵ Society of American Archivists, 'Glossary of Archival and Records Management Terminology', available at <<http://www2.archivists.org/glossary/terms/practical-obscure>> [05.03.2014].

At least in my country our public access frameworks are fundamentally breaking down under the digital onslaught. Archival legislation provides rights of access after the passage of elapsed time (30 years being reduced to 15 year by ministerial fiat). Privacy protection and the rights of individuals in their own information is managed under different legislation. Freedom of information regimes bring another slew of legislative provisions. Open data initiatives are actively being pursued in most governments without any particular mandate. Sharing data and cross data border flows are managed under a variety of jurisdictional arrangements. The gaps in the access frameworks are beginning to show quite badly, resulting in an environment that is potentially so complex and fragmented that both liberal and restrictive actions relating to access can simultaneously be supported in a single case. The digital world is making the fractures obvious. The issues are being dealt with as social issues of some concern. To assume or demand a position of authority in the digital world, archival institutions should be active participants in such debates.

As we observe the digital environment can we see the beginnings of a shared ownership and responsibility for personal information? An emerging capacity to distribute rights of ownership and determinations on management of personal information shared between the person who is the subject of the information and the organisation that asks for that information, makes the record. We haven't begun to think this through professionally, but it is a likely prospect for a proportion of the record that we need to consider 'archival'.

Arrangement and description

How well do our existing models of arrangement and description support digitisation. Are we leveraging off our valuable intellectual constructs in descriptive systems to provide the contextual information to those that want to pursue it. Are we able to reconceptualise our archival descriptive systems as more than passive and static constructs, to enable the connective relationship based web of records. Linking to the horizontal connections, not simply the hierarchical, again positions archival descriptive data at the core of interesting semantic web possibilities. That arrangement and description is needing a shake up, is recognised partially by the establishment of the Expert Group on Archival Description of the ICA. Many await its recommendations with great interest.

In the absence of good solid conceptual models which enable relationships and linking to our archival systems, we often find images disconnected from their meaning and isolated from their connections. At best we can see archives linking images to their lowest level item descriptions, which certainly maintains them in their contextual stream, but fails to present the material in user friendly ways. Indeed a member of National Archives of Australia staff described watching users interact with their 'records search' application as akin to 'watching a train wreck'⁶. So is it one or the other, records digitised and isolated from their context OR presented within a system the general users find arcane and difficult. Surely we can be more creative than this!

⁶ Zoe D'Arcy, tweet from Digital Treasures Symposium, Canberra, 21 June 2013.

And all this leads of course to the capacity of our archival descriptive systems to manage born digital records. Recent estimates within the federal government of Australia estimated that there is, in 2014, 10.7 million gigabytes of digital records in Australian government.⁷ We know that there are vast accumulations of metadata being collected and processed in various government agencies. How well then, are our expectations that such records will be transferred to archival custody holding up. If we transferred even the 1-5% rule of thumb for records of archival retention, would our systems cope? Are our systems capable of scaling? Are our professional methods capable of scaling. This problem has been raised to us as a key professional challenge for the past 25 years, but we seem to be able to just keep on doing what we always did.

These questions of scale, of presentation, of linkage to systems and relevance to society become core existential questions for archives. If archival institutions do not really focus on what they do, what they provide, what value they add, and how archives are different (or the same) as other information resources, and articulate these with vigour and passion, they are at risk from misplaced public perception. Without this clear articulation, clear strategic approaches and embracing of different ways of delivering our services or materials, archives and archivists will be under threat and seen as unable to participate appropriately in the digital environment. Are archives to be considered as just like other resource providers – just putting

⁷ National Archives of Australia 'Changing to a digital environment' Media Release 14.03.2012 available at <www.naa.gov.au/about-us/media/media-releases/2012/5.aspx> [05.03.2014].

stuff out there? What is distinct about providing digitised archival records and how well are we articulating this?

We see illustrations of just this type of external determination of the role of archives in the 'forced' amalgamations of the National Archives and Libraries in Canada, New Zealand, the Netherlands, Ireland and more. That many of these forced amalgamations have been unsuccessful provides us with a renewed opportunity to articulate what is the same between members of the GLAM (Galleries, Libraries, Archives and Museums Sector) and what is different. Focussing on strengths of various professional approaches seems obvious, and research is emerging to differentiate archival strengths in the areas of accountability, evidential records, research and preservation.⁸ We possess these strengths not to the exclusion of others, nor uniquely. Research is showing however, that these are our professional strengths, so how should we use them?

Being strategic with commercial partnerships

As a profession we failed to understand and grasp an opportunity with digitisation and genealogists. That has now become an amazingly lucrative business operated by Ancestry.com. Were archives ever positioned to do what Ancestry.com has succeeded in doing? It is a moot point, as we certainly didn't. In fact, we willingly give away our records to be digitised to such organisations, with varying levels of cleverness and forward thinking in relation to the rights to

⁸ Shannon Wellington, 'Building GLAMour: Converging practice between Gallery, Library, Archive and Museum entities in New Zealand Memory Institutions', PhD thesis, Victoria University, Wellington, 2013.

publicly access such records as should be available to taxpayers. Basically archivists are bad negotiators of commercial rights and while there are significant discussions to be had on pros and cons of some of the arrangements for commercial digitisation in exchange for exclusivity on the web, by and large, we need to professionally revisit whether these arrangements are in archives best interests, in the best interests of the public and whether we can explore other, perhaps more consistently applied models across the profession. Its big business that we are enabling here, with Ancestry.com valued at 1.6 billion dollars in 2010.⁹ There is absolutely nothing wrong with enabling new and innovative big business utilising government data. But do archivists really understand the models? Perhaps because of confidentiality restrictions in contracts with such companies there is dearth of professional discussion.

Let it go: but what then, is role of the archives?

We need to explore and understand what the digital age does to archival holdings. Digitised material once placed on the web is out there. There is no taking it back, nor any controlling of what is done with the material. That is the nature of the digital age. So what does this do to our closely held professional belief in custody and possession? If others can create more attractive portals or front ends to digitised archives is this something to be encouraged and supported? It seems ridiculous to suggest that it is not, but it raises significant

⁹ 'Ancestry.com sets \$1.6 Billion Deal' Wall Street Journal October 22, 2012 available at Wall Street Journal <<http://online.wsj.com/news/articles/SB10001424052970203406404578071390645955994>> [November 2013].

questions of just what the archival role is in the digital world. If others can digitise better, and others can provide reference services better, what then should we be concentrating on? Should we give up those roles which are assumed by others? Or do we provide an organisation struggling to provide minimum levels of service across a broad range of areas, constantly being measured by yardsticks that derive from specialist service providers.

There are reasons to be optimistic in this digital age. But if archives treat the digital age as the same as the paper age and just keep doing what they have always done, then the future does not look so bright. If we can reconceptualise ourselves as providers of context, as authenticators of records, as custodians of last resort, as metadata suppliers to a distributed web of digital and digitised resources, we may look quite differently on the future.

How well we approach digitisation of our holdings in the face of the intense pressure to reprioritise archival activity to support digitisation, is perhaps a critical moment for archival institutions. The reappraisal inherent in working out what to digitise can, and how to link digitised material to our existing systems, how to scale our descriptive system to deal with vast quantities of digitised and born digital records all provide us with huge professional challenges. We must craft and tell a credible and sustainable tale that is bigger than the simple equation that digitisation = access. Doing so is critical for our professional future.

Analoge Bildarchive auf dem Weg ins digitale Zeitalter – Chancen und Herausforderungen für die Bereitstellung und Benutzung bildhafter Materialien

von Elke Bauer

1. Seit nunmehr einem Vierteljahrhundert rücken Bilder zunehmend ins Zentrum wissenschaftlichen Interesses der Kultur- und Geschichtswissenschaften. Zunächst eher zögerlich mehren sich in den letzten Jahren sowohl die einschlägigen Tagungen wie Publikationen zum Thema. Bilder allgemein, aber insbesondere Fotografien finden nicht mehr nur als Illustrationen zu Texten Verwendung, sondern sind immer öfter selbst Untersuchungsgegenstand. Parallel zu dieser Entwicklung hielt die Digitalisierung Einzug in die Welt der Bilder. Die analoge Fotografie gehört in der Zwischenzeit beinahe zur Gänze der Vergangenheit an, die Digitalfotografie beherrscht den Markt. Die institutionellen Einrichtungen wie Archive und Sammlungen, die Bildzeugnisse bewahren, werden zunehmend mit diesen digital produzierten Bilddateien konfrontiert. Hinzu kommt die Digitalisierung der analogen Bestände. Alle diese Entwicklungen führen und führten bereits zu großen Umbrüchen innerhalb der Bildzeugnisse aufbewahrenden Einrichtungen.¹

¹ Wenn im weiteren Verlauf von Bildarchiven die Rede ist, so meine ich Institutionen allgemein, die Bilder aufbewahren, die für vor allem geistes- und gesellschafts- und kulturwissenschaftliche Fragestellungen relevant sind und deren Zielgruppen in ihrer „analogen“ Form bisher vor allem die Wissenschaft und thematisch Interessierte bildeten.

Dieser Beitrag basiert auf meiner derzeitigen Untersuchung zum Übergang analoger Bildarchive ins digitale Zeitalter – berichtet also aus einer „Work-in-progress“-Situation. Zu den wichtigsten Fragen des Projekts gehören: Hat der so genannte *Visual Turn* innerhalb der Geisteswissenschaften Auswirkungen auf die Onlineangebote institutioneller Bildarchive? Was sollen und können Onlinebildangebote aus Sicht der Institutionen leisten, was sollten und können sie aus Sicht der Nutzerinnen und Nutzer leisten? Decken sich die Ansprüche der Onlinebildangebote mit denen, die sie nutzen, oder sind die Ansprüche konträr zueinander? Verändern Onlineangebote das Nutzerverhalten, und wenn ja, hat dieses veränderte Verhalten wiederum Auswirkungen auf die Gestaltung der Onlineangebote? Entwickeln sich allmählich Standards zur Bilderschließung in der digitalen Welt?

1. Von der Karteikarte zum digitalen Wissensspeicher?

Meine Ausgangsthese ist, dass die Institutionen, die Bilder ins Netz stellen, zunächst an eine digitale Karteikarte dachten. Onlinekataloge wurden und werden als digitales weltweit schnell zugängliches Nachweisinstrument konzipiert. Die Einrichtungen werden dadurch weltweit sichtbar und wahrgenommen. User können schnell erkennen, ob ein Bestand für sie von Interesse ist oder nicht. Zu untersuchen wäre, ob dieser Wandel das Nutzungsverhalten maßgeblich ändert. Kommen nun weniger Menschen in die analogen Archive, da sie die Bestände bequem von zuhause aus betrachten können? Werden die Onlinekataloge zunehmend dem Archivbesuch gleichgesetzt? Dies müsste untersucht werden. Möglicherweise finden sogar mehr Personen den Weg ins Archiv, da sich der Bekanntheitsgrad der

einzelnen Einrichtung durch die Internetpräsenz erhöht. Und auch in Zeiten vor dem Internet gab es – der heutigen Onlinerecherche entsprechend – schriftliche Anfragen, denen z.B. von Archivseite mit einer Liste der vorhandenen Bilder oder Papierkopien derselben zur Voransicht geantwortet wurde. Auch dieser Personenkreis kam im Zweifelsfall nicht ins Archiv, sondern bestellte Abzüge aus der Ferne. Hat sich somit nur die Quantität der Nutzung, aber nicht die Qualität geändert?

Die Textinformationen zum Bildinhalt und der digitalisierten Bildvorlage, die in den meisten Onlinebilddatenbanken zu finden sind, gehen nicht über die spärlichen Informationen der klassischen Karteikarten, der Zettelkataloge hinaus. Textliche Zusatzinformationen wie Bildrückseiten oder nur mittelbar mit der Vorlage verbundene Texte, wie z.B. Diavorträge oder Schriftstücke, in deren Kontext die Aufnahme entstand oder verwendet wurde, sind nur in wenigen Ausnahmefällen zu finden. Auf möglicherweise vorhandene weitere Ausfertigungen einer Bildvorlage wird nicht verwiesen. Der Nutzer erfährt somit nicht, ob neben einem digitalisierten Negativ noch zeitgenössische Abzüge existieren, die zum Beispiel durch Beschriftungen Hinweise auf die Bildentstehung oder -rezeption geben. Informationen, die man bei Betrachtung eines zeitgenössischen Abzugs im Archiv erhalten würde. Es ist fast paradox: Während man im Archiv selten die Negative zu Gesicht bekommt, werden einem im Netz eher die Abzüge, zumindest die Rückseiten oder Untersatzkartons vorenthalten. Sind wir also wirklich auf dem Weg von der analogen Karteikarte zu digitalen Wissensspeichern, die rein theoretisch möglich und sicherlich wünschenswert wären? Oder bleiben die Möglichkeiten, die ein Onlineangebot beispielsweise mit

Verlinkungen, Hypertexten oder PDF-Beigaben bieten könnte, doch zum Großteil blanke Theorie, da die Umsetzung an den geringen Personalkapazitäten und damit in letzter Instanz am Geld scheitert? Bildarchive gehören nicht zu den finanziell üppig ausgestatteten Einrichtungen. Dabei gibt es von Seiten der Wissenschaft in Zeiten des *Visual, Pictorial* oder auch *Icon Turns*² den durchaus begründeten Vorwurf, Bilder würden in den Archiven prinzipiell zu flach erschlossen und bedürften der besseren Kontextualisierung.³

2. Digitalisierung: Vom dreidimensionalen Objekt zur zweidimensionalen Bilddatei

Was passiert aber nun bei der Onlinestellung von bildhaften Materialien? Zunächst werden analoge Vorlagen digitalisiert und für die Onlinedatenbank aufbereitet, sprich komprimiert. So wie die analoge Fotografie den fotografierten Gegenstand seiner Haptik, seines Geruchs, seiner Umgebung beraubt, so beraubt die Digitalisierung

² William John Thomas Mitchell, *The Pictorial Turn*, in: Gustav Frank (Hrsg.), *Bildtheorie*, Frankfurt 2008, S. 101-135; Gerhard Paul, *Von der Historischen Bildkunde zur Visual History. Eine Einführung*, in: Gerhard Paul (Hrsg.), *Visual History. Ein Studienbuch*, Göttingen 2006, S. 7-39; Horst Bredekamp, *Drehmomente – Merkmale und Ansprüche des iconic turn*, in: Hubert Burda/Christa Maar (Hrsg.), *Die neue Macht der Bilder*, Köln 2005, S. 15-26.

³ Vgl. Jens Jäger, *Fotografie und Geschichte*, Frankfurt am Main 2009, S. 193; Christoph Hamann, *Visual History und Geschichtsdidaktik. Bildkompetenz in der historisch-politischen Bildung*, Berlin 2007, S. 18, <http://opus.kobv.de/tuberlin/volltexte/2007/1536/pdf/hamann_christoph.pdf> [04.03.2014]; Miriam Y. Arani: *Fotografische Selbst- und Fremdbilder von Deutschen und Polen im Reichsgau Wartheland 1935-1945. Unter besonderer Berücksichtigung der Region Wielkopolska*, 2 Bde. Hamburg 2008, Bd. 1, u.a. S. 72, 77.

wiederum das Objekt Fotografie, Grafik, Illustration etc. genau dieses Kontextes. Übrig bleibt eine digitale Datei, die bei entsprechender Soft- und Hardware auf dem Bildschirm als zweidimensionales Bild sichtbar gemacht werden kann, die Dreidimensionalität der Bildvorlage verschwindet. Die Größe relativiert sich, ob die digitalisierte Vorlage ein Kleinbildnegativ oder ein großformatiger Abzug war, ist nicht mehr ersichtlich. Es gibt zunächst keinerlei zusätzliche Information, die Bildvorlage wird auf den Bildinhalt reduziert, also genau auf das, was der „neue“, quellenkritischere Umgang mit Bildern gerade mühsam zu ändern versucht. Gleichzeitig oder auch zeitlich versetzt mit der Herstellung des Digitalisats werden Daten zur Bildvorlage und Bildinhalt in einer Datenbank erfasst. Leider gibt es hierfür noch keine verbindlichen Standards, weder, welche Daten dringend erforderlich wären, noch in welcher Form diese aufgenommen werden sollten – obwohl es entsprechende Handbücher zum Aufbau von Bilddatenbanken schon recht lange gibt.⁴ Angaben zum Digitalisat

⁴ Genannt seien hier: Kathryn Pfenniger, *Bildarchiv digital*, Stuttgart 2001 (Museumsmagazin 08) und Edwin Klijn (Hrsg.), *SEPIADES – Recommendations for cataloguing photographic collections. Advisory report by the SEPIA Working Group on Descriptive Models for Photographic Collections*, Amsterdam 2003, <<http://www.ica.org/download.php?id=1593>> [04.03.2014].

Zur Vorgabe wie digitalisiert werden sollte, gab es auch recht früh Handreichungen wie z.B. das von der Minerva Arbeitsgruppe 6 herausgegebene *Good Practice Handbuch für Digitalisierungsprojekte* Version 1.3, o.O. 2004, <http://www.minervaeurope.org/structure/workinggroups/goodpract/document/handbuch1_3.pdf> [03.03.2014]. In der Schweiz bemüht sich die staatlich finanzierte Einrichtung „Memoriav“ (<<http://de.memoriav.ch/>> [04.03.2014]) einheitliche Standards durchzusetzen und Bildarchiven bei Digitalisierungsprojekten behilflich zu sein. In Deutschland versucht u.a. die DFG mit ihren Praxisregeln

selbst finden erst langsam Eingang in die Bilddatenbanken, dabei handelt es sich in den allermeisten Fällen um die Angabe zur Auflösung der Voransicht und/oder der Auflösung, in der ein Bild bestellt bzw. direkt heruntergeladen werden kann. Die Angaben zur analogen Bildvorlage sind sehr unterschiedlich, je nachdem, ob es sich um dokumentarische Fotografie oder um künstlerische Vorlagen wie Druckgrafik, Gemälde oder Zeichnungen handelt. Bei künstlerischen Vorlagen finden wir z.B. Angaben zur Herstellungstechnik, zur Größe, zum Material, zum Urheber, bei fotografischen Vorlagen reduzieren sich die Notate mehrheitlich auf Größe, Fotograf (falls ermittelbar) sowie den Hinweis, ob es sich um ein Negativ oder Positiv handelt. Nähere Hinweise zur Herstellungstechnik und Trägermaterial werden selten angegeben – am häufigsten in Museumsdatenbanken im Falle von Fotokunst. Zeigt die digitalisierte Fotografie ein Kunstwerk, finden sich oft noch Hinweise zur Größe, Herstellungstechnik und Urheber desselben. Bei all diesen verschiedenen Angaben ist es für die Nutzung eines Angebots wichtig, auf den ersten Blick zu erkennen, ob die jeweiligen Einträge dem Digitalisat, der analogen Vorlage oder dem Bildinhalt zuzuordnen sind.

3. Bilder online: Findmittel oder „virtueller Archivbesuch“

Bilder sind allgegenwärtig und durch das Internet in bisher nicht gekanntem Maße auffindbar, sichtbar, auswertbar. Wir können in einer vor wenigen Jahren noch unvorstellbaren kurzen Zeit weltweit

Standards zu setzen, zumindest bei den von ihr geförderten Projekten, siehe: DFG-Praxisregeln „Digitalisierung“ 02/13 (DFG-Vordruck 12.151 – 02/13), <http://www.dfg.de/formulare/12_151/12_151_de.pdf> [03.03.2014].

Archivbestände recherchieren und im Falle der Bilder zudem betrachten. Da Onlineangebote somit auch als Werbefenster fungieren und auf Bildbestände aufmerksam machen, steigen die Nutzungszahlen. Dies führt dazu, Bestellvorgänge zu rationalisieren – automatisierte Bestellvorgänge oder Downloadmöglichkeiten sind in der Zwischenzeit eher der Normalfall, denn die Ausnahme. Damit wird aber der direkte Kontakt mit dem Archivpersonal (und sei es per E-Mail) weitestgehend gekappt und die Bildarchive müssen sich spätestens jetzt fragen, welchen Anspruch sie allgemein und mit ihren Onlineangeboten im Speziellen verfolgen. Wollen sie ihre Bildinhalte einfach nur vermarkten (was in Zeiten des *open access* immer schwerer werden wird) – dann sind die spärlichen Metadaten, die den Digitalisaten beigelegt sind, sicherlich ausreichend. Oder ist es wie im analogen Zeitalter ihr primäres Ziel, die Bilder der Forschung und einem allgemein interessierten Publikum zur Verfügung zu stellen, – dann wäre über ein wissenschaftsbasiertes Onlineangebot nachzudenken.

Die Karteikarte war der Zugang zum Original, der Onlinedatensatz ist der Zugang zum Digitalisat einer Bildvorlage. Die Karteikarte diente somit einem anderen Zweck als der zum Digitalisat gehörende Datensatz, zumindest wenn der Zugang zur digitalen Kopie immer häufiger den Gang ins analoge Archiv ersetzt und die Onlinedatenbank *peu à peu* – ob ursprünglich beabsichtigt oder nicht – zum „virtuellen Archiv“ konvergiert. Wenn die Nutzer sich zunehmend auf die Bilddatenbanken beschränken, dann wird neben der Relevanz der Qualität der Metadaten auch bedeutsam, dass in der Regel nur ein Bruchteil der real existierenden Bildbestände der einzelnen Archive bzw. Sammlungen online gezeigt werden, sei es aus rechtlichen Gründen oder einfach, weil die große Masse noch gar nicht digital

vorliegt. Es besteht die Gefahr, dass nur noch die online gestellten Bestände gefunden und nachgefragt werden. Die Möglichkeit, sich Bilder direkt herunterzuladen oder zumindest online bestellen zu können, wird zu dieser Entwicklung ihren Teil beitragen.

Aufgrund der genannten Problemfelder sollten sich Bildarchive ihrer Verantwortung bewusst sein und genau überlegen, welche Bestände sie zunächst ins Netz stellen und mit welchen Informationen, welche später und welche vielleicht nie. Wünschenswert wäre, wenn die Onlinedatenbanken neben der Einzelbilderschließung auch Beschreibungen derjenigen Bestände enthielten, die noch nicht online gefunden werden können. Das SEPIADES-Handbuch⁵ stellt in dieser Hinsicht ein hilfreiches, wenngleich leider nicht häufig angewendetes Instrument dar. Es bietet durch sein gestaffeltes Vorgehen der Erschließung die Möglichkeit, der Bildermasse in den Archiven Herr zu werden. Beginnend von der allgemeinen Beschreibung eines Gesamtbestandes (zum Beispiel einer Institution, die neben Bildern auch andere Bestände beherbergt), über einzelne Sammlungsbeschreibungen bis hin zur Einzelbilderschließung. Bei der Einzelbilderschließung wird klar zwischen Angaben zur Bildvorlage und zum Bildinhalt unterschieden. SEPIADES gibt auch Vorschläge zur Erfassung von Fotografien, die bereits digital erzeugt wurden. Zudem wird unterschieden, ob eine Bilddatenbank nur den Kontakt nach außen regeln soll oder eben für alle Vorgänge vorgesehen ist, die mit einem Bestand zu tun haben, inklusive Akzession, Rechteverwaltung, Bestellvorgängen und Rechnungsstellung.

⁵ SEPIADES (wie Anm. 4).

4. Wer nutzt die Onlinedatenbanken? – Der User, das unbekannte Wesen

Museen nutzen ihre Onlinedatenbanken, um die Schätze ihrer Depots der Öffentlichkeit zugänglich zu machen.⁶ Schätze, die sie in ihren normalen Ausstellungsräumen nur zu einem kleinen Teil präsentieren können. Archivbestände ruhen per se im Magazin, sind jedoch durch Findmittel für Interessierte auffindbar und werden bei entsprechender Anfrage vorgelegt, dem normalen Museumsbesucher bleibt der Blick ins Depot hingegen verwehrt. Trotz dieser unterschiedlichen Ausgangslage ähneln sich die Bilddatenbanken im Aufbau: Sowohl Archive als auch Museen zeigen Bilder, fügen ihnen Datensätze hinzu, um sie auffindbar und nutzbar zu machen.

„Online-Kataloge stehen vor der Aufgabe, ihr Publikum aufzuspüren und gezielt anzusprechen“, sagte Regina Doppelbauer, Leiterin der „Albertina-Sammlungen online“ im November 2013 auf der EVA-Tagung in Berlin.⁷ Für den Aufbau eines Angebots scheint es deswegen unabdingbar, dass sich die Archive wie die Museen über ihre Zielgruppe(n) bewusst sind. Im Normalfall macht es keine

⁶ Als Beispiele seien genannt: Staatsgalerie Stuttgart <<http://www.staatsgalerie.de/onlinekatalog/>> [04.03.2014], Tate Gallery London <<http://www.tate.org.uk/art>> [04.03.2014], Rijksmuseum Amsterdam <<https://www.rijksmuseum.nl/en/explore-the-collection>> [04.03.2014] oder die Albertina Wien (<<http://sammlungenonline.albertina.at/>> [04.03.2014]).

⁷ Regina Doppelbauer, Vom Publikum her denken. „Sammlungen online“ der Albertina. Aktuelles Selbstverständnis und Zukunft von online-Katalogen, in: Konferenzband EVA-Berlin 2013, „Elektronische Medien & Kunst, Kultur, Historie“, Berlin 2013, S. 79.

Schwierigkeiten, sich über die ursprünglich avisierte Zielgruppe einer Einrichtung im Klaren zu sein, oftmals ist diese sogar in den Satzungen nachzulesen: Archive sind zumeist, wie bereits eingangs erwähnt, für die Wissenschaft und thematisch interessierte Personen da, beiläufig verirrt sich kaum jemand in sie. Gänzlich anders verhält es sich im Internet, hier kommen neue Nutzergruppen hinzu. Zum Fachpublikum gesellt sich nunmehr der so genannte „Flaneur“, der User, der sich mehr oder minder zufällig auf die Seite verirrt. Soll das Onlineangebot auch ihn abholen? Letzterer erwartet einen leichten Zugang zu den Materialien, der nicht bei allen Datenbanken gewährleistet ist. Es heißt also eine Entscheidung treffen, ob man nur seine Zielgruppe erreichen möchte oder aber den Personenkreis insgesamt, der das Onlineangebot findet. Dieser Personenkreis, der mit der alten Zielgruppe nur noch zum Teil identisch ist, ist sehr schwer zu analysieren, denn er bleibt mehr oder minder anonym.⁸ Die Londoner Tate Gallery⁹ fährt zweigleisig: Sie bietet neben der Expertensuche ein Flaneurangebot mittels Bildkacheln an, die man durchstreifen und bei Interesse anklicken kann, um mehr zu erfahren. Dieses „Mehr“ ist wiederum in zwei Stufen unterteilt, neben umfangreichen Metadaten werden bei entsprechendem Klick zum Beispiel Katalogeinträge angezeigt. Zusätzlich gibt es bei jedem Bild das Angebot einer weiteren

⁸ Umfragen werden nur selten beantwortet, die Feedbackfunktionen viel zu selten benutzt, automatische Abfragen und Social-Media-Funktionen bieten auch keine wirkliche Lösung des Problems.

⁹ Vgl. <<http://www.tate.org.uk/art>> [04.03.2014].

Leiste mit Bildkacheln, für die sich der Besucher darüber hinaus interessieren könnte – ähnlich wie bei Amazon.¹⁰

Neben der Definition der Ziel- oder eben auch Nutzergruppen, stellt sich die Frage, inwieweit ein Onlineangebot auch auf die Bedürfnisse der sich nun formierenden *Digital Humanities* reagieren sollte, sprich: die Auslesbarkeit der Daten? Kontrolliertes Vokabular und kontrollierte Metadaten, die sowohl von den *Digital Humanities* genutzt werden können als auch der Einbindung in größere Portallösungen wie der Deutschen Digitalen Bibliothek und der Europeana dienen, sind hier gefordert.

5. Vom Wünschenswerten zum Machbaren

Doch können die Einrichtungen, die hinter den Angeboten stehen, all diese Dinge berücksichtigen? Ist all dies überhaupt realistisch, wenn Bestände, die in die Millionen Bilder gehen können, von wenigen Personen gehandhabt werden müssen. Archive wie zum Beispiel das Aargauer Staatsarchiv in der Schweiz, das ein gesamtes analoges Pressebildarchiv mit rund sieben Millionen Bildträgern unterschiedlicher Materialität übernommen hat,¹¹ müssen vermehrt über Standards der Kassation bei Bildbeständen nachdenken.¹² Aber auch

¹⁰ Vgl. <www.amazon.de> [04.03.2014].

¹¹ Nora Mathys, Ringier Bildarchiv. Eine Chance für die Geschichtsschreibung der Schweiz, in: *Traverse*. 18, 3, 2011, S. 7-14.

¹² Nora Mathys/Andrea Voellmin/Walter Leimgruber (Hrsg.), *Über den Wert der Fotografie. Zu wissenschaftlichen Kriterien für die Bewahrung von Fotosammlungen*, Baden 2013; Michel Pfeiffer, *Visuelle Überlieferungsbildung – Neue Sammlungs- und Bewertungsperspektiven oder nur alter Wein in neuen Schläuchen?*, in: Irene Ziehe/Ulrich Hägele (Hrsg.), *Fotografie und Film im Archiv. Sammeln,*

wenn von den sieben Millionen Bildern letztendlich „nur“ vier Millionen übrig blieben, dass auch diese nicht annähernd vollständig in der Tiefe erschlossen werden können, versteht sich von selbst.

Die Digitalisierung schont die Originale, sie macht Teile der Bestände weltweit recherchierbar, sie ermöglicht die Automatisierung der Bestellvorgänge, aber sie erfordert auch – zusätzlich zur Neuerschließung – die Retrokatalogisierung analog bereits erschlossener Bestände, sie nimmt den Archiven die konservatorische Betreuung der Materialien nicht ab, sie benötigt neben dem physischen Aufbewahrungsraum Speicherkapazitäten für die digitalen Abbilder, die bei hochauflösenden Scans nicht unerheblich sind und sie stellt die Frage nach der Langzeitarchivierung. Gerade die Langzeitarchivierung stellt für digital erzeugte Fotografien ein noch größeres Problem dar als für die analoge, denn ist die Datei nicht mehr lesbar, ist die Fotografie verschwunden. So kommt es, dass die analogen Bestände hochauflösend digitalisiert und die digitalen Bestände analog hochwertig ausgedruckt werden.

Digitalisierung und Erschließung sollten im Idealfall synchronisiert werden, nur ein Digitalisat ist oftmals wesentlich schneller hergestellt als ein vernünftiger Datensatz. Hinzu kommt die Rechtklärung, die vor allem im Bereich der Dokumentar- und Alltagsfotografie eine Herausforderung darstellt, denn die Urheber und deren Nachkommen sind äußerst schwer zu ermitteln. Die Fotos im Archiv zu betrachten und auszuwerten, war auch bei urheberrechtlich geschützten Materialien kein Problem, die Onlinestellung und damit Publikation der

Bewahren, Erforschen, Münster u.a. 2013 (Visuelle Kultur. Studien und Materialien 6).

bildhaften Materialien ist es umso mehr, denn nicht einmal die kleinen Voransichten sind erlaubt. Und selbst bei Beständen, bei denen die aufbewahrende Einrichtung über die Nutzungsrechte verfügt, wird es problematisch, wenn sie vor dem 1. Januar 1966 in die Einrichtung kamen,¹³ denn in diesem Fall muss der Urheber ermittelt werden, da die damalige Übertragung der Nutzungsrechte die seinerzeit unbekannte Nutzungsart „Internet“ nicht mit einschloss.¹⁴ Eine schier unlösbare Aufgabe. Viele Fotografien sind nicht mit dem Namen ihrer Urheber verknüpft. Und wer fotografiert hat, ist häufig kaum mehr zu ermitteln. Alles also verwaiste Werke? – Ja, aber der Gesetzgeber schreibt vor, dass ein Werk erst nach gründlicher und entsprechend dokumentierter Recherche als verwaist deklariert werden kann – und dies bei hunderttausenden Fotografien? Onlinebilddatenbanken (auch mit öffentlichen Geldern geförderte) scheinen somit häufig am Rande der Illegalität zu agieren.

6. Und die Lösung aus dem Dilemma? – Think positive

Möglichst offen zu benennen, an wen sich ein Onlineangebot richtet und welchem Zweck es dient, könnte ein Lösungsweg aus dem Dilemma „Wunschdenken kontra Machbarkeit“ sein, d. h. offen zu benennen, welche Bestände wie tief erschlossen sind, welche digitalisiert und welche nur im Archiv einsehbar sind. Die Erschließung sollte über Beschreibungen des Gesamtbestands, einzelner

¹³ UrhG §137I Übergangsregelung für neue Nutzungsarten <http://www.gesetze-im-internet.de/urhg/_137I.html> [04.03.2014].

¹⁴ UrhG §32c <http://www.gesetze-im-internet.de/urhg/_32c.html> [04.03.2014].

Sammlungen und Sammlungsgruppen bis hin zur Einzelbilderschließung mit entsprechender Kontextualisierung gehen, damit sich Nutzerinnen und Nutzer über den gesamten Bestand, den eine Einrichtung bietet, einen Eindruck machen kann. Es kann sinnvoll sein, von einzelnen Beständen nur beispielhafte Bilder für das Onlineangebot auszuwählen, um eben diesen ersten Eindruck zu vermitteln.

Das die Onlinebilddatenbanken das Nutzerverhalten verändern werden und zum Teil bereits verändert haben, ist unbestritten. In welchem Maße genau, dazu fehlen die empirischen Untersuchungen. Eine sorgfältigere Kontextualisierung der Bildmaterialien wäre in jedem Falle wünschenswert. Es ist allerdings äußerst unwahrscheinlich, dass dies rückwirkend für die vielen Millionen Bilder, die von den einschlägigen Institutionen bereits ins Netz gestellt wurden, praktikabel ist.

Doch ist es überhaupt dramatisch, wenn nur rudimentäre Informationen zu den Bildern existieren? – Eigentlich nicht, wenn, wie oben bereits gesagt, die Angebote offenlegen, was sie leisten können und was eben nicht. Nutzerinnen und Nutzer müssen bei der Recherche rasch merken, dass sich weitere Informationen zu einem digitalisierten Bild im Archiv finden lassen und dass es möglicherweise weiteres Material zum gesuchten Thema gibt, sich also der Gang ins Archiv lohnt, denn wer nie ein analoges Archiv besucht hat, kann auch nicht einschätzen, was ihm an Wissen in der digitalen Welt entgeht.

The View from Outside: Amerikanische Erwartungen an das digitale Angebot deutscher Archive

von Rainer Hering

Die Archive in Deutschland werden nicht nur von historisch Interessierten aus dem eigenen Land, sondern gerade auch von vielen aus dem Ausland besucht. Während personenkundlich orientierte Benutzer in der Regel mehr über ihre aus dem deutschsprachigen Raum ausgewanderten Vorfahren erfahren möchten, stehen für die Wissenschaftlerinnen und Wissenschaftler Sachthemen oder Biographien der deutschen Geschichte im Vordergrund. Dabei bringen sie z.T. neue Fragestellungen ein, die die deutsche Geschichte unter anderen Aspekten und aus anderen Blickwinkeln mit zum Teil neuen Ansätzen erforschen und somit bereichern. In vielen Fällen wird auf Archivquellen zurückgegriffen, da sie gern exemplarisch lokale und regionale Themen verfolgen und in übergreifende Fragestellungen integrieren. *The View from Outside* bietet vielfach einen anderen Blick auf die Quellen und ist mit spezifischen Erwartungen an die deutschen Archive verbunden. Am Beispiel Nordamerika möchte ich das verdeutlichen.

1. German Studies

Zum Hintergrund ist es wichtig zu wissen, dass das interdisziplinäre Arbeiten in den USA und in Kanada in der universitären Praxis eine Selbstverständlichkeit ist. Deutlich wird das im Feld der German Studies, das alle zum deutschsprachigen Raum arbeitenden Disziplinen, die in Deutschland akademisch getrennt sind, zusammenfasst:

Germanistik, Geschichtswissenschaft, Politikwissenschaft, Soziologie, Archivwissenschaft, historische Musikwissenschaft, Kunstgeschichte, Kirchengeschichte, Medizingeschichte und Wirtschaft. Geographisch liegt der Schwerpunkt auf Deutschland, Österreich und der Schweiz, zeitlich auf dem 20. Jahrhundert, wobei das 18. und 19. Jahrhundert auch in den Blick geraten. In vielen Publikationen, aber gerade auch in Lehrveranstaltungen spielen jüdische Studien und Holocaust Studies eine besondere Rolle. Für deutsche Verhältnisse ungewöhnlich, werden vielfach Überblicksvorlesungen und -seminare angeboten, die ein breites thematisches Spektrum oder einen langen Zeitraum, z.B. Deutsche Geschichte vom 16. bis 20. Jahrhundert, umfassen.¹

Stilistisch werden auf Schreibweisen, die auf breitere Lesbarkeit abzielen, Wert gelegt. Historische Monographien, insbesondere Dissertationen oder das „zweite Buch“, das ansatzweise mit der deutschen Habilitationsschrift verglichen werden kann, sowie Sammelwerke werden zumeist als Bücher von renommierten (Universitäts-) Verlagen publiziert. Dabei unterliegen sie – ebenso wie Beiträge in

¹ Die englischsprachige Wikipedia definiert den Begriff so: "German studies is the field of humanities that researches, documents, and disseminates German language and literature in both its historic and present forms. Academic departments of German studies often include classes on German culture, German history, and German politics in addition to the language and literature component. Common German names for the field are Germanistik, Deutsche Philologie, and Deutsche Sprachwissenschaft und Literaturwissenschaft. In English the terms Germanistics or Germanics are sometimes used (mostly by Germans), but the subject is more often referred to as German studies, German language and literature, or German philology." (<http://en.wikipedia.org/wiki/German_studies> [11.11.2013]). Vgl. Peter U. Hohendahl, *German Studies in the United States. A Historical Handbook*, New York 2003.

Fachzeitschriften – einem Peer Review-Verfahren, bei dem für die Autorinnen und Autoren anonym bleibende externe Gutachter letztlich über die Aufnahme und über inhaltliche Veränderungen entscheiden. Die Rolle der akademischen Betreuenden tritt demgegenüber im Publikationsprozess zurück. Druckkostenzuschüsse sind – anders als in der Bundesrepublik – zumeist nicht erforderlich, das finanzielle Risiko tragen die Verlage.

2. German Studies Association

Die wichtigste Organisation im Bereich der German Studies in Nordamerika ist die German Studies Association (GSA).² Gegründet 1976 als Western Association for German Studies (WAGS), wurde sie 1984 in German Studies Association umbenannt. Der Mitbegründer Professor Gerald Kleinfeld (State University Arizona) war bis zu seiner Emeritierung 2005 Executive Director, seitdem hat diese zentrale Funktion Professor David Barclay (Kalamazoo College) inne. Die Präsidentschaft wechselt alle zwei Jahre. Entscheidungen trifft das GSA Board, dessen neun Mitglieder auf drei Jahre gewählt werden.

Die permanenten Gremien der German Studies Association sind das Nominating Committee, das Kandidaten für die Positionen innerhalb der Organisation ermittelt, das Program Committee, das aus den eingereichten Vorschlägen das Programm der jährlichen Konferenz

² GSA: <<https://www.thegsa.org/about/index.html>> [19.11.2013]. Eine Geschichte der German Studies Association liegt noch nicht vor. Die folgenden Angaben entstammen der Internetseite der Organisation sowie Gesprächen mit führenden Repräsentanten, wie z.B. dem Gründer Gerald Kleinfeld und dem derzeitigen Executive Director David Barclay. Der Verfasser ist seit 1993 Mitglieder der German Studies Association und nimmt seitdem an den Jahrestagungen teil.

zusammenstellt, die Prize Committees, die für die Vergabe der Preise zuständig sind (GSA/DAAD Book Prize Committee, Sybil Halpern Milton Memorial Book Prize Committee, GSA/DAAD Article Prize Committee). Die Amtszeit dieser Committees beträgt ein Jahr. Darüber hinaus gibt es das Investment Committee und das Berlin Program Selection Committee. Die Mitgliedschaft im Archives Committee, auf das im Folgenden noch näher eingegangen wird, ist zeitlich nicht befristet.³

Die Organisation versteht sich selbst wie folgt: "The German Studies Association is the premier multi- and interdisciplinary association of scholars focused on German, Austrian, and Swiss history, literature, culture studies, political science, and economics. A non-profit educational organization, GSA promotes research and study of Germany, Austria, and Switzerland. Membership is open to anyone interested in German studies. The membership generally consists of professors and advanced students at universities and colleges in North America, although several hundred members reside in Europe and Asia. Guidelines for curricula in German Studies are offered by GSA for assistance to universities, colleges, and high schools in planning programs in German Studies."⁴

Die German Studies Association ist dem American Council of Learned Societies (ACLS) angeschlossen. Zurzeit hat sie etwa 2.000 Mitglieder. Sie gibt dreimal im Jahr eine Fachzeitschrift, die German Studies Review (GSR), sowie zweimal im Jahr einen Newsletter heraus. Die German Studies Association veranstaltet jährlich eine

³ GSA: <https://www.thegsa.org/about/by_laws.html> [19.11.2013].

⁴ GSA: <https://www.thegsa.org/about/by_laws.html> [19.11.2013].

Konferenz, an der 2013 ca. 1.350 Personen aus aller Welt, vor allem aus den USA, Kanada, Deutschland und Österreich, teilnahmen. Diese umfasste 330 Sektionen mit je drei bis vier Vorträgen und einem Kommentar.⁵

Anfang Oktober 2013 gelangte die German Studies Association in Deutschland in die Schlagzeilen, weil die USA einen Festredner, den in Bulgarien geborenen deutschen Schriftsteller Illija Trojanow (Jahrgang 1965), nicht zur Jahrestagung nach Denver einreisen ließen – ohne Angaben von Gründen. Interessanterweise wurde die Tagung in den Meldungen als „Germanistenkongress“ bezeichnet, was zeigt, dass der Begriff der German Studies in der Öffentlichkeit immer noch zu eng verstanden wird.⁶

Schon bald nach der Gründung der Organisation wurde das Archives Committee eingerichtet, das das Präsidium, den Geschäftsführer und die Mitglieder in archivischen Fragen berät. Aufgabe dieses Komitees ist es, jährlich auf der Mitgliederversammlung über die Situation und aktuelle Themen des Archivwesens im deutschsprachigen Raum zu informieren. Dieser Bericht wird im Newsletter

⁵ GSA: <<https://www.thegsa.org/about/index.html>> [19.11.2013].

⁶ Illija Trojanow, Willkür und Freiheit, in: Frankfurter Allgemeine Zeitung vom 2.10.2013; David E. Barclay, GSA Speaker Denied Entry to the United States: Response of the German Studies Association, in: German Studies Association Newsletter 38, 2013, Nr. 2 (Winter 2013/14), S. 8-10. Interessanterweise erhielt Trojanow kurz darauf ein Zehnjahresvisum für die USA und konnte nach New York reisen, vgl. Alexander Osang, Volksfeinde, in: Der Spiegel 47/2013 vom 18.11.2013, S. 140f. Auch in diesem Artikel ist bezüglich der GSA-Tagung in Denver von einem „Germanistenkongress“ die Rede (S. 140). Zu Trojanow vgl. Julian Preece (Ed.), Illija Trojanow (Contemporary German Writers & Filmmakers 2), Oxford u.a. 2013.

publiziert. Darüber hinaus werden konkrete Fragen der Mitglieder geklärt und ggf. Probleme bei der Archivbenutzung besprochen, um einvernehmliche Lösungen zu finden.⁷

Mehrfach hat die German Studies Association aufgrund der Arbeit des Archives Committee durch öffentliche Verlautbarungen und/oder Briefe, z.B. an das Außenministerium oder andere öffentliche Stellen in den USA und in der Bundesrepublik, zu Problemen des Archivwesens Stellung genommen. Über viele Jahre stand die Öffnung der Unterlagen des Internationalen Suchdienstes in Bad Arolsen für die wissenschaftliche Forschung im Mittelpunkt. Aber auch zu anderen Fragen, wie der Zukunft des Tourismusarchivs in Berlin, hat die Organisation Stellung genommen und ihr internationales Gewicht für die Sicherung und die Zugänglichkeit von Unterlagen eingesetzt.

Deutlich wird die hohe Bedeutung der Archivarbeit für die Mitglieder der German Studies Association – sowohl für Germanisten wie für Politikwissenschaftler, Historiker und andere historisch Interessierte. Dies spiegelt den besonderen Stellenwert der Archivarbeit im deutschsprachigen Raum für die Forschung in Nordamerika wider.

3. Rahmenbedingungen der Archivarbeit in Übersee

Die Arbeit in deutschsprachigen Archiven gehört zum Handwerk vor allem von Historikern und Germanisten auch in den USA und Kanada. Um diese Nutzergruppe angemessen unterstützen zu können, ist

⁷ Zuletzt: Rainer Hering, GSA Archives Committee Report 2013, in: German Studies Association Newsletter 38, 2013, Nr. 2 (Winter 2013/14), S. 28-35. Der Verfasser ist seit 2001 Mitglied und seit 2006 Vorsitzender des Archives Committee. Das Landesarchiv Schleswig-Holstein ist institutionelles Mitglied der German Studies Association.

es für die Archive wichtig, deren Rahmenbedingungen für die archivische Quellenarbeit in Übersee zu kennen. Der Lehrbetrieb an nord-amerikanischen Universitäten läuft in der Regel von September bis Mai, d.h. im Sommer von Mai bis August bzw. September gibt es für Professoren Zeit für Forschungen in Archiven. In der Regel bewerben sie sich um Stipendien, um die Reisekosten und ggf. entstehende Kopierkosten etc. aufbringen zu können. Doktoranden haben – sofern sie über ein Stipendium, z.B. des Deutschen Akademischen Austauschdienstes (DAAD), verfügen – zumeist neun bis zwölf Monate Zeit für einen Forschungsaufenthalt. Aufgrund der finanziellen und zeitlichen Begrenzungen sind sie darauf angewiesen, ihre Archivbesuche effektiv vorzubereiten und konzentriert durchzuführen.

Einige wenige der potentiellen Nutzer haben zuvor an Archivkursen des Deutschen Historischen Instituts oder des United States Holocaust-Memorial Museum in Washington teilgenommen und verfügen über Grundinformationen zum deutschsprachigen Archivwesen.⁸ Im Allgemeinen sind diese aber nicht vorauszusetzen. Daher ist kompetente Beratung vor Ort entscheidend, weil in der Regel ein Archivbesuch nur einmal erfolgen kann. Wünschenswert sind ebenfalls kurze Bestellzeiten für Archivalien, möglichst keine Begrenzungen bei der Zahl der zu bestellenden Einheiten und lange Öffnungszeiten der Lesesäle. Um Reisekosten zu sparen, besteht der Wunsch, viele Unterlagen zu reproduzieren – Selbstbedienungsscanner bzw. kurze Zeiten bei der Anfertigung von kostengünstigen (digitalen) Reproduktionen sind gern gesehen.

⁸ Mark Stoneman, Archival Summer Seminar in Germany, in: Bulletin of the German Historical Institute 53, Fall 2013, S. 139-141.

Ein weiterer Aspekt ist wichtig: Das deutsche System der Schutzfristen mit seinen Verkürzungsmöglichkeiten ist für Forschende aus dem Ausland oft unbekannt. Daher ist es erforderlich, dass sie rechtzeitig darüber informiert werden, sofern ihr Forschungsvorhaben davon betroffen ist, damit sie zeitnah Anträge stellen und ggf. Bescheinigungen ihrer Betreuer beibringen können.

Amerikanische Erwartungen an das digitale Angebot deutschsprachiger Archive: Zentral für die Vorbereitung eines Archivbesuches ist das digitale Angebot der jeweiligen Archive, zunächst sind es deren Internetseiten.⁹ Hier werden erwartet:

- Übersichten zum archivischen Angebot in Deutschland und in den einzelnen Bundesländern in allen Sparten. Hier wird das Archivportal-D eine bedeutende Rolle spielen.
- Informationen über die Archive und ihre Nutzungsmöglichkeiten (Öffnungszeiten, Anreisen, Reproduktionsmöglichkeiten),
- gern gesehen sind Hinweise auf günstige Unterbringungsmöglichkeiten,
- Informationen zur Archivgeschichte und anderen Archiven des Sprengels,
- kommentierte Beständeübersichten,
- Bestandsverzeichnisse,

⁹ Die folgenden Ausführungen basieren auf Angaben der Mitglieder des Archives Committee der German Studies Association, langjährigen Erfahrungen und einer Befragung auf der Jahrestagung der Organisation in Denver im Oktober 2013. Allen Beteiligten, vor allem Prof. Meike Werner PhD (Vanderbilt University, Nashville, Tennessee), danke ich auch an dieser Stelle für Ihre Auskünfte und Hinweise.

- Recherchemöglichkeiten über Bestände hinweg,
- Bestellmöglichkeiten im Voraus,
- Erläuterungen zur Rechtslage der Benutzung und
- Schriftproben gerade bei Nachlässen zum Einstellen auf die in den Quellen vorkommenden wichtigsten Handschriften.

4. Amerikanische Digitalisierungserwartungen

Interessanterweise steht nicht die Digitalisierung von Beständen im Vordergrund des nordamerikanischen Interesses. Der Schwerpunkt liegt eindeutig auf den Findmitteln, wobei PDF-Dateien schon als ausreichend angesehen werden, und Kontextinformationen. Das mag für gut ausgestattete Archive eine Selbstverständlichkeit sein – für viele, vor allem kleinere Archive ist es aufgrund der fehlenden Ressourcen bis dahin noch ein weiter Weg. Wenn es um die Online-Bereitstellung von Archivalien geht, sind in erster Linie Fotos und Karten gefragt. Darüber hinaus stehen kleinere Druckschriften, sogenannte „graue Literatur“ etc., die sich nicht in Bibliotheken befinden, auf der Wunschliste für eine digitale Bereitstellung. Das gilt auch für Zeitungsartikelsammlungen oder in Archiven vorhandene Zeitungen.

Für die akademische Lehre werden Quellenbeispiele mit Transkriptionen und Erläuterungen im Internet als sehr hilfreich angesehen.¹⁰

¹⁰ In diesem Kontext kann z.B. auf das Projekt des Bundesarchivs in Zusammenarbeit mit dem Verband der Geschichtslehrer Deutschlands hingewiesen werden, Quellen aus dem Bundesarchiv auszuwählen, zu kommentieren und mit didaktischen Hinweisen zu ergänzen, damit Lehrende sie ohne großen zeitlichen Aufwand im Unterricht einsetzen können. Diese Dokumente werden als „Original“ zu

Geht es um Archivalien, so werden komplette Bestände digital gefordert, nicht herausragende Einzelstücke oder Teile von Akten. Für die personenkundliche und die Emigrationsforschung sind Unterlagen zur Auswanderung zentral, wie z.B. Passagierlisten – hier ist beispielsweise das Hamburger Projekt „Link to your roots“ zu nennen, in dessen Rahmen diese Listen transkribiert und überprüft im Internet angeboten werden.¹¹ Ansonsten dominiert das Interesse an zentraler Überlieferung der Entscheidungsträger, also vor allem Gremienprotokolle und Finanzunterlagen, die über nahezu jedes Vorhaben Auskunft geben. Bei Personalakten und Nachlässen stehen prominente Persönlichkeiten im Vordergrund. Personen oder Kontakte,

sehen sein mit allen Bearbeitungsspuren, in denen oft eine eigene Information steckt und die neugierig machen auf die Spurensuche in Archiven <<http://www.bundesarchiv.de/oeffentlichkeitsarbeit/meldungen/03792/index.html>> [27.12.2013].

¹¹ Paul Flamme, Das Projekt "Link To Your Roots". Projekt zur digitalen Erschließung der Auswandererlisten im Staatsarchiv Hamburg, in: Jens Murken (Hrsg.), Die Archive am Beginn des 3. Jahrtausends – Archivarbeit zwischen Rationalisierungsdruck und Serviceerwartungen. Referate des 71. Deutschen Archivtages 2000 in Nürnberg, Siegburg 2002 (Archivar Beiband 6), S. 203-220; ders., Herstellung und Vermarktung einer Auswandererdatenbank durch das Staatsarchiv Hamburg. Das Projekt "Link to your Roots", 2010, in: Archivar 63, 2010, S. 151-157; Peter Gabrielsson, Das Projekt "Link To Your Roots". Die digitale Erschließung der Auswandererlisten im Staatsarchiv Hamburg, in: Archivar 53, 2000, S. 326-329; ders., Das Projekt "Link To Your Roots". Die digitale Erschließung der Auswandererlisten im Staatsarchiv Hamburg, 2000, in: Auskunft. Zeitschrift für Bibliothek, Archiv und Information in Norddeutschland 20, 2000, S. 378-387; Jürgen Sielemann/Ulf Bollmann/Rainer Hering (Hrsg.), Überseeische Auswanderung und Familienforschung, Hamburg 2002 (Veröffentlichungen aus dem Staatsarchiv der Freien und Hansestadt Hamburg XVIII); amerikanische Ausgabe: Jürgen Sielemann/Ulf Bollmann/Rainer Hering (Ed.), Overseas Emigration And Family Research, Bergenfield, N.J. 2003.

die die deutsch-amerikanischen Beziehungen betreffen, sind ebenfalls gefragt.

Zeitlich liegen die Schwerpunkte auf dem Kaiserreich, der Weimarer Republik, dem „Dritten Reich“ und der DDR, zunehmend auch auf der Bundesrepublik, während Forschungsvorhaben zum Mittelalter oder zur Frühen Neuzeit in den letzten Jahrzehnten an Bedeutung verloren haben. Im Moment gewinnt jedoch das Thema Reformation – und Reformationsjubiläen – im Blick auf das Jubiläumsjahr 2017 an Interesse. David Blackbourn hat auf der Tagung der German Studies Association in Denver 2013 erneut ein Plädoyer für die stärkere Berücksichtigung des 19. Jahrhunderts im deutschsprachigen Raum gehalten, das noch immer auf Tagungen, in Zeitschriften und auf dem akademischen Buchmarkt Nordamerikas unterrepräsentiert ist.¹² Die Zeiträume Erster Weltkrieg, Weimarer Republik und „Drittes Reich“ haben nach wie vor einen herausragenden Stellenwert in der nordamerikanischen Forschung und Lehre. Erkennbar ist ein Trend zu Regional- und Lokalstudien. Die Nachfrage nach den Themen Auswanderung, Emigration und Remigration sowie jüdische Geschichte ist nach wie vor hoch.

Hilfreich für die Digitalisierungsentscheidungen in den deutschsprachigen Archiven sind sicherlich bisherige Erfahrungen in der Benutzung. Darüber hinaus ist es sinnvoll, die Entwicklung der historisch ausgerichteten Forschung diesseits und jenseits des Atlantiks zu beobachten, um angemessen auf Nutzerinteressen reagieren zu können. Der Blick von außen ist für die Selbstwahrnehmung

¹² David Blackbourn, „Honey, I Shrunk German History“, in: German Studies Association Newsletter 38, 2013, Nr. 2 (Winter 2013/14), S. 44-53.

unverzichtbar und bereichert die archivische Aufgabenerfüllung der Sicherung, Erschließung und Bereitstellung von Rechts- und Kulturgut. Auch aus diesem Grunde ist es sinnvoll, wenn Archivarinnen und Archivare in engem interdisziplinären Austausch mit den German Studies, insbesondere der Geschichtswissenschaft, der Germanistik und der Politikwissenschaft, stehen.

Die automatisierte Volltexterkennung historischer Handschriften als gemeinsame Aufgabe von Archiven, Geistes- und Computerwissenschaft. Das Modell einer zentralen Transkriptionsplattform als virtuelle Forschungsumgebung

von Guenter Muehlberger

1. Einleitung

Die massenhafte Digitalisierung archivalischer Dokumente ist in Deutschland und Österreich noch keine Selbstverständlichkeit. Während die National Archives London bereits 2011 mehr als 80 Millionen historische Dokumente digitalisiert haben und auch das Norwegische Nationalarchiv oder die US National Archives über große Digitalisierungsprogramme verfügen, in denen jährlich hunderttausende von Dokumenten gescannt, indexiert, elektronisch archiviert und in digitaler Form der Öffentlichkeit zur Verfügung gestellt werden,¹ haben die Vorträge und Diskussionen am 18. Archivwissenschaftlichen Kolloquium in Marburg deutlich gemacht, dass die Vorstellung „alles zu digitalisieren“ immer noch bei vielen deutschen Archivaren auf ungläubiges Kopfschütteln oder gar heftigen Widerspruch stößt. Dabei bietet aus Sicht des Verfassers nur die systematische und

¹ In Beantwortung einer „Freedom of Information“ Anfrage teilte das National Archives mit, dass 80 Millionen Dokumente digitalisiert wurden, davon 75% mit kommerziellen Partnern. Vgl.: <<http://www.nationalarchives.gov.uk/foi/breakdown-of-digitised-documents.htm>> [21.03.2014] Das Digitalisierungsprogramm in Norwegen und den Vereinten Staaten von Amerika: <<http://nordichistoryblog.hypotheses.org/1169>> [21.03.2014] bzw. <<http://www.archives.gov/digitization/>> [21.03.2014].

flächendeckende Digitalisierung großer historischer Aktenbestände die reale Chance eine Technologie zu entwickeln, die den Zugriff auf Archivmaterialien in Zukunft revolutionieren könnte: Eine automatisierte Handschriftenerkennung, bei der historische Handschriften durch den Einsatz von Mustererkennung im Volltext erfasst werden – ganz so wie dies im gedruckten Bereich mit der Optical Character Recognition (OCR) bereits gang und gäbe ist. Der vorliegende Aufsatz beruht auf Forschungsarbeiten, die im Rahmen des EU Projekts tranScriptorium durchgeführt werden.² Das Projekt läuft von Januar 2012 bis Dezember 2015 und wird von der Technischen Universität Valencia koordiniert. Das hier vorgestellte Modell einer interdisziplinären Zusammenarbeit zwischen Archiven, Geistes- und Computerwissenschaften soll als prototypische Anwendung von der Universität Innsbruck, Institut für Germanistik, realisiert werden. Erste Ergebnisse liegen bereits vor und werden 2014 der Öffentlichkeit vorgestellt.

2. Handschriftenerkennung – die Technologie im Hintergrund

2.1. Allgemeiner Hintergrund

Die automatisierte Erkennung von Handschriften geht in Ansätzen bereits auf Arbeiten aus den frühen 90er Jahren des 20. Jahrhunderts zurück.³ Entscheidende Impulse für die Handwritten Text Recognition (HTR) kamen dann später aus dem Feld der automatisierten

² Webseite des Projekts: <<http://www.transcriptorium.eu/>> [21.03.2014].

³ R. Plamondon & S. N. Srihari, Online and off-line handwriting recognition: a comprehensive survey, in: IEEE Transactions on Pattern Analysis and Machine Intelligence, 22/1, 2000, S. 63–84. doi:10.1109/34.824821.

Spracherkennung. Im Gegensatz zum gedruckten Text, der trotz aller unterschiedlichen Schrifttypen durch einen hohen Grad der Normalisierung auszeichnet, sind die Ausdrucksmöglichkeiten sowohl bei der gesprochenen Sprache als auch bei handschriftlichen Dokumenten wesentlich größer und die Verarbeitung der vorhandenen Daten daher entsprechend anspruchsvoller. Bei der Erkennung von individuellen Handschriften unterscheidet man zwischen Online- und Offline-Verfahren: Bei Online-Verfahren wird eine elektronische Schreibunterlage verwendet, sodass nicht nur das Schriftbild als Graphik, sondern auch die zeitliche Abfolge während der das Schriftbild zustande gekommen ist, für die Datenauswertung zur Verfügung steht. Die Online-Erkennung von Handschriften gilt heute als weitgehend gelöste Forschungsaufgabe, kommerzielle Anwendungen sind dementsprechend bereits vorhanden.

Anders sieht es hingegen bei der Offline-Erkennung aus: Hier gibt es „nur“ den Schriftzug als zweidimensionales Bild, die zeitliche Dimension fehlt. Diese ist jedoch die Voraussetzung für die Anwendung statistischer Verfahren, die darauf ausgelegt sind, wiederkehrende Muster – also in diesem Fall Buchstaben und Wörter – in Verlaufskurven zu erkennen. Um diesem Problem zu begegnen wird üblicherweise ein kleiner Bildausschnitt („Sliding Window“) verwendet, der diesen zeitlichen Ablauf simuliert und dazu führt, dass der Schriftzug in eine lange Reihe sich überlappender Abschnitte zerlegt wird.⁴

⁴ T. Plötz & G. Fink (2009), Markov models for offline handwriting recognition: a survey, in: International Journal on Document Analysis and Recognition (IJDAR), 12(4), 2009, S. 269–298. doi:10.1007/s10032-009-0098-4. Vgl. auch B. Gatos

Im Einzelnen sieht der typische Ablauf dann wie folgt aus: Nach der Binarisierung des Bildes und einer ersten groben Klassifizierung des Inhalts einer Seite in inhaltlich bedeutsame (Text, Abbildungen) und „störende“ Elemente wie Seitenränder oder durchscheinende Schrift, werden Textblöcke und Zeilen eines Dokuments erkannt. Innerhalb der Zeilen wird die Schrift in einem weiteren Schritt normalisiert indem die oftmals bei Handschriften zu beobachtende Neigung ausgeglichen wird. Dieses normalisierte Abbild einer Zeile wird dann mit dem schon erwähnten „Sliding Window“ Verfahren „vermessen“, d.h. das Vorkommen der Pixel innerhalb des jeweiligen Abschnitts nach bestimmten Kriterien ausgewertet. Auf dieser Basis wird nun eine Verlaufskurve erstellt, die mit statistischen Verfahren ausgewertet wird. Als besonders effizient haben sich dabei Hidden Markov Modelle (HMM) erwiesen. Das Besondere an Hidden Markov Modellen ist die Tatsache, dass die Segmentierung und Klassifikation als integrierter Vorgang zu betrachten sind. Es muss daher im Vorhinein keine Segmentierung einzelner Wörter oder Buchstaben durchgeführt werden, sondern das mathematische Modell liefert die besten Hypothesen, welche Wörter und Buchstaben für eine bestimmte Zeile insgesamt am wahrscheinlichsten sind. Somit wird keine Vorentscheidung getroffen, ob nun 5 oder 8 Wörter in einer Zeile zu finden sind, sondern es werden unterschiedliche Möglichkeiten präsentiert, wie eine Zeile mit Wörtern „aufgefüllt“ werden kann, und diese unterschiedlichen Optionen werden wiederum mit

u.a., D3.1.2, Description and evaluation of tools for DIA, HTR and KWS (2013), <<http://transcriptorium.eu/pdfs/deliverables/transcriptorium-D3.1.2-31December2013.pdf>> [21.03.2014].

Wahrscheinlichkeiten versehen. Die Darstellung dieser Hypothesen oder Optionen erfolgt in einem Wortgraphen, der alle vom mathematischen Modell gebildeten Hypothesen zu einer bestimmten Zeile, sowie den damit verbundenen Wahrscheinlichkeiten enthält. Der wahrscheinlichste Weg durch den Wortgraphen sollte im besten Fall auch das richtige Ergebnis für die jeweilige Zeile darstellen. Wird dem System ein Fehler mitgeteilt, so ändert sich damit auch die Ausgangsbasis im Wortgraphen, d.h. der zuerst schlechter bewertete Knoten erhält eine höhere Wahrscheinlichkeit und wird so zum Ausgangspunkt für die weitere Berechnung. Dementsprechend erhöhen sich die Wahrscheinlichkeiten für den restlichen Pfad und das Ergebnis sollte besser werden. Ein Aspekt, der für die Entwicklung einer interaktiven Benutzeroberfläche eine wichtige Rolle spielen kann.⁵

Wir haben bisher allerdings einen ganz wesentlichen Aspekt noch unerwähnt gelassen: Damit ein sinnvoller Wortgraph gebildet werden kann, muss vorher das Modell trainiert werden. Die Mustererkennung beruht auf einem maschinellen lernenden Ansatz, bei dem das mathematische Modell mit einer Reihe von korrekten Daten gefüttert werden muss, um daraus die für die jeweilige Handschrift typischen Muster (Worte, Buchstaben) errechnen zu können. Diese Muster dienen dann als Grundlage für den Vergleich mit der neu zu erkennenden Handschrift. Die Trainingsdaten für diesen Vorgang

⁵ V. Romero u.a., The ESPOSALLES database: An ancient marriage license corpus for off-line handwriting recognition, in: Pattern Recognition, 46, 6, 2013, S. 1658–1669, <http://dx.doi.org/10.1016/j.patcog.2012.11.024> [21.03.2014].

werden als Referenzdaten bzw. „Ground Truth“ bezeichnet.⁶ Sie bestehen im Wesentlichen aus einer korrekten Transkription des Textes sowie der topographischen Repräsentation dieser Transkription am digitalen Bild. Diese topographische Repräsentation ist – wie oben bereits kurz erwähnt – jene Fläche am digitalen Bild, die alle Zeilen eines Textblocks bzw. alle Buchstaben einer Zeile enthält. Innerhalb des Projekts ist die Forschungsgruppe des Nationalen Forschungszentrum Demokritos in Athen für die Entwicklung einer Segmentierungssoftware zuständig, mit der die Aufgabe der Block- und Zeilenerkennung weitgehend automatisiert erfolgt.⁷ Klarerweise erfordert die Erstellung von detaillierten Trainingsdaten einen erheblichen manuellen Aufwand, insbesondere die genaue Zeilentrennung ist arbeitsaufwändig. Eine wesentlich einfachere Methode wäre etwa das Einzeichnen von Grundlinien, d.h. jener gedachten Linie auf der die Buchstaben „sitzen“. Ein wichtiger Teil des Projekts besteht daher auch darin Erfahrungswerte zu gewinnen, welchen Einfluss die Segmentierung auf die Erkennung besitzt.

Aber auch die Erstellung des Textes ist keine selbstverständliche Aufgabe: Während in früheren Jahren Herausgeber von Transkriptionen relativ stark in den Text eingegriffen haben, wird heutzutage der diplomatischen Transkription meist der Vorzug gegeben. Trotzdem ist die Behandlung von Sonderzeichen, Abkürzungen, Groß- und Kleinschreibung von Eigennamen keineswegs einheitlich. So

⁶ D. Kondermann, Ground Truth Design Principles. An Overview, in: VIGTA '13, St. Petersburg, Russland 2013, S. 1–4.

⁷ National Centre for Scientific Research (NCSR). Webseite: <<http://www.demokritos.gr/>> [21.03.2014].

berücksichtigen nur wenige Transkriptionen das Vorhandensein des langen „f“, das ja auch tatsächlich für die Benutzung eines Dokuments kaum einen Unterschied macht. Für den Prozess des Modelltrainings bedeutet dies jedoch, dass zwei verschiedene Grapheme mit nur einer Repräsentation versehen werden und somit eine Unschärfe vorliegt. Ähnlich schwierig ist der Umgang mit den für frühneuzeitliche oder mittelalterliche Handschriften typischen (morphologischen) Abkürzungen, die keineswegs einheitlich behandelt werden können. Hier bedarf es klarer Übereinkünfte und einer transparenten Vorgangsweise. Für die Erstellung der Trainingsdaten wird im Projekt das PAGE Format verwendet, das im Rahmen des EU-Projekts IMPACT entwickelt wurde und das auch für wissenschaftliche Wettbewerbe bei der Bildsegmentierung, der OCR und HTR Verarbeitung Verwendung findet.⁸

2.2. Erste Ergebnisse der automatisierten Handschriften-erkennung

Im ersten Jahr des Projekts wurde an verschiedenen Dokumenten gearbeitet, darunter mehr als 400 Seiten aus der Sammlung des Bentham Archivs. Es handelt sich hierbei um Schriften und Notizblätter aus dem Nachlass von Jeremy Bentham (1748-1832),

⁸ S. Pletschacher, A. Antonacopoulos, The PAGE (Page Analysis and Ground-Truth Elements) Format Framework, in: 20th International Conference on Pattern Recognition, 2010, S. 257–260. doi:10.1109/ICPR.2010.72.

A. Antonacopoulos, C. Clausner, C. Papadopoulos, C. & S. Pletschacher, ICDAR2013 Competition on Historical Book Recognition – HBR2013, in: 12th International Conference on Document Analysis and Recognition ICDAR, 2013, S. 1491–1495. doi:10.1109/ICDAR.2013.294.

die von einem weiteren Projektpartner, dem University London College herausgegeben werden.⁹ Das Layout dieser Notizblätter ist oftmals sehr komplex. Im jetzigen Stadium der Handschriftenerkennung, bei der der Fokus auf der Basistechnologie liegen soll, war es daher sinnvoll eine Auswahl zu treffen und nur Seiten mit einer fortlaufenden Schrift und einem „einfachen“ Layout zu verarbeiten. Eine Einschränkung auf bestimmte Schreiber wurde jedoch nicht durchgeführt, d.h. es finden sich durchaus unterschiedliche „Hände“ innerhalb des Testdatensets. Im von der Universität Valencia durchgeführten Experiment wurden 350 Seiten als Trainingsmaterial benutzt, 50 Seiten als Evaluierungsdatensatz. Die Ergebnisse werden mittels Word Error Rate (WER) dargestellt. Das bedeutet, dass die Anzahl der richtig erkannten Wörter mit den korrekten Evaluierungsdaten verglichen wird. Wie aus den Ergebnissen hervorgeht, wurden rund ein Drittel aller Wörter (34,5%) falsch erkannt, wobei die Ergebnisse von Seite zu Seite sehr unterschiedlich sind und von unter 10% WER bis über 60% WER reichen können.¹⁰ Die Ergebnisse sind somit vergleichbar mit jenen für die Erkennung historischer Bücher der Zeit vor 1800, bei denen ja ebenfalls mit einer Wortfehlerquote von 20-50% zu rechnen ist. Es ist zu erwarten, dass – sobald die zugrundeliegenden Mechanismen und Modelle an mehreren

⁹ University College London, About the Bentham Project, <<http://www.ucl.ac.uk/Bentham-Project/about>> [21.03.2014].

¹⁰ Jafar Tanha, u.a., D4.1, Tools for development and deployment of linguistic resources for HTR. Deliverable 4.1., 2013, S. 24, <<http://transcriptorium.eu/pdfs/deliverables/tranScriptorium-D4.1.2-31December2013.pdf>> [21.03.2014].

Testdatensets ausprobiert wurden, ein besseres Verständnis vorliegen wird und jene Parameter gefunden werden können, die für eine gute Erkennung besonders wichtig sind.

Für den deutschsprachigen Raum wurden drei Testdatensets ausgewählt, alle drei befinden sich derzeit in der Vorbereitungsphase für die eigentliche Erkennung. Zum einen handelt es sich um die Ratsprotokolle aus Zwettl und Bozen, zum anderen um eine Auswahl von Urteilen des deutschen Reichsgerichts. Die Ratsprotokolle Zwettl liegen in 23 transkribierten Bänden vor und umfassen ca. 5500 Seiten. Sie beginnen im Jahr 1553 und sind durchgehend transkribiert bis 1882. Für die Stadt Bozen liegen ca. 30.000 digitalisierte Seiten vor, allerdings sind bisher nur 500 Seiten aus dem 16. und 17. Jahrhundert transkribiert. Die Reichsgerichtsurteile, die 2014 prioritär behandelt werden sollen, umfassen ca. 1500 digitalisierte und transkribierte Seiten der Jahre 1901 bis 1909. Während die ersten beiden Testdatensets sowohl von der Lesbarkeit der Handschriften als auch der sprachlichen Grundlage eine echte Herausforderung darstellen, handelt es sich bei den Reichsgerichtsurteilen um saubere Reinschriften, die von professionellen Gerichtsschreibern angefertigt wurden. Diese juristischen Texte sind gespickt mit Fachvokabular doch andererseits dem modernen Deutsch sehr ähnlich. Wir erwarten daher grundsätzlich gute Ergebnisse. Das bereits oben kurz dargestellte „Bentham Set“ ist auch die Grundlage für einen wissenschaftlichen Wettbewerb, der im Rahmen der 14. International Conference on Frontiers in Handwriting Recognition (ICFHR-2014) ausgeschrieben wurde.

Die Ergebnisse des Wettbewerbs werden bei dieser Konferenz der Öffentlichkeit vorgestellt.¹¹

2.3. Linguistische Komponenten

Eine Besonderheit des oben beschriebenen Verfahrens ist die Unmöglichkeit sogenannte Out-of-Vocabulary Wörter erkennen zu können, da wie bereits ausgeführt die Erkennung auf der Basis von Wortgraphen durchgeführt wird. Wörter, die nicht im Trainingsset vorgekommen sind, werden daher grundsätzlich nicht korrekt erkannt, da sie mit keinen vorhandenen Vorlagen abgeglichen werden können. Um diese Schwäche auszugleichen, werden lexikon- und korpusbasierte Methoden angewendet. Diese Aufgabe wird im Projekt vom Niederländischen Lexikographischen Institut (INL) in Leyden übernommen.¹² Für ein effektives sprachliches Modell werden dabei vor allem zwei Ressourcen benötigt: Erstens ein möglichst vollständiges Lexikon, das auch die historischen Schreibvarianten eines Lemmas¹³ enthält. Aufgrund der mangelnden Normalisierung finden sich unter Umständen sehr viele historische aber auch regionale Varianten ein und desselben Wortes in einem Text, was die Wahrscheinlichkeit, dass eine bestimmte Variante beim Trainieren nicht vorgefunden wurde, drastisch erhöht. Zweitens spielen aber auch Informationen, die nur auf Basis eines Corpus erstellt werden

¹¹ Website der Konferenz: <<http://www.icfhr2014.org/>> [21.03.2014] sowie des Wettbewerbs: <http://www.icfhr2014.org/?page_id=519> [21.03.2014].

¹² Website des Instituut voor Nederlandse Lexicologie: <<http://www.inl.nl/>> [21.03.2014].

¹³ Die Grundform eines Wortes.

können, eine wichtige Rolle. Die Häufigkeit eines Wortes (oder einer Variante) sowie das gemeinsame Auftreten von Wörtern (Konkurrenz, n-Gramme) liefern zusätzliche Daten, die für die Verbesserung der Erkennungsgenauigkeit genutzt werden können. Als besonders aussichtsreich hat sich dabei im ersten Jahr ein Ansatz erwiesen, der auf der Auswertung von domainspezifischen Korpora beruht. Die oben präsentierten Ergebnisse für das Bentham Testdatenset konnten signifikant um rund 20% verbessert werden.¹⁴

2.4. Zwischenbilanz

Fassen wir kurz zusammen: Auch wenn die hier präsentierten Ergebnisse keineswegs repräsentativ sind und die tatsächlichen Faktoren für den Erfolg und Misserfolg der Handschriftenerkennung nur ansatzweise bekannt sind, so ist doch klar, dass es sich um eine Technologie handelt, die in den nächsten 5-10 Jahren deutliche Fortschritte machen wird und in der einen oder anderen Weise auch für Anwendungen im Archiv und der Wissenschaft zur Verfügung stehen wird. Die entscheidenden Wegmarken hin zu einer produktiven Anwendbarkeit liegen allerdings weniger in der Technologie selbst, als vielmehr in der Verfügbarkeit von geeignetem Testmaterial. Zu diesem Schluss kommt auch Plötz, der in seinem Forschungsbericht zur Handschriftenerkennung folgende Schlüsse zieht: *The most prominent one [problem, G.M.], which can be considered a universal problem of any area of statistical pattern recognition, is the problem of limited data. Though some notable data collection efforts exist and some quite substantial datasets have also been made publicly*

¹⁴ Wie Anm. 10, Jafar, linguistic resources, S. 24.

*available already, these sample sets are still far too small [...] for training a statistical recognizer that might be able to show close to human performance in automatically reading handwritten script.*¹⁵

Auch hier ist der Vergleich mit der Spracherkennungstechnologie aufschlussreich: Erst seitdem wirklich große Testdatensets zur Verfügung stehen, konnte auch die Forschung entscheidenden Fortschritte erzielen. Anders allerdings als bei Erkennung gesprochener Sprachen, bei der massive industrielle und auch politische Interessen (Telefonüberwachung) dahinter standen und stehen, unterliegt das Feld der historischen Handschriftenerkennung keinen derartigen Triebkräfte. Umso mehr ein Grund, dass sich die verschiedenen Interessensgruppen zusammenschließen und eine gemeinsame Basis für die weitere Entwicklung schaffen.

3. Modell einer Transkriptionsplattform

3.1. Vorüberlegungen

Wir haben bisher die rein technologischen Aspekte der Handschriftenerkennung beleuchtet. Im Folgenden geht es nun darum, in Ansätzen ein Modell vorzustellen, wie die Entwicklung der Technologie, die Digitalisierungsanstrengungen in Archiven sowie die Bedürfnisse von Geisteswissenschaftlern vereint werden – und auf dieser gemeinsamen Basis die Entwicklung der automatisierten Handschriftenerkennung einen ganz wesentlichen Schritt nach vorn machen könnte. Als Schlüssel sehen wir die Errichtung einer zentralen Transkriptionsplattform an, die sich an alle drei Benutzergruppen richtet

¹⁵ Wie Anm. 4, S. 294.

und die jeweiligen Ansprüche innerhalb einer virtuellen Forschungsumgebung bedient.

Zwei Überlegungen seien hier an den Anfang gestellt: Mit der Digitalisierung wird die Gebundenheit des Archivmaterials obsolet. Die digitalen Images können mit geringem Aufwand ständig verfügbar gehalten werden – nicht nur für einen allfälligen Benutzer eines Archivs, sondern eben auch für automatisierte Prozesse, wie sie im Rahmen der Handschriftenerkennung notwendig sind. Wichtig ist dabei der Umstand, dass für alle Verfahren wie die Block- und Zeilensegmentierung, die Binarisierung und die eigentliche Handschriftenerkennung keineswegs die sogenannten Master- oder Archivdateien benötigt werden, wie sie etwa in den Praxisregeln der Deutschen Forschungsgemeinschaft (DFG) definiert sind (TIFF unkomprimiert, 300 ppi, 24 Bit Farbe), sondern „Arbeitsdateien“ mit gleicher Auflösung, aber verlustbehafteten, komprimierten Formaten (JPEG, JPEG2000) völlig ausreichend sind.¹⁶ Damit wird nicht nur der Übermittlungsaufwand, sondern auch der Speicherbedarf deutlich reduziert. Zieht man die oben geforderten größeren Mengen von Testmaterial in Betracht, dann benötigen selbst eine Million Seitenbilder nicht mehr als 5 Terabyte Speicherbedarf. Mit anderen Worten: Eine zentrale Sammelstelle für digitalisierte Archivmaterialien, die als Arbeitsdateien der Forschung zur Verfügung stehen, kann mit relativ geringem Aufwand errichtet werden.

¹⁶ Deutsche Forschungsgemeinschaft (Hrsg.): DFG-Praxisregeln. „Digitalisierung“. DFG-Vordruck 12.151–02/13 Webseite: <http://www.dfg.de/formulare/12_151/12_151_de.pdf> [21.03.2014].

Die zweite wichtige Überlegung richtet sich an die Geisteswissenschaftler: Sie stellen jene Gruppe dar, die im Regelfall und oftmals in Kooperation mit den Archiven die eigentliche Transkription handschriftlicher Dokumente durchführt. Das kann im Rahmen einer historischen Edition erfolgen, oder aber als Teil einer historischen Recherche, bei der unter Umständen nur einzelne wichtige Teile eines Dokuments oder einer Sammlung transkribiert werden. Im Bereich der Textedition hat sich inzwischen die in den 90er Jahren des 20. Jahrhunderts entwickelte Auszeichnungssprache der Text Encoding Initiative (TEI) als de-facto Standard durchgesetzt.¹⁷ TEI stellt ein ausgefeiltes Set an Merkmalen zur Verfügung, wie sie für die möglichst einwandfreie Transkription eines Textes benötigt werden. Darüber hinaus können auch Metadaten mit dem Dokument verknüpft werden, darunter auch wichtige Angaben zu den detaillierten Transkriptionsregeln („editorial declaration“), wie sie für eine wissenschaftliche Edition zwingend notwendig sind. Während also die Voraussetzungen für eine standardisierte und hochentwickelte Transkription sehr gut sind, so ist der Aspekt der Verknüpfung mit dem zugrundeliegenden digitalen Bild bei TEI unterentwickelt. Zwar werden die Seitenbilder typischerweise mittels dem Element „faksimile“ verknüpft und auch Abbildungen innerhalb eines Textes können mit dem Element „zone“ und entsprechenden Koordinaten genau platziert werden, aber dies wird nur in sehr eingeschränktem Maß auch tatsächlich in der Praxis durchgeführt. Das eigentliche Credo lautet immer noch, dass eine gute Transkription den Text vollständig

¹⁷ Text Encoding Initiative. Webseite: <<http://www.tei-c.org/>> [21.03.2014].

abbildet und somit ersetzt. Das mag bei einfachem Layout auch zutreffend sein, aber schon bei etwas komplexeren Akten, bei denen diverse Notizen angefügt wurden und vielleicht auch noch eine zeitliche Abfolge durch die Verwendung unterschiedlicher Schreibstifte und Farben ersichtlich ist, wird die reine Texttranskription zu einer mühsamen Angelegenheit. Stattdessen wäre eine engere Verknüpfung mit dem digitalen Faksimile des Dokuments und damit der Gewinn an Augenscheinlichkeit in vielen Fällen wünschenswert und würde den Wert einer (digitalen) Edition ohne Zweifel erhöhen.

An dieser Stelle könnten sich nun die Bedürfnisse der Computerwissenschaftler nach korrekten Trainingsdaten und das Bedürfnis der Geisteswissenschaftler nach einer kombinierten Edition von Text und Seitenbild eines Dokuments treffen. Wer sich also der ohnehin mühsamen und zeitraubenden Tätigkeit einer Transkription unterzieht, könnte dies somit gleich in einer Art und Weise tun, dass Text und Bild miteinander in Beziehung gesetzt sind und damit auch Referenzdaten (Ground Truth) für die Weiterentwicklung der HTR geschaffen werden.

3.2. Modellbeschreibung

In Abbildung 1 werden diese allgemeinen Überlegungen graphisch dargestellt und miteinander in Beziehung gesetzt. Als vierter Mitspieler ist in der Abbildung auch noch die Öffentlichkeit vertreten, die einen verbesserten Zugang zu den automatisch erkannten Handschriften erhält.

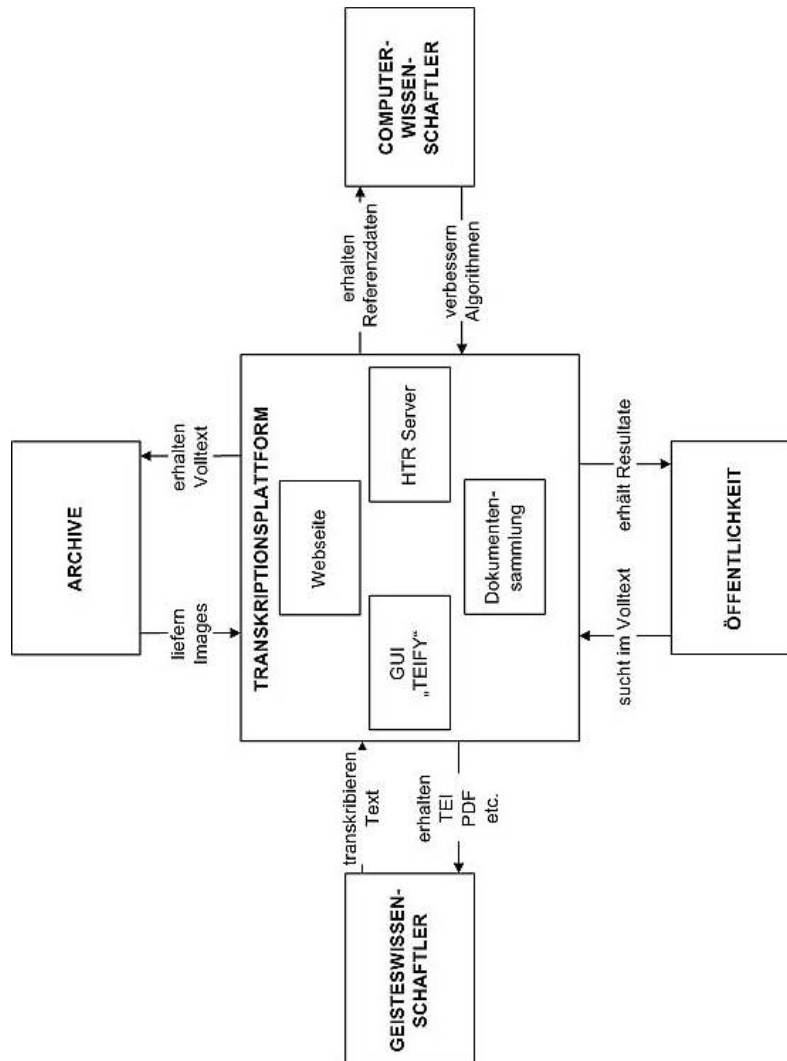


Abb.1: Transkriptionsplattform – Schema.

Die wesentlichen Merkmale der Transkriptionsplattform für die einzelnen Interessensgruppen sollen im Folgenden kurz dargestellt werden.

Archive

Archive können Arbeitskopien ihrer digitalisierten Archivmaterialien in die Plattform einspielen. Die Übermittlung der Daten kann dabei weitgehend automatisiert erfolgen und auf Standardschnittstellen zurückgreifen. Archive besitzen innerhalb der Transkriptionsplattform die volle Kontrolle über ihre Materialien und können bestimmen, wer mit welchen Rechten darauf zugreifen kann, wer einen Export der Dokumente durchführen darf und welche Formate ausgeliefert werden.

Für Archive kann der Nutzen einer derartigen Transkriptionsplattform in dreierlei Hinsicht definiert werden: Erstens können Archive die Plattform benutzen, um die standardisierte Transkription bestimmter Materialien innerhalb einer kontrollierten Umgebung durchzuführen, ohne dass sie selbst ein derartiges System aufbauen müssen. Zum zweiten erhalten sie korrigierte Transkriptionen ihrer Dokumente in standardisierten Formaten und drittens ermöglicht die integrierte Handschriftenerkennung auch die Verarbeitung großer Dokumentenmengen. Um eine möglichst effiziente Anbindung an bestehende Repositorien und Infrastrukturen zu ermöglichen, wird hier auf METS/ALTO/TEI/PDF Pakete gesetzt.¹⁸ Auf den kürzesten

¹⁸ METS (Metadata Encoding and Transmission Standard) und ALTO (Analyzed Layout and Text Object) sind XML Schemata, die von der Library of Congress verwaltet werden. Webseite: <<http://www.loc.gov/standards/>> [21.03.2014].

Nenner gebracht könnte man sagen, dass Archive als Gegenleistung für ihre Bereitschaft digitale Seitenbilder zur Verfügung zu stellen, standardisierte Textdateien zurück erhalten – und dabei niemals die Kontrolle über ihre Dokumente verlieren.

Geisteswissenschaftler

Betrachtet man die Werkzeuge, mit denen Geisteswissenschaftler derzeit Transkriptionen erstellen, dann werden überraschenderweise vor allem zwei genannt: Word und Oxygen. Mit beiden lässt sich sicherlich ein korrekter Text erstellen und auch die Umwandlung in ein TEI Dokument sollte keine größeren Schwierigkeiten bereiten, die Verknüpfung der Texte mit den Seitenbildern auf der Basis von Teilen einer Seite, oder auf der Ebene von Zeilen und Wörtern, ist damit jedoch nur in sehr mühsamer und wenig standardisierter Weise möglich. Die Plattform bietet daher ein spezielles Tool, mit dem eine Transkription möglichst standardisiert und „normgerecht“ durchgeführt werden kann und die Verknüpfung zwischen Text und Bild weitgehend automatisiert erfolgt. Mit dem in JAVA implementierten Tool (Arbeitstitel: „transdignify“) können nicht nur die Ergebnisse der automatisierten Erkennung korrigiert, sondern auch Dokumente auf den Server hochgeladen und diverse Prozesse angestoßen werden.

Der Mehrwert für den Geisteswissenschaftler besteht in einem qualitativ wesentlich verbesserten Resultat seiner Arbeit, bzw. den Exportformaten, die ihm angeboten werden. Dazu gehört in erster Linie ein korrektes TEI-Dokument, das das Standardset an Tags enthält, wie sie für jede Transkription bedeutsam sind. Dieses TEI-Dokument kann dann auch die Ausgangsbasis für die weiterführende

Verarbeitung bzw. das Einspielen in entsprechende Repositorien sein.

Zu den Vorteilen gehört andererseits aber auch die Erstellung neuer Formate, wie sie im Handschriftenbereich noch nicht verbreitet sind: Einmal ein PDF, das das Bild der gesamten Seite im Vordergrund anzeigt und den Text in einer zweiten Schicht im Hintergrund hält, sodass eine Volltextsuche automatisch die gefundenen Begriffe im Bild hervorhebt und zum andern ein PDF, das den transkribierten Text in gleicher Weise auf der Seite anordnet, wie im Original und der Leser somit sofort die Funktion eines Textes als Notiz, Marginalie, Fußnote, etc. wahrnehmen kann. Wahlweise kann die Anordnung auch mit dem Originalbild auf der linken und der im Layout gleich dargestellten Transkription auf der rechten Seite erfolgen.

Zusätzlich ist hier auch noch die soziale Komponente zu erwähnen. Mittels der Transkriptionsplattform kann die eigentliche Transkription als Gemeinschaftsarbeit durchgeführt werden. Nicht nur dass mehrere Personen gleichzeitig an einem Dokument arbeiten können, es lassen sich auch verschiedene Rollen definieren, die an die Fähigkeiten der Mitarbeiter und die Verantwortlichkeiten innerhalb eines Projekts angepasst werden können. Die Behandlung besonders schwierig zu lesender Passagen oder Wörter kann systematisiert werden, indem diese als „unsicher“ markiert werden und andere Teammitglieder eingeladen werden können, sich nur diese Stellen anzusehen und gegebenenfalls zu korrigieren. Zuletzt ermöglicht auch noch ein eingebautes Evaluierungstool die automatische Erstellung einer Zufallsstichprobe und die Erstellung einer Fehlerquote für die durchgeführte Transkription.

Computerwissenschaftler

Während die Geisteswissenschaftler in erster Linie mit der Client-Komponente des Systems arbeiten sollen, finden die Computerwissenschaftler die für sie relevanten Daten am Serversystem. Da alle Dokumente zentral vorgehalten werden und auch alle mittels der lokalen Software durchgeführten Transkriptionen zentral gespeichert werden, wächst mit jeder neu durchgeführten Transkription die Anzahl der verfügbaren Test- und Referenzdaten. Somit geht im besten Fall keine – wo und von wem auch immer – erstellte Transkription „verloren“, sondern dient der weiteren technologischen Entwicklung. Darüber hinaus können auch die Schritte, die ein Geisteswissenschaftler bei der Transkription eines Textes wählt mit protokolliert werden und die Benutzeroberfläche von transdignify dementsprechend optimiert werden.

Bei dem im Rahmen des Projekts entwickelten Prototypen der Transkriptionsplattform ist allerdings ein automatisiertes Verfahren für das Training der Schriftmodelle noch nicht enthalten. Derzeit müssen noch zu viele Parameter manuell und von Fachleuten eingestellt werden, sodass eine weitergehende Automatisierung noch nicht sinnvoll wäre.

3.3. Eine zentrale Transkriptionsplattform (TRP)

Die zentrale Transkriptionsplattform bietet darüber hinaus noch einige weitere wichtige Funktionen, die die einzelnen Komponenten miteinander in Beziehung setzen: So können sich Benutzer auf der Webseite der Plattform registrieren, bzw. um alle Funktionalitäten nutzen zu können, werden sie dazu auch verpflichtet. Als registrierte Benutzer können sie das Tool herunterladen und mit Unterstützung

der Plattform ihre Transkription durchführen. Als JAVA Programm kann transdignify auch lokal eingesetzt werden, viele Funktionen sind allerdings nur nutzbar, wenn eine Verbindung mit dem Server besteht.

Mit dem Tool können beliebig viele digitalisierte Dokumente hochgeladen werden, die dann am Server weiterverarbeitet werden. Dazu gehört die Segmentierung der Images in Blöcke und Zeilen. Wenn die Trainingsphase für eine Handschrift durchgeführt wurde, was konkret bedeutet, dass rund 50 Seiten korrekt transkribiert werden müssen, dann kann die HTR Engine für diese Schrift trainiert werden. Wie bereits erwähnt ist das derzeit (und für die Dauer des Projekts) ein manueller Prozess, der offline erfolgt. Sobald das Training abgeschlossen ist, kann innerhalb von transdignify auch ein interaktiver Modus bei der Transkription gewählt werden. Beim interaktiven Modus soll vermieden werden, dass der Benutzer falsche Wörter ausbessern muss, sondern es soll vielmehr darum gehen, dass er die richtigen innerhalb einer Tabelle präsentiert bekommt und somit eine positive Selektion durchführen kann. Hier werden die geplanten Tests mit den Benutzern von entscheidender Bedeutung für die weitere Entwicklung einer effektiven Benutzeroberfläche sein.

Eine wichtige Rolle spielt dabei die Tatsache, dass es sich um ein zentrales System handelt. So wissen Computerwissenschaftler jederzeit ganz genau, welche Methode Benutzer für ihre Transkription auswählen, wie rasch sie die Transkription erstellen, wo es offenbar Schwierigkeiten gibt und wie viele Segmentierungsfehler verbessert werden müssen – um nur einige Aspekte zu nennen, die von Interesse sein könnten. Wir sehen darin die größte Chance transdignify in

einer Art und Weise zu optimieren, dass dieses Tool tatsächlich den Wünschen und Bedürfnissen der Benutzer gerecht wird.

Ein zusätzlicher Vorteil der zentralen Transkriptionsplattform liegt darin, dass mittel- und langfristig immer mehr trainierte Handschriftenmodelle zur Verfügung stehen, sodass die Erkennung der Handschriften einen immer höheren Grad der Perfektion erreichen wird. Das bedeutet, dass keine Transkriptionsarbeit „vergebens“ ist, sondern immer auch dazu beiträgt, die HTR Engine im Hintergrund zu verbessern. Das ist – vom Modell her – auch der größte Unterschied zu den OCR Anwendungen, bei denen versucht wird, eine möglichst leistungsfähige Software zu entwickeln, die lokal eingesetzt werden kann. Jeder Trainingseffekt bleibt aber so im lokalen Umfeld stecken, eine Verbesserung des Gesamtsystems ist damit nicht gegeben. Die Transkriptionsplattform will diese Einschränkung vermeiden und das System als solches soll kontinuierlich weiterwachsen und – wie erwähnt – durch jede neu gelernte Handschrift verbessert werden. Daraus leitet sich nun allerdings auch ab, dass die Handschriftenerkennung als ein sich wiederholender Prozess zu betrachten ist. Da ein ständiger Vergleich zwischen den vorhandenen korrekten Transkriptionen und den vorliegenden Wort- und Buchstabenmustern möglich ist, kann automatisch entschieden werden, für welche Dokumente eine Wiederholung der Handschriftenerkennung sinnvoll ist und eine signifikante Verbesserung der Ergebnisse erwartet werden kann. Somit werden alle in der Transkriptionsplattform enthaltenen und noch nicht korrekt transkribierten Dokumente ständig neu „umgewälzt“ und verarbeitet, um die Ergebnisse der automatisierten Erkennung zu optimieren.

3.4. Geschäftsmodell

Es ist natürlich noch zu früh, um ein konkretes Geschäftsmodell für die Transkriptionsplattform präsentieren zu können, doch sollen hier einige allgemeine Überlegungen angestellt werden. Die Tatsache, dass der Bereich der historischen Handschriftenerkennung für eine kommerzielle Ausbeutung einen viel zu kleinen Markt darstellt als dass er für kommerzielle Firmen von Interesse wäre, muss sich dabei nicht unbedingt als Nachteil erweisen, sondern kann im Gegenteil in einen Vorteil umgemünzt werden. Grundvoraussetzung dafür ist, dass alle drei beteiligten Gruppen – Archivare, Geistes- und Computerwissenschaftler – ihre Interessen in dem jeweiligen Geschäftsmodell gewahrt sehen und der oben beschriebene konkrete Nutzen innerhalb des Geschäftsmodells transparent und effektiv abgebildet wird.

Werfen wir also kurz einen Blick auf die Interessenslagen der drei beteiligten Gruppen: Archive werden in erster Linie an der ständigen Verbesserung und Erweiterung ihres Volltextkorpus interessiert sein. Es ist zu erwarten, dass sie für einen derartigen Service grundsätzlich bereit sind zu zahlen, ähnlich wie dies bereits bei der Digitalisierung oder der OCR Verarbeitung erfolgt. Allerdings wird es sich um keine großen Summen handeln, auch bei der OCR historischer Schriften ist die Schmerzgrenze sehr, sehr tief angesetzt, was wohl auch darauf zurückzuführen ist, dass das Ergebnis einer Volltexterkennung ganz grundsätzlich mit Fehlern behaftet ist und daher die Einsicht, dass trotzdem ein erheblicher technologischer Aufwand betrieben werden muss, um einen derartigen „fehlerhaften“ Text zu erzeugen, sich noch nicht allgemein durchgesetzt hat. Andererseits

würde in dem oben beschriebenen Modell die Transkriptionsplattform auch Funktionen wie eine Volltextsuche übernehmen, sodass der eigentliche Mehrwert im Vergleich zu einer reinen OCR Erkennung deutlich höher wäre.

Überblickt man die Lage in den Geisteswissenschaften, dann ist klar, dass die Transkriptionsplattform keine ernsthaften finanziellen Hürden aufbauen sollte, was die Zugänglichkeit ihres Materials und ihrer Services für die Geisteswissenschaftler betrifft. Vorrangig wird es vielmehr darum gehen, eine möglichst große Anzahl an Geisteswissenschaftlern zu überzeugen, die Transkriptionsplattform für ihre wissenschaftliche Arbeit zu nutzen und von ihren vertrauten Textverarbeitungsprogrammen und XML-Editoren auf transdignify umzusteigen, bzw. mindestens die Basistranskription innerhalb der Plattform durchzuführen. Wenn diese Rechnung aufgeht und tatsächlich mehrere hundert oder gar tausende Benutzer das Tool für ihre Arbeit verwenden, dann könnte man durchaus ein Subskriptionsmodell ins Auge fassen, das allerdings sehr niederschwellig anzusetzen wäre und nur wenige EURO pro Monat umfassen könnte. Trotzdem könnte aufgrund der großen Zahl an Benutzern ein wahrnehmbarer Beitrag zum Erhalt der Transkriptionsplattform geleistet werden.

Die Interessenslage bei Computerwissenschaftlern hingegen sieht völlig anders aus: Sie benötigen große Mengen gut aufbereiteten Test- und Referenzmaterials, sowie herausfordernde Aufgaben, um möglichst innovative Methoden und Experimente durchführen und – dieser Aspekt wird oftmals unterschätzt – publizieren zu können. Im Gegenzug kann davon ausgegangen werden, dass sie bereit sind, der Transkriptionsplattform ein kostenloses Nutzungsrecht an

den von ihnen entwickelten Tools und Algorithmen zu gewähren. Dieses kostenlose Nutzungsrecht muss keineswegs bedeuten, dass die Software an sich als Open Source vorhanden wäre. Tatsache ist, dass bei der Frage nach der Open Source Verfügbarkeit von Software die Interessen vieler Universitäten und Forschungsinstitute diametral den Vorstellungen der Forschungsförderungsagenturen entgegenstehen und die Forderung nach einer generellen Open Source Stellung viele Forschungspartner abschrecken würde. Aus Sicht der Transkriptionsplattform ist diese Frage allerdings zweitrangig, entscheidend ist vielmehr, dass es gelingt, exzellente Forschungsteams für die Aufgabe der Handschriftenerkennung zu begeistern und – als Gegenzug für die Verwendung des umfangreich und bestens aufbereiteten Testmaterials – eine nicht exklusive Nutzungsbewilligung für die entwickelte Software im Rahmen dieser virtuellen Forschungsumgebung zu erhalten.

3.5. Rechtliche Implementierung

Betrachtet man die verschiedenen Rechtsmodelle, die insbesondere auch im Bibliotheksbereich in den letzten Jahren für ähnliche Einrichtungen genutzt wurden, wie eingetragene Vereine, Gesellschaften mit beschränkter Haftung, oder Stiftungen, dann fällt auf, dass ein naheliegendes und äußerst erfolgreiches Modell bisher kaum eine Rolle in den Diskussionen gespielt hat: die Genossenschaft. Es würde an dieser Stelle zu weit führen, um eine detaillierte Beschreibung der rechtlichen und organisatorischen Eckpunkte einer

Genossenschaft zu liefern, einige wesentliche Charakteristika können jedoch kurz angeführt werden.¹⁹

Das wahrscheinlich wichtigste Merkmal jeder Genossenschaft ist, dass sie nicht auf Gewinnmaximierung ausgerichtet ist, sondern der allgemeine Nutzen für die Genossenschaftsmitglieder, die zugleich auch Anteilseigner sind, im Mittelpunkt steht. Umgekehrt tritt allerdings die Genossenschaft gegenüber ihren Mitgliedern privatwirtschaftlich auf, d.h. es werden die Leistungen konkret verrechnet und unterliegen Angebot und Nachfrage. Die Mitglieder sind also keineswegs gezwungen, eine bestimmte Leistung von „ihrer“ Genossenschaft einzukaufen, sofern diese Leistung am Markt billiger angeboten wird. Andererseits haben alle Mitglieder ein starkes Interesse daran, dass die Genossenschaft wirtschaftlich positiv bilanziert, da sie nicht nur Anteilseigentümer sind, sondern mit ihrem Beitritt auch signalisiert haben, dass sie am Gesamtnutzen der Genossenschaft – in diesem Fall der möglichst umfangreichen, effizienten und technologisch fortschrittlichen Erstellung von Volltext für historische Handschriften – ein vitales Interesse besitzen. Als weitere Besonderheiten kommen die basisdemokratische Verfassung einer Genossenschaft sowie die rechtliche Absicherung der Mitglieder im Falle eines Konkurses hinzu. Hier sieht z.B. das österreichische Genossenschaftsrecht vor, dass nicht mehr als der doppelte Betrag des erworbenen

¹⁹ Vgl.: Webseite des Österreichischen Raiffeisenverbands: <<http://kooperieren.at>> [21.03.2014]. Die rechtlichen Grundlagen für Österreich finden sich hier: Bundesrecht konsolidiert: Gesamte Rechtsvorschrift für Genossenschaftsgesetz, Fassung vom 20.03.2014: <<https://www.ris.bka.gv.at/GeltendeFassung.wxe?Abfrage=Bundesnormen&Gesetzesnummer=10001680>> [21.03.2014].

Anteilsscheins im Falle eines Konkurses fällig wird. Schließlich sei auch noch erwähnt, dass die Anteilsscheine an einer Genossenschaft, ganz im Gegensatz zu anderen Rechtsformen nicht zum Gegenstand der Spekulation werden können, da ein Verkauf immer nur auf Basis des Nominalwerts erfolgen kann, nicht aber auf Basis des tatsächlichen Werts einer Genossenschaft.

Wendet man diese allgemeinen Überlegungen auf das Modell der Transkriptionsplattform an, dann sieht man rasch, dass sich hier alle beteiligten Institutionen wiederfinden könnten: Der allgemeine Anteilsschein an der Genossenschaft wäre gering und würde wohl über eine einmalige Mitgliedsgebühr im zweistelligen Eurobereich nicht hinausgehen. Jeder der aus der Transkriptionsplattform einen Nutzen ziehen möchte, wäre verpflichtet Mitglied der Genossenschaft zu werden (d.h. einen Anteilsschein zu erwerben). Konkrete Leistungen, die die Genossenschaft gegenüber ihren Mitgliedern erbringt, würden darüber hinaus allerdings gesondert verrechnet werden. So wäre zu erwarten, dass Archive für die erbrachten Leistungen bereit wären einige hundert bis einige tausend EURO pro Jahr zu bezahlen, je nach der Anzahl der innerhalb der Plattform verarbeiteten Dokumente. Geisteswissenschaftler würden mit einer kleinen Subskriptionsgebühr belastet werden, sofern sie Dokumente hochladen und wissenschaftlich bearbeiten möchten und HTR Forscher würden voraussichtlich für die Verwendung der Testdaten einen kleinen Beitrag leisten bzw. würden vertraglich verpflichtet werden sozusagen als Naturalleistung die entwickelte Software für die Plattform kostenlos nutzbar zu machen.

Ob die mit diesem Geschäftsmodell zu erzielenden Summen tatsächlich auch eine Abdeckung der Gesamtkosten ermöglichen würde, kann hier nicht gesagt werden. Klar ist allerdings, dass auch die öffentliche Hand ein Interesse an der Errichtung einer derartigen virtuellen Forschungsumgebung besitzen müsste – und das Modell der Genossenschaft ohne Zweifel die Voraussetzungen für eine transparente und marktnahe Bewertung der erbrachten Leistungen bieten würde. Im dritten Jahr des tranScriptorium Projekts sollen diese Überlegungen fortgeführt und ein auch mit Zahlen unterfüttertes Geschäftsmodell vorgelegt werden.

Wir sind überzeugt, dass mit dem Modell einer Transkriptionsplattform als virtuelle Forschungs- und Serviceumgebung ein gangbarer Weg vorgezeichnet ist, der nicht nur den Interessen von Archiven, Geistes- und Computerwissenschaftlern sondern auch der breiten Öffentlichkeit und der öffentlichen Hand entgegenkommt und so Archivmaterialien in völlig neuer Art und Weise erschlossen werden können.

4. Danksagung

An dieser Stelle möchten wir uns ganz herzlich für die freundliche Unterstützung durch Prof. Dr. Martin Scheutz (Universität Wien), Herrn Friedel Moll (Stadtarchiv Zwettl), Herrn Dr. Hannes Obermair (Stadtarchiv Bozen) sowie Prof. Dr. Jan Thiessen (Universität Tübingen) und den Mitarbeiterinnen und Mitarbeitern der Bibliothek des Bundesgerichtshofs in Karlsruhe bedanken, die uns in großartiger Weise bei der Digitalisierung und Rechteklärung für die jeweiligen Datensets unterstützt haben.

Das Projekt tranScriptorium (Projektnummer: 600707) wird teilweise aus den Mitteln des 7. Rahmenprogramms für Forschung und technologische Entwicklung der Europäischen Union finanziert.

Wenn ich mir was wünschen dürfte. Wunsch(t)raum Archiv für NutzerInnen im digitalen Zeitalter

von Sylvia Necker

Das digitale Archiv verheißt scheinbar unbeschränkte Möglichkeiten und wird gleichermaßen zum Wunschraum wie zum Wunschtraum eines jeden Nutzers. Doch was wünschen sich Nutzerinnen und Nutzer eines digitalen Archivs eigentlich? Die Veranstalter des 18. Archivwissenschaftlichen Kolloquiums in Marburg haben mich gebeten, diese Frage aus der Sicht von Forscherinnen und Forschern zu beantworten. Archive stehen in den nächsten Jahren vor allem vor der Aufgabe bei der Digitalisierung von Archivgut die „richtigen“ Prioritäten zu setzen. In diesem Beitrag soll es um Empfehlungen gehen, welche Bestände aus der Sicht der Forschung bei der Digitalisierung priorisiert werden könnten. Doch diese vermeintlich einfache Frage nach einer Priorisierung und nach den Wünschen der Nutzer ist nicht eindeutig zu beantworten. Wenn ich mir was wünschen dürfte, ja was eigentlich dann? Je länger ich über die Frage nachdachte sowie Kolleginnen und Kollegen befragte, desto weniger hatte ich eine Idee, wie diese Frage sinnvoll zu beantworten wäre. Diese Ratlosigkeit erinnerte mich an aktuelle städtebauliche Partizipationsprozesse, die ich als Stadt- und Architekturhistorikerin analysiere: auch hier werden Bürger gefragt, was sie sich wünschen, einen Baum vor der Tür, eine Bank hinter dem Haus. Mitte der 1990er Jahre gab es in Hamburg Proteste gegen eine vollständige Bebauung des Elbufers. Die Bewohner der Viertels St. Pauli schlossen sich zu einer Initiative

zusammen und starteten eine "Wunschproduktion".¹ Jeder sollte aufschreiben und aufzeichnen, was er oder sie sich für einen gemeinsam zu nutzenden Park wünscht. Auch hier wurde nach dem Wunsch gefragt, statt die Nutzung der Stadt zum Ausgangspunkt zu machen. Und doch ist die Sache mit der Wünscherei nicht so einfach. Denn, um zur geschichtswissenschaftlichen Forschung zurückzukehren, welche Bestände zu digitalisieren sinnvoll wären, hängt zunächst einmal sehr stark von der Forschungsperspektive ab. Und tatsächlich verändert sich nichts schneller als Forschungsfragen, blickt man auf die vielen „Turns“ in der Geschichtswissenschaft der letzten Jahre. Durch den iconic turn wurden auf einmal Photographien, Gemälde und Bilder für Historiker interessant und dementsprechend auch in den Archiven gesucht.² Ebenso gab es die Einsicht, dass Karten und Pläne nicht nur für Geographen wichtige Objekte der Forschung sein können.³ Der „material turn“ lässt

¹ Vgl. <www.parkfiction.org> [27.02.2014]; Uwe Lewitzky, Kunst für alle? Kunst im öffentlichen Raum zwischen Partizipation, Intervention und Neuer Urbanität, Bielefeld 2005.

² Zur Bedeutung von Photographien und Bildern als Quellen für die Geschichtswissenschaft vgl. Gerhard Paul, BilderMACHT. Studien zur "Visual History" des 20. und 21. Jahrhunderts, Göttingen 2013; Gerhard Paul, Von der historischen Bildkunde zur Visual History. Eine Einführung; Ders. (Hrsg.), Visual history. Ein Studienbuch, Göttingen 2006, S. 7-36.

³ Vergleiche verschiedener Studien aus der Wissenschafts- und der Kulturgeschichte: Ute Schneider, Kartographie als imperiale Raumgestaltung. Alexander (Sándor) Radós Karten und Atlanten, in: Zeithistorische Forschungen/Studies in Contemporary History, 3, 1, 2006, S. 1-10, <www.zeithistorische-forschungen.de/16126041-Schneider-1-2006> [27.02.2014]; Bo Strath, Karten - Repräsentationen Europas aus vier Jahrhunderten, in: Rüdiger Hohls/Iris Schröder/Hannes Siegrist (Hrsg.), Europa und die Europäer. Quellen und Essays zur modernen europäischen Geschichte. Festschrift für Hartmut Kaelble zum

gegenwärtig Geschichtswissenschaftler, Kulturwissenschaftler, aber auch Kunsthistoriker, Denkmalpfleger und Bauhistoriker zusammenrücken. Modelle, Objekte, Alltagsgegenstände, Dinge aller Art und bauliche Hinterlassenschaften rücken so stärker in den Fokus der Forschung.⁴ Deshalb kann aus meiner Sicht die Frage auch nicht lauten: "Was wünschst Du Dir?", sondern wie nutzt Du Dein Archiv. Und so möchte ich die Blickrichtung auf die unterschiedlichen Funktionen und Nutzungen des Archivs lenken und Beispiele für einige Herausforderungen und Vorteile digitalisierter Bestände im Archiv skizzieren.

1. Archive generieren Forschungsfragen

Auf dem Archivwissenschaftlichen Kolloquium wurde mehrfach diskutiert, ob immer nur Forschungsfragen Ausgangspunkt für Digitalisierungsprojekte im Archiv sein können, oder ob nicht auch der umgekehrte Weg gegangen werden kann, indem ein herausragender Bestand digitalisiert und zum Ausgangspunkt für neue Forschungsfragen wird. So einen Fall möchte ich Ihnen vorstellen.

Das Leibniz-Institut für Regionalentwicklung und Strukturplanung (IRS) in Erkner ist ein sozialwissenschaftlich fundiertes Institut, hat

65. Geburtstag, Stuttgart 2005, S. 237-249; Sylvia Necker, Eingeschriebene Ideen. Zur besonderen Bedeutung von Karten und Plänen als Quellen für die Geschichtswissenschaft, in: Blätter für deutsche Landesgeschichte 147, 2011, S. 171-182.

⁴ Hans Peter Hahn, Materielle Kultur. Eine Einführung, Berlin 2005; Andreas Ludwig, Materielle Kultur, Version: 1.0, in: Docupedia-Zeitgeschichte, 30.05.2011, <http://docupedia.de/zg/Materielle_Kultur?oldid=88033> [27.02.2014].

aber seit knapp zwei Jahren eine neue Abteilung: die Historische Forschungsstelle/Wissenschaftliche Sammlungen.⁵ Sie beherbergt eine Forschungsabteilung, die sich hauptsächlich Fragen der DDR-Bau- und Planungsgeschichte widmet, sowie ein Archiv mit Teilbeständen aus der DDR-Bauakademie und etlichen Nachlässen von DDR-Architekten. Auf diese Weise sind Forschung und Archiv in der neu geschaffenen Abteilung sehr eng verknüpft, was Auswirkungen auf die Funktion und Nutzung des Archivs hat. Die Historische Forschungsstelle ist durchaus bemüht das Archiv nicht nur als „Backup“ zu nutzen, auf das leicht zurückgegriffen werden kann, sondern die stetig wachsenden Bestände im Archiv sollen auch neue Forschungsprojekte generieren und ermöglichen.

Seit 2012 sind die Wissenschaftlichen Sammlungen Projektpartner in einem vom Deutschen Museum München geleiteten und vom Senatsausschuss Wettbewerb (SAW) der Leibniz-Gemeinschaft bis 2015 finanzierten Projekt. Im Projekt DigiPortA - Digitalisierung und Erschließung von Porträtbeständen in Archiven der Leibniz-Gemeinschaft werden Porträtbestände digitalisiert und erschlossen. Der Bestand von ca. 33.000 Fotografien, Zeichnungen und Gemälden mit Portraits u.a. von Wissenschaftlern, Industriellen, Architekten, Künstlern, Technikern wird zum Abschluss des Projekts recherchierbar online gestellt. Im Teilprojekt des IRS werden Aufnahmeanträge des Bundes der Architekten der DDR erschlossen und in die gemeinsame Datenbank eingegeben. Bis auf wenige verlorene Teilbestände, wurden die Anträge mit Angaben zu Lebenslauf,

⁵ Vgl. <<http://www.irs-net.de/profil/wissenschaftliche-sammlungen/index.php>> [27.02.2014].

beruflichem Werdegang und teilweise Werklisten komplett erschlossen.⁶ Das Sample der digitalisierten Aufnahmeanträge ist groß genug, um daran auch quantitative Fragen zu stellen: Wer hat bei wem studiert? Wer hat bei wem eine Anstellung gefunden? Wer wurde von wem empfohlen? Wie war die berufliche Herkunft der Eltern zusammengesetzt? Wer hat an welchen Wettbewerben teilgenommen? Bestanden Parteimitgliedschaft oder Mitgliedschaften in Berufs- und Interessensverbänden?⁷ Wie sah die räumliche Streuung der Bauaufgaben wie z.B. der Wohnungstypen aus? Wo etwa wurden die meisten Kulturhäuser errichtet, wo die meisten Ferienarchitekturen? Auch Fragen der Rationalisierung des Bauwesens können recherchiert werden, da Pilotprojekte in den angehängten Werkverzeichnissen angegeben wurden.

⁶ Vgl. <<http://www.irs-net.de/forschung/forschungsabteilung-5/digiporta/>> [27.02.2014].

⁷ Vgl. Abb. 1 und 2.

1355

Bund der Architekten der DDR

Bezirksgruppe Erfurt

Aufgenommene: 1971
 Anweis: 01842
12.14.72
Karteibogen

(Bitte mit Schreibmaschine ausfüllen und 2 Paßbilder hinzufügen)

7/2

1. Name: [redacted] Vorname: [redacted]
 (bei Frauen auch Geburtsname)

2. Geburtsdatum: [redacted] 3. Geburtsort und Kreis: Greitz-Gr.Schmalkalden

4. Staatsangehörigkeit: DDR 5. Familienstand: verh.

6. Jetzige Wohnung: [redacted]
 Telefonnummer pr.: [redacted]

7. Soziale Herkunft (Beruf des Vaters bzw. der Mutter) Malermeister

8. Parteizugehörigkeit: keine Gesellschaftliche Organisationen: FDGB, DSP, Gesellschaft

9. Ausbildung: Krankenhauswesen der DDR

a) Erlernter Beruf: (Dr.-Ing Architekt) Maler
 Art des Abschlusses: Pacharbeiter

b) Fachschulstudium: nein von [redacted] bis [redacted]
 Art des Abschlusses: [redacted]

c) Hochschulstudium: ja
 Wo studiert? HAB Weimar von [redacted] bis [redacted]

Fakultät oder Abteilung: Architektur

Art des Abschlusses: Dipl.-Ing

Welche akademischen Grade? Dr.-Ing des Wissenschaftszweiges

Wann und durch wen verliehen: 1970 TU Dresden Sektion Architektur

d) Fernstudium: nein von [redacted] bis [redacted]
 Art des Abschlusses: [redacted]

e) Abendstudium: nein von [redacted] bis [redacted]
 Art des Abschlusses: [redacted]

f) Sonstiges Studium: FDV-Intensivkurs(14 Tage); Weiterbildungslehrgang in Städtebau
und Architektur (4 Wochen)

g) Besuch von Weiterbildungsmaßnahmen des BdA: [redacted]

Abb. 1 Formular Aufnahmeantrag des Bundes der Architekten der DDR, Blatt 1, Quelle: IRS/Wiss.Samml., B2-1355.

16. Verzeichnis Ihrer wichtigsten Wettbewerbsfolge:

Art und Bezeichnung des Wettbewerbs	Ort	Auslober	Jahr der Auslobung	Kollektiv oder Einzelteilnehmer
Ledigenwohnheim	Berlin		1957	

16 a. Entscheidung der Jury:

17. Auszeichnungen:

Staatliche Auszeichnungen:	Wann?	Auszeichnungen gesellsch. Organisationen:	Wann?
Aktivist der Sozialistischen Arbeit	1970		

den 22. 3. 1972

34.3.72

(Unterschrift)

157 Bm 1524 71 3.5

Abb. 2 Formular Aufnahmeantrag des Bundes der Architekten der DDR, Blatt 4, Quelle: IRS/Wiss.Samml., B2-1355.

Es handelt sich also um einen geschlossenen Bestand, der über eine quantitative Auswertung der Daten neue qualitative Forschungsfragen etwa über kollektivbiographische Ansätze oder über eine historische Netzwerkanalyse ermöglicht. Natürlich haben Architekturhistoriker das bisher auch schon immer gemacht. Aber die Suche nach Daten zu Schüler-Lehrer-Beziehungen, zu einzelnen Werken und zu Stationen im beruflichen Werdegang war mühsamer und oft von Zufällen und der Zugänglichkeit und dem Überlieferungszustand einzelner Bestände abhängig. Die Auswertung der Daten aus dem Aufnahmeantrag ermöglichen jetzt einen systematischen Zugriff auf Lebensläufe von Architekten und Architektinnen in der DDR.

2. Ordnungssysteme im digitalen Dschungel

Diese Art der Nutzung leuchtet sofort ein, zumal der Bestand in einer überschaubaren Größe ist. Doch wie sieht es aus, wenn wir wirklich große Teile des Archivguts digital zur Verfügung hätten? Schließlich träumen doch viele Historikerinnen und Historiker von der Akte per Mausklick. Doch stellt sich nicht nur für die Nutzer, sondern auch für die Archive die Frage, was eigentlich verfügbar gemacht werden sollte und welches Archiv bei nicht gerade wachsender Personaldecke die Ressourcen hat, Archivgut in diesen Mengen digital zu erfassen und zu erschließen. Und vor allem: ist für den Nutzer ein browsen durch das Archiv ein sinnvoller Zugang zu Archivgut? Quasi lieber Bindehautentzündung als Staublunge?

Als Beispiel sei das Archiv des Leo Baeck Instituts New York angeführt. Auf der Website des Instituts ist es möglich, geschlossene

Nachlässe, Konvolute, Photographien und auch handschriftliche Dokumente abzurufen.⁸ Der Einstieg in die Bestände des Archivs ist nahezu barrierefrei – durch ein Suchfeld kann man zu allen angebotenen Ressourcen gelangen. Der direkte Zugriff befreit die Forscherinnen und Forscher von Einschränkungen wie Öffnungszeiten der Archive sowie Kopier- und Reisekosten, die durch digitalisierte Bestände erheblich gemindert werden können. Auch ist eine Nachnutzung von Akten für andere Forschungszwecke durch den 24-Stunden-Zugriff jetzt möglich. Das sind alles unschätzbare Vorteile der Archive im digitalen Zeitalter. Und doch möchte ich einen ganz kleinen Wehrmutstropfen anbringen: Die Tektonik des Archivs bildet sich auf manchen Portalen – dazu gehört aus meiner Sicht auch das Leo Baeck Archiv – nur sehr schemenhaft ab. So kann das Suchen schnell frustrierend werden. Denn ohne genau zu wissen, nach was man sucht, kommt man über eine Stichwortsuche nicht weiter. Portale wie das des Leo-Baeck-Instituts sind also dann besonders hilfreich, wenn konkrete Nachlässe oder Bestände gesucht werden. Ein digitales Durchstreifen des Archivs über das Suchfeld ist nach meiner Erfahrung leider nicht so zielführend wie die Durchsicht eines Findbuchs, wobei dies auch an der noch nicht genug ausgeprägten Kompetenz von Historikerinnen und Historikern im Umgang mit digitalen Ordnungssystemen liegen kann.⁹ Die digitale Haptik will eben erst erlernt sein. Insofern fände ich es umso wichtiger, dass Archive

⁸ Vgl. <<http://www.lbi.org/digibaeck/>> [27.02.2014].

⁹ Zur Problematisierung digitaler Ordnungssysteme in der geschichtswissenschaftlichen Forschung vgl. Peter Haber, *Digital Past. Geschichtswissenschaft im digitalen Zeitalter*, München 2011.

im Rahmen der zukünftigen Digitalisierungsprojekte, auf eine möglichst gute Ablesbarkeit ihrer Tektonik achten.¹⁰ Zudem die wichtige Vermittlerfigur "Archivar" im digitalen Archiv wegfällt.¹¹ In den meisten Stadt-, Kreis- und Landesarchiven sowie im Bundesarchiv ist die Ablesbarkeit der Tektonik durch Benutzeroberflächen wie ARGUS oder den über Midosa bereitgestellten digitalen Findbüchern der Fall. Auch die schon in der Deutschen Digitalen Bibliothek eingespeisten Archivalien, lassen sich über einen „Tektonik-Baum“ eindeutig zuordnen.¹²

3. Digitale Haptik

Der Aspekt der digitalen Haptik hat aber nicht nur mit Organisationsstruktur zu tun, sondern auch mit der Frage der Präsentation. Nicht nur die Qualität der gedruckten Produkte ist sehr unterschiedlich, da nur noch wenige in Bildrechte, Papier und Druckverfahren investieren möchten, auch nicht jede Website und jedes Webportal lässt sich mit dem gleichen Vergnügen durchklicken. Aus meiner Sicht werden die interaktiven Möglichkeiten und Verknüpfung von Wissen, Objekten, Webseiten und Institutionen noch viel zu wenig genutzt. Und so

¹¹ Außer die Idee eines Live-Chats mit Archivaren würde sich durchsetzen. Zumindest eine Vorstufe zur virtuellen Kommunikation wird an Bibliotheken schon länger praktiziert. Vgl. u.a. <<http://www.sub.uni-goettingen.de/kontakt/frag-die-sub/>> [27.02.2014]. Ein interaktives Beispiel ist „Stella“, die einen auf der Startseite der Staats- und Universitätsbibliothek Hamburg mit folgenden Worten begrüßt: „Guten Tag! Mein Name ist Stella. Ich ahne, Sie haben eine Frage. Kann ich helfen?“. Vgl. <<http://www.sub.uni-hamburg.de/home.html>> [27.02.2014].

¹² Vgl. <<https://www.deutsche-digitale-bibliothek.de/item/VUNDJEJ7HJXSESCNJCZNWJ65O5O4LI3B>> [27.02.2014].

bleibt es manchmal bei einer schnöden digitalen Kopie der Archivalie. Natürlich bedeutet Digitalisierung im ersten Schritt Bestandsicherung und Bewahrung der Kulturgüter – teuer genug –, aber es bedürfte eben gerade auch an Investitionen in die digitalen Umgebungen, also in die Präsentation der Bestände. Oder anders gesagt: neben der technischen Seite müssen wir uns auch der Ästhetik widmen und für Nutzer Anreize schaffen, die Digitalisate auch zu finden und zu benutzen. Die Befürchtung ist auch, dass die vielen innovativen Projekte der letzten Jahre sich in den kommenden Jahren aus Personal- und Finanzierungsmangel auf das Nötigste beschränken müssen und die digitalen Portale und Produkte nicht weiter ergänzt und erweitert werden können. Gerade weil das Internet dynamisch ist, braucht es permanente Betreuung.

4. Digitale Sammelbecken

Bleiben wir bei den innovativen Projekten. In den letzten Jahren wurden mehrere Meta-Portale vorgestellt, die als digitale Sammelbecken von Kulturgut und Wissen aller Art fungieren, wie etwa das Portal „europeana“¹³ oder die auf nationaler Ebene angesiedelte „Deutsche Digitale Bibliothek (DDB)“¹⁴, die sich noch in der beta-Phase befindet.¹⁵ Versucht man die digitalen Fundgruben auf die

¹³ Vgl. <<http://www.europeana.eu/>> [27.02.2014].

¹⁴ Vgl. <<https://www.deutsche-digitale-bibliothek.de/>> [27.02.2014].

¹⁵ Zur Diskussion um Archivierung, Bewahrung und Nutzung von Kulturgut im digitalen Zeitalter vgl. Hedy Graber/Dominik Landwehr/Veronika Sellier (Hrsg.), *Kultur digital. Begriffe, Hintergründe, Beispiele* (Migros-Kulturprozent), Basel 2011; Alfredo M. Ronchi, *ECulture. Cultural Content in the Digital Age*, Berlin, Heidelberg 2008; Gerald Maier, *Kulturgut aus Archiven, Bibliotheken und Museen*

eigenen Forschungsschwerpunkte abzuklopfen, entsteht ein ambivalentes Bild: einerseits reizt die Fülle des unterschiedlichen Materials, das dort präsentiert wird. Schier unendliche Photobestände aus allen Winkeln Europas lassen sich nun sichten und für die eigenen Forschungen nutzen. Doch wie lassen sich die Funde gewichten? Gibt man beispielsweise die Schlagworte „Architektur“ und „Nationalsozialismus“ ein, zeigt mir die DDB gleich mehrere Objekte, darunter ein schöner Zufallsfund aus dem Landesamt für Denkmalpflege Hessen. Dokumentiert ist eine Kfz-Halle des ehemaligen Führerhauptquartiers 'Adlerhorst' in Ober-Mörlen im hessischen Wetteraukreis.¹⁶ Sehr positiv ist die beigefügte genaue Beschreibung des Objektes, die leider nicht allen in der DDB aufzufindenden Objekten angehängt ist. Trotz des Auszugs aus dem Denkmal-Inventar ist die Relevanz dieser Kfz-Halle schwer einzuschätzen. Kann sie als typischer NS-Bau in Hessen gelten? Wie viele Bauten dieser Art gibt es? Ist der Architekt für die architekturgeschichtliche Forschung relevant? Und in welcher Beziehung steht der Bau zu anderen NS-Bauten in Hessen bzw. im Deutschen Reich? Der Kontext – also die entscheidende

im Internet. Die Digitalisierung des kulturellen Erbes in Deutschland auf nationaler und europäischer Ebene, in: Matthias Buchholz (Hrsg.), *Samisdat in Mitteleuropa. Prozeß, Archiv, Erinnerung*, Dresden 2007, S. 127-139. Für die geschichtswissenschaftliche Forschung stellt sich die Frage nach der digitalen Haptik auch in anderer Weise, ob das Digitalisat die Quelle adäquat repräsentieren kann. Vgl. Daniel Jeller, *Die historische Quelle im Zeitalter ihrer digitalen Reproduzierbarkeit. Eine Frage nach der Wechselwirkung zwischen historischen Quellen und Identität und nach der Veränderung dieses Verhältnisses durch die Digitalisierung*, in: Adelheid Krah, *Die virtuelle Urkundenlandschaft der Diözese Passau. Vorträge der Tagung vom 16./17. September 2010 in Passau*, Passau, 2011, S. 45-56.

¹⁶ Vgl. Abb. 3.

Kategorie für geschichtswissenschaftliche Forschung – ließe sich nur über eine viel, viel größere Datenbasis einschätzen.¹⁷



Abb. 3 Screenshot Deutsche Digitale Bibliothek, 27.02.2014.

¹⁷ Dieser Kritikpunkt ist insofern zu entkräften, als sich die DDB noch im beta-Modus befindet und die Zielstellung des Portals noch nicht erreicht ist: „Ziel der Deutschen Digitalen Bibliothek (DDB) ist es, jedermann über das Internet freien Zugang zum kulturellen und wissenschaftlichen Erbe Deutschlands zu eröffnen, also zu Millionen von Büchern, Archivalien, Bildern, Skulpturen, Musikstücken und anderen Tondokumenten, Filmen und Noten. Als zentrales nationales Portal soll die DDB perspektivisch die digitalen Angebote aller deutschen Kultur- und Wissenschaftseinrichtungen miteinander vernetzen. (...) Bis diese Vision in vollem Umfang verwirklicht ist, muss noch ein weiter Weg gegangen werden. Die Digitalisierung des deutschen kulturellen und wissenschaftlichen Erbes ist ein Prozess, der auf Jahre angelegt ist“. Vgl. <<https://www.deutsche-digitale-bibliothek.de/content/about/>> [27.02.2014].

5. Archivportale und Massendigitalisierung

Lässt sich daraus also folgern, dass man nur warten muss, bis alle Bestände in die regionalen, nationalen und internationalen Archiv-, Bibliotheks- und Museumsportale eingespeist sind oder bleibt die Vorstellung, einer totalen Erfassung aller Kulturgüter der Welt eine Utopie? So verlockend die unbegrenzte Zugänglichkeit auch ist, bleibt dennoch die Frage nach der Bezahlbarkeit und der Sinnhaftigkeit einer Massendigitalisierung von Archivalien. Vielmehr könnte das u.a. vom Deutschen Museum München propagierte dezentrale Sammeln bzw. das "Sammeln im Verbund"¹⁸ ebenso interessant sein wie eine Bündelung in thematisch orientierten Fachportalen, die auf dem Archivwissenschaftlichen Kolloquium immer wieder thematisiert wurde. Abzuwarten bleibt, welche Aufgabe in diesem Zusammenhang das in der Entwicklung befindliche „Archivportal-D“¹⁹ übernehmen kann, um digitalisierte und digitale Quellen aus Archiven für die Öffentlichkeit und für die Forschung bereit zu stellen.²⁰ Eine

¹⁸ Vgl. <<http://www.deutsches-museum.de/archiv/sammeln-im-verbund/>> [27.02.2014].

¹⁹ Vgl. <<http://www.landesarchiv-bw.de/web/54267>> [27.02.2014], <<http://www.landesarchiv-bw.de/web/54699>> [27.02.2014], <<http://www.archivschule.de/DE/forschung/archivportal-d/>> [27.02.2014].

²⁰ Dass es die Sache nicht zwingend einfacher macht, wenn Bestände schon digital überliefert werden, zeigen die Problemaufrisse von Rainer Hering zum Umgang mit E-Mails als Quellen für die geschichtswissenschaftliche Forschung: Rainer Hering, Digitale Welt - digitale Quellen. Herausforderungen für die historische Forschung, in: Blätter für deutsche Landesgeschichte 147, 2011, S. 149-158; Rainer Hering, Ohnmächtig vor Bits and Bytes? Archivische Aufgaben im Zeitalter der Digitalisierung, in: Ders. (Hrsg.), Wie mächtig sind Archive? Perspektiven der Archivwissenschaft, Hamburg 2013, S. 83-97.

Digitalisierung in Masse und Gänze scheint mir jedoch für zwei Gattungen sehr sinnvoll zu sein: zum einen sind die Findbücher – ob über einen digitalen oder analogen Zugriff – immer noch unser zentraler Zugang ins Archiv. Sie sollten, wenn Ressourcen und Zeit es zulassen, so tief wie möglich erschlossen werden. Zum anderen sind als historische Quelle Zeitschriften und Zeitungen ein sehr wichtiger Bestand, der realistischer Weise in Masse digitalisiert werden kann.²¹ Für die Forschung bedeutet ein möglichst vollständiger Zugang zu digitalisierten Zeitschriften und Zeitungen eine große Erleichterung der Arbeit, zumal, wenn sich die Digitalisate entsprechend über OCR im Volltext durchsuchen ließen.

6. Präsentation von Digitalisaten in Online-Ausstellungen

Zuletzt möchte ich noch einen Blick auf das wachsende Angebot von Präsentationsplattformen richten, die Quellen – seien es Photos, Objekte oder Akten – in thematischen Zusammenhängen zugänglich machen. Für Archive ist es die Möglichkeit, besondere Stücke aus dem Bestand zu zeigen. So ist auf der Startseite des Landesarchivs Baden-Württemberg jeweils die „Archivalie des Monats“²² zu finden. Das Bundesarchiv zeigt „thematische Galerien aus Bildern und Dokumenten“²³. Für Museen ergibt sich die Möglichkeit, Ausstellungen zumindest in Teilen zu archivieren und als virtuelle Ausstellung

²¹ Vgl. Birgit Seiderer, Die Digitalisierung von Zeitungen im deutschsprachigen Raum - ein Zustandsbericht, in: Zeitschrift für Bibliothekswesen und Bibliographie 57(2010)3/4, S. 165-171.

²² Vgl. <<http://www.landesarchiv-bw.de/web/56169>> [27.02.2014].

im Internet zu präsentieren.²⁴ Oftmals sind mühsam aus den Archiven zusammengetragene Originale nur für die Dauer einer Ausstellung zu sehen. Danach verschwinden die Dokumente, Photos und Objekte wieder in den Magazinen und den Depots. Mit einer Übertragung der analogen Ausstellung in eine virtuelle Museumsplattform lässt sich das Archivierte in seiner Ausstellungsanordnung aufs Neue archivieren. Die zeigende Institution – das Museum –, wie auch die speichernde Institution – das Archiv – können so gewinnbringend verknüpft werden.

Im September 2012 wurde die von mir kuratierte Ausstellung "Hitlerbauten in Linz. Wohnsiedlungen zwischen Alltag und Geschichte. 1938 bis zur Gegenwart" im NORDICO Stadtmuseum in Linz eröffnet. Für vier Monate waren teilweise bisher noch vollkommen unbekannte Karten und Pläne aus dem Archiv der Stadt Linz und privaten Sammlungen zu sehen. Schon während der Konzeption der Ausstellung hat mich das forum oberösterreichische geschichte²⁵ auf die Möglichkeit der Übertragung möglichst großer Teile der Ausstellung auf die Seiten „Virtuelles Museum Oberösterreich“²⁶ angesprochen. Letztlich konnten wir alle Saaltexte sowie die Beschreibungen zu den Objekten „retten“ und als Online-Ausstellung

²³ Vgl. <http://www.bundesarchiv.de/oeffentlichkeitsarbeit/bilder_dokumente/index.html.de> [27.02.2014].

²⁴ Gerald Maier, Thomas Fricke (Hrsg.), Kulturgut aus Archiven, Bibliotheken und Museen im Internet. Neue Ansätze und Techniken, Stuttgart 2004.

²⁵ Vgl. Abb. 4.

²⁶ Vgl. Forum OÖ Geschichte. Virtuelles Museum Oberösterreich: <<http://www.oogeschichte.at/>> [27.02.2014].

präsentieren. Nutzergruppen wie Schülerinnen und Schüler sowie die "interessierte" Öffentlichkeit können so an diesen Digitalisaten „nachhaltig“ partizipieren.

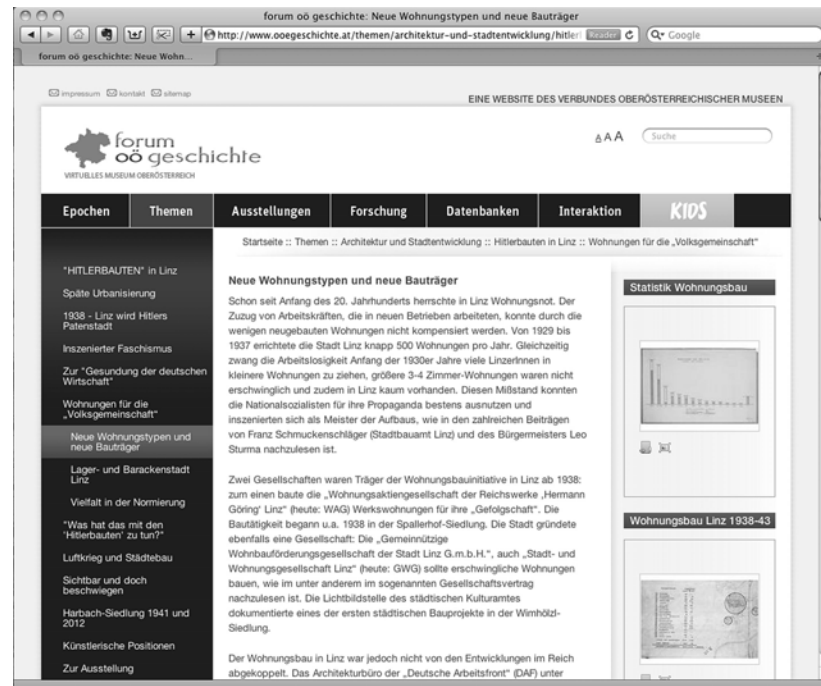


Abb. 4 Screenshot Virtuelles Museum Oberösterreich, 27.02.2014.

7. Digitales Archiv was nun?

Was also ist das Wunscharchiv im digitalen Zeitalter? Und welche Bedeutung hat das Verhältnis von analogen und digitalen Recherchen? Zunächst einmal scheint es mir nicht zwingend, eine vollständige Digitalisierung aller Bestände anzustreben, vielmehr sollten zunächst Bestandsverzeichnisse und Findbücher detailliert über vorhandene Archivalien Auskunft geben. Ob dann Photographien, Karten, Protokolle oder Nachlässe wichtiger sind als Behördenkorrespondenz und welche Archivalien priorisiert digitalisiert werden, hängt letztlich von den Forschungsfragen ab, die sich je nach Generation und Epoche stark verändern können. Deshalb kann streng genommen eine Empfehlung, was priorisiert werden soll, gar nicht gegeben werden. Zudem rechtliche Bestimmungen einen Rahmen für die Digitalisierung vorgeben, die z.B. für mein Forschungsfeld "Zeitgeschichte" sehr ungünstig sind.

Trotz aller vorgebrachten Bedenken: die Digitalisierungsbestrebungen und intensiven Debatten um eine Priorisierung sind auch aus der Perspektive der Forschung erstrebenswert. Dennoch sollte in Archiven die Zukunft nicht ausschließlich in digitalen Kategorien gedacht werden. Immer noch harren in den Depots vieler Archive unerschlossene Bestände, die über den Digitalisierungs-Hype nicht in Vergessenheit geraten sollten. Mein Wunsch für das Archiv der Zukunft lautet also digitales und analog-archivalisches Arbeiten zu verbinden und nicht das eine über das andere zu vergessen.

Nutzen und Gefahren des digitalen Dokuments. Erfahrungsbericht aus Sicht der Unternehmensgeschichte

von Andrea H. Schneider

I. Einleitung

Im 21. Jahrhundert scheint der Umgang mit digitalen Dokumenten gewiss als ein Standard. Dennoch gibt es eine Vielzahl von Aspekten, die man sowohl bei der Erstellung als auch bei der Verfügbarmachung von historischem Material in Form von Digitalisaten berücksichtigen muss. Gleiches gilt jedoch auch für die Nutzerperspektive. Dieser Aufsatz versucht weniger in einer theoretischen Überlegung, sondern vielmehr als Ergebnis praktischer Erfahrungen Denkanstöße zu liefern, die beim bevorstehenden oder bereits begonnenen Prozess der Digitalisierung in Archiven von Interesse sein könnten.

Die Unternehmensgeschichte als Beispiel für diese Überlegungen zu nehmen, liegt insofern nahe als Überlieferungen aus diesem historischen Bereich sowohl in privat geführten Unternehmensarchiven bei Firmen, Banken und Institutionen existieren. Darüber hinaus aber auch Archivbestände zum Thema Unternehmens- und Wirtschaftsgeschichte in öffentlichen Archiven liegen.

Die Gesellschaft für Unternehmensgeschichte (GUG) wiederum unterhält kein Unternehmensarchiv, steht aber in engem Kontakt zur Vereinigung der deutschen Wirtschaftsarchivare. Dort wiederum finden sich die Vertreter der einzelnen Unternehmen, die ein historisches Archiv unterhalten, zusammen. Die GUG berät und begleitet

seit den 1980er Jahren Unternehmen zunächst vorwiegend bei der Erforschung ihrer eigenen Geschichte. In den letzten Jahren ist jedoch häufiger im Umfeld einer solchen Studie auch der Wunsch in den betreffenden Unternehmen entstanden, das gesammelte historische Material einmal in ein Archiv einzubringen und bestandswahrend zu lagern. Auch beobachten wir häufig, dass Unternehmen durch die Aufarbeitung ihrer Geschichte die Bedeutung von historischem Material erkennen und dadurch sich veranlasst sehen, über das bereits gesammelte historische Material hinaus, zukunftsweisend auch das aktuelle Aktenmaterial einer Vorschrift zur historischen Bewahrung zu unterziehen. In diesem Prozess berät und begleitet die GUG verschiedene Unternehmen. In diesem Zusammenhang ist das Thema „Digitales Archiv“ auch ein praktisches und aktuelles für die GUG.

Wenden wir uns also der Frage des Nutzens von Digitalisaten zu. Naheliegend und nicht von der Hand zu weisen ist die Antwort, dass dies von sehr großem Nutzen für die Forschung ist. Zukunftsutopien lebend muss der Historiker in dieser Vorstellung einer digitalen Welt nicht mehr in die Archive dieser Welt reisen, sondern sitzt in seinem Büro und erhält per Knopfdruck über die Datenautobahnen die Dokumente zum Ausdruck am eigenen PC. Ein enormer Nutzen in Punkto Zeit- und Kostenmanagement. Das Digitalisat entspricht in dieser Welt der Vorstellung Aldous Huxleys von der „schönen neuen Welt“.

Mittlerweile beschäftigt sich die GUG seit Jahren mit Fragen rund um das digitale Dokument. Wir haben erfahren, dass das Verständnis von Historikern/Nutzern und den IT-Experten zum Thema Digitalisierung große Unterschiede aufweist; ein gemeinsames

Verständnis jedoch am Ende wichtig ist, um eine digitale Archivierung sinnvoll zu gestalten.

II. Perspektiven

a. Unternehmen

Grundsätzlich gibt es in Unternehmen keine Rechtsvorschrift, historisches Aktengut aufzubewahren. Lediglich die steuerlichen Zehnjahres-Fristen sind dort de jure maßgebend. Nichtsdestotrotz gibt es eine ganze Reihe (auch teilweise schon über 100 Jahre alte) Unternehmensarchive. Mit dem Forschungsboom in den 1990er Jahren im Rahmen der Studien zur Rolle der Unternehmen im Nationalsozialismus kam eine ganze Reihe weiterer Archive hinzu. Gerade in den letzten Jahren lässt sich ein wachsendes Bewusstsein verschiedener Unternehmen beobachten in Bezug auf die Bedeutung von Geschichte und die dafür notwendigen historischen Materialien. Bemerkenswert ist hierbei, dass vor allem kleinere und mittlere Unternehmen ihre Geschichte „entdecken“.

Die Motivationslage ist überaus unterschiedlich. Manchmal ist es ein bevorstehendes Jubiläum (so z.B. im Fall von Gerolsteiner Brunnen, das sich selbst zum 125jährigen Bestehen/Jubiläum ein Archiv schenkte). Zu diesem komme ich später noch, da die GUG den Aufbau des Archivs seit nunmehr drei Jahren bewerkstelligt. Im Falle der Linde SE ist es ein Forschungsprojekt für das Material gesammelt wurde und das heute in seiner Wertigkeit so eingeschätzt wird, dass für eine Professionalisierung des Umgangs mit dem historischen Material Geldmittel in die Hand genommen werden.

Zum Verständnis der privaten Unternehmensarchive ist wichtig zu sehen, dass in der Regel ein Archiv in einem Wirtschaftsbetrieb um Geldmittel werben muss bzw. seine Existenzberechtigung nachweisen muss. Es ist nicht selbstverständlich, dass es auf Dauer existiert und es muss seinen Mehrwert unter Beweis stellen. Dies ist allein mit der Bewahrung von Archivgut nicht getan, d.h. ein Archiv muss also aktive Beiträge im Zusammenhang mit dem tragenden Unternehmen leisten. In vielen Unternehmen gibt es jedoch heute noch kein professionelles historisches Archiv. Häufig sind Registraturen Ausgang oder auch Bewahrungsort von ungeordneten historischen Dokumenten.

Sind keine professionellen Archivare oder Historiker vor Ort, ist der Schritt naheliegend zu sagen: wir versuchen möglichst viel zu digitalisieren, damit wir aus Kostengründen Platz sparen und damit wir aus Dienstleistungsgründen schnell Digitalisate zur Verfügung stellen können.

Gerade für Markenartikelhersteller sind digitale Welten sehr angenehm, weil ein schneller Zugriff auf historisch unbedenkliches Material wie Werbebroschüren, Produktfotographien, etc. gewährleistet werden kann. Zum Beispiel lassen sich intern Bildmaterial, Werbekampagnen, etc. verfügbar halten, was für aktuelle Anwendungen bzw. Nachfrage von externen wie internen Nutzern wichtig sein kann. Auch lassen sich in einem digitalen Archiv teilweise schneller Antworten finden auf historische Fragen. Gerade auch für kleinere Unternehmen oder Unternehmen mit begrenzten Geldmitteln, die sich keine permanente Betreuung ihres Archivs leisten

wollen oder können, sind diese Vorstellungen des schönen, leicht erstellbaren und leicht benutzbaren digitalen Archivs verführerisch.

Aber das digitale Archiv ist auch mit großen Gefahren verbunden. Einmal kommt es darauf an, dass die digitalen Dokumente, Bilder oder Filme sehr gut verzeichnet sind und einem nach allen Regeln der Archivierung verfahrenen Verzeichnungssystem angelegt sind und die Herkunft und die Zusammenhänge der Akten nicht verloren gehen. Die größte Problematik liegt darin verborgen, dass einzelne Dokumente eben keinen Kontext liefern. Ein einzelnes, aus dem Kontext gerissenes Dokument ohne historischen Bezug kann sehr schnell ein falsches Ergebnis liefern. D.h. die Fachkompetenz ist auf allen Ebenen immer zu bedenken. Besonders gefährlich wird es auch dann, wenn digitale Dokumente versendet werden können, da sich dann die Gefahr von Fehltritten etc. potenziert.

Es stellt sich auch immer wieder die Frage, nach welcher Priorisierung man vorgeht in Punkto Digitalisierung und ob wirklich alle Akten, Bilder, etc. digitalisiert werden müssen. Wir haben dabei einen kleinen Kriterienkatalog aufgestellt, der sagt:

1. Wesentliche Unternehmensakten wie Protokolle der Leitungsgremien, Unternehmenskennzahlen und -berichte werden auch aus Gründen der zusätzlichen Sicherung digitalisiert. Die Originalakten werden jedoch unter allen Umständen archivgerecht und bestandswährend gelagert.
2. Weitere Aktenbestände werden nach Nutzungsanfragen digitalisiert, d.h. stellt sich heraus, dass mit gewissen Aktengruppen regelmäßig gearbeitet wird, werden diese digitalisiert bzw. Einzeldokumente bei Anfrage direkt digitalisiert.

3. Jüngere Aktenbestände, die ohnehin in digitaler Form vorliegen (z.B. Mitarbeiter- oder Kundenzeitschriften, Werbematerial, etc.) werden in digitaler Form ins Archiv überspielt (Belegexemplare werden aufbewahrt).
4. Filme werden in der Regel digitalisiert, da ihre Aufbewahrung kostenintensiv ist und sich dieser Aufwand meist bei kleineren Beständen nicht lohnt.
5. Fotos und Dias werden digitalisiert wegen der damit erzielten leichteren Verwendbarkeit.

Das Kernstück der Archivierung ist und bleibt allerdings die Verzeichnung und gute Verschlagwortung. Häufig verfolgen die IT-Abteilungen und mit ihnen auch die zuständigen Personen im Unternehmen das Motto, alles zu digitalisieren mit der Begründung des kostengünstigen Speicherplatzes. Das vermeintlich schnelle Einscannen aller Dokumente, die dann jeweils mit einem eindeutigen Code versehen werden und über OSCR Suchen auffindbar gemacht werden, scheint eine schnelle, kostengünstige und alle Belange befriedigende Lösung. In der Regel benötigt es jedoch langer Diskussionen, um den Spezialisten verständlich zu machen, dass für den Historiker nicht NUR das Finden des einzelnen Dokuments von Belang ist, sondern dass es für den Historiker darauf ankommt zu sehen, woher das Dokument stammt, welche Dokumente noch dazu gehören; der Historiker braucht die Akte und nicht das Dokument. Selbst automatische Querverweise sind nur begrenzt hilfreich. Und so schön die Meta-Suchen auch sind, so werfen sie doch nur, wenn nicht fachgerecht verzeichnet wurde, ein unzusammenhängendes

Konvolut von Dokumenten aus. In der Regel haben sie darüber hinaus nur eine Trefferquote, wenn sie gut sind, von 90 Prozent.

Natürlich ist auch dies sehr vorteilhaft, nur gibt es dabei ein sehr großes Problem: man kann nicht sicher sein, dass man ALLE Dokumente erhält. Denn bei der Meta-Suche wird eben auch nur gefunden, was der Nutzer namentlich kennt und damit eingeben kann. Es kann aber durchaus passieren, dass auf diesem Wege Dokumente nicht erscheinen.

Fazit ist keinesfalls, dass eine Digitalisierung nicht begrüßenswert wäre. Worauf es ankommt ist, dass eine Digitalisierung nur ein ergänzender Schritt sein kann. Dass dies nicht die Lösung oder gar der Ersatz der klassischen Archivierung sein darf.

Dies trifft dann auch auf den Punkt der Sicherung zu. Auch hier ist die Digitalisierung und vor allem der Speicherplatz sehr viel kostengünstiger geworden, nichtsdestotrotz unterliegen die digitalen Archive anderen Gefahren bei der Langzeitspeicherung als dies mit klassischem Papier der Fall ist. Aber diese Diskussion führen die Archive sicherlich an anderer Stelle.

Die Frage der Priorisierung der Digitalisierung von Aktenmaterial lässt sich aus Sicht der Unternehmen und der historischen Nutzung relativ eindeutig beantworten: eine relevanzorientierte (Kernakten des Unternehmens) gepaart mit einer nutzungsorientierten (intern Kommunikationsabteilungen etc.). Im Archiv der Porsche SE verfährt man z.B. so, dass bei jeder externen Anfrage, die angefragten Dokumente digitalisiert werden. Dies ist zwar eine weniger systematische Herangehensweise, ist aber extrem nutzerfreundlich, weil eben jeder Nutzer unmittelbar bedient wird. Dies wird zwangsläufig über

die Zeit dazu führen, dass ein großer Teil des Bestandes digitalisiert wird.

Eine solche Herangehensweise bedeutet natürlich, dass der Bestand erst im Laufe von vielen Jahren digitalisiert ist und wird deswegen möglicherweise von den anbietenden Digitalisierungsfirmen nicht positiv gesehen. Denn auf diesem Wege kann die Digitalisierung in-house durchgeführt werden. Die dafür nötigen Scanner sind erschwinglich und haben eine hohe Geschwindigkeit, so dass dies ein relativ pragmatischer Ansatz zur Frage der Priorisierung sein kann.

b. Historiker

Wie aber ist die Sicht des Historikers? Des Unternehmenshistorikers? Die moderne Unternehmensgeschichte beschränkt sich nicht auf das Innenleben eines Unternehmens: z.B. kontextualisiert eine Fallstudie das Unternehmen und seine Rolle in Wirtschaft und Gesellschaft und die wechselseitigen Einflüsse werden ebenfalls mit untersucht. Politische Rahmenbedingungen und ein grundlegendes Marktverständnis sind heute Grundvoraussetzungen für wissenschaftliches Arbeiten in dieser Spezialdisziplin. Darüber hinaus befasst sich die moderne Unternehmensgeschichte aber häufig auch nicht nur mit einem Unternehmen, sondern untersucht Branchen oder Cluster. Ein jüngster Trend nimmt das Verhältnis von Regionen und internationalen Unternehmen in den Fokus.

Darüber hinaus hat sich in den letzten 20 Jahren auch der Schwerpunkt geändert. Wurden früher sehr häufig Unternehmen der Großindustrie und dabei in der Regel Schwerindustrie untersucht,

sind in jüngster Zeit gerade die kleinen und mittelständischen Unternehmen in den Fokus gerückt. Der neueste Trend – zumindest international – ist es, auch Bereiche „beyond the firm“ mit einzubeziehen, d.h. Non-Profit-Organisationen, Vereine, etc.

Wenn man also versucht eine Priorisierung vorzunehmen, wird sehr schnell klar, dass der Historiker für die verschiedenen historischen Perspektiven, Fragestellungen und Methoden sehr unterschiedliche Akten benötigt. Soll eine Fallstudie zu einem Unternehmen durchgeführt werden, dann benötigt der Historiker möglichst umfassende und tiefgehende Akten dieses Unternehmens. Sie finden sich nicht nur in den Archiven von Unternehmen selbst, sondern häufig auch in den regionalen Wirtschaftsarchiven oder auch im Bundesarchiv. Ein Historiker mit einem solchen Forschungsansatz würde, gefragt nach der Priorisierung, wohl antworten, es sei zielführend, möglichst einen Bestand nach dem nächsten komplett zu digitalisieren, da ihm nur auf diesem Weg umfassend geholfen wäre.

Was für ihn jedoch sehr hilfreich wäre, wären z.B. komplette Digitalisierungen von Behördenakten, die dann mit einer Metasuche durchsucht werden könnten. Dabei würde es jedoch nicht nur darauf ankommen, die Texte, sondern auch z.B. die Briefköpfe etc. mit in die Suchfunktion zu integrieren. Wenn wir Historiker nach Behörden dokumenten suchen – und gerade im Hinblick auf kleinere und wenig bekanntere Unternehmen – dann tauchen diese in der Regel in den Findbüchern kaum auf. Dennoch mag es Schriftwechsel geben oder zumindest einzelne Dokumente, die bislang aufgrund des immensen Suchaufwandes gar nicht erst gesucht, oder eben auch einfach nicht gefunden werden. Sie können aber dennoch relevant sein.

Dabei ist aus der Sicht der Unternehmensgeschichte natürlich ganz klar, dass uns die Akten aus dem Wirtschafts- und Finanzministerium am wichtigsten sind. Auf regionaler und kommunaler Ebene sind es in der Regel dann Akten der Industrie- und Handelskammern (IHK), in denen sich interessante Akten befinden könnten, die sich in der Regel nicht durch direkte Suche mittels Findbüchern finden lassen.

Ein Historiker, der eine Branche untersucht, würde wohl weniger an diesen Akten interessiert sein. Er fände es vor allem spannend, im Querschnitt möglichst viele Kennzahlen von Unternehmen und – wo vorhanden – die Protokolle, Geschäftsberichte etc. – zu finden. Gerade für quantifizierende Ansätze wäre diese Ebene überaus begrüßenswert.

Wären Aktenmaterialien unter diesem Aspekt digitalisiert worden, so wäre es zwingend nötig den Nutzer darauf hinzuweisen, dass dies nur eine Auswahl der existierenden Akten ist und weitere Akten in physischer Form verfügbar sind. Die große Gefahr des Digitalisierens besteht grundsätzlich darin, dass bei einem digitalen Angebot von Akten (grundsätzlich in Teilen oder aufgrund des Suchergebnisses in Teilen) die Forscher vielleicht dazu neigen, nicht nach mehr zu suchen durch einen ergänzenden Archivbesuch. Gerade mit Blick darauf, dass ein solches Angebot natürlich die Forschung im Ausland stimulieren würde. Gerade auch die Unternehmensgeschichte erfreut sich einer beachtlichen Zahl von Forschern aus den USA, die sich sehr regelmäßig mit deutschen Unternehmen, Branchen befassen. Man muss daher dafür Sorge tragen, dass ein solches Angebot nicht zum Verflachen der Interpretationstiefe führt.

Historiker, die sich mit Clustern befassen, würden wohl beide Ansätze hilfreich finden. Die Aktensuchen zu Fragen der Klein- und Mittelstandsunternehmen (KMUs) lassen sich möglicherweise gleichfalls am besten durch Metasuchoptionen quer durch die Archivlandschaft ergänzen.

Wie aber soll man dies systematisieren? Möglicherweise wäre tatsächlich ein Vorgehen, wie in Privatarchiven gelebt, gar nicht so unangemessen und man könnte nach Anfrage vorgehen. Hier ist die Grundüberlegung die einer Aufwandsoptimierung: da die Akte aufgrund einer Anfrage ohnehin gehoben werden muss, ist der zusätzliche Aufwand des Scannens nur gering. Ob dies als Option tatsächlich für die öffentlichen Archive in Frage kommt, hängt davon ab, wie in den Häusern die Digitalisierung durchgeführt wird und ob sich ein solcher Modus integrieren ließe.

Ansonsten ließe sich vielleicht auch ein provokanter Alternativweg einschlagen: nämlich die Aktenbestände aus dem Bereich „Unternehmen“ zunächst zu digitalisieren, die noch nicht häufig bewegt wurden. Dies könnte nämlich dazu führen, dass Forschung hierzu stimuliert wird und neue Bestände bekannt würden. Auch dies hielte ich für einen interessanten Ansatz bei der Digitalisierung.

c. Forschungsfragen

Neben der Problematik grundsätzlicher Natur dieser unterschiedlichen Themenfelder stellt sich die Frage der Relevanz von Forschungsfragen in Bezug auf die deutsche Unternehmensgeschichte.

Seit dem 21. Jahrhundert gibt es eine neue Forschungsrichtung, die sich mit dem Thema kleine und mittlere Unternehmen und/oder

Familienunternehmen befasst. Nicht nur deutsche Forscher sind hier aktiv. Weltweit blickt man gerade mit diesem Forschungsschwerpunkt auch nach Deutschland – im Amerikanischen Sprachgebrauch hat sich mittlerweile das Wort „Mittelstand“ für „Small and Medium Sized Enterprises“ eingebürgert. Neben der Frage nach der Überlebensstärke und der Unternehmensbesonderheit von Familienunternehmen gilt das Interesse vor allem der Frage nach dem Grund für die große Zahl von „Hidden Champions“¹ in Deutschland. Es gibt dieses Phänomen zwar durchaus auch in anderen Ländern, aber nicht in diesem Umfang und in dieser klaren Ausprägung. Jüngste Vergleiche haben gezeigt, dass z.B. in Großbritannien die KMUs weniger Familienunternehmen sind, sondern Start-ups, die sehr schnell an den Kapitalmarkt gebracht werden. Diesen Fragen wird sich die Forschung noch für eine längere Zeit stellen.

Ein anderer Forschungstrend ist die Frage nach dem Verhältnis von Unternehmen und Umwelt. Auch hierzu gibt es eine Reihe von Ansätzen. International bereits sehr etabliert und in Deutschland in Ansätzen vorhanden sind die Forschungen zu internationalen bzw. multinationalen Unternehmen. Hier gibt es auch neuere Ansätze, die hinterfragen, was den Zusammenhang von Regionen und multinationalen Unternehmen ausmacht. Es sind also neue Analyseebenen mit denen man sich befasst. Gehofft hatte die Forschung, dass 2014 als die „Hundert-Jahre-Erster-Weltkrieg“ Untersuchungen zu diesem Zeitraum beflügelt. Eine jüngste Tagung hat hier jedoch gezeigt, dass

¹ Kleine Unternehmen mit Weltmarktführerschaft, deren Namen weitestgehend unbekannt ist (bspw. Zulieferunternehmen).

auf dem Gebiet der Unternehmensgeschichte nicht außergewöhnlich viel geforscht wird. Das mag unter Umständen tatsächlich auch an weniger bekannten oder nicht gut zugänglichen Quellenbeständen liegen. Aber hier würde ein Digitalisierungsprojekt sicherlich zu spät kommen. Auch wenn die 100 Jahre für die Dauer des Weltkrieges noch etwas anhalten wird.

Was für die Unternehmensgeschichte kein neuer Trend ist, aber nach wie vor Thema ist und mit Sicherheit auch noch bleiben wird, ist die Zeit des Nationalsozialismus. Jede Studie, zumindest wenn es um ein Unternehmen geht, kann und wird diese Zeit intensiv bearbeiten müssen. Im Zentrum steht hier natürlich die Frage des Handlungsspielraums der Unternehmer und daher geht es nicht nur um die Unternehmen selbst, sondern hier ganz besonders um das Zusammenspiel von staatlichen und Partei-Organisationen und den Unternehmen. Wir haben als GUG unzählige Studien zu diesem Zeitraum durchgeführt. Akten finden sich in sehr vielen, sehr unterschiedlichen Archiven, da nicht nur Unternehmen, sondern auch handelnde Personen Spuren hinterlassen. Die Arbeits- und Gewerbeaufsichtsämter dokumentieren Sachverhalte im Zusammenhang mit Zwangsarbeitern; neben den bekannten Staatsstellen. Würde man nach einer Priorisierung in diesem Zusammenhang fragen, ist die Antwort eigentlich ganz klar: man kann hier nicht priorisieren. Nach zwei Jahrzehnten Forschung ist eines sehr deutlich: die Fragen rund um das Thema „Drittes Reich“ bedarf einer sehr exakten und sehr detaillierten Analyse. Es gibt kein schwarz und weiß in diesem Fall. Wir haben Staatsunternehmen, unter deren Mäntelchen „Widerstand“ geleistet wurde und wir haben Privatunternehmen, die weit über die

staatlichen Vorgaben hinaus gegen die Moral verstoßen haben. Und dazwischen gibt es alles Denkbare und Udenkbare in sämtlichen Abstufungen. Da aber gerade an solchen Punkten die Forschung keine pauschalen Urteile abgeben kann, braucht es Akten in der tiefstmöglichen Ausprägung.

In Zusammenhang mit der Frage der Digitalisierung kommt beim Themenbereich „Nationalsozialismus“ hinzu, dass diese Akten – wenn unkommentiert zur Verfügung gestellt – leicht zu Fehleinschätzungen führen könnten. Gerade der Umgang mit den Akten aus der Zeit des „Dritten Reiches“ erfordert doch eine professionelle Ausbildung und ein gutes Verständnis dieser Zeit. Allzu schnell sind hier Dokumente falsch interpretiert. Und allzu schnell kann dies zu fatalen Folgen führen. In welche Richtung auch immer. Natürlich können die Archive nicht der Forschung Qualitätsstandards aufdrängen; dafür muss die Forschung selbst sorgen. Aber gerade in diesem Fall muss bei einer angedachten Digitalisierung dieser Aspekt der Qualitätssicherung der Forschung mitgedacht werden, um einen angemessenen Weg zu finden. Hier geht es um die tatsächliche Verfügbarkeit der Akten: darf jeder diese Akten sehen oder wird es auch für die digitalen Akten eine Zugangsbeschränkung geben und wer kontrolliert dies? Denn der Laie mag schnell ein erkennbares Hakenkreuz oder andere Symbole des „Dritten Reiches“ völlig falsch verstehen und falsche Schlüsse ziehen. Auch die hochproblematischen Akten der Entnazifizierungsverfahren könnten, ungefiltert an eine breite Öffentlichkeit gegeben, rasch historisch zumindest fragwürdige Aussagen hervorbringen.

III. Rückschlüsse

Welche Hilfestellung kann die Unternehmensgeschichte also bei der Frage der Priorisierung der Aktenbestände in Bezug auf die Digitalisierung geben? Eine klare Antwort gibt es wohl nicht. Möglicherweise bieten sich für unterschiedliche Bestände verschiedene Ansätze bei der Priorisierung an. Und möglicherweise – je nachdem wie bei der Digitalisierung grundsätzlich vorgegangen wird – wäre ein „on demand“ Verfahren vielleicht sogar am zielführendsten.

Daneben scheint wichtig, dass die Archive als „Anbieter“ und die Historiker als „Nutzer“ sich den Gefahren und auch der Verführungen der Digitalisate bewusst sind. Wenn dieses Bewusstsein beim Umgang mit Digitalisaten integriert wird und die Nutzer lernen damit umzugehen, kann man die Digitalisierung nur begrüßen, da sie durch Zeit- und Kosteneinsparungen die Forschung definitiv beflügeln kann.

Instrumente, Skizzen, Briefe... das besondere Handwerkzeug der Wissenschaftsgeschichte?

von Susan Splinter

Im Folgenden werde ich als Vertreterin einer spezifischen Nutzergruppe – der Wissenschaftshistoriker und -historikerinnen – Wünsche für die Priorisierung des zu digitalisierenden Archivguts formulieren. Dazu wird einführend umrissen, welches Spektrum die Wissenschaftsgeschichte umfasst. In einem zweiten Schritt wird vorgestellt, mit welchen Quellen Wissenschaftshistoriker und -historikerinnen arbeiten und welche Besonderheiten es in der Methodik der Wissenschaftsgeschichte gibt, bevor im letzten Teil auf die Ausgangsfrage eingegangen wird: Welche Prioritäten sind für eine Digitalisierungsstrategie aus Sicht der Wissenschaftsgeschichte notwendig? Dabei wird die Frage aufgeworfen, ob eine bessere Verfügbarkeit, stärkere Vernetzung und konsequente Verschlagwortung der Findmittel einer Digitalisierung des Archivguts vorzuziehen sind.

Naturwissenschafts-, Medizin- und Technikgeschichte

Während im 18. Jahrhundert Wissenschaftsgeschichte hauptsächlich dazu diente, sich als Forscher einen Überblick über den Stand der Forschung anzueignen, löste sich gegen Ende des 19. Jahrhunderts die historische Beschäftigung mit dem eigenen Fach von der Forschungsarbeit eines Naturwissenschaftlers oder Technikers. Im Laufe des 19. Jahrhunderts etablierten und institutionalisierten sich zahlreiche naturwissenschaftliche Fächer; eine historische Verortung zu Beginn einer naturwissenschaftlichen Schrift wurde obsolet. Eine

historische Einordnung wissenschaftlicher Erkenntnisse war angesichts des positivistischen Denkens nicht erstrebenswert. Vor allem durch die enormen wissenschaftlichen, technischen und wirtschaftlichen Umwälzungen im 19. Jahrhundert entstand zudem eine Fortschrittsgläubigkeit. Die Geschichte der Naturwissenschaften, die erfolgreichen Errungenschaften der Medizin, Naturwissenschaften und Technik, wurden monumentalistisch inszeniert; Irrtümer und Sackgassen ignoriert. Die Wissenschaftsgeschichte wurde internalistisch betrieben.¹

Erst im Laufe der zweiten Hälfte des 20. Jahrhunderts hat sich die Geschichte der Naturwissenschaften, der Medizin und der Technik von ihren jeweiligen Fachdisziplinen gelöst. Sie wurde zu einer Reflexionsinstanz, die naturwissenschaftliches, technisches und medizinisches Forschen als Aktivität in einer kulturell und sozial prägenden Umgebung ernst nahm und in diesem Kontext eigene Modi der Bearbeitung entwickelte. Heute wird die Geschichte der Naturwissenschaften, der Medizin und der Technik nicht mehr als akkumulierende Errungenschaftsgeschichtsschreibung verstanden, die „als Stufen von Treppchen“ „in die Höhe gegenwärtiger Wissenschaften hinaufführen“, sondern es werden Denkweisen und Praktiken von Wissenschaftlern und Wissenschaftlerinnen im zeitgenössischen Kontext betrachtet.² Statt einer Kontinuität in der

¹ Hans Jörg Rheinberger, Wozu Wissenschaftsgeschichte?, in: Rudolf Seising/Menso Folkerts/Ulf Hashagen (Hrsg.), Form, Zahl, Ordnung, Studien zur Wissenschafts- und Technikgeschichte, Stuttgart, 2004, S. 51-62.

² Ute Daniel, Kompendium Kulturgeschichte. Theorien, Praxis, Schlüsselwörter, Frankfurt/M. 2001, S. 363 f.

Wissensproduktion stehen nun Fragen zur Verflechtung von wissenschaftlichen, religiösen, ästhetischen und sozialen Aspekten im Mittelpunkt. So wird beispielsweise in der Medizingeschichte berücksichtigt, dass die zeitgenössischen Sichtweisen auf die biologischen Körperabläufe im Kontext der kulturellen und sozialen Lebenswelt betrachtet werden müssen; das Erleben, Wahrnehmen und Beschreiben körperlicher Gegebenheiten ist kulturell determiniert und dies muss in der historischen Analyse beachtet werden.

Obgleich die Zwitterstellung der Naturwissenschafts- und Technikgeschichte zwischen den Geistes- und Naturwissenschaften auch als Chance begriffen werden kann, eine verbindende und grenzübergreifende Position einzunehmen, bleibt sie sowohl methodisch-inhaltlich als auch fachpolitisch problematisch. Vielen NaturwissenschaftlerInnen und TechnikerInnen ist eine historische Auseinandersetzung – zumal nicht als positivistische Fortschrittsgeschichte – suspekt. Dagegen benötigen Historiker und Historikerinnen gerade für die Beschäftigung mit der Naturwissenschaftsgeschichte aus der Zeit nach 1830 eine entsprechende Fachkompetenz in dem zu untersuchenden Thema; das macht die Konzeption und Beantwortung von Forschungsfragen für Geisteswissenschaftler und –wissenschaftlerinnen schwierig. Thematische Offenheit und über das Ausgangsstadium hinausführende Kenntnisse in den Geschichts- bzw. Naturwissenschaften sind für den/die Wissenschafts-, Technik- oder MedizinhistorikerIn unerlässlich. Diese Zwischenstellung zeigt sich auch bei der Verortung entsprechender Lehrstühle. So findet sich die Medizingeschichte immer in der medizinischen Fakultät, während Technik- und Naturwissenschaftsgeschichte an philosophischen,

historischen und naturwissenschaftlichen Instituten angegliedert war bzw. ist.

Mit dem Aufkommen der Kulturgeschichte in den letzten Jahren entstanden neue Forschungsfragen auf den Gebieten der Wissens- und Wissenschaftsgeschichte. In der Geschichte der Naturwissenschaften (history of science) ist es heute durchaus üblich, sich auch mit vermeintlich „irrationalen“ Themen, wie Wunder, Monster und Religiosität sowie deren Verflechtungen zum wissenschaftlichen Denken zu beschäftigen. Ebenso spielen die Verquickungen zwischen politischen und wissenschaftlichen Aspekten eine Rolle.³ Die Wissensgeschichte (history of knowledge) fragt nach den verschiedenen Formen des Wissens (akademisch, populär, Wissen der Praktiker), nach der Zirkulation von Wissen zwischen gesellschaftlichen Gruppen und der Produktion von Wissen.⁴

Obwohl die Geschichte der Medizin, Naturwissenschaft und Technik ein disparates, verschiedene Forschungen umfassendes Gebiet ist, existieren in Deutschland zwei Fachgesellschaften, in denen das gesamte Themenspektrum im kollegialen Austausch, durch (gemeinsame) Tagungen und eigene Zeitschriften vertreten

³ Daniel, Kulturgeschichte (wie Anm. 2). Lorraine Daston/Katherine Park, *Wonders and the Order of Nature. 1150-1750*, New York 1998, die 96. Jahrestagung der DGGMNT in Jena zu „Die Wissenschaften, Technik und Medizin im Kalten Krieg“ und Wolfgang Uwe Eckart, *Medizin in der NS-Diktatur. Ideologie, Praxis, Folgen*, Wien u. a. 2012 sind Beispiele.

⁴ Philipp Sarasin, Was ist Wissensgeschichte?, in: *Internationales Archiv für Sozialgeschichte der deutschen Literatur* 36, 2011, S. 159-172; Daniel Speich Chassé/David Gugerli, Wissensgeschichte. Eine Standortbestimmung, in: *Traverse* 19, 2012, S. 85-100.

wird. Neben den beiden Gesellschaften der Deutschen Gesellschaft für die Geschichte der Medizin, Naturwissenschaft und Technik (DGGMNT) sowie der Gesellschaft für Wissenschaftsgeschichte (GWG) existieren noch einige spezielle Verbände und Vereinigungen, wie die Gesellschaft für Technikgeschichte (GTG) und der Fachverband Medizingeschichte. Des Weiteren besteht die Möglichkeit über Mailinglisten, sich über aktuelle Fragen auszutauschen, Tagungen bekannt zu machen und das Fach betreffende Informationen zu kommunizieren.⁵ Ein interdisziplinärer Austausch ist u.a. bei den jährlichen stattfindenden Tagungen der DGGMNT, der GWG und der GTG gewährleistet.

Quellen und Methoden

Diese organisatorische und fachliche Besonderheit spiegelt sich auch in der Auswahl und Nutzung der Quellen sowie in der Methodik wider. So nutzt die Geschichte der Medizin, Naturwissenschaften und Technik analog zu anderen Feldern der historischen Forschung die große Bandbreite an schriftlichen Quellen. Aktenmaterial, seien es Briefe, Gerichtsprotokolle, Rechnungen, Sitzungsniederschriften jeder Art, Pro Memoriae o.ä. sind für den/die Wissenschaftshistoriker und -historikerin je nach Fragestellung genauso bedeutsam wie für den

⁵ Für die Geschichte der Naturwissenschaften in Deutschland ist hier die Oldenburg-Mailingliste <http://www.uni-regensburg.de/Fakultaeten/philo_Fak_IV/Philosophie/Wissenschaftsgeschichte/OLDENBURG.htm> [20.03.2014] zu nennen, für die Technikgeschichte die Mailingliste der GTG <<http://www.gtg.tu-berlin.de/ws/index.php/wir-ueber-uns/maillingliste>> [20.03.2014] und für die Medizingeschichte die Mailingliste Medizingeschichte <<http://www.fachverband-medizingeschichte.de/6links/4maillist.html>> [20.03.2014].

Wirtschafts-, Landes- oder Rechtshistoriker bzw. die -historikerin. Dieses Material wird meist von Stadt- oder Staatsarchiven verwaltet und bereitgestellt. Auch bei der Interpretation der Quellen ähneln sich die methodischen Abläufe. Inhaltliche, personelle und räumliche Kontexte zu erfahren, sind von elementarer Bedeutung.

Wichtig für den/die Wissenschaftshistoriker und -historikerin sind die Materialien, die speziell an den Handlungsorten von Wissenschaft entstanden und oft auch dort verwahrt werden. Universitäts- und Akademiearchive besitzen Personalakten, Sitzungsniederschriften und Akten, die Einblicke in die wissenschaftlichen und institutionellen Arbeitsabläufe gewähren. Darüber hinaus sind sie für das Verständnis sozialer Netzwerke bedeutsam. Ebenso finden sich hier Überreste wissenschaftlicher Produktion, quasi Zwischenschritte auf dem Weg zur Erkenntnis und/oder Publikation. Sie schärfen das Verständnis für die Entstehung wissenschaftlicher Einsichten und helfen Fragen nach der Entstehungsgeschichte zu beantworten. Dabei zeigen sie den Aufbau experimenteller Anordnungen, die Art und Weise des Zusammenarbeitens von unterschiedlichen Personengruppen (LaborassistentInnen, WissenschaftlerInnen usw.), die materiellen Gegebenheiten oder das Lese- und Diskussionsverhalten. In Notiz- oder wissenschaftlichen Tagebüchern und (Teil-)Nachlässen, die oft an den Wirkungsstätten der WissenschaftlerInnen aufbewahrt werden, findet man Informationen zur sozialen und geistigen Heimat der Forscher und Forscherinnen, zu Praktiken von Wissenschaft und ähnlichem mehr.

Die Arbeits- und Forschungsbedingungen sind an Archiven von Forschungseinrichtungen sehr heterogen. Neben Universitäten und

Akademien besitzen sehr häufig Bibliotheken (umfangreiche) Handschriftenabteilungen. Museen bewahren Archivgut auf, das zum Teil nicht erschlossen ist. Natürlich ist es nicht die vordergründige Aufgabe von Museen, archivarische Strukturen aufzubauen und bereitzuhalten; dazu fehlen personelle und finanzielle Ressourcen, doch findet sich hier Archivgut, was sowohl für die Forschungs- als auch die Archivlandschaft von Interesse sein kann. Es gibt gut ausgestattete und erschlossene Archive; gleichermaßen existieren auch Einrichtungen, an denen es nur wenig Personal, daraus resultierend unterschiedlich aufbereitete Hilfsmittel und mangelnde Räumlichkeiten (auch für die Benutzung der Archivalien) gibt. Während kommunale und staatliche Archive über eigenes Personal und Räumlichkeiten verfügen, kann es durchaus vorkommen, dass Akademie- oder Universitätsarchive dies nicht haben. Um eine adäquate Forschungsinfrastruktur zu schaffen, wäre es nötig, die verschiedenen Kultureinrichtungen für ihre archivalischen Bestände und deren Aufbereitung zu sensibilisieren, das Archivgut nach den üblichen Regeln zu erschließen und analog der bestehenden Angebote kommunaler und staatlicher Archive zugänglich zu machen.

Um Forschungen der Industrie zu untersuchen, sind die Quellenbestände von Firmen von großer Bedeutung. In diesen Fällen liegt die Aufbewahrung und Aufbereitung der Archivalien in privaten Händen. Damit sind Arbeiten zu naturwissenschaftlichen und technischen Forschungen in Unternehmen von der Überlieferungssituation, dem Erschließungszustand und der Zugänglichkeit der Quellen abhängig. Diese Archivgruppe einzubinden, ist für Historiker und Historikerinnen der Technik-, Naturwissenschafts- und Unternehmensgeschichte wesentlich.

Die oben erwähnten Handlungsabläufe, wie Handgriffe, Experimentalpraktiken, Memorier- und Schreibtechniken sowie die materiellen Quellen von Wissenschaft, also dreidimensionale Objekte, wie Gerätschaften o.ä. stehen seit einigen Jahrzehnten im Fokus des Forschungsinteresses. Die Materialität von wissenschaftlicher Produktion wird sehr intensiv untersucht. Abseits einer Geschichte der Begriffe, Ideen und Theorien geraten mit dem „material turn“ gegenständliche Zeugnisse in den Blick.⁶ Dabei interessieren Aufbau, Hersteller und Funktionsweise genauso wie die Forschungspraxis, also die Handhabung, die Handhabbarkeit und das dazugehörige Handlungswissen. Das Studium von Labortagebüchern wird also ergänzt durch das Studium der Objekte, anhand derer man Gebrauchsspuren, Veränderungen, Skalierungen, Herstellerangaben u.a. erkennen kann. Des Weiteren gibt es Wissenschaftshistoriker und –historikerinnen, die alte Instrumente nachbauen und historische Experimente mittels der Replikationsmethode nachvollziehen. Durch das Studium entsprechender Aufzeichnungen werden die Objekte soweit als möglich mit originalen Materialien und Arbeitsprozessen hergestellt, um dann damit zu experimentieren. Sowohl der Herstellungsprozess als auch der Versuchsablauf offenbaren faszinierende Kenntnisse der wissenschaftlichen Kultur. Durch den baulichen und experimentellen Nachvollzug gewinnt man Einblick in experimentelle Praktiken, in die materielle Kultur der Wissenschaften, in die Existenz „unsichtbarer

⁶ Christoph Meinel (Hrsg.), *Instrument – Experiment. Historische Studien*, Berlin 2000.

Hände“ und spürt „tacit knowledge“ auf.⁷ Erst der praktische Nachvollzug macht deutlich, dass nicht-schriftlich überliefertes Wissen bzw. seinerzeit selbstverständliche Fähigkeiten, an Forschungsprozessen ebenso beteiligt waren wie Assistenten und Assistentinnen oder Gehilfen und Gehilfinnen. Die allgemein bekannten Handgriffe und die Hilfestellungen durch andere Personen lassen sich selten in schriftlichen Quellen finden, so dass sich weitere Facetten des Forschungsalltags zur wissenschaftlichen Untersuchung eröffnen. Gleichzeitig korrigiert die Kenntnisnahme von Gehilfen und Gehilfinnen ein „hagiographisches Zerrbild“ der Wissenschaftspraxis, weil neben den Wissenschaftlern und Wissenschaftlerinnen nun Mechaniker, Techniker, Zeichner oder die helfende Ehefrau zur Kenntnis genommen werden.⁸

Wünsche und Bedürfnisse bei der Digitalisierung

Neben Bibliotheken finden Wissenschaftshistoriker und –historikerinnen vor allem in Stadt- und Staatsarchiven, in Archiven von Forschungseinrichtungen und in Objektsammlungen ihre Quellen. Welche besonderen Wünsche und Bedürfnisse ergeben sich daraus?

Der wichtigste und erste Schritt für ein digitales Angebot ist die Online-Verfügbarkeit der Repertorien. Zur gründlichen Vorbereitung

⁷ H. Otto Sibum, Experimentelle Wissenschaftsgeschichte, in: Meinel, Instrument, (wie Anm. 3), S. 61-73.

⁸ Klaus Hentschel, Wie kann Wissenschafts- und Technikgeschichte die vielen „unsichtbaren Hände“ der Forschungspraxis sichtbar machen? in: Ders. (Hrsg.): Unsichtbare Hände. Zur Rolle von Laborassistenten, Mechanikern, Zeichnern u.a. Amanuenses in der physikalischen Forschungs- und Entwicklungsarbeit, Berlin u.a. 2008, S. 11-25, hier S. 13.

eines Archivbesuchs hilft es, Umfang und Art der archivalischen Materialien bereits zuvor zu kennen. Gegebenenfalls vermeidet eine Vorbestellung längere Wartezeiten.

Daraus ergibt sich, dass der Archivbesuch durch Online-Angebote nicht ersetzt, sondern besser planbar und effizienter werden soll. Die kontemplative Arbeitsatmosphäre der Archive ermöglicht ein konzentriertes Forschen, da man den üblichen Unterbrechungen des Büroalltags entzogen ist. Diese (mehrtägige) Fokussierung auf den Forschungsgegenstand, der Austausch mit Fachleuten – hier sind sowohl andere Archivnutzer als auch Archivare gemeint – und die Vielzahl an parallelen Recherchemöglichkeiten (Fachliteratur, Findmittel u.ä.) steigert die Produktivität.

Entscheidend für die Vorbereitung ist das umfassende Durchsuchen potentieller Quellen nach verschiedenen Aspekten. Biographische, geographische, institutionelle und semantische Kategorien erleichtern die Suche je nach Fragestellung. Dass diese Kategorien bzw. Zuordnungen persistent, also dauerhaft und eindeutig, sein sollten, ist evident. In diesem Falle wäre es sinnvoll, dass bereits bei der Erschließung der Archivalien eine Versorgung mittels Normdaten, z.B. der Gemeinsamen Normdatei (GND), durchgeführt wird. Die GND wird von der Deutschen Nationalbibliothek, allen Bibliotheksverbünden, der Zeitschriftendatenbank, Museen und Redaktionen geführt und verwendet. Die GND enthält eindeutige Datensätze für Personen, Körperschaften, Sachschlagwörter u.ä. und steht

kostenfrei zur Verfügung.⁹ Darüber hinaus existieren weitere Angebote.¹⁰ Damit sind persistente Identifier möglich.

Eine umfassende Versorgung von Findmitteln mittels Normdaten ist nicht nur für die eindeutige Durchsuchbarkeit, sondern auch für die Weiterverarbeitung der Daten Voraussetzung. Demnach profitieren nicht nur AnwenderInnen, sondern auch AnbieterInnen, die ihre Suchangebote und Findmittel erweitern bzw. verbessern wollen. Dank der persistenten Identifier lassen sich weitere Angebote für die Nutzer und Nutzerinnen generieren. So wäre in weiteren Schritten eine Vernetzung zu übergreifenden Recherchen denkbar. Dazu müssten die Findmittel frei zugänglich und auch rechtlich nachnutzbar sein. Einerseits bestünde so die Möglichkeit, Angebote verschiedener Archive miteinander zu verbinden. Hierfür wären – früher oder später – Absprachen bei den Klassifikationen und Normierungen notwendig, um institutionsübergreifende Angebote unterbreiten zu können. Verbindungen zwischen Standards unterschiedlicher Institutionsgruppen (Museen, Archive, Bibliotheken) wären hilfreich. Mit einer solchen Vernetzung ließe sich quasi auf einen Klick in Universitäts- und Akademiearchiven nach bestimmten Themen suchen. Man erhielte gegebenenfalls Material zu Personen, zu institutionellen Abläufen, Quellen zu Forschungen, wie Tagebücher, Skizzen u.ä. Im besten Falle würde eine solche Suche auch Objekte

⁹ GND: <http://www.dnb.de/DE/Standardisierung/GND/gnd_node.html> [20.03.2014].

¹⁰ Ich danke Matthias Reinert, der mich auf „The Social networks and Archival Context Project“ aufmerksam machte <<http://socialarchive.iath.virginia.edu/>> [20.03.2014].

bzw. Aufbewahrungsorte von gegenständlichen Quellen anzeigen. Neben Schriftquellen könnte man nun auch Objekte viel stärker in die Untersuchung einbeziehen. Kommunikationswege und Netzwerke ließen sich so viel leichter nachzeichnen, weil eine Vielzahl von bisher verstreuten Quellen zusammengetragen werden kann und Verbindungen sowie Querverweise besser wahrzunehmen sind.

Dieses vernetzte Angebot könnte zum einen völlig neue Forschungsfragen ermöglichen und zum anderen würde es bereits in der Planungsphase des Archivbesuchs helfen, die potentielle Tiefe und Tragweite von Forschungsfragen abzuschätzen.

Andererseits bestünde die Möglichkeit der Vernetzung mit anderen Informationsangeboten des Internets. Die Sichtbarkeit einer Internetpräsenz wird erhöht und gleichzeitig wird ein Mehrwert geschaffen, weil neben den bereits vorhandenen Daten zusätzliche Informationen angeboten werden könnten, z.B. vorhandene Literatur im Karlsruher Virtuellen Katalog (KVK)¹¹ und biographische Angaben aus Lexika.

Solche Vernetzungsbestrebungen existieren bereits; erinnert sei an „www.bam-portal.de“, „Europeana“ und „Archivportal-D“, an die „Deutsche Digitale Bibliothek“ oder „museum digital“; hier wird versucht die Erschließungsleistung kulturbewahrender Einrichtungen zusammenzuführen und spartenübergreifend nutzbar zu machen. Bei diesem (aufwändigen) Prozess des Zusammenführens von Angeboten ist eine Standardisierung und Klassifizierung elementar, da sonst Treffer bzw. Trefferlisten generiert werden, die nicht den

¹¹ Vgl. <<http://www.ubka.uni-karlsruhe.de/kvk.html>> [20.03.2014].

Nutzerintentionen entsprechen. Bei einer Suche nach Personen mit häufig vorkommenden Namensbestandteilen, wie „Müller“, „Wilhelm“ oder „Friedrich“ wird das Ergebnis enorm verfälscht, wenn mittels einer Volltextsuche und einer „oder-Verknüpfung“ die Datenbanken durchsucht werden. Dann erhält man eine umfangreiche Liste, in der sämtliche Einträge aller Datenbanken enthalten sind, in der eines der gesuchten Worte vorkommt. Ob man sich die Mühe macht, die relevanten Ergebnisse herauszufiltern oder das Such- bzw. Ergebnisangebot komplett verwirft, hängt sicher vom Umfang und weiteren einschränkenden Suchfunktionen ab.

Ein weiteres Problem tritt beispielsweise bei der Nennung von Personennamen in lateinischen Buchtiteln auf. Hier verändert sich der Name, in dem ein –us oder –i angefügt wird. Diese Wortveränderung wird nicht zwangsläufig erkannt, so dass Informationen zu dieser Person nicht gelistet werden. Damit steigen die Fehlerrate und die Frustration auf Seiten des Nutzers bzw. der Nutzerin. Der Wert des Suchergebnisses wird minimiert.

Ein Beispiel für ein präzises und vielseitiges Rechercheangebot ist die Seite „www.universitaetssammlungen.de“ des Hermann von Helmholtz-Zentrums für Kulturtechnik (Humboldt-Universität Berlin). Sie bietet eine Objektdatenbank zur Erfassung, Dokumentation und Präsentation von Universitätssammlungsobjekten. Eindeutige, relevante Suchergebnisse einschließlich der Namensvarianten erhält man durch eine spezielle Personensuche. Des Weiteren kann auch eine Volltextsuche genutzt und nach anderen Kategorien unterschieden werden. Die Suchergebnisse können geclustert präsentiert werden.

Neben einer großen Sichtbarkeit dank Einbindung und Vernetzung mit anderen Informationsangeboten ist die Programmierung einer Suchseite, die Ergebnisse nach Relevanz sortiert, entscheidend. Nicht die Quantität der Einträge bzw. die Länge der Ergebnisliste, sondern deren qualitativer Bezug zur Suchanfrage ist für den Erfolg eines Internetangebots ausschlaggebend. Die Bildung von Gruppen ist hilfreich. Zu denken wäre hier einerseits an mediale Typen, wie Bilder, Gegenstände, Bücher, Briefe, Protokolle usw., andererseits an semantische und geographische Cluster. So kann der Nutzer das Suchergebnis individuell eingrenzen, eventuell auch kombinieren.

Mittels Standardisierung und Klassifizierung lassen sich in einem weiteren Schritt Angebote im Sinne des „semantic web“ generieren. „Das semantische Netz ist kein neues Netz, sondern eine Erweiterung des bestehenden, in der Informationen mit maschinenlesbarem Bedeutungsinhalt versehen sind.“ Mittels strukturierter Informationen und über Schlussregeln können Informationen automatisch miteinander verknüpft werden. So können Inhalte von Internetseiten nicht nur angezeigt, sondern auch weiterverarbeitet werden und dank der aufgestellten Beziehungen neue Erkenntnisse sichtbar machen.¹² Demnächst bietet die „Deutsche Biographie“ auf ihrer Seite eine strukturierte, vorausgewählte Linkliste an, die mögliche Literaturbestände in Bibliotheken, Angaben aus Professorenkatalogen, weitere biographische Nachschlagewerke, Quellennachweise, Nennungen in Editionen sowie Bildmaterial mittels GND vereint. Zurzeit arbeitet die

¹² Tim Berners-Lee/James Hendler/Ora Lassila, Mein Computer versteht mich, in: Spektrum der Wissenschaft, 8/2001, S. 42-49.

„Deutsche Biographie“ an Angeboten, die formalisierte semantische Aussagen ermöglicht, zum Beispiel Schüler-Lehrer-Beziehungen, gemeinsame Wirkungsstätten u.ä. So sollen sich komplexe Anfragen formulieren und auch visualisieren lassen. Darüber hinaus werden durch den Abgleich gegen Ortsnormdaten geographische Bezüge darstellbar. Es können beispielsweise lokale wissenschaftliche Ballungszentren deutlich werden.

Erst in einem zweiten Schritt halte ich die Digitalisierung von Archivgut für angebracht. Da für den Wissenschaftshistoriker und die Wissenschaftshistorikerin vor allem Universitäts- und Akademiearchive von Interesse sind, denke ich in erster Linie an Sitzungsprotokolle sowie Nachlässe und personenbezogene Quellen wie Personal- oder Wahlakten dieser Archive. Diese Materialien lassen sich für unterschiedliche Fragestellungen nutzen und gewähren Einblick in die Aktivitäten der entsprechenden Institution. Mit Sitzungsprotokollen lassen sich Fragen der Sozial-, Alltags-, Kultur- und Wissenschaftsgeschichte beantworten. Personal- und Wahlakten sind bei sozialhistorischen Untersuchungen und bei biographischen Zugängen von Interesse. Mit diesen Quellentypen lassen sich also verschiedene Forschungsfragen bei diversen methodischen Herangehensweisen bearbeiten.

Aber nicht nur Bilder und Texte sind interessant, sondern auch dreidimensionale Objekte. So stellt das Smithsonian Teile seiner Objektsammlungen digital zur Verfügung. Die Objekte, 3D präsentiert, können von allen Seiten und mit enormer Vergrößerung am

Bildschirm dargestellt werden; teilweise sollen sich mit Hilfe eines 3D-Druckers Repliken anfertigen lassen.¹³

Bei der Digitalisierung der Archivalien stellt sich die Frage nach der inhaltlichen Erschließung, nur mittels derer man die passenden Quellen angezeigt bekommt. Auch hier wäre – wie bei der Erschließung der Findmittel – eine eindeutige Durchsuchbarkeit nach Personen, Orten, Jahreszahlen und möglicherweise nach Schlagworten wünschenswert. Bei all diesen Suchfeldern wäre eine sich selbst ergänzende Eingabemöglichkeit hilfreich, wie man sie beispielsweise bei Google kennt. Mittels dieser Autocomplete-Funktion lässt sich schnell erkennen, welche Wörter mit Schreibtoleranz standardisiert sind.

Zusammenfassung

Die Bedürfnisse von Wissenschaftshistorikern und –historikerinnen unterscheiden sich von denen anderer Historiker und Historikerinnen zwar nicht grundlegend, doch interessieren sie sich v.a. für Quellen aus Universitäts- und Akademiearchiven, für wissenschaftshistorische Objektsammlungen und Quellen aus Firmenarchiven. Daher sind diese Archive und Kultureinrichtungen bei der Diskussion der Digitalisierungsstrategien zu berücksichtigen. Das teilweise sehr spezielle Quellenmaterial (z.B. Objekte) verlangt ebenso wie die verschiedenen beteiligten Institutionen intensive Gespräche, in denen gemeinsame Standards und die detaillierte Zusammenarbeit geklärt werden müssen. Dieser Prozess ist angesichts einer langfristigen

¹³ Smithsonian <<http://3d.si.edu/>> [20.03.2014].

Zusammenführung der digitalen Angebote und im Hinblick auf die Nutzerfreundlichkeit sehr lohnenswert.

Als erstes wünscht man sich gut erschlossene, online zugängliche Findmittel der Einrichtungen, die für Wissenschaftshistoriker und –historikerinnen von Bedeutung sind. Eine Suchseite, die Archiv- und Museumsbestände einschließlich universitärer Sammlungen vereint, würde die Sichtbarkeit und den Nutzen erhöhen. Eventuell bietet sich hier das Anknüpfen an bestehende Angebote an, bei denen mitunter zahlreiche kulturbewahrende Institutionen ihre Daten allgemein zugänglich präsentieren. Neben einer offensiven Internetpräsenz muss das Angebot dauerhaft zugänglich sein und gepflegt werden. Bereits in der Planungsphase sollte dies berücksichtigt werden. So bieten drittmittelfinanzierte Projekte eine gute Möglichkeit, Netz-Angebote zu erstellen, doch sollte der personelle Aufwand zur Pflege einer entsprechenden Internetseite nach Erstellung mitgedacht werden.

Die Repertorien und Archivalien sollten leicht durchsuchbar, mit anderen Angeboten vernetzbar sein und der Nutzer sollte in die Lage versetzt werden können, relevante von irrelevanten Treffern mittels Filter zu separieren. Diese Ziele sind nur durch Standardisierungen und Kooperationen erreichbar. Ohne gute Aufbereitung der Daten ginge die archivalische Information in der Datenfülle des Internets verloren. Daher plädiere ich dafür, zuerst den Ausbau und die Tiefenschärfe bei der Präsentation der Findmittel mit Hilfe von standardisierten Indexierungen in Angriff zu nehmen. Darauf aufbauend könnten komplexe Suchanfragen mittels „semantic web“ bearbeitbar werden. Dies würde über die bisherige Forschungsinfrastruktur hinausweisen, weil neue Zusammenhänge zwischen einzelnen

Informationen zutage treten können. Letztendlich steht also erst einmal weniger der Wunsch nach digital verfügbaren Archivalien als vielmehr der Wunsch nach einer leicht aufzufindenden und verschiedene Institutionen übergreifenden Suchseite, mit der dank standardisierter und klassifizierter Daten relevante Quellenbestände aus Archiven, Museen, Akademien und Universitäten generierbar werden. Ohne eine solche Internetpräsenz besteht die Gefahr, dass die später digitalisierten Quellen in dem riesigen Angebot des Internets nicht zur Kenntnis gelangen.

HerBalt, Hereditas Baltica – „Virtueller Lesesaal für baltisches Archivgut“. Ein länderverbindendes Digitalisierungsprojekt

von Peter Wörster

1.

Das Thema des Kolloquiums „Digitalisierung im Archiv. Neue Wege der Bereitstellung des Archivguts“ kann durch ein Projekt aus der Dokumentensammlung des Herder-Instituts Marburg, wichtige Aspekte hinzugewinnen, geht es dabei doch um Digitalisierung und Bereitstellung und um internationale Zusammenarbeit vor dem Hintergrund gemeinsamer Geschichte.

Das Herder-Institut ist das Zentrum für die historische Ostmitteleuropaforschung in Deutschland. Es ist ein Institut der Wissenschaftsgemeinschaft Leibniz, deren Spezifikum darin besteht, dass die Mitgliedsinstitutionen halb vom Bund und halb von den Ländern finanziert werden.¹

Das Herder-Institut verfügt über bedeutende Forschungsmöglichkeiten in der großen Forschungsbibliothek und in den Wissenschaftlichen Sammlungen (Bilder, Karten, Dokumente). Zu diesen Wissenschaftlichen Sammlungen gehört auch die Dokumentensammlung, die ein Archiv im klassischen Sinne ist, mit der Besonderheit allerdings, dass ihr Archivsprengel 1000 bis 1500 Kilometer weit entfernt ist:

¹ Vgl. den Internetauftritt: <www.herder-institut.de> [20.03.2014].



Abb. 1: Das Arbeitsgebiet der Dokumentensammlung des Herder-Instituts: Estland, Livland, Kurland (oben), Estland, Lettland, Litauen (unten).

Die Dokumentensammlung ist das größte Archiv zur baltischen Geschichte in Deutschland. Warum das so ist und so sein kann, muss wegen der Kürze der Zeit an dieser Stelle unerörtert bleiben.² Wichtig ist, dass die bedeutenden Archivbestände zur Geschichte der baltischen Region im Herder-Institut von der deutschen und internationalen Baltikumsforschung intensiv genutzt werden. Wichtig ist außerdem, dass sie zur Grundlage vielfältiger Zusammenarbeit mit den Archiven in den baltischen Staaten Estland und Lettland geworden sind. Dazu gehört das langfristig und kooperativ angelegte Projekt HerBalt³ zur Digitalisierung ausgewählter Bereiche baltischen Archivguts und zur gemeinsamen Präsentation in Archivportalen.

² Vgl. Csaba János Kenéz und Peter Wörster, *Archivbestände zur Geschichte Est-, Liv- und Kurlands in der Dokumentensammlung des Herder-Instituts*. Marburg 2000 (Sammlungen des Herder-Instituts zur Ostmitteleuropa-Forschung, Bd. 9). Auch als e-Buch: <http://www.herder-institut.de/fileadmin/user_upload/pdf/Literatur_DSHI/Archivbestaende_Est_Liv_Kurl_Bd_9.pdf> [20.03.2014]; Dorothee M. Goeze und Peter Wörster: *Die Dokumentensammlung im Herder-Institut. Geschichte und Profil*. 5. Aufl., Marburg/Lahn 2013 <http://www.herder-institut.de/fileadmin/user_upload/pdf/Literatur_DSHI/Sammlung-1xx-3a-Auflage.pdf> [20.03.2014].

³ Der Begriff Hereditas Baltica (HerBalt) wurde in einer Arbeitsbesprechung 2010 in einem Hotel in Dorpat geprägt.

2.

HerBalt begann 2011 in Zusammenarbeit mit Estland: Es war das erste, auf drei Jahre angelegte Teilprojekt zwischen dem Estnischen Nationalarchiv und dem Herder-Institut (mit finanzieller Unterstützung durch den Bundesbeauftragten für Kultur und Medien sowie privatrechtlicher, meist deutschbaltischer Gesellschaften und Stiftungen). Es ging um die Digitalisierung und Internetbereitstellung grundlegend wichtiger Materialien zur Personenkunde und zentral bedeutsamer Quellen zur politischen Geschichte des Baltikums. Einbezogen wurden folgende Bestände:

- das Archiv der Matrikelkommission der Estländischen Ritterschaft und des Estländischen Gemeinnützigen Verbandes (wie die Ritterschaft nach ihrer Auflösung genannt wurde); die Akten beider Kommissionen bis 1939,
- das Archiv des deutschbaltischen Genealogen und Archivars Gottfried Carl Georg von Törne (1854-1918), ca. 8.000 Familien aller Stände,
- Landtagsprotokolle der Estländischen Ritterschaft 1800-1851.

Bei diesem ersten Teilprojekt ging es um zusammen rund 130.000 Digitalisate.

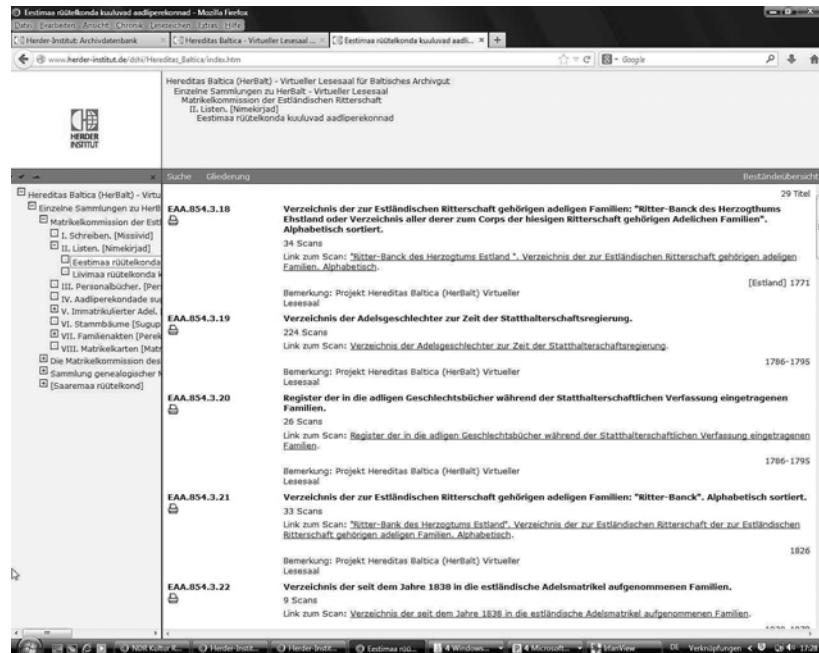


Abb. 2: Einblick in die DSHI-Archivdatenbank „Hereditas Baltica“, HerBalt (Screenshot).

Alle Archivbestände, die in gemeinsamer deutsch-estnischer Arbeit digitalisiert wurden, sind in der DSHI-Archivdatenbank (Midosa) verzeichnet worden, wo sie unter dem Stichwort *HerBalt* auf der Homepage des Herder-Instituts recherchiert werden können.

Es wurde mit den estnischen Kollegen verabredet, dass wir die Verzeichnung, die in Estland schon vorliegt, überarbeiten. Wir haben uns der durchaus zeitaufwendigen Durchsicht und Überarbeitung aus folgenden Gründen unterzogen: Einiges musste korrigiert werden –

falsche Lesungen etwa oder Tippfehler; manches konnten wir tiefer erschließen, weil wir mehr Zeit in diese vertiefte Erschließung unter Einbeziehung einschlägiger Literatur investieren konnten; estnische Beschreibungen haben wir ins Deutsche übersetzt. So wurde eine Anreicherung der Verzeichnungsdaten (*catalogue-enrichement*) erreicht, zu der auch folgendes gehörte: Heißt es im estnischen System beispielsweise bei alphabetisch geordneten Familienakten pauschal „A-B“, so dass alle Familiennamen, die dazwischen liegen, durch „virtuelles“ Blättern mit großem Zeitaufwand ermittelt werden müssen, haben wir uns bemüht, alle tatsächlich auftretende Familiennamen mit all ihren möglichen Schreibvarianten in die Archivdatenbank einzugeben. Wir haben auch Umfang und Laufzeiten in die Verzeichnung aufgenommen, die im estnischen System ebenfalls meistens fehlen. All das ist sicher als nutzerfreundliche Anreicherung zu bewerten.

Alle unsere Verzeichnungseinheiten sind mit dem estnischen Archivportal *Saaga*, das auch eine interessante Lupenfunktion hat, verlinkt. So kann jeder Nutzer von unserer Archivdatenbank direkt zum gesuchten Digitalisat gelangen, das in Estland auf dem Server liegt, wobei lediglich eine Registrierung im estnischen Archivportal *Saaga* notwendig ist.

Das Projekt HerBalt beruht auf dem Prinzip, dass im Herder-Institut Marburg keine eigenen Speicherkapazitäten zur Verfügung gestellt werden müssen und auch für die laufende Pflege des Archivportals nicht zu sorgen ist.

Noch zwei Beispiele der Zusammenarbeit mit Lettland:

- A. In Riga gab es bis zur Umsiedlung der Deutschbalten 1939 die Compagnie der Schwarzen Häupter als eine der alt-ehrwürdigen Korporationen der Stadt. Sie unterhielt seit Jahrhunderten das prächtige Schwarzhäupterhaus direkt im Zentrum der Altstadt, das ein Zentrum des Kulturlebens und der städtischen Repräsentanz war.

Die Compagnie besaß bis 1939 ein höchst bemerkenswertes Archiv, das Bestände vom Anfang des 15. Jhs. bis zum 20. Jh. umfasste. Bei der Umsiedlung der Deutschbalten 1939 durften 42 Bände der ältesten Archivstücke nach Deutschland mitgenommen werden, die 1952 ins Archiv im Herder-Institut Marburg gelangten. Die jüngeren Teile des Archivs, insbesondere umfangreiche Akten aus der Zeit vom 18. bis 20. Jh. blieben in Riga und befinden sich heute im Historischen Staatsarchiv Lettlands.

2013 ist es gelungen, die Finanzierung eines Projekts durch den Bundesbeauftragten für Kultur und Medien zu erreichen. Dieses Projekt zielt auf eine virtuelle Zusammenführung der durch Umsiedlung und Kriegsfolgen getrennten Bestände: Die in Lettland und Deutschland befindlichen Teile des Schwarzhäupterarchivs werden digitalisiert und über das Portal des Lettischen Nationalarchivs nach 75 Jahren erstmals wieder geschlossen der Forschung zur Verfügung gestellt.

- B. Bei der Umsiedlung der Deutschbalten 1939 konnte die deutsche Seite ausgewähltes Archivgut verfilmen, das sich in estnischen

und lettischen Archiven befand.⁴ Dazu gehörten auch große Teile der Akten der Kurländischen Seelenrevisionen 1797-1834, die die Grundlage für jegliche historische Demographie zu Kurland bilden.⁵

Der Bestand Kurländische Seelenrevisionslisten hat 1945 erhebliche Verluste an Originalen erlitten. Diese Verluste können zum Teil durch die Kopienüberlieferung in Marburg ersetzt werden. Daher wurde mit dem Lettischen Nationalarchiv 2011/12 vereinbart, die noch vorhandene Originalüberlieferung in Riga mit der in Marburg vorhandenen kopialen Überlieferung virtuell zusammenzuführen. In Kürze wird dieser einigermaßen vollständige Quellenkorpus im Archivportal Raduraksti des Lettischen Nationalarchivs gemeinsam der internationalen Forschung zur Verfügung stehen – den zahlreichen Familienforschern ebenso wie den Erforschern der historischen Demographie Kurlands und des Baltikums.

Auch das Archiv der Compagnie der Schwarzen Häupter zu Riga und die Akten der Kurländischen Seelenrevisionen werden in Zukunft in unserer Archivdatenbank recherchierbar und durch entsprechende Verlinkung direkt mit dem Digitalisat auf dem Server in Lettland verbunden sein, wie es im Falle der Zusammenarbeit mit Estland etwas ausführlicher vorgestellt wurde.

⁴ Vgl. Kenéz/Wörster, Archivbestände (wie Anm.2), S. 8-18.

⁵ Vgl. Arthur Hoheisel/Peter Wörster, Die Kurländischen Seelenrevisionslisten 1797-1834 (1858). Marburg 1997 (Sammlungen des Herder-Instituts zur Ostmitteleuropa-Forschung, Bd. 2).

3.

Vieles lag und liegt zu diesem Thema natürlich „in der Luft“ – in Deutschland, aber erst recht auch in den baltischen Partnerinstitutionen. Doch lässt sich sagen, dass die *Dokumentensammlung* mit dem von uns 2010/2011 entwickelten Projekt *HerBalt* einige besondere Anliegen verbindet. Das wichtigste Anliegen ist sicher die virtuelle Zusammenführung von Beständen, die durch Krieg und Kriegsfolgen getrennt wurden. *HerBalt* ist zur Zeit das größte Projekt der Zusammenarbeit im Archivbereich zwischen Deutschland und den baltischen Staaten.

Wir möchten durch HerBalt die gute Zusammenarbeit mit baltischen Archiven langfristig und „nachhaltig“ institutionalisieren. Das vom Bund und den Ländern unterhaltene Herder-Institut kann für die Einrichtungen im Baltikum der prädestinierte Vertragspartner sein.

Gegenstand des Projektes sind:

- die Digitalisierung und Internetbereitstellung ausgewählter Bereiche baltischen Archivguts, das in den großen Archiven Estlands, Lettlands und Litauens sowie in Deutschland oder anderen Staaten aufbewahrt wird,
- das geschieht in loser Folge – nach den Interessen der Forschung und mit Partnereinrichtungen, die von Teilprojekt zu Teilprojekt wechseln, und einer ebenso wechselnden Finanzierung.

Ziele sind:

- die Präsentation von archivischen Quellen, die für einen großen Kreis von Nutzern interessant sind, vor allem Quellen zu

Personen, Familien oder zur Bevölkerungsgeschichte insgesamt,

- die virtuelle Zusammenführung der durch Kriegs- und Nachkriegsschicksal getrennten Bestände,
- der leichte und rasche Zugang zu diesem kulturellen Erbe unabhängig vom Ort der Bearbeitung und Aufbewahrung,
- damit die Bewahrung und Pflege des kulturellen Erbes durch Schonung der Originale.

Einige Teilprojekte unter dem Dach von HerBalt sind schon abgeschlossen, einige wurden begonnen oder werden vorbereitet:

- Matrikel der Estländischen Ritterschaft, Törne-Archiv: Zusammenarbeit mit dem Estnischen Historischen Staatsarchiv in Dorpat/Tartu (lief ab 2011, 2013 Fachtagung in Estland, 2013 abgeschlossen, steht online),
- Kurländische Seelenrevisionslisten: Zusammenarbeit mit dem Historischen Staatsarchiv Lettlands in Riga und dem Max-Planck-Institut für Demographie in Rostock (bereits abgeschlossen, wird 2014 online gestellt),
- Archiv der Compagnie der Schwarzen Häupter zu Riga: Zusammenarbeit mit dem Historischen Staatsarchiv Lettlands und mit der Compagnie der Schwarzen Häupter aus Riga e.V. in Bremen (wird 2014 abgeschlossen und 2015 online gestellt),
- Deutschbaltisches (Dialekt-)Wörterbuch: Zusammenarbeit mit der Universität Dorpat/Tartu (online-Fassung soll 2015

abgeschlossen werden; weitere Arbeiten folgen, insbesondere Abschlusstagung).

4.

Neben der Digitalisierung möglichst aller Findmittel und deren Bereitstellung im Internet betreibt die Dokumentensammlung des Herder-Instituts Marburg die Digitalisierung von Archivalien insbesondere im Bereich der Nutzerversorgung und seit 2007 in ausgewählten Stücken für die Internetplattform Archivale des Monats.⁶ Außerdem verfolgt die archivische Arbeit im Herder-Institut seit ca. drei Jahren die Digitalisierung von Archivmaterialien in Zusammenarbeit mit Partnereinrichtungen in den baltischen Staaten und, so erhoffen wir es uns für die Zukunft, mit Archiven auch in anderen Ländern, in denen man über baltische Bestände verfügt, wie es in Schweden und Finnland, in Russland, Polen und Dänemark insbesondere der Fall ist, also rundherum um das Mare Balticum.

⁶ Vgl. <<http://www.herder-institut.de/servicebereiche/dokumentensammlung/archivale-des-monats.html>> [20.03.2014].

Digitalisierung und die DDB: Standards und Perspektiven

von Reinhard Altenhöner

Der vorliegende Beitrag basiert auf einem Vortrag, der beim 18. Archivwissenschaftlichen Kolloquium der Archivschule Marburg im November 2013 gehalten wurde. Er stellt das Projekt Deutsche Digitale Bibliothek (DDB) vor, beschreibt Ausgangslage, Umsetzungsstand und bestehende Herausforderungen sowie Ziele.¹ Von besonderer Bedeutung für die DDB sind Digitalisierungsinitiativen in den verschiedenen Sparten des Kulturerbes, denn hier entsteht ja das Material, das über die DDB zugänglich gemacht wird. Ihre Intensivierung und standardisierte Vorgehensweisen sind der Mutterboden, auf dem die DDB gedeiht. Daher hat die Digitalisierung für die DDB eine hohe Relevanz.

In diesem Beitrag wird der Aktualität halber der funktionale Stand der DDB zugrunde gelegt, der Ende März 2014 mit einem öffentlichen Launch in Betrieb geht. Damit wird die DDB nach 16 Monaten Beta-Betrieb in einen neuen Status überführt und enthält nun viele der Funktionalitäten, die nach der ursprünglichen Planung bereits früher hätten realisiert sein sollen.

1. Deutsche Digitale Bibliothek (DDB) - Ziele und Beginn

Nach mehrjährigen Vorarbeiten startete die DDB formal mit den „Gemeinsame[n] Eckpunkte[n] zur Errichtung einer Deutschen

¹ Vgl. <www.deutsche-digitale-bibliothek.de> [24.03.2014].

Digitalen Bibliothek (DDB)“,² die Bund, Länder und Kommunen Ende 2009 verabschiedeten. Flankiert wurde dieses Papier durch ein „Verwaltungs- und Finanzabkommen über die Errichtung und den Betrieb der Deutschen Digitalen Bibliothek“.³ Beide Dokumente bestimmen den Auftrag des Vorhabens, legen wesentliche Zielgruppen fest und definieren außerdem die organisatorischen Bedingungen, unter denen die DDB handelt.

Festgelegt wird in den Papieren, dass die DDB einen „wesentlichen Beitrag zur Förderung der Wissens- und Informationsgesellschaft“ leisten und den „kulturellen und wissenschaftlichen Reichtum Deutschlands in seiner ganzen Vielfalt national und international (...) präsentieren“ soll.⁴ Mit der Bereitstellung der digitalisierten Inhalte werden die Bedingungen für die Information in Forschung, Lehre und Wirtschaft grundlegend verbessert.

Ein ganz wesentliches Merkmal der DDB ist ihr spartenübergreifender Ansatz: Digitale Objekte aus Archiven, Museen und Bibliotheken werden ebenso einbezogen wie Material aus dem Bereich des Denkmalschutzes, der Medienarchive und auch der Wissenschaft selbst.⁵ Die DDB sorgt also für einen harmonisierten, über eine Oberfläche erreichbaren Zugang auf diese Objekte aus verschiedenen

² URL: <https://www.deutsche-digitale-gemeinsame-eckpunkte-finale-fassung-bibliothek.de/static/de/sc_documents/div/02122009.pdf> [24.03.2014].

³ URL: <https://www.deutsche-digitale-bibliothek.de/static/de/sc_documents/div/verwaltungs_und_finanzabkommen_finale%20Fassung02122009.pdf> [24.03.2014].

⁴ Vgl. Eckpunkte, S. 4.

⁵ Vgl. Eckpunkte, S. 8f.

Sparten (in der Regel digitale Abbilder von physischen Objekten aus den Sammlungen, aber auch reine Metadateninformationen), sie ist aber nicht der Zugang selbst: Eine Recherche in der DDB führt den Nutzer in der Regel zu den Präsentationsplattformen der jeweiligen Kultureinrichtung.

Ausdrücklich fordern die Gründungsdokumente der DDB innovative technische Werkzeuge ein, die umfassende Recherchen möglich werden lassen. Insbesondere sollen die heterogenen Datenbestände der verschiedenen Partnereinrichtungen ineinander verwoben, also in eine semantische Beziehung gebracht werden, sodass zum Beispiel eine Suche nach Personen übergreifend Objekte aus allen Sparten und in unterschiedlichen Funktionen gegenüber dem jeweiligen Objekt zu Tage fördert.⁶

Zusammenfassend lässt sich sagen, dass die DDB die digitalen Bestände aus Kultur und Wissenschaft in Deutschland zusammenführt. Sie eröffnet einen zentralen Zugangspunkt auf das kulturelle Erbe und stellt die dazu notwendige technische und organisatorische Infrastruktur bereit. Die DDB ist auf diese Weise einerseits das spartenübergreifende und interdisziplinäre Zugangsportal zu den digitalen Angeboten der deutschen Kultur- und Wissenschaftseinrichtungen und sorgt dafür, dass Kultur und Wissen aus Deutschland in digitaler Form innerhalb des Landes und global sichtbar, zugänglich und erfahrbar wird. Darüber hinaus vernetzt und kontextualisiert sie die digitalen Bestände und Objekte aus verschiedenen Sparten und öffnet sie für vielfältige Möglichkeiten der Nachnutzung, Anreicherung

⁶ Vgl. Eckpunkte, S. 10f.

und Erweiterung durch private Nutzer, Wissenschaftler und Institutionen. Sie bietet so nicht nur neue Einsichten in verschiedenste Materialgruppen, sondern eröffnet auch neue Möglichkeiten der Zusammenarbeit und der interaktiven Teilhabe an diesen Beständen und ihrer Vernetzung.

Aufgabe der DDB ist es also, zum zentralen nationalen Baustein für die Transformation deutscher Kultur- und Wissensschätze ins digitale Zeitalter zu werden. Sie ist so gleichzeitig ein wesentlicher Katalysator für die spartenübergreifenden Bemühungen zur Digitalisierung des Kulturerbes. Darüber hinaus führt sie Kultur- und Wissenschaftseinrichtungen zusammen und eröffnet Potentiale für den Austausch von Erfahrungen, Technologien und Diensten. Innovative Entwicklungen im Bereich der Digitalisierung und Zugänglichmachung von Kulturgut bilden den Rahmen für erweiterte Aufgabenschwerpunkte, zu deren wichtigstem die Standardisierung in diesem Feld gehört. Als von der öffentlichen Hand getragenes und kooperativ aufgestelltes Vorhaben erfüllt die DDB damit letztlich einen nationalen kulturpolitischen Auftrag von Bund, Ländern und Gemeinden.

Die DDB versteht sich außerdem als nationaler Beitrag zur Europeana. Hier werden die Kulturgüter bzw. deren digitale Abbilder aller Mitgliedstaaten der Europäischen Union zusammengeführt und weltweit zugänglich gemacht. Dieser Auftrag der DDB ist additiv zu verstehen und lässt ihre Funktionen nicht in der Europeana aufgehen: Die Grundlagenarbeit des Zusammenführens und Harmonisierens der Datensets erfolgt am besten vor Ort. Zum anderen gibt es aber auch einen wesentlichen Unterschied zur Europeana: Die DDB macht auch Informationen zugänglich, die auf noch nicht

digitalisierte Inhalte verweisen und bedient damit ein wichtiges Interesse der Wissenschaft. Basis dafür sind die im Vergleich zur European reichhaltigeren Metadaten. Der daraus resultierende semantische Wert, den sie in den einzelnen Sparten erhalten, wird durch komplexe Prozessschritte genutzt.

Organisatorisch bildet das Kompetenznetzwerk mit insgesamt 13 Partnereinrichtungen den Kern der DDB, ein aus diesem Kreis gebildeter Vorstand vertritt das Kompetenznetzwerk nach außen. Als Aufsichtsgremium wurde ein Kuratorium gebildet, das mit Vertretern von Bund und Ländern sowie eines Vertreters der kommunalen Spitzenverbände besetzt ist. Für das operative Handeln sahen die Gründungsdokumente die Etablierung einer Geschäftsstelle vor, die bei der Stiftung Preußischer Kulturbesitz als Bund-/Ländereinrichtung angesiedelt wurde.

Der ursprünglichen Planung nach sollte die DDB bereits 2011 in Betrieb gehen. Mit der technischen Umsetzung wurde durch den Bundesbeauftragten für Kultur und Medien (BKM) 2010 das Fraunhofer Institut für Intelligente Analyse- und Informationssysteme (IAIS) beauftragt. Die Finanzierung dieses sogenannten Aufbauprojekts erfolgte über erhebliche Sondermittel des Bundes aus dem Konjunkturprogramm 2.

Neben Sondermitteln zum initialen Aufbau war und ist für die DDB wesentlich die Etablierung eines Etats, der von Bund und allen Ländern paritätisch getragen, insgesamt ein Volumen von 2,6 Mio. € umfasst.

2. Deutsche Digitale Bibliothek (DDB) - Aufbau und erreichter Stand

2011, mit der Bereitstellung der Mittel für den regulären Etat, nahm das Kompetenznetzwerk seine Arbeit auf. Wichtige Schritte waren der Aufbau der Geschäftsstelle und die Auswahl eines Partners für den technisch-operativen Betrieb der DDB. Hier fiel die Wahl auf das FIZ Karlsruhe – Leibniz-Institut für Informationsinfrastruktur. Das FIZ nahm dabei nicht nur Aufgaben der technischen Bereitstellung wahr, sondern zunehmend auch Entwicklungsleistungen.

Aus dem Kreis des Kompetenznetzwerks übernahm die Deutsche Nationalbibliothek die Funktion der Koordination zwischen den Partnern des Kompetenznetzwerks (KNW) und dem damaligen Entwicklungspartner IAIS. 2012, nach Übergabe der Ergebnisse des Aufbauprojekts durch IAIS an das Kompetenznetzwerk übernahm die Deutsche Nationalbibliothek (DNB) die Projektleitung und die technische Gesamtkoordination mit dem Ziel, die DDB in eine erste betriebs- und veröffentlichungsfähige Stufe zu überführen. Dies erfolgte mit dem sogenannten Beta-Launch Ende November 2012.

Wesentliche Aufgabe der DNB ist damals wie heute die Koordination aller Beteiligten – der fachlich zuständigen KNW-Arbeitsgruppen⁷, der unterschiedlichen Umsetzungspartner⁸ sowie der Geschäftsstelle⁹. Zu den Arbeitsinstrumenten gehören das

⁷ KNW-Arbeitsgruppen vor allem für Präsentation, Daten, Innovation sowie weiterer ad hoc eingerichteter Arbeitsgruppen zu speziellen Themen.

⁸ Umsetzungspartner: FIZ Karlsruhe, Fraunhofer IAIS sowie weitere externe Auftragnehmer.

⁹ Geschäftsstelle für den Bereich Kommunikation von Projektplanungen und -ergebnissen, für die Gremienbetreuung und andere Aufgaben mehr.

Erstellen und Umsetzung von Projektplänen, von Aufwandsabschätzungen, Anforderungsdefinitionen, ferner die Auswahl, Beauftragung und Überwachung von Umsetzungspartnern und deren Ergebnissen. 2013 kam nach vorbereitenden Planungsschritten als neue Einrichtung des KNW die Servicestelle hinzu, eine Funktion, die ebenfalls von der DNB wahrgenommen wird.¹⁰ Damit hatte der organisatorische Aufbau der DDB einen stabilen Stand erreicht.¹¹

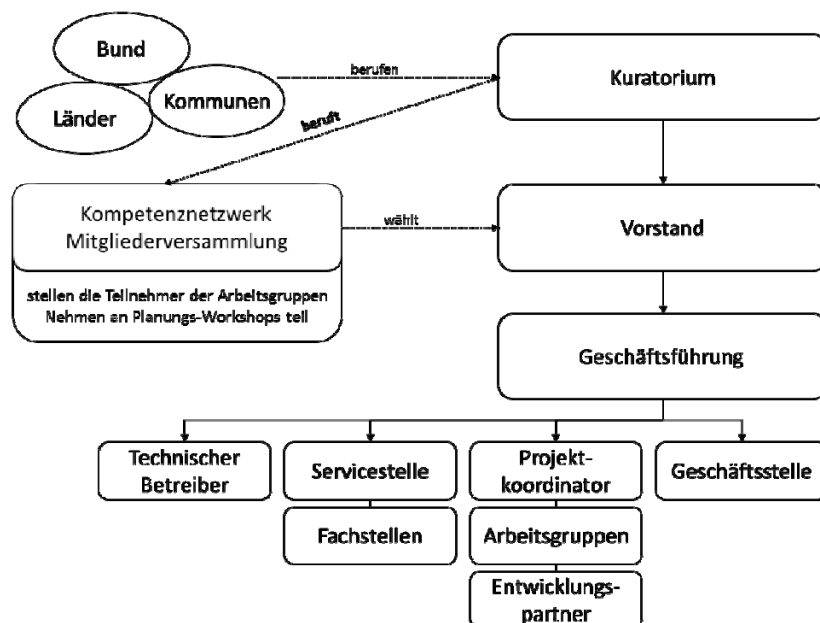


Abb. 1: Organigramm der DDB.

¹⁰ Denise Baumgart/Francesca Schulze, Die Servicestelle der Deutschen Digitalen Bibliothek, in: Dialog mit Bibliotheken 24,2, 2012, S. 8 - 12.

¹¹ Vgl. Abb.1.

Aufgabe der Servicestelle ist zunächst, im Rahmen einer mittelfristig angelegten Aufbauphase einen leistungsfähigen Arbeitsprozess für die Aufnahme neuer digitaler Bestände in die DDB aufzubauen. Dafür entwickelt die Servicestelle ein kooperatives Arbeitsverfahren, das die spezifischen Kompetenzen in den unterschiedlichen Einrichtungen innerhalb und außerhalb des KNWs effizient bündeln soll. In fünf definierten Arbeitsbereichen bzw. Phasen¹² agiert sie gegenüber den Kultur- und Wissenschaftseinrichtungen als ein „Single Point of Contact“ – vom Erstkontakt bis zur Veröffentlichung der digitalen Bestände im Portal der DDB und in der Europeana.

Zusätzlich wurden – koordiniert durch die Servicestelle – dezentrale Fachstellen eingerichtet, die im Auftrag der Servicestelle definierte Aufgaben übernehmen. Als Bindeglieder in ihre jeweiligen Communities haben sie beratende Funktionen und leisten konzeptionelle Beiträge, insbesondere im Bereich der sparten- und material-spezifischen Lieferformate¹³. Weitere Aufgaben liegen in der Akquise neuer Datenlieferanten sowie die Unterstützung der jeweiligen Einrichtung bei der Datenklärung/-analyse.

Die Harmonisierung und Aufbereitung der Daten für die DDB hat sich bislang als eine der entscheidenden Aufgaben der DDB erwiesen. Inhaltlich geht es dabei um die Etablierung von Standards und die spartenübergreifende Abbildung der verschiedenen Lieferformate

¹² Arbeitsbereiche / Phasen: Community-Arbeit, Registrierung, Content-Akquise, Datenclearing und Datenübergabe.

¹³ Spezifische Lieferformate: u.a. METS-MODS-DDB, EAD-DDB, LIDO-DDB, siehe unten.

auf ein mächtiges Zielformat. Im Fall der DDB kommt das European Data Model (EDM) zum Einsatz.¹⁴ Dieser Ansatz einer spartenübergreifenden Zusammenführung von Daten in eine gemeinsame semantische Datenmodellierung hat auch Rückwirkungen auf die Sparten selbst: Die systematische Ausrichtung von Daten unterschiedlicher Provenienz auf ein Zielformat führt nämlich auch dazu, dass Entstehungsprozesse von Daten in den einzelnen Sparten hinterfragt und vereinheitlicht werden. Das, was früher als Erschließungsleistung nur Relevanz für die einzelne Einrichtung hatte, gewinnt an Bedeutung für die gesamte Sparte und auch darüber hinaus.¹⁵

Weitere Aufgaben der DNB im Rahmen ihrer Arbeit als Koordinator beziehen sich auf die Außendarstellung der DDB in der Fachöffentlichkeit sowie die Kooperation im Bereich der technischen Entwicklung und der Standardisierung mit anderen Partnern – insbesondere mit der Europeana.

Schließlich gehören Aufgaben der mittelfristigen Strategieentwicklung zum Aufgabenkatalog der DNB. Ein hier zum Einsatz kommendes Instrument sind die jährlich durchgeführten Planungsworkshops unter Einbeziehung des gesamten KNW und Externer, in denen grundsätzliche Entwicklungstendenzen und die Planung des jeweiligen Folgejahres abgestimmt werden.

Neben diesen besonderen Teilnahmemöglichkeiten an der Deutschen Digitalen Bibliothek stehen die regulären Optionen: Prinzipiell

¹⁴ Das EDM ist unter: <http://pro.europeana.eu/edm-documentation> [23.04.2014] dokumentiert.

¹⁵ Vgl. Abs. 4 dieses Beitrags.

steht die DDB allen deutschen Kultur- und Wissenschaftseinrichtungen offen. Nach einer Registrierung werden sie auf der Kulturlandkarte der DDB verzeichnet und so mit einer Reihe von Basisinformationen im Netz sichtbar. Um auch mit ihren Daten sichtbar zu werden, ist der Abschluss des Kooperationsvertrages erforderlich, der einerseits zwischen DDB und der jeweiligen Einrichtung eine stabile Basis für die Zusammenarbeit begründet und andererseits inhaltlich die Nutzungsrechte an Metadaten und den sogenannten Derivaten, also Ableitungen von den digitalen Objekten (vor allem Vorschaubilder) klärt. Als Standardfall wird dabei eine Lizenzierung der Metadaten unter CC0 angesehen – als Voraussetzung auch dafür, dass die Daten an die Europeana weitergegeben werden können. Außerdem werden Festlegungen für den gegenseitigen Verantwortungs- und Haftungsbereich getroffen. Zurzeit sind bereits über 2.000 Einrichtungen registriert, mehr als 100 liefern Daten an die DDB.¹⁶

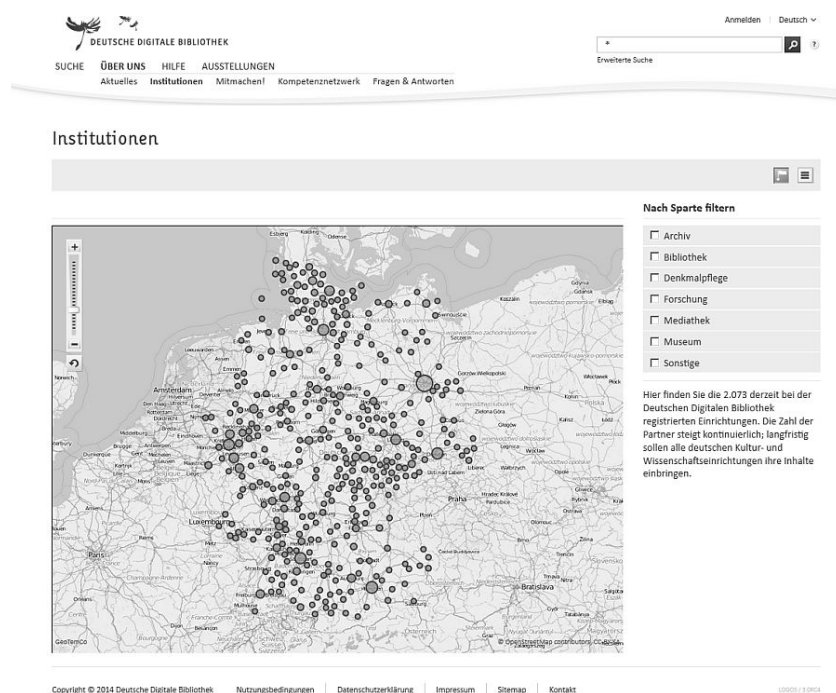


Abb. 2: Kulturlandkarte der DDB.

Seit dem Beta launch 2012 wurden nicht nur Software und Betrieb stabilisiert, sondern wesentliche Umbauten in der Gesamtarchitektur vorgenommen. So wurde das proprietäre interne Datenmodell auf das Datenmodell der Europeana (EDM) umgestellt, die Benutzeroberfläche der DDB softwaretechnisch komplett erneuert sowie die La-deprozesse erweitert und optimiert.

¹⁶ Vgl. Abb. 2.

Ende März 2014 erreicht die DDB mit nun rund 8 Mio. nachgewiesenen Objekten und einem konsolidierten Funktionsumfang einen neuen Leistungsstand, zu dem auch die reguläre Belieferung der Europeana mit Daten gehört. Damit ist die Entwicklung der DDB allerdings keinesfalls abgeschlossen, denn noch sind viele Bestände von interessierten Datengebern nicht in die DDB integriert.

3. Deutsche Digitale Bibliothek (DDB) - Funktionalitäten und weiterer Ausbau

Der hier folgende kurze Überblick fokussiert besonders den allgemein im Internet sichtbaren Teil der DDB.

Wichtigstes Moment der Sichtbarkeit der DDB ist ihr Auftritt im Web und hier vor allem das Portal. Der zentrale Einstieg in die DDB erfolgt über einen Suchschlitz¹⁷, der eine leistungsfähige Suchmaschine anspricht. Treffer lassen sich in Listen- oder Galerieform darstellen und mittels einer Reihe von Facetten¹⁸ weiter eingrenzen. Auch innerhalb der Facetten selbst kann weiter gesucht werden, auch eine rollen- resp. funktionsbasierte Differenzierung von Personen und Orten ist möglich. Wählt der Benutzer einen Suchtreffer aus, gelangt er zur sog. Detailansicht, also einer Seite, auf der die Metadaten des zum Suchtreffer gehörenden Objekts gezeigt werden. Darüber hinaus bietet die Detailansicht einen Medienviewer¹⁹, der die Anzeige oder das Abspielen von Objekten erlaubt, sowie einen Link

¹⁷ Vgl. Abb. 3.

¹⁸ Vgl. Abb. 4.

¹⁹ Vgl. Abb. 5.

zum Angebot des Datengebers sowie Informationen zum Rechtestatus des Objekts. Bei Objekten, die Teil einer Hierarchie sind, z.B. Archivalien, die in eine Findbuch-Tektonik eingebettet sind, wird diese Hierarchie als navigierbare Baumstruktur im unteren Teil der Detailansicht eingeblendet.²⁰ Für angemeldete Benutzer ist es in der Detailansicht möglich, sich den Inhalt des zugrundeliegenden Datenpakets als XML anzeigen zu lassen, was insbesondere für professionelle Benutzer wie z.B. auch die Datenlieferanten selbst interessant ist.

²⁰ Vgl. Abb. 6.

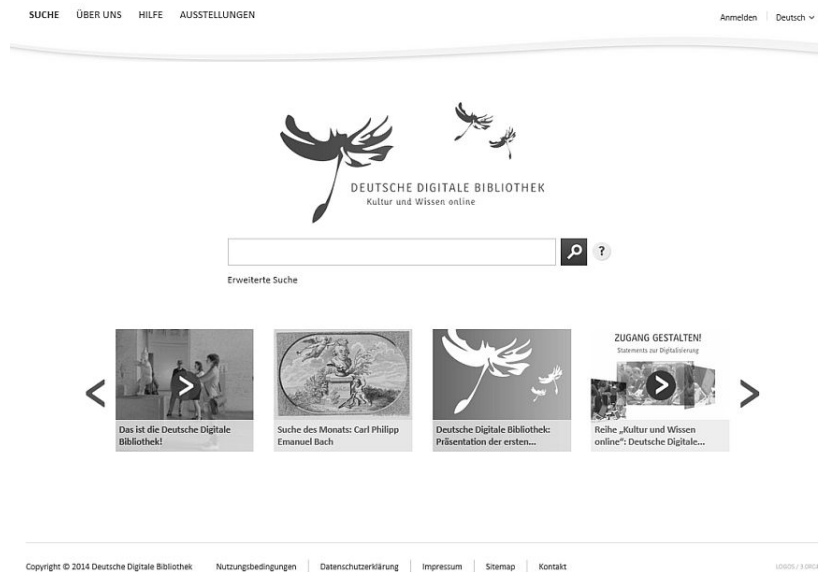


Abb. 3: Einstiegsseite der DDB.

DEUTSCHE DIGITALE BIBLIOTHEK

SUCHE ÜBER UNS HILFE AUSSTELLUNGEN

Anmelden Deutsch

Ergebnisse pro Seite: 20 Seiten reich: 100

1-20 von 154 Ergebnissen

Nur Ergebnisse mit Multimediale

Ergebnisse filtern

Von Tag Monat Jahr

Bis Tag Monat Jahr

Erweiterung: Erweitert Genau

Ort

Person/Organisation

Bach, Johann Sebastian

Beteiligt an

Filter hinzufügen

Stichwort

Sprache

Medien

Sparte

Dokument

Alle Filter aufheben

Alle Filter beibehalten

Objekte vergleichen

Erstes Objekt

Zweites Objekt

Johannespassion: für Soli, Chor, Orchester mit Cembalo und Orgel, von Johann Sebastian Bach

2000

Johannespassion: für Soli, Chor, Orchester mit Cembalo und Orgel, von Johann Sebastian Bach...

Bach, Johann Sebastian; Korne, Barbara; Springer, Ingeborg; Menzel, Peter; Politz, Hermann...

Orgel-Büchlein <in dulci jubilo>

1949

Erhält es werden: Passover, BWV 590, Teil 4 (Schluss) / Bach, Johann Sebastian (Einheitsrechte...

...Passover, Org. BWV 590, in der 1. Fassung, Choralbearbeitung BWV 615 / Bach, Johann Sebastian...

Bach, Johann Sebastian; Ande-Conelli, Hans...

Orgel-Büchlein <in dulci jubilo>

1949

Erhält es werden: Passover, BWV 590, Teil 4 (Schluss) / Bach, Johann Sebastian (Einheitsrechte...

...Passover, Org. BWV 590, in der 1. Fassung, Choralbearbeitung BWV 615 / Bach, Johann Sebastian...

Bach, Johann Sebastian; Ande-Conelli, Hans...

Ave Maria

1925

Bach, Johann Sebastian; Gould, Charles; Davis, Erinn; Hoffman, Joseph... Ave Maria / I...

Bach-Gould, Ave Maria...

Ergebnisse pro Seite: 20 Seiten reich: 100

Seite 1 von 8 Weiter > Ende >

Abb. 4: Ergebnisliste in der DDB, Blick auf „aufgeblähte“ Facetten.

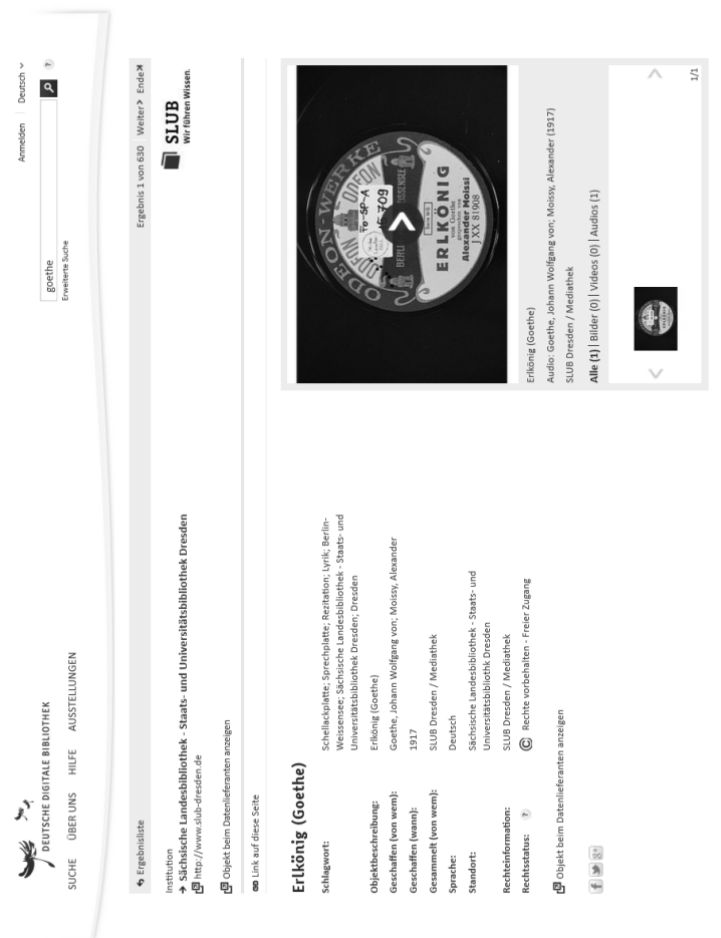


Abb. 5: Treffer in der DDB mit Viewer.

DEUTSCHE DIGITALE BIBLIOTHEK

ANMELDEN | Deutsch

SUCHE ÜBER UNS HILFE AUSSTELLUNGEN

Erweiterte Suche

Ergebnisse 4.999 von 8.267 Weiter Ende

Kontext < Zurück Ergebnis 4.999 von 8.267 Weiter > Ende

Institution
→ Landesarchiv Baden-Württemberg
http://www.landearchiv-bw.de

Objekt beim Datenlieferanten anzeigen

Link auf diese Seite

Flößerei auf der Enz

Archivsignatur: Landesarchiv Baden-Württemberg, Abt. Hauptstaatsarchiv Stuttgart, N 220 T 31

Kontext: Nachlass Heinrich Schickhardt, Architekt und Ingenieur >> 2. Technische Zeichnungen und Pläne >> 2.2. Wasserbau >> 2.2.4. Flößerei

Laufzeit: 1612-1613

Umfang: 1 Bl

Indesbegriff Ort: Enz (Kuss), Flößerei
Enklösterle Cw, Flößerei

Digitalisat im Angebot des Archivs: kein Link zu einem Digitalisat vorhanden

Bestand: Landesarchiv Baden-Württemberg, Abt. Hauptstaatsarchiv Stuttgart, N 220 Nachlass Heinrich Schickhardt, Architekt und Ingenieur

Online-Einblick im Angebot des Archivs: http://www.landearchiv-bw.de/plink/71-1-29318a-b

Rechtsinformation: Es gelten die Nutzungsbedingungen des Landesarchivs Baden-Württemberg.

Objekt beim Datenlieferanten anzeigen

Facebook Twitter YouTube

Verbundene Objekte

2. Technische Zeichnungen und Pläne

- 2.2. Wasserbau
 - 2.2.4. Flößerei
 - Flößerei auf der Enz
 - Sonstiges
 - Gutachten über das Flößen bei Enklösterle von M. Johann Ottinger, geschrieben von Schickhardt (3 Folio)
 - Bericht über das Flößen von "geschnitten zeugs" und Brennholz bei Enklösterle aus dem Altensteiger Forst (Poliohaft)
 - Nachlass Schickhardt über das Flößen von Holz auf der Enz, Holz aus Schöllkopf und dem Steinswald zu Flößen auf der Lauter unter Preudenstadt bis zum Vecher bei Würtingen; Flößen auf der kleinen Enz, bei Enklösterle und bei Enswaltingen, mit Bemerkungen über die Vorteile des Flößens gegenüber den Fuhrn (2 Folio, 3 Quart, 2 Oktav, teilweise...)
 - Flößen von Götteffingen, Simmersfeld, Fünfborn, Enklösterle und Poppelmühl; Berechnungen über Flößlasten (2 Folio Doppelblatt)
 - Auszug über die Ausgaben für Flößholz in Einzingen an der Enz (Folio Doppelblatt)

Abb. 6: Hierarchie von Objekten in der DDB.

Suchanfragen können gespeichert und Suchtreffer (Vorschaubilder) paarweise verglichen werden.

Als weitere Funktionen zur Verbesserung individueller Such- und Arbeitsmöglichkeiten insbesondere auch für Wissenschaft und Lehre bietet die DDB die Speicherung von Objekten als Favoriten. Die Verwaltung und Annotation der gespeicherten Favoriten in persönlichen Kollektionen („Favoritenlisten“) ist ebenso möglich wie das Teilen von URLs zur Objekt-Detailansicht via Email und Soziale Medien. Über die Veröffentlichung von Favoritenlisten als persönliche Themenseiten werden anderen Benutzern individuell gestaltete bzw. entdeckende Einstiege in die DDB ermöglicht.

Der Anspruch der DDB, auch Querbezüge zwischen verschiedenen Objekten aus unterschiedlichen Einrichtungen und Sparten sichtbar werden zu lassen, materialisiert sich über Personenseiten, die in die DDB integriert, die in die Gemeinsamen Normdatei (GND)²¹ geflossene Standardisierungsarbeit ausnutzen: Für in der GND erfasste Personen werden auf einer gesonderten Seite Informationen aus der GND sowie in der DDB verfügbare Objekte angeboten. Sie werden mit der Person verknüpft und mit externen Links aus anderen Informationsquellen angereichert.

²¹ <http://www.dnb.de/DE/Standardisierung/GND/gnd_node.html>
[23.04.2014].

DEUTSCHE DIGITALE BIBLIOTHEK

SUCHE ÜBER UNS HILFE AUSSTELLUNGEN

Anmelden Deutsch

goethe johann wolfgang von

Erweiterte Suche

Link auf diese Seite

Johann Wolfgang von Goethe

Schriftsteller, Publizist, Politiker, Jurist, Naturwissenschaftler, Theaterintendant, Maler, Zeichner
 Geboren: 28. August 1749, Frankfurt am Main
 Gestorben: 22. März 1832, Weimar

Objekte:

Gedichte von Johann Wolfgang von Goethe...
 Gedichte von Johann Wolfgang von Goethe...
 Gedichte von Johann Wolfgang von Goethe...
 Gedichte von Johann Wolfgang von Goethe...
 Gedichte von Johann Wolfgang von Goethe...

[Alle Audios \(85\)](#)

Beteiligt an:

- Hermann und Dorothea : in neun Gesängen
Goethe, Johann Wolfgang von. - Berlin : Weidmann-Bücher-Verl.-Ges., 1927, [1927]
- Die Wahlverwandtschaften
Goethe, Johann Wolfgang von. - Leipzig : Bibliogr. Inst., 1938, [1938]
- Faust : eine Tragödie ; Teil I
Goethe, Johann Wolfgang von. - Leipzig : Amelang, 1919
- Reineke Fuchs
Goethe, Johann Wolfgang von. - München : Recht, 1922

[Alle Objekte \(8\)](#)

Thema in:

- Zwischen den Zeiten : zur gesellschaftlichen Modernität in Goethes späten Erzählungen ; eine literatursoziologische Analyse
Voss, Alexandra (1975-). - 2004
- Goethe und die Architekturtheorie
Büchenschul, Jan. - 2009

[Alle Objekte \(2\)](#)

Johann Wolfgang von Goethe
 Quelle: [Wikimedia](#)

Suche:

☒ "Johann Wolfgang von Goethe" in der DDB

Externe Links:

- Wikipedia (Deutsch)
- Bibliothèque nationale de France | National Library of France
- Wikipedia (English)
- Library of Congress / NACO
- Wikisource
- VIAF - Virtual International Authority File
- Historische Kommission bei der Bayerischen Akademie der Wissenschaften und Bundesarchiv
- Historische Kommission bei der Bayerischen Akademie der Wissenschaften und Bundesarchiv
- Historisches Lexikon der Schweiz (HLS)

Copyright © 2014 Deutsche Digitale Bibliothek Nutzungsbedingungen Datenschutzerklärung Impressum Sitemap Kontakt

UDN / URN

Abb. 7: Entitätenseite in der DDB.

Voraussetzung für diese Funktion ist die Verknüpfung der unterschiedlichen Objektmetadaten mit den entsprechenden GND-Identifiern. Über diese Brückenlegungen werden erweiterte Zugangsmöglichkeiten und Bezüge möglich. Ein entdeckender Zugang erlaubt neue Einsichten in Zusammenhänge zwischen Objekten. Gezielt werden solche Verbindungen auch durch die virtuellen Ausstellungen der DDB vermittelt, die Objekte aus der DDB in dem Kontext eines Ausstellungsthemas zusammenführen und redaktionell aufbereitet präsentieren.

Neben der Ausgestaltung der Suchoberfläche bietet die DDB auch einen anderen Zugang auf ihren Datenindex: Externe Anwendungen, die auf die Daten der DDB zugreifen wollen, können ein sogenanntes Application Program Interface (API) nutzen. Also eine Programmierschnittstelle, die den Zugriff auf Daten und Methoden der Deutschen Digitalen Bibliothek (DDB) ermöglicht.²² Die API erlaubt die Entwicklung vielfältiger Anwendungen, die die in der DDB vorgehaltenen Inhalte nutzen und sie jeweils nach den eigenen Wünschen darstellen und in unterschiedliche Kontexte einbetten. Ein besonderes interessantes und fachlich relevantes Beispiel für eine solche Anwendung ist das Archivportal-D, das datenseitig vollständig auf dem Index der DDB basiert.²³

Neben diesen beiden grundlegenden Funktionalitäten verfügt die DDB über weitere Werkzeuge, die im sog. Serviceportal zusammengefasst sind: Einerseits über eine Dokumenten-, Informations- und Kommunikationsplattform, auf der beispielsweise ein Glossar, das in der DDB verwendeten (technischen) Begriffe enthält, andererseits eine Reihe von Administrationswerkzeugen. Beispiele hierfür sind das Tool zur Registrierung von Kultur- und Wissenschaftseinrichtungen und eine Software für Nutzerumfragen. Darüber hinaus wird zur Unterstützung der Geschäftsabläufe der Servicestelle eine Helpdesk-/CRM-Software eingerichtet, die die Kommunikation zwischen der Servicestelle und den Fachstellen erleichtert. Ziel ist es, zentral relevante Informationen wie die Kontaktdaten der Einrichtungen und

²² Siehe auch <<https://api.deutsche-digitale-bibliothek.de/>> [23.04.2014].

²³ Vgl. <<http://www.landesarchiv-bw.de/web/54267>> [23.04.2014].

Metadaten zu den Sammlungen der Kultur- und Wissenschaftseinrichtungen zusammen mit Statusangaben zur Lieferung, zum Stand des konzeptionellen und technischen Mappings, zum Datenclearing und Ingest der Sammlungen zu pflegen und aktuell zu halten.

4. Dateneinspielungen und Standards in der DDB

Für die Dateneinspielungen ist die Transformation der per FTP geladenen oder über eine OAI-PMH-Schnittstelle geharvesteten Metadaten von grundlegender Bedeutung. Metadaten wie auch die Binärdateien (Bilddaten) werden zunächst validiert und DDB-spezifisch aufbereitet: Zum Beispiel werden komplexe Hierarchien in Einheiten aufgeteilt, die später je ein digitales Objekt in der DDB bilden werden. Anschließend erfolgt eine Datennormalisierung. Diese werden in einem von der DDB definierten Eingangsformat vorliegenden Originalmetadaten nach den vom konzeptionellen Mapping festgelegten Regeln in das interne Datenformat der DDB überführt. In diesem Prozess werden zudem verschiedene Sichten der Daten generiert, die die Grundlage für die Funktionalitäten der DDB bilden. Als Bsp. lassen sich die Darstellung im Portal, die Suchindizes und -facetten sowie den Datenexport an Europeana anführen. Auch eine Anreicherung und Verknüpfung der Metadaten findet statt. Hinzu können statistische Auswertungen oder Qualitätssicherungsmaßnahmen treten und eindeutige Identifikatoren für die erzeugten digitalen Objekte vergeben werden. Schließlich werden die Metadaten und die Binärdateien im Archiv gespeichert.

Basis der technischen Transformation ist ein konzeptionelles Mapping, in dem die Umsetzung der einzelnen Elemente des

Ausgangsformats²⁴ detailliert dokumentiert ist. Dies erfolgt in Form von Konkordanztabellen, in denen Entitäten eines zugelassenen Eingangsformats den Entitäten des Zielformats zugeordnet werden. Das interne Metadatenformat der DDB basiert auf dem Europeana Data Model (EDM), das in Form eines DDB-spezifischen Anwendungsprofil zum Einsatz kommt und besonders die spartenübergreifende Suche, semantische Vernetzungsmöglichkeiten und die strukturierte Darstellung der (digitalen) Inhalte und ihrer Metadaten in der DDB optimal unterstützt. Das EDM selbst basiert auf den Relationsprinzipien von Linked Data, die Daten sind in der Syntax des Resource Description Frameworks (RDF) modelliert. Insgesamt besteht EDM aus 15 Klassen und ermöglicht es, ein Objekt des Kulturerbes aus unterschiedlicher Perspektive zu beschreiben: Zum einen als kulturwissenschaftliches Objekt selbst, zum anderen seine digitale Repräsentation sowie schließlich die dazugehörigen Metadaten.²⁵ Es erlaubt die nachhaltige Verknüpfung von Objekten mit ihren Kontexten, beispielsweise mit Personen, Orten oder Ereignissen. Auf dieser Grundlage werden zukünftig im Portal der DDB weitere explorative Recherche- und Präsentationsformen über verschiedene Datenbestände hinweg angeboten.

Zugelassene Lieferformate für Metadaten sind zurzeit Dublin Core (DC), DenkXWeb (Denkmalpflege), indexMeta (Filmarchive),

²⁴ Das Ausgangsformat ist in der Regel ein in der jeweiligen Sparte maßgebliches Metadatenformat.

²⁵ Stephan Bartholmei/Francesca Schulze: Die Weiterentwicklung der Deutschen Digitalen Bibliothek, in: Dialog mit Bibliotheken, 2, 2013, S. 7 – 12, S. 8.

ESE²⁶, EAD(DDB)²⁷, METS/MODS²⁸, MARCXML²⁹ und schließlich LIDO³⁰ für die Museumswelt. Diese Formate sind in den jeweiligen Sparten weit verbreitet. Für die Ablieferung der Metadaten wird das XML-Format bevorzugt, da die Transformation mittels XSLT-basierter Skripte erfolgt. Können die genannten gängigen Standardformate nicht geliefert werden, bieten die Servicestelle der DDB und ihre Fachstellen Unterstützung an, um das jeweils in der einzelnen Einrichtung genutzte Format auf eines der zugelassenen Lieferformate abzubilden.³¹

Allerdings sind solche Anpassungen aufwändig. Sie benötigen Zeit und viele Abstimmungsrunden und bergen das Risiko von Informationsverlusten oder auch Verzerrungen, da nicht immer ein Sachverhalt wirklich adäquat gemappt werden kann. Häufig ist dies nicht ein Problem der mangelnden inhaltlichen Granularität im Zielformat, sondern von uneinheitlichen oder nicht stringent strukturierten Daten beim Ausgangsbestand. Aus dem Gesagten wird deutlich, dass das Wirken der DDB auch dahin geht, den Einsatz von Standardformaten

²⁶ ESE: Europeana Semantic Elements als früheres Format der Europeana.

²⁷ EAD (DDB): Encoded Archival Description (Deutsche Digitale Bibliothek) - Standardformat der Sparte Archiv.

²⁸ METS/MODS: Metadata Encoding & Transmission Standard/Metadata Object Description Schema (Standardformat im Digitalisierungsprojektumfeld in Bibliotheken).

²⁹ MARCXML: Machine-Readable Cataloging – das internationale Standardmetadatenformat der Bibliotheken in XML.

³⁰ LIDO: Lightweight Information Describing Objects.

in den jeweiligen Sparten zu forcieren. Damit gehen letztlich auch gemeinsame Beschreibungsplattformen wie die GND oder auch die Nutzung von Persistent Identifier Diensten einher.

Nur auf diesem Weg kann die DDB den quantitativen³² und qualitativen³³ Anforderungen entsprechen. Diese Überlegung trifft letztlich nicht nur für die Metadaten, sondern auch für die Digitalisierungsmaßnahmen in den verschiedenen Sparten zu, deren Ergebnisse in hoffentlich steigender Zahl in die DDB eingespeist werden.

5. Digitalisierung – die zentrale Herausforderung der Kulturerbeerichtungen

Der Stand der Digitalisierung in den verschiedenen Sparten der DDB ist sehr unterschiedlich. Insbesondere in Bibliotheken wird vor allem mit Förderung der DFG seit Jahren systematisch digitalisiert, sodass Digitalisierungsaktivitäten aus der Sicht der DDB nicht ohne einen Blick auf diese Initiative gewürdigt werden können.

Ausgangspunkt war ab Mitte der neunziger Jahre vorbereitet ein entsprechendes Förderprogramm der DFG, ab 1997 mit der erklärten Absicht, den Zugang auf für Wissenschaft und Forschung relevantes Material durch die Retrodigitalisierung von gemeinfreiem schriftlichem Kulturgut zu verbessern. Daran knüpften – nach einer

³¹ Vgl. <<http://www-t1.deutsche-digitale-bibliothek.de/content/faq/>> [23.04.2014].

³² Quantitativen Anforderungen: Ausweitung des Volumens recherchierbarer Objekte.

³³ Qualitative Anforderungen: spartenübergreifende Recherchen in angemessener Tiefe.

intensiven Explorationsphase - Aktivitäten wie der Aufbau von Digitalisierungszentren und die Etablierung technischer Standards und Workflowtools an, die als Grundlage für alle weiteren Projekte galten. Der umfassende Anspruch des Förderprogramms gilt nahezu unverändert fort: So beschreibt das Positionspapier zu den Schwerpunkten der wissenschaftlichen Literaturversorgungs- und Informationssysteme (LIS) bis 2015 die großflächige Digitalisierung der nachgewiesenen Drucke des 16.-18. Jahrhunderts und der Sondersammelgebietsbestände sowie der Findmittel der deutschen Archive als eigenständige Maßnahmen.³⁴

Insgesamt beträgt das Fördervolumen durch die DFG bis 2012 über 150 Millionen Euro und wird mit neuen Schwerpunkten (hist. Zeitungen, unikales Material wie Archivgut, Handschriften und museale Objekte) fortgeführt.³⁵ Zunehmend bedeutsam wurden in den vergangenen Jahren auch Förderprogramme der Länder, die allerdings oft nur befristet als Sonderprogramm aufgelegt werden.

³⁴ Rolf Griebel, Die Förderung der wissenschaftlichen Informationsinfrastruktur durch die DFG: Zwischenbilanz zum DFG-Positionspapier, Wissenschaftliche Literaturversorgungs- und Informationssysteme: Schwerpunkte der Förderung bis 2015, in: Zeitschrift für Bibliothekswesen und Bibliographie (ZfBB), 57,2, 2010, S. 71-85, S. 77ff.

³⁵ Frank Simon-Ritz: Kulturelles Erbe im digitalen Zeitalter: Der Weg der Bibliotheken, in: Forum Bibliothek und Information 64, 2012, 03, S. 218–222, S. 219. Vgl. a. das Positionspapier der DFG „Die digitale Transformation weiter gestalten“, Bonn, 2012, S. 10. URL: http://www.dfg.de/download/pdf/foerderung/programme/lis/positionspapier_digitale_transformation.pdf [23.04.2014].

Insgesamt ist durch die Förderung der DFG im Bereich der Bibliothekssparten eine leistungsfähige Infrastruktur entstanden, die aus Digitalisierungszentren an verschiedenen Orten besteht. Diese Zentren übernehmen vorwiegend regional ausgerichtet neben der Durchführung von Projekten auch Beratungs- und Koordinationsaufgaben, um insbesondere kleinere Einrichtungen bei der Durchführung entsprechender Projekte zu unterstützen. Ergänzt werden solche regionalen Digitalisierungszentren durch Einrichtungen wie die Servicestelle Digitalisierung (digiS) am Konrad-Zuse-Institut in Berlin, mit der 2012 erstmals in Deutschland eine spartenübergreifende Beratungsstelle für das Land Berlin geschaffen wurde.³⁶ Aufgabe dieser Stelle ist neben der Beratung und dem Wissenstransfer die Vernetzung und Koordinierung sowie die Erbringung bzw. Vermittlung von IT-Dienstleistungen für Digitalisierungsprojekte. Ähnlich angelegt ist die Koordinierungsstelle Brandenburg-digital.³⁷

Auch in der hinsichtlich Trägerschaft, Aufgaben und Ausstattung sehr differenzierten Archivlandschaft gibt es vergleichbare Aktivitäten: Allerdings steht hier zunächst die Digitalisierung von Bestandsübersichten und Findmitteln im Mittelpunkt, Archivgut selbst wurde bislang nur in wenigen Projekten digitalisiert. Allerdings soll sich diese Situation mit einem 2013 neu aufgelegten Förderschwerpunkt der DFG zur Digitalisierung archivalischer Quellen verändern. Das entsprechende Pilotprojekt bindet viele relevante Beteiligte ein und

³⁶ Vgl. <www.servicestelle-digitalisierung.de/> [23.04.2014].

³⁷ Vgl. <<http://www.brandenburg-digital.org/>> [23.04.2014].

wirkt auf eine koordinierte Vorgehensweise und den Einsatz standardisierter Verfahren hin.³⁸

Während sich also in den Archiven ein systematischer Einstieg in die Digitalisierung abzeichnet, ist die Situation in den Museen und auch in den anderen Sparten der DDB ungleich offener: Zwar existieren hier eine ganze Reihe nennenswerter Initiativen zur Digitalisierung, darunter auch kommerziell ausgerichtete Initiativen wie die Bildagentur bpk³⁹, das Vorgehen ist aber nicht einheitlich und angesichts des vorsichtig geschätzten Volumens von rund 450 Mio. Objekten an digitalisierungswürdigem Material erst am Beginn.⁴⁰ Eine wesentliche Schwierigkeit besteht darin, dass der Erschließungsstand für die entsprechenden Sammlungen häufig unbefriedigend ist. Nach allen Erfahrungen benötigen aber Digitalisierungsprojekte eine gute Erschließungsbasis und die Verfügbarkeit der entsprechenden Metadaten. Daher müssen vor entsprechenden Digitalisierungsprojekten in Archiven und Museen häufig die Voraussetzung durch Erstellung maschinenlesbarer qualitativ hochwertiger Metadaten erst hergestellt werden.

Die Frage der öffentlichen Finanzierung und Finanzierbarkeit großer Digitalisierungsinitiativen wurde unter dem Eindruck des Google Library Project, das 2004 gestartet rund 15 Mio. Bände in Bibliotheken digitalisiert, öffentlich insbesondere auf Ebene der EU diskutiert.

³⁸ Vgl. <<http://www.archivschule.de/DE/forschung/digitalisierung/>> [23.04.2014].

³⁹ Vgl. <http://bpkgate.picturemaxx.com/webgate_cms/> [23.04.2014].

Das Projekt selbst schien zu zeigen, dass die private Hand – entsprechende Bedingungen vorausgesetzt – den Einstieg in eine wirkliche Massendigitalisierung zu leisten imstande ist. In der Folge allerdings stieß die Google Initiative auf verschiedene Schwierigkeiten und das Modell der Public-Private-Partnership, das Google mit seinen Partnern entwickelte, hat sich weder in dieser Konstellation, noch mit anderen kommerziellen Partnern wesentlich ausweiten lassen. Entsprechende Verhandlungen der DDB mit Google haben bis heute nicht zu konkreten Ergebnissen geführt. Allerdings entstand im Angesicht des Google-Projekts auch ein wichtiger Impuls für Aktivitäten der Europäischen Union zur Gründung der Europeana, die mit zahlreichen Projekten zur Digitalisierung von Material verknüpft sind. Dieser Impuls verschafft auch eine rechtlich klarere Situation im Umgang mit vergriffenen und verwaisten Werken - zumindest für den Buchbereich ein wichtiger Schritt. Die dahinter stehende Idee, über die öffentliche Hand für einen breiten Zugang zur Nutzung digitalisierter gemeinfreier Materialien zu sorgen und Anreize zur Digitalisierung und Zugänglichkeit von urheberrechtlich geschütztem Material zu schaffen, wurde in einem Report einer von der EU-Kommission eingesetzten Expertengruppe formuliert. Bestehende Initiativen wie die Europeana sollen dabei verstetigt werden.⁴¹

⁴⁰ Schätzung der Stiftung Preußischer Kulturbesitz, Institut für Museumsforschung.

⁴¹ Maurice Lévy/Elisabeth Niggemann/Jacques de Decker, The new renaissance. Report of the "comité des sages" on bringing Europe's cultural heritage online, Brüssel 2011, URL: <http://ec.europa.eu/information_society/activities/digital_libraries/doc/refgroup/final_report_cds.pdf> [23.04.2014].

Im deutschen Bibliotheksbereich erscheint heute zumindest das Ziel, den Druckbestand in einem Zeitraum von 1800 bis 2020 zu digitalisieren, erreichbar.⁴² Die Mehrzahl der Bestände in deutschen Bibliotheken liegt aber danach, nimmt man den kaum abschätzbaren Bestand an grundsätzlich (resp. rechtlich) digitalisierbaren Material aus Archiven und Museen bzw. der Kulturdomäne insgesamt in den Blick, so erscheint nicht absehbar, dass eine private Finanzierung das gewaltige Investitionsvolumen aufbringen kann, dass für die Digitalisierung des nationalen Kulturgutes erforderlich ist. Angesichts des Volumens ist ein flankierendes übergreifendes Programm erforderlich, das von Bund und Ländern über einen langen Zeitraum hinweg finanziert wird und eine entsprechende Infrastruktur aufbauen hilft. Dabei darf nicht vergessen werden, dass die Digitalisierung erhebliche Folgekosten im Bereich der Bereitstellungs- und Speicherinfrastruktur nach sich zieht, die auf Dauer abgesichert werden muss. Perspektivisch werden hier die Aufwände für die Sicherung der Langzeitverfügbarkeit der digitalisierten Objekte deutlich ansteigen.

6. Digitalisierung - Perspektiven der DDB

Die bisher dokumentierten Überlegungen verdeutlichen die Bedeutung der Koordination in nationaler und internationaler Perspektive und eröffnen die Frage, welche Rolle die DDB hier spielen kann.

Bislang wurde in Zusammenhang mit der DDB vor allem durch ein Thesenpapier des Deutschen Bibliotheksverbandes (DBV) die

⁴² Thomas Bürger, Die Digitalisierung der kulturellen und wissenschaftlichen Überlieferung – Versuch einer Zwischenbilanz, in: Den Wandel gestalten – Informations-Infrastrukturen im digitalen Zeitalter, ZfBB 58, 3-4, 2011, S. 133-141.

Forderung nach einer nationalen Digitalisierungsstrategie artikuliert und vor allem mit entsprechenden Mittelbedarfsberechnungen unterlegt.⁴³ Auch an anderen Stellen wird die Förderung der Digitalisierung und der entsprechenden infrastrukturellen Voraussetzungen gefordert.⁴⁴

Bei der Umsetzung sollen inhaltliche Cluster helfen, die die prioritär zu digitalisierenden Korpora festlegen und redundante Aktivitäten vermeiden helfen. Die DDB soll dabei als organisatorischer Kern der nationalen Digitalisierungsinfrastruktur dienen und die Digitalisierung selbst im Wesentlichen in einem Netzwerk verteilter Einrichtungen erfolgen. Auch in der Politik wurde das Thema diskutiert, allerdings sieht der Bund über seine Beteiligung am Aufbau der DDB hinaus vor allem die Einrichtungen bzw. die Länder in der Pflicht, über Kooperationen Realisierungswege zu finden. Immerhin wurden über die Digitalisierungsoffensive des Beauftragten der Bundesregierung für Kultur und Medien 2013 fünf Millionen Euro aus dem Bundeshaushalt für die Digitalisierung von Büchern, Archivalien und Museumsgut zur Verfügung gestellt und über die DNB verteilt. Zu Grunde gelegt wurden für die Ergebnisse aus den verschiedenen über die Projektmittel

⁴³ Deutscher Bibliotheksverband: Deutschland braucht eine nationale Digitalisierungsstrategie!: Thesenpapier des Deutschen Bibliotheksverbands e.V. Berlin: 2011. URL: http://www.bibliotheksverband.de/fileadmin/user_upload/DBV/positionen/ThesenpapierDigitalisierung_dbv_Papier.pdf [23.04.2014].

⁴⁴ Kommission Zukunft der Informationsinfrastruktur: Gesamtkonzept für die Informationsstruktur in Deutschland (Stand: April 2011). URL: http://www.allianzinitiative.de/fileadmin/user_upload/KII_Gesamtkonzept.pdf [23.04.2014], S. 37ff. und S.55 ff. sowie in den verschiedenen Veröffentlichungen des Wissenschaftsrates der Jahre 2011 und 2012.

finanzierten Projekte die Standards, die die DDB für die Einspielung von Daten und Objekte als Voraussetzung definiert hat – als weiteren Weg zur Verfahrensstandardisierung.

Insgesamt aber existiert weder auf europäischer noch auf nationaler Ebene ein übergreifendes und langfristig finanziertes Programm zur Digitalisierung von Kulturgut in Bibliotheken, Archiven und Museen sowie den anderen Sparten der Kulturdomäne. Insofern besteht nach wie vor eine große Abhängigkeit von den Projektfördertöpfen der DFG und der weiteren Finanziere, die entsprechende Programme auflegen. Zu einer planmäßigen nationalen Digitalisierungsstrategie aber gehört ihre finanzielle Absicherung – für die Planung selbst, noch mehr aber für ihre Durchführung – beides aber ist derzeit nicht absehbar. Insofern sind in nächster Zeit eher kleine Schritte zu erwarten und es wird schwer, den wachsenden Erwartungen aus Wissenschaft und Forschung Rechnung zu tragen.

Dennoch versteht sich die DDB auch als Katalysator für Digitalisierungsvorhaben: Durch die Koordination verschiedener Initiativen und durch die Unterstützung von Standardisierungsprozessen für Metadaten, Abläufe und Qualitätskriterien sorgt sie für verlässliche Planungsdaten und eine erhöhte Transparenz sowie Sicherheit im Umgang mit Förderungen - insbesondere bei den Einrichtungen im Kompetenznetzwerk. Hierzu hat die DDB 2013 einen Workshop mit potentiellen Kooperationspartnern und bereits aktiven Einrichtungen durchgeführt, in dem der Bedarf für eine übergreifende Koordination der einzelnen Aktivitäten klar geworden ist. Deutlich wurde dabei auch, dass die Erarbeitung eines umfassenden Masterplans zur Digitalisierung des in deutschen Einrichtungen liegenden Kulturguts

schon angesichts des Volumens als wenig realistisch und wohl auch nicht zielführend angesehen wird. Angesichts der Feststellung, dass die Digitalisierungsaktivitäten in Deutschland trotz einiger kooperativer Projekte insbesondere überregional eher unkoordiniert ablaufen, wird eine aktivere Rolle der DDB gefordert, um für eine abgesicherte Finanzierung, allgemein akzeptierte technische Standards und eindeutige rechtliche Rahmenbedingungen zu sorgen. Die Entscheidung, welche Bestände in welcher Reihenfolge digitalisiert werden, sollte in den einzelnen Einrichtungen selbst fallen, entsprechende Planungen sollten vorgehalten und kontinuierlich fortgeschrieben werden. Eine koordinierende Rolle für die DDB sehen die Teilnehmer auch im Hinblick auf Werkzeuge zur Datenoptimierung und -anreicherung sowie deren Angebot auf einer technischen Plattform.

Die DDB wird diesen „Auftrag“ aufarbeiten und in die Planungsüberlegungen einbeziehen – sie selbst wird 2014 evaluiert, um die weiteren Perspektiven nach Auslauf der ersten Förderphase Ende 2015 abzustecken. Es bleibt zu hoffen, dass mit der Neubeauftragung für den Zeitraum ab 2016 die Digitalisierung als eigentliche Triebfeder für die Dienste der DDB – insbesondere für die Datenplattform und das Portal – einen angemessenen Eingang in die Programmplanung der DDB findet.

Scanner, Standards, Strategien – Digitalisierung im Digitalisierungszentrum der Niedersächsischen Staats- und Universitätsbibliothek Göttingen

von Martin Liebetruth

Die Deutsche Forschungsgemeinschaft (DFG) hat 1997 zwei Digitalisierungszentren in deutschen Bibliotheken mit einer Startfinanzierung für den Verlauf von vier Jahren ins Leben gerufen: Das Münchner Digitalisierungszentrum¹ der Bayerischen Staatsbibliothek sowie das Digitalisierungszentrum der Niedersächsischen Staats- und Universitätsbibliothek Göttingen (SUB)². Im Vorfeld hatten sich die technischen Möglichkeiten (spezielle Buchscanner mit gut auflösenden Kamerachips und leistungsfähige Arbeitsplatzrechner) so weit entwickelt, dass nun der Aufbau einer digitalen Bibliothek ins Auge gefasst werden konnte. Dabei ging es zunächst darum, wichtige und bedeutende Bestände in den Bibliotheken retrospektiv zu digitalisieren, also einzuscannen. Den Nutzerinnen und Nutzern sollte über das damals häufig noch mit Modems ins heimische Büro gelangende Internet der freie Zugang auf historische Titel der Bibliotheken ermöglicht werden.

Neben den beiden Zentren wurden in den Folgejahren an vielen Standorten in den Bibliotheken technische Möglichkeiten zur Digitalisierung eingerichtet, zum Teil wiederum durch finanzielle

¹ Vgl. <www.digitale-sammlungen.de> [16.08.2014].

² Vgl. <gdz.sub.uni-goettingen.de> [16.08.2014].

Unterstützung seitens der DFG. Das führte zu einer Reihe von Digitalisierungsprojekten, die Titel ins Internet brachten, die

- von besonderem Forschungsinteresse waren
- einen besonderen Bestand verfügbar machten
- in einem konservatorisch bedenklichen Zustand waren und durch die Digitalisierung überhaupt erst dem Publikum verfügbar gemacht wurden.

Schon bald aber – und spätestens seit den umfangreichen Tätigkeiten des Internet-Konzerns Google Inc. im Bereich der Buchdigitalisierung – wurde evident, dass diese Form der sogenannten „Boutiquen“-Digitalisierung³ für den genannten Bereich der Literatur ihre Berechtigung hat, dass aber im Zuge der technischen Weiterentwicklung eine Massendigitalisierung von großen Buchbeständen in greifbare Nähe gerückt und sowohl technisch, finanziell und auch organisatorisch als machbar angesehen werden konnte.

Die SUB Göttingen konnte inzwischen mehr als zwanzig Digitalisierungsprojekte erfolgreich einwerben und durchführen, u.a.

- Itineraria
- Nordamerica
- Wissenschaftsgeschichte
- Mathematica
- Zoologica

³ Ein Begriff, den Ronald Milne 2006 in der Diskussion um die Digitalisierung geprägt hat [R. Milne, The Google Mass Digitisation Project at Oxford. *Liber Quarterly*, Volume 16, Issue ¾, 2006.].

- Blumenbach Online
- VD 17
- VD 18.

Projektbeispiele

Auf zwei dieser Projekte sei im Folgenden im Detail eingegangen, da sie prototypisch die unterschiedlichen Projektarbeiten beleuchten.

Objekte der Digitalisierung im Göttinger Digitalisierungszentrum sind in der Regel gemeinfreie Druckwerke aus dem Bestand der SUB Göttingen. Diese können für ein Projekt nach unterschiedlichsten Kriterien zusammengestellt werden. Eines dieser Kriterien kann die inhaltliche Zugehörigkeit zu einem wissenschaftlichen Thema sein. Im Projekt EZOLO bzw. EZOLO2 war genau dies der Fall: Zoologische Literatur aus den Jahren 1550 bis 1820 stand im Fokus der Arbeiten.

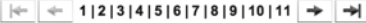
Kolleginnen und Kollegen des Johann-Friedrich-Blumenbach Instituts der Georg-August-Universität Göttingen trugen die Idee zu diesem Digitalisierungsprojekt an die Bibliothek heran. Für die Taxonomen ist es nicht immer leicht möglich, auf frühe wissenschaftliche und deshalb seltene Literatur zuzugreifen. Dies über das Internet zu ermöglichen, war eines der Ziele dieses Projektes. Ein zweites Ziel bestand darin, die in den Texten ausgewiesenen Tierarten für Taxonomen such- und nutzbar zu machen. Eine Arbeitsgruppe des Zoologischen Instituts machte dies möglich: Aus der historischen Literatur wurden Artnamen identifiziert, Schreibvarianten recherchiert und dokumentiert sowie Fundstellen in der Literatur aufgeführt. Die

Ergebnisse lassen sich in der frei zugänglichen Fachdatenbank animalbase.org⁴ nachvollziehen und stehen der wissenschaftlichen Öffentlichkeit zur Verfügung.

COLLECTION: ZOOLOGICA

Suchen in dieser Kollektion 

2181 Treffer

 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11  

	1 <i>Observationum circa vermes intestinales pars II.</i> Autor: Weigel, Christian Ehrenfried Monograph
	2 <i>Fragments of the natural history of Pennsylvania</i> Autor: Barton, Benjamin Smith Mehrbändiges Werk / Band
	3 <i>Icones rerum naturalium quas in itinere orientali depingi</i> Autor: Forsk^{al}, P. Monograph
	4 30 Autor: Russell, Alexander Mehrbändiges Werk / Band
	5 <i>Historia naturalis testaceorum Britanniae, or the British Conchology contenant les descriptions & autres particularités d'histoire naturelle des coquilles de la Grande Bretagne & de l'Irlande</i> Autor: Mendes da Costa, Emanuel Monograph
	6 <i>Das Faulthier. Der Ameisenfresser. Das Schuppenthier. Das Gürtelthier. Das Nashorn. Der Elefant. Das Wallroß</i> Autor: Schreber, Johann Christian Daniel

Abb. 1: Digitalisierte Titel des Projektes EZOOL.

⁴ Vgl. <www.animalbase.org> [16.08.2014] sowie Abb. 1.

Der Name Johann Friedrich Blumenbach als Namenspatron des Zoologischen Instituts ist oben bereits gefallen. Blumenbach (1752 – 1840) war ein bedeutender Naturforscher und ein wesentlicher Begründer der Zoologie als wissenschaftlicher Disziplin. Seine Grundlagenwerke erreichten weite Verbreitung und wurden vielfach übersetzt.

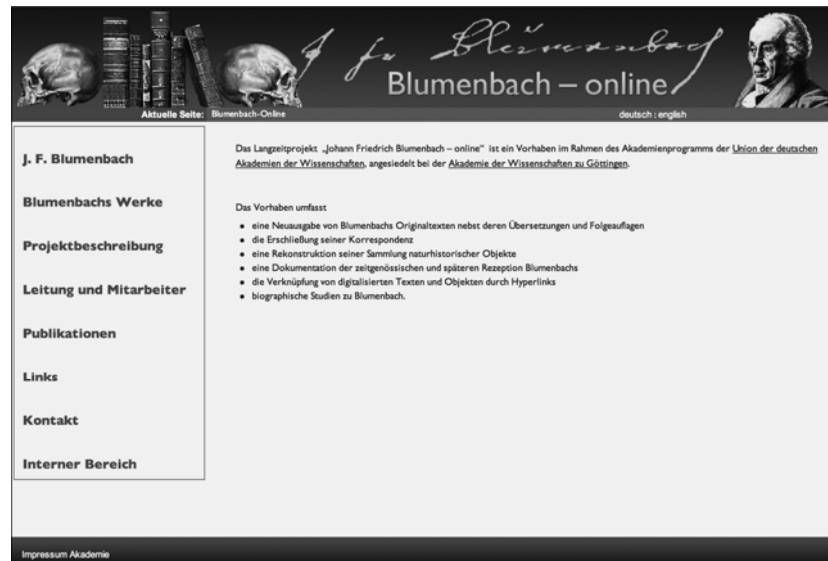


Abb. 2: Homepage des Projektes Blumenbach Online.

In dem Gemeinschaftsprojekt »Blumenbach online«,⁵ das die Akademie der Wissenschaften zu Göttingen, die Sammlungen der

⁵ Vgl. < www.blumenbach-online.de > [16.08.2014] sowie Abb. 2.

Georg-August Universität Göttingen und die SUB Göttingen gemeinsam durchführen, werden nicht nur die schriftlichen Zeugnisse Blumenbachs aufbereitet und ediert, hier werden auch die zu Lebzeiten des Namensgebers von ihm gesammelten und erforschten Objekte, die sich in den universitären Sammlungen der Georgia Augusta erhalten haben, in digitaler Form präsentiert und – wo sich inhaltliche Verweise ergeben – mit den Schriften Blumenbachs verknüpft.

Beiden knapp skizzierten Projekten gemein ist die Konzentration auf ein inhaltlich recht begrenztes Spektrum: die Werke eines bedeutenden Forschers und die wichtigsten zeitgenössischen Publikationen zu einem bestimmten Gebiet der Wissenschaft, hier der Zoologie.

Die DFG startete 2006 ein Programm zur Massendigitalisierung von Bibliotheksbeständen, in dem die deutschsprachige Literatur des 16., 17. und 18. Jahrhunderts jeweils in Jahrhundertblöcken möglichst umfassend digitalisiert werden soll. Kooperativ bearbeiten mehrere große Bibliotheken in Deutschland die Titel eines bestimmten Jahrhunderts. Für das Projekt VD18, das für die folgende Betrachtung der Massendigitalisierung als Beispiel angeführt sei, sind dies:

- Die Staatsbibliothek zu Berlin – Preußischer Kulturbesitz
- Die Sächsische Landesbibliothek, Staats- und Universitätsbibliothek Dresden
- Die Niedersächsische Staats- und Universitätsbibliothek Göttingen
- Die Universitäts- und Landesbibliothek Sachsen-Anhalt, Halle

- Die Bayerische Staatsbibliothek München.

Grundlage der Titelauswahl ist das retrospektiv erstellte „Verzeichnis der im deutschen Sprachraum erschienenen Drucke des 18. Jahrhunderts (VD18)“⁶, einer Nationalbibliographie, die ebenso für die anderen Jahrhunderte entsteht.

Standards in der Massendigitalisierung

Um die große Menge der zu digitalisierenden Materialien mit vertretbarem Aufwand und den dazu erforderlichen einzusetzenden personellen und finanziellen Mittel zu stemmen, ist es unbedingt sinnvoll und notwendig, den gesamten Produktionsprozess soweit als möglich zu standardisieren. Das beginnt bei der bibliographischen Erschließung und setzt sich fort bei der normierten Erfassung von Druckorten, Druckern und Verlegern. Personennamen werden normiert angesetzt und über Identifier mit der Gemeinsamen Normdatei (GND) der Deutschen Nationalbibliothek (DNB) verknüpft.

Die standardisierte Bearbeitung der Digitalisate geht weiter in der Präsentation der digitalisierten Werke. Hier wurde mit dem DFG-Viewer eine einheitliche Präsentationsform für alle von der DFG geförderten Projekte geschaffen.⁷

⁶ Vgl. <www.vd18.de> [16.08.2014].

⁷ Vgl. <www.dfg-viewer.de> [16.08.2014] sowie Abb. 3.

Die Säugthiere in Abbildungen nach der Natur mit Beschreibungen
< Das Faulthier. Der Ameisenfresser. Das Schuppenthier. Das Gürtelthier. Das Nashorn. Der Elefant. Das Wallroß
< Cover
TitlePage
TitlePage
Der Säugathiere zwote Abteilung.
Fünftes Geschlecht. Das Faulthier.
Sechstes Geschlecht. Der Ameisenfresser.
Siebentes Geschlecht. Das Schuppenthier.
Achtes Geschlecht. Das Gürtelthier.
Neuntes Geschlecht. Das Nashorn.
Zehntes Geschlecht. Der Elefant.
Elftes Geschlecht. Das Wallroß.
Folgendes ist im II. Theile zu verbessern und zuzusetzen.

Abb. 3: Ansicht eines elektronischen Inhaltsverzeichnisses im DFG-Viewer.

Im DFG-Viewer werden auch die sogenannten Strukturdaten mit angezeigt. Dabei handelt es sich um elektronische, maschinenlesbare Inhaltsverzeichnisse, die den im Buch abgedruckten

Inhaltsverzeichnissen entsprechen können. Bände ohne solche gedruckten Inhaltsverzeichnisse, erhalten über die Strukturdaten einen Mehrwert bei der Präsentation: Elemente des Druckwerkes werden benannt, die im Werk gedruckte Überschrift oder Bezeichnung abgetippt, so dass dann Überschriften von Kapiteln oder anderen relevanten Struktureinheiten auch direkt am Bildschirm angezeigt werden können. Damit gelingt eine schnellere Orientierung im digitalisierten Werk.

Workflow

An dieser Stelle sei ein Blick in die Arbeitsabläufe im Göttinger Digitalisierungszentrum gestattet, um die Arbeit in einem Massendigitalisierungsprojekt wie VD18 zu verdeutlichen.

Nachdem ein Titel anhand der VD18-Liste aus dem Magazin gehoben worden ist, wird er auf seinen konservatorischen Zustand geprüft. Diese Prüfung kann zum Ergebnis führen, dass ein Band sich in einem so desolaten physischen Zustand befindet, dass eine Digitalisierung zu größeren Schäden führen würde. In diesem Fall erfolgt keine Digitalisierung am Göttinger Exemplar, die Titeldaten werden an andere Projektbibliotheken weitergeleitet, um dort etwaig vorhandene Exemplare (und im besserem physischen Zustand) für die Digitalisierung zu übernehmen. Im Göttinger OPAC wird ein entsprechender Vermerk angebracht („aus konservatorischen Gründen nicht digitalisierbar“).

Bei positiver konservatorischer Prüfung des Titels wird festgestellt, bis zu welchem Winkel sich der Band ohne Schädigung öffnen

lässt. Möglich sind in der Bearbeitung Öffnungswinkel von 90 oder 180 Grad. Der ermittelte Öffnungswinkel wird vermerkt und im weiteren Workflow beachtet.

Für beide ermittelten Öffnungswinkel stehen dann verschiedene Buchscanner bzw. Aufnahmemöglichkeiten bereit. Die vorbereiteten und mit Laufzetteln versehenen Bände werden in den Produktionsbereich gebracht und entsprechend des ermittelten Öffnungswinkels aufgestellt. Die erfahrenen Scanoperatorinnen bedienen in der Regel ein Gerät, kennen sich also mit den Besonderheiten und Möglichkeiten der ausgewählten Station sehr gut aus. Der vorgesehene Band wird dann in das Aufnahmegerät gelegt und der Vorgang im System angelegt. Nun erfolgt die eigentliche Digitalisierungstätigkeit.

Ziel ist es, jede einzelne Druckseite später im Internet präsentieren und referenzieren zu können. Dabei erlaubt es die Software beim 180 Grad-Digitalisieren, zwei Buchseiten in einem ‚Schuss‘ aufzunehmen und automatisch einzeln abzuspeichern. Für die 90 Grad-Lösung gibt es vergleichbare Angebote auf dem Markt, in Göttingen wird jedoch ein System eingesetzt, das mittels einer handelsüblichen hochwertigen DSLR-Kamera nur einzelne Seiten aufnimmt. Diese werden dann später automatisch in der richtigen, dem Buch folgenden Reihenfolge zusammengebracht. Falttafeln, eingebundene Karten oder Stiche werden besonders sorgfältig behandelt und gegebenenfalls auf besonders geeigneten Geräten separat aufgenommen.

Hinsichtlich der technischen Parameter gelten die von der DFG empfohlenen Mindestanforderungen⁸:

- Mindestauflösung 300 ppi bei Farbdigitalisaten
- Höhere Auflösungen, sofern die Vorlage dies erforderlich macht (kleine Vorlagen, Stiche, Abbildungen etc.)
- Bildformat während der Aufnahme, Bearbeitung und für die Archivierung der Masterdateien ist unkomprimiertes TIFF, das als eines der robusteren Bilddateiformate angesehen werden kann und eine entsprechend weite Verbreitung hat.

Nach Abschluss der Aufnahmen durch die Scanoperatoren erfolgt – nach dem Vier-Augen-Prinzip – eine Qualitätssicherung durch eine zweite Person. Etwaige Korrekturwünsche werden durch den gleichen Scanoperator durchgeführt, der das Buch aufgenommen hat. Korrekturen können sich beispielsweise auf vergessene Seiten oder angeschnittene Bildelemente beziehen.

Erst nach erfolgreicher und bestätigter Abnahme des digitalisierten Bandes erfolgt die weitere Bearbeitung der erzeugten digitalen Bilder. Diese werden nun halbautomatisch für die Präsentation im Netz vorbereitet. Dazu wird zunächst eine digitale Kopie des Bilddatensatzes erstellt. Die digitalen ‚Originale‘ bleiben in der weiteren Bearbeitung unangetastet und bilden den archivierfähigen „Digitalen Master“.

⁸ DFG-Praxisregeln „Digitalisierung“, Fassung 02/13, vgl. <http://www.dfg.de/formulare/12_151/12_151_de.pdf> [16.08.2014].

Mit der Vorbereitung der kopierten Bildsätze für die Bildschirmpräsentation ist der Bereich des Scannens abgeschlossen. Im weiteren Verlauf werden die Titel nun tiefererschlossen: mit Hilfe eines Metadateneditors (vgl. Abb. 4) wird die im Band vorhandene Paginierung in die Strukturdaten geschrieben. Dabei kann zwischen arabisch, römisch oder nicht paginierten Seiten unterschieden werden. Bildnummer (physische Zählung) und Seitennummer (logische Zählung) stimmen in der Regel nicht überein. Erst durch den Eintrag im System wird es dem Nutzer später möglich sein, auch nach aufgedruckten Seitennummern zu suchen.

Paginierung | Strukturdaten | **Metadaten** Bild au

Metadaten bearbeiten

Personen

Vorname:

Nachname:  

Rolle: 

Metadaten

Alle Autoren u. Beteiligte (orig.)	Schreber, Johann Christian Daniel  
Digitale Kollektion	Zoologica  
Erscheinungsjahr Band	1775  
Erscheinungsort	Erlangen  
HauptTitel	Die Säugthiere in Abbildungen nach der Natur mit Beschreibungen  
Haupttitel (Sortierung)	Säugthiere in Abbildungen nach der Natur mit Beschreibungen  

Abb. 4: Metadateneditor in Goobi.

Schließlich erfolgt die eigentliche Tiefenerschließung: Die einem Titel eigenen Strukturelemente (z.B. Titel, Kapitel, Abbildungen) werden eingetragen, und die vorhandenen Überschriften in das System transkribiert. Aus dieser Arbeit resultiert dann eine METS/MODS-Datei, die die Grundlage für die Präsentation und Rechercheoption des digitalisierten Titels bildet.

Die Säugthiere in Abbildungen nach der Natur mit Beschreibungen	
Johann Christian Daniel Schreber;	
Das Faulthier. Der Ameisenfresser. Das Schuppenthier. Das Gürtelthier. Das Nashorn. Der Elefant. Das Wallroß	uncounted
Schreber, Johann Christian Daniel	
Einband	uncounted
Titelseite	uncounted
Titelseite	uncounted
Der Säugathiere zwote Abteilung.	uncounted
Fünftes Geschlecht. Das Faulthier.	196
Sechstes Geschlecht. Der Ameisenfresser.	202
Siebentes Geschlecht. Das Schuppenthier.	208
Achstes Geschlecht. Das Gürtelthier.	213
Neuntes Geschlecht. Das Nashorn.	228
Zehntes Geschlecht. Der Elefant.	241
Elftes Geschlecht. Das Wallroß.	259
Folgendes ist im II. Theile zu verbessern und zuzusetzen.	278
Verzeichniß der Kupferplatten des zweeten Theils...	279

Abb. 5: Elektronisches Inhaltsverzeichnis in der GDZ-Präsentation.

Nachdem Bilddatensatz und begleitende METS/MODS-Datei in das Dokumentensystem eingespielt sind, steht der Titel allgemein zur Verfügung, die Metadaten können durchsucht werden und bieten durch die tiefere Erschließung einen höheren Nutzwert für die Forschung (vgl. Abb. 5).

Softwareunterstützung im Workflow

Wie oben dargestellt, sind die Arbeitsabläufe im Göttinger Digitalisierungszentrum sehr granular eingerichtet. Zugleich muss in den Massendigitalisierungsprojekten eine große Anzahl von Titeln in einem definierten Zeitraum mit hohen Qualitätsansprüchen bearbeitet werden. Deswegen wurde initial an der SUB Göttingen eine Software-Lösung entwickelt, die den gesamten Workflow der Digitalisierung

abbildet, kontrolliert und steuert. In der Open-Source-Softwaresuite GOOBI⁹, die inzwischen federführend von der SLUB Dresden betreut und deren Entwicklung in einem gemeinnützigen Verein (Goobi e.V.) vorangetrieben wird, ist eine solche Lösung gefunden worden. Inzwischen wird das System in vielen deutschen und internationalen Bibliotheken eingesetzt, und auch Archive sowie andere Gedächtnisinstitutionen nutzen GOOBI als System.¹⁰

Nr.	Titel	Status
1	Bibliographische Aufnahme	☒
2	Scannen	☒
3	Qualitätskontrolle	☒
4	Automatischer Kopierschritt	☒
5	Imaging	☒
6	Struktur- und Metadaten	☒
7	DMS Import	☒
8	Archivierung	☐

Abb. 6: Beispielmodule für die Abbildung von Workflowschritten in Goobi.

Da GOOBI modular aufgebaut ist, können die unterschiedlichsten Anforderungen, die einzelne Projekte für die erfolgreiche Durchführung mitbringen, in der Software aufgefangen und umgesetzt werden.

⁹ Vgl. <www.goobi.org> [16.08.2014] sowie Abb. 6.

¹⁰ Vgl. die Liste der unterschiedlichsten Materialien sowie Institutionen, wie sie unter <www.goobi.org/referenzen> [16.08.2014] ausführlicher dargestellt werden.

Zentrales Verzeichnis Digitalisierter Drucke

Da die in den VD-Projekten erschlossenen Strukturdaten in aller Regel nicht in den allgemeinen Online-Katalogen der Bibliotheken indiziert werden, wurde mit dem Zentralen Verzeichnis Digitalisierter Drucke (zvdd)¹¹ ein einheitliches Rechercheinstrument geschaffen – vergleichbar mit dem DFG-Viewer für die Präsentation von digitalisierten Werken.

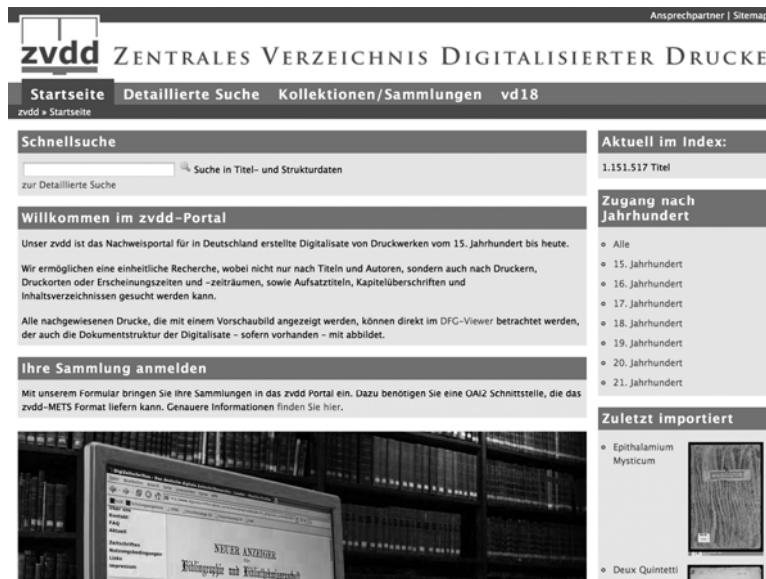


Abb. 7: Homepage des zvdd.

¹¹ Vgl. <www.zvdd.de> [16.08.2014] sowie Abb. 7.

Im zvdd werden die, in den von der DFG geförderten Digitalisierungsprojekten, erzeugten Digitalisate aggregiert. Außerdem werden die erzeugten Struktur- und Metadaten aus den METS/MODS-Dateien indiziert und ermöglichen damit an zentraler Stelle die Recherche über inhaltlich tief erschlossene Retrodigitalisate. Eine zusätzliche Facetierung bei den Suchergebnissen erleichtert das Auffinden des eigentlich Gesuchten.

Eine zentrale Recherchemöglichkeit über alle in Deutschland retrodigitalisierten Drucke war lange Zeit ein ersehntes Desiderat im deutschen Digitalisierungswesen. Es ist wünschenswert, dass dies für andere Dokumentenformen und Sammlungskontexte ebenso realisiert wird. Die Deutsche Digitale Bibliothek der Deutschen Nationalbibliothek kann hier nur eine erste Stufe sein.

Inhalterschließung und Volltexte

Werden bislang in deutschen Gedächtnisinstitutionen schon über eine Million Druckwerke online angeboten¹², so steht eine komplette Volltextfassung der digitalisierten Bände noch aus. Dies gilt vor allem für Druckwerke, die bis zur Mitte des 19. Jahrhunderts erschienen sind.¹³ Die komplexen Probleme bei der softwaregesteuerten Volltexterkennung von früheren Drucken haben noch recht hohe

¹² Vgl. Statistik <zvdd.de> [16.08.2014].

¹³ Für von der DFG geförderte Digitalisierungsprojekte, die Druckwerke ab Erscheinungsjahr 1850 in den Fokus nehmen, ist die Erstellung von Volltexten seit der neuesten Fassung der „Praxisregeln“ verpflichtend. (DFG-Praxisregeln 02/13, S. 30).

Fehlerquoten zur Folge. Die fehlerbehafteten, maschinell erkannten Texte können zwar mittels Suchalgorithmen, die eine Toleranz gegenüber falsch erkannten Wortvarianten besitzen, für eine Volltextrecherche hinzugezogen werden. Eine wissenschaftliche, zitierfähige und dem ursprünglich gedruckten Text zu 100 Prozent entsprechende digitale Textform kann von den derzeit erhältlichen Softwarelösungen noch nicht erwartet werden. Hier werden zukünftige Versionen und Lösungen weitere Fortschritte bringen.

Einzelne Projekte, die eine erhöhte Textrichtigkeit in den Fokus nehmen, betreiben über verschiedene Verfahren (sogenanntes Double-Keying oder eine Kombination aus maschineller Erkennung und menschlicher Korrektur) einen großen Aufwand, der für digitale Editionen inzwischen Standard geworden ist. Volltexte werden dort zusätzlich in XML/TEI ausgezeichnet, im Text erwähnte Personen, Orte oder andere Entitäten werden mit Normdatenverzeichnissen verknüpft und werden so selbst zu Teilen des sogenannten Semantic Web.

Bibliotheken und andere Einrichtungen der Digital Humanities haben die Anforderungen der Wissenschaftler an solche digitalen Editionen aufgenommen und bieten inzwischen mit Virtuellen Forschungsumgebungen wie z.B. Textgrid¹⁴ Lösungen an, die den gesamten Workflow von der Retrodigitalisierung bis zur digitalen oder gedruckten Publikation abdecken.

¹⁴ Vgl. <www.textgrid.de> [16.08.2014].

Ein Blick von den heute schon möglichen Ergebnissen zeigt auf, dass die Berücksichtigung verbreiteter Standards in den einzelnen Schritten von der Digitalisierung bis zum Datenaustausch, die zur Präsentation, Visualisierung und Arbeit mit digitalisierten Kulturgütern führen, unabdingbar ist, um heutigen und zukünftigen Forschern Quellen und andere Objekte wissenschaftlichen Interesses zur Verfügung zu stellen.

Mit langem Atem zu gutem Ende

– Erfahrungen bei „museum-digital“ –

von Stefan Rohde-Enslin

Die „Digitale Revolution“ ist allumfassend. Sie ändert unsere Umwelt, unsere Werkzeuge, unseren Blick auf Sachverhalte und Dinge, die Dinge selbst, schließlich uns und auch das, was von uns bleibt. Es ist also selbstverständlich, dass auch Museen von diesem Umbruch betroffen sind. Sie sind es in vielfältiger Weise.

Was im Folgenden vorgestellt werden soll nennt sich „museum-digital“ und es gehört in dieses Themenfeld: „Museum und Digitalisierung“ - aktuell aber nur in einen Teilbereich davon. Bei „museum-digital“ werden Dinge wie digitale Präsentationstechnik im Museum, die Digitalisierung der Verwaltung der Einrichtung, digitale Medien zur Ausstellungsvermarktung zunächst nicht behandelt. Vielmehr geht es erstmal um das, was letztlich ein Museum ausmacht: Museumsobjekte und deren Digitalisierung.

Das Thema, mit dem sich „museum-digital“ beschäftigt ist also relativ leicht zu umreißen. Schwieriger ist es, einen passenden Begriff für die strukturelle Form von „museum-digital“ zu benennen. Es handelt sich nicht um ein „Projekt“, denn ein „Projekt“ hat einen definierten Anfang und ein festgeschriebenes Ende, hat Träger und Bearbeiter. Es ist auch keine „Initiative“, auch wenn dieser Begriff schon eher den Kern der Sache beschreibt - aber eine „Initiative“ ist auf ein bestimmtes Ziel hin ausgerichtet und trägt das Erreichen des Ziels als Ende ihrer Existenz bereits in ihrer Konzeption.

„museum-digital“ ist eher ein „Vorhaben“, es ist mehr Tat und weniger Struktur. Es ist eine Absicht, ein Wollen, ein Wirken, ein Werden – und genau deshalb gibt es keine fest gefügte, starre Struktur – die Dinge sind im Fluss, was große Vor- aber auch große Nachteile hat, dazu später mehr. „museum-digital“ passt damit hervorragend in die Internet-Zeit – auch das Internet selbst wächst und lebt ohne „Masterplan“. In gewisser Hinsicht: Manifestationen des kantschen „Sapere Aude“.

Am leichtesten wird verständlich, was und wie „museum-digital“ ist, wenn man den Blick auf seine Entstehung und die Entwicklung der letzten Jahre lenkt. Dabei zeigt sich Einiges, das spezifisch für Museen ist, in den meisten Aspekten aber unterscheiden sich Museen nur marginal von anderen Kulturerbe bewahrenden Einrichtungen. Die folgende Schilderung sollte Erfahrungen und Gedanken enthalten, die in den Diskurs der Digitalisierung der Archive passen – und sei es nur als Anregung.

Der Startschuss für „museum-digital“ fiel vor ziemlich genau fünf Jahren im Kreismuseum in Bitterfeld. Schon im Vorfeld hatte der Museumsverband Sachsen-Anhalt das Thema „Digitalisierung“ in mehreren Veranstaltungen für seine Mitglieder thematisiert. Woraufhin einige Museen dort vorstellig wurden und vorschlugen eine „Arbeitsgruppe Digitalisierung“ im Museumsverband zu gründen. Die Geschäftsführung des Museumsverbandes kam dem sehr gerne nach und besorgte sich fachlichen Beistand vom Institut für Museumsforschung der Staatlichen Museen zu Berlin (Stiftung Preußischer Kulturbesitz).

Auf diesem ersten Treffen, Ende Januar 2009, wurde ziemlich schnell deutlich, dass zwar alle Anwesenden etwas zu Digitalisierung in der neuen AG besprechen wollten, dass aber dieses „Etwas“ für jeden etwas Anderes umfasste. Die einen wollten bestehende Datenbanken durch neuere ersetzen, andere wollten sich erst eine Datenbank anschaffen und fragten welche, wieder andere waren besonders am Thema „Digitale Objektfotografie“ interessiert.

Am Ende dieses Treffens fiel der Beschluss sich zunächst einmal der Problematik digitaler Publikation von Museumsobjekt-Informationen anzunehmen. Andere Themen sollen folgen. Die Ausgangsfrage der Arbeit der AG Digitalisierung lautete: „Wie muss ein System gestaltet sein, welches ALLEN Museen AUF EINFACHE WEISE ermöglicht GEMEINSAM ihre Objekte im Internet zu präsentieren und diese an Portale aller Art weiter zu reichen?“ Das, was online unter www.museum-digital.de verfügbar ist, ist der jeweils aktuelle Stand der Beantwortung dieser Frage.

Aber der Reihe nach.

Die Lage erfassen

Der erste Schritt bestand darin zu schauen, was es bereits gab. Am Anfang der konkreten Arbeit stand eine Betrachtung dessen, was im Internet von Museen über ihre Objekte veröffentlicht wird und wie dieses geschieht. Webseiten und Datenbanken weltweit wurden besucht und analysiert. Im Ergebnis zeigte sich, dass nur wenige mehr als „Kerninformationen“ pro Objekt veröffentlichten – plus (wenigstens) einer Abbildung. Die dargebotenen Informationen umfassten zumeist genau das, was auch Interpol, Unesco und ICOM (Internationaler Museumsrat) zur Bekämpfung illegalen Kunsthandels,

allen Museen empfehlen zu ihren Objekten schriftlich und bildlich festzuhalten:

- Art des Gegenstandes
- Material und Techniken
- Maße
- Inschriften und Markierungen
- Unterscheidende Merkmale
- Titel
- Objekt
- Datum oder Periode
- Hersteller
- Beschreibung¹

Wenige allgemeine Informationspartikel pro Objekt, dazu ein Bild: Keine gehobenen Ansprüche - es dürfte ohne viel Aufwand möglich sein, eine einfache Software zu schreiben und Objekte ins Netz zu stellen. Wirklich? Wie leicht oder schwer sind diese wenigen Angaben in den verschiedenen Museen zu erhalten?

„Gib mir Fünf“

Der zweite Schritt bestand in einer kleinen „Forschung“. Jedes in der AG Digitalisierung vertretene Museum sollte zu den ermittelten Kerninformationen möglichst fünf verschiedene seiner Objekte

¹ ObjektID, <<http://archives.icom.museum/object-id/checklist/german.pdf>> [12.04.2014].

zusammentragen und dazu mindestens je ein Bild erstellen. Es sollte aber auch die Zeit messen, die dafür pro Objekt nötig war.

Gleichzeitig wurde beschlossen, dass wir, sollten wir online gehen, besonderes Augenmerk auf (Teil-)Sammlungen legen würden. Wem es möglich war, der sollte die Gesamtheit seiner Objekte in Gruppen (Teilsammlungen) einteilen und zu jeder Teilsammlung Informationen für fünf verschiedene, die jeweilige Teilsammlung charakterisierende Objekte samt Foto und Zeitmessung ermitteln. Die verfolgte Idee war, dass ja keineswegs alles auf einen Schlag digitalisiert oder digital publiziert werden muss. Bei einer Struktur „Museum“ – „Teilsammlungen“ – „Objekt“ würde durch die Publikation charakteristischer Objekte mit passenden Sammlungsbeschreibungen recht schnell ein „Gesamtbild“ des Museums, seiner Zielsetzung und seiner Objekte entstehen. „Fünf Objekte-pro-Sammlung“ war erben, ernsthaft angestrebt wurde „Fünf Objekte-pro-Museum“.

Der Museumsverband Sachsen-Anhalt lud auch Museen außerhalb der AG Digitalisierung ein, sich an dieser kleinen Recherche zu beteiligen. Die Beteiligung war recht groß, knapp 50 Museen machten mit. Auch der Museumsverband Rheinland-Pfalz fand die Untersuchung interessant genug um eine parallele Umfrage zu starten. Dies war gut, konnten so doch Unterschiede in den Erschließungstraditionen zwischen Ost- und West aufgedeckt werden.

Das Ergebnis war jedoch in Ost und West das Gleiche: Es dauerte zwischen 3 Minuten und 3 Stunden die entsprechenden wenigen Informationen zusammenzutragen. Große und kleine Museen unterschieden sich darin nicht und auch die Ausrichtung (Kunstgeschichte, Naturgeschichte, Technikgeschichte) spielte keine Rolle.

Es war leicht zu erkennen, das Nadelöhr beim Vorhaben „digital Publizieren“ würde nicht das Online stellen, sondern das Erzeugen publizierbarer Informationen sein.

„Praktische Versuche“

Die Informationen und Abbildungen zu den Objekten wurden per E-Mail zusammengetragen und auf dieser Grundlage wurde eine Datenbank entwickelt. Diese Datenbank war an das damals übliche Datenaustausch-Format für Museen namens „museumdat“ angelehnt und sollte späteres Exportieren vereinfachen. Aus „museumdat“ ist mittlerweile „LIDO“ geworden – im Kern ist es aber geblieben, wie es war. Auch heute ist es leicht international nutzbare (da im gängigen Standard verfügbare) Daten aus der museum-digital-Datenbank zu erhalten, die Datenbank selbst ist gemäß dem Exportformat gestaltet.

Die Datenbank selbst war dann Ausgangspunkt für eine Internet-ausspielung. Ein erster Entwurf für eine Darstellung wurde erstellt und wieder wurden alle beteiligten Museumsmitarbeiter gebeten das Design und die Gestaltung zu kommentieren. Sie sollten aber auch Museumsbesucher, eigene Bekannte, Nachbarn, Väter, Mütter, Kinder, zu Kommentaren einzuladen. Einige Anpassungen, vor allem in Schrift- und Bildgröße, waren nötig. Am Ende der Einarbeitung der Kommentare stand ein Basisdesign, das bis heute Bestand hat.

Die erste Version einer Online-Datenbank, zunächst für Sachsen-Anhalt, wenig später dann auch für Rheinland-Pfalz, konnte ins Netz gehen und damit öffentlich sichtbar werden.

Redakteure unterwegs

Dass etwas Vorzeigbares entstanden war, hatte einen großen Vorteil, man konnte den Entscheidungsträgern im Land anschaulich demonstrieren, wohin die Reise gehen sollte. Tatsächlich kam von dort eine finanzielle Unterstützung in Höhe von knapp 20.000 Euro. Mit diesem Geld wurden „Daten-Redakteure“ angestellt, welche in verschiedene bisher nicht involvierte Museen gingen, um dort Informationen zu einigen Objekten zusammenzutragen und aufzubereiten.

Auch die Redakteure hatten den Auftrag zu ermitteln, wie viel Zeit und Aufwand in den einzelnen Museen nötig war, die jeweiligen Informationen zusammenzutragen. In diesen Museen zeigte sich das gleiche Bild: Entweder es gab gleich mehrere Datenbanken (meistens pro Sammlungsbereich eine) oder – häufiger der Fall – es gab überhaupt keine Datenbank und man arbeitete noch mit Karteikarten, Listen oder Eingangsbüchern. Objektfotos waren entweder nicht oder nur in Schwarzweiß als Arbeitsfotos vorhanden, die gerade noch zur Identifizierung der Objekte, nicht aber zur Publikation tauglich waren. Objektbeschreibungen waren entweder nicht vorhanden oder bei Vorhandensein einer Datenbank, von Praktikanten eingetippt worden. Teilweise war dies vor vielen Jahren geschehen und es war bisher keine Zeit gewesen das Eingetippte auf Fehler hin zu überprüfen. Entsprechend viele Tippfehler waren dort zu finden. Die Beschreibungen selbst waren zum Teil 40 Jahre alt und auch inhaltlich möglicherweise nicht mehr im Einklang mit dem heutigen Stand der wissenschaftlichen Kenntnis. Die Redakteure hatten entsprechend viel zu tun. Insgesamt bestätigte sich das Bild aus der „Gib mir Fünf“-Initiative: Es dauert drei Minuten bis drei Stunden um

für ein einzelnes Objekt alle notwendigen Informationen (inklusive publizierbarer Abbildung) zu erhalten.

Direktes Eingeben wird möglich

Auch die überarbeiteten Datensätze der Redakteure wurden per E-Mail an den Eingebenden übermittelt. Der hatte alle Hände voll zu tun – zu viel für ein Ehrenamt. Es wurde also umgehend ein Eingabe- und Bearbeitungsmodul für die Datenbank geschaffen.

Auch in Rheinland-Pfalz hatte „museum-digital“ Fuß gefasst. Dort wurden Fördermittel beim Land beantragt und in Höhe von 5.000 Euro bewilligt. Mit diesem Geld wurde das Eingabe- und Bearbeitungsmodul weiter ausgebaut und eine stabile Nutzerverwaltung etabliert. Fortan war und ist es möglich dass ein Museum selbst neue Nutzer (weitere „Museumsdirektoren“ oder „Eingebende“) anlegt. Die Kontrolle über die Objektinformationen wurde also vollständig an die Museen zurückgegeben, die jederzeit alles ändern oder auch löschen können. Jede Änderung kommt sofort im Internet zur Anzeige.

„Frag den Fachmann“

Die nächste Neuerung in „museum-digital“ beinhaltete eine Erweiterung des Eingabe- und Bearbeitungswerkzeuges um eine „Frag den Fachmann“-Funktion. In vielen Museen ist es so, dass verschiedene Objektgruppen vorhanden sind, die Zahl und Ausbildung der Mitarbeiter dem aber nicht entspricht. Beispielsweise gibt es nicht wenige Regional- und Stadtmuseen, die über kleinere Ethnologische Sammlungen verfügen, weil jemand aus dieser Ortschaft in deutsch-kolonialen Diensten gestanden hatte und seine Sammlungen vor

langer Zeit ans Museum gegangen sind. Das gleiche Museum kann auch über eine Gemäldesammlung von lokaler oder gar internationaler Bedeutung verfügen. Ist dieses Museum, wie es oft der Fall ist, von einem Kunsthistoriker geleitet, so liegt es nahe, dass die Erforschung und Erschließung der ethnologischen Objekte ins Hintertreffen gerät, und andersherum, wenn ein solches Museum von einem Ethnologen geleitet wird, dann steht nicht unbedingt die Mal- und Zeichenkunst im Vordergrund der Erschließungsaktivitäten. Die Idee, welche in der AG Digitalisierung geboren wurde, war zugleich genial und einfach: Wer immer sich berufen fühlte nannte sein Thema und seine E-Mail-Adresse und erklärte sich bereit Kollegen bei der Erschließung zu helfen. Im Eingabe- und Bearbeitungswerkzeug ist die entstehende Liste jederzeit abrufbar. Das digitale Arbeiten führt auf diese Weise zu neuen Kooperations- und Unterstützungsmöglichkeiten, denn vorher hatten vielfach die Kollegen nicht voneinander gewusst, wer Ethnologe, Archäologe oder Entomologe war.

„Frag den Nutzer“

So, wie das Internet die Zusammenarbeit und Kommunikation unter Museumskollegen erleichtert, so erleichtert es ebenfalls die Kommunikation zwischen Museum und Internetpublikum. Bei „museum-digital“ wurde ein Feedback-Button eingeführt, der auf jeder Objektseite erscheinen kann. Dort heißt es „Ist Ihnen etwas aufgefallen? Womöglich ein Fehler? Oder wissen Sie vielleicht mehr?“

Etliche Kollegen hatten ein Unbehagen mit dieser Präsentation. Steht da doch eigentlich, dass die Auskunft des Museums (die Objektseite also) Fehler enthalten könnte und dass jemand aus dem Publikum mehr wissen könnte als das Museum selbst. Ein solcher

Text verlangt ein Selbstverständnis als offene und das Publikum einbeziehende Einrichtung. Es ist etwas grundsätzlich anderes, ob ein Museum eine Ausstellung erstellt und dem Publikum die Welt erläutert oder ob das gleiche Museum dem Publikum zugesteht, dem Museum seine Welt zu erläutern. Der Text für diesen Feedback-Button war vom Stadtgeschichtlichen Museum in Leipzig entlehnt, in Englisch fand er sich aber auch auf den Seiten des British Museum.

Allem Unbehagen zum Trotz - und in der Folge großer Vorbilder - wurde der Feedback-Button eingeführt und jedem Museum zugleich die Möglichkeit eingeräumt ihn zu aktivieren oder zu deaktivieren. Die Kollegen haben den Button nicht deaktiviert! Im Gegenteil: Die zahlreichen eingehenden Kommentare haben zwar, wie befürchtet, zu Kaufangeboten (vom Typ: „Schreibmaschine auf Dachboden gefunden, was gibt man mir dafür?“) geführt, in vielen Fällen aber haben sich Universitätsprofessoren, Sammler oder andere Spezialisten aus aller Welt gemeldet und dem jeweiligen Museum ihr Hintergrundwissen zur Verfügung gestellt. Mittlerweile gibt es bei den an „museum-digital“ beteiligten Museen keinerlei Zweifel mehr daran, dass dieses ein sehr hilfreicher Button ist, der nicht nur die Kommunikation zwischen Publikum und Museum verbessert, sondern auch der Objekterschließung und -dokumentation hilft.

Auf dem Weg in die Europeana

Die Idee Objektdaten mit „museum-digital“ ins Netz zu stellen, um sie auch in Portalen wie der damals noch jungen Plattform europeana zu präsentieren war schon seit den ersten Tagen von „museum-digital“ präsent. Über das Institut für Museumsforschung (SMB-PK)

kam es zu einer engen Kooperation mit dem EU-Projekt Athena, dessen Absicht es war Museumsobjekt-Informationen aus ganz Europa zusammenzutragen und in der Europeana verfügbar zu machen. Das Athena-Projekt ist im Europeana-Kontext bis heute wegweisend, denn zwei Dinge wurden damals geschaffen, mit denen auch viele heutige Europeana-bezogene Projekte arbeiten: Zum einen wurde das LIDO-Austauschformat (korrekter: Harvesting Format) für Objektinformationen geschaffen, zum anderen entwickelten die Projektpartner der National Technical University Athens ein Software-Werkzeug mit dem es vergleichsweise einfach ist Daten aus einem Metadatenformat in ein anderes zu überführen („mappen“), den Daten also eine Form zu geben, in welcher sie sehr leicht in die Europeana integrierbar sind.

Bei der Entwicklung dieses Werkzeugs spielte „museum-digital“, das damals noch im LIDO-Vorgänger-Format „museumsdat“ exportierte, eine wichtige Rolle. Die Objektdaten aus „museum-digital“ waren das Testmaterial mit dem die griechischen Kollegen ihr Werkzeug entwickelten, welches heute auch von Europeana selbst zur Datentransformation eingesetzt wird.

Auch publizistisch wurde dieser Schritt bekannt. In der Mitteldeutschen Zeitung erschien ein Artikel unter dem Titel „Fahnen aus Anhalt flattern als Testversion in Griechenland.“² Später waren dann die Objektinformationen der beteiligten Museen tatsächlich in der Europeana zu finden. Die meisten der Kollegen waren allerdings

² Vgl. <<http://www.mz-web.de/bernburg/fahnen-aus-anhalt-flattern-als-testversion-in-griechenland,20640898,17930174.html>> [12.04.2014].

enttäuscht über die Form, in welcher die Informationen dort gezeigt wurden. Vor allem löste die Tatsache, dass dort zwar die Aggregationsprojekte (hier also „Athena“) und die Sub-Aggregatoren (also „museum-digital“) prominent genannt werden, das jeweilige Museum, also der Platz, an dem die Arbeit der Objekterfassung und –beschreibung geleistet wurde, nur sehr versteckt zur Anzeige kam. Bei „museum-digital“ zielt alles darauf ab, mit den Objektinformationen immer auch die Kenntnis zu verbreiten, dass es das entsprechende Museum gibt – bei Europeana schien/scheint etwas anderes im Vordergrund zu stehen. Vielleicht rührte ein Teil der Enttäuschung auch daher, dass zwischen Datenweitergabe und Veröffentlichung in der Europeana ein ganzes Jahr verstrichen war!

Es wurde ein neuer Export aus „museum-digital“ erstellt, der an vielen Stellen den Namen des Museums weitertrug. Die Objekt-Präsentation in der Europeana verbesserte sich daraufhin etwas. Auch dieses „Update“ kam nicht von heute auf morgen. Bis die neuen Informationen die alten ersetzt hatten, war mehr als ein halbes Jahr verstrichen.

Später dann, als die Europeana ihre Einstellung zu Publikationsrechten dahingehend festschrieb, dass nur noch Objekte unter CC0-Lizenz (vom Typ: „Jeder darf alles mit den Objekten machen – inklusive Geld verdienen“) akzeptiert werden, weigerten sich einige der Museen weiterhin daran teilzunehmen. Unerwartet war die Frage: Wer, in einem städtischen Museum, darf eigentlich einer solcher Lizenz zustimmen? Der Direktor des Museums? Der Kulturausschuss der Stadt? Vielen Kollegen, die ja nicht mehr viel von der Europeana

erwarteten, war die aufwändige Suche nach einer Antwort zu mühselig.

Importieren wird möglich

Einiges war ja schon unternommen worden, um die Objektinformationen in die „museum-digital“-Datenbank hinein zu bekommen. Der Schritt von zentraler Eingabe und Übermittlung per E-Mail zur Eingabe in Kontrolle des Museums war vollzogen. Es zeigte sich aber, dass es Museen gab, die durchaus auch in ihrer Inventarisierungsdatenbank Informationen in einer Form zur Verfügung hatten, die diese mit wenig Überarbeitung publizierbar machen würden. Es lag also nahe ein Import-Werkzeug zu entwickeln. Da beim Datenimport sehr vieles schief gehen kann, wird der Import zentral ausgeführt. Anfänglich konnte das Importwerkzeug nur CSV-Dateien importieren, aber es wurde ständig weiterentwickelt und ist mittlerweile in der Lage mehr oder weniger direkt aus allen gängigen Museumsdokumentations-Systemen zu importieren.

Viele Kollegen, die diesen Weg, Export aus hausinterner Datenbank – Publikations-Aufbereitung in „museum-digital“, gehen wollten, mussten die Erfahrung machen, dass – so überhaupt vorhanden – die Export-Fähigkeiten der hauseigenen Software sehr begrenzt und zumeist nicht ausreichend dokumentiert waren. Beim Importieren zeigte sich, dass auch die Exportformate, also die Form, in welcher die Daten aus den einzelnen Software-Produkten herauskommen, nicht nur von Software zu Software, sondern sogar von Version des Programms zur nächsten Version des Programms variieren!

Das Importierwerkzeug ist entsprechend flexibel gestaltet. Wenn auch stellenweise mit einigem Aufwand verbunden – bisher ist noch kein Import fehlgeschlagen.

Auf nationaler Ebene

Bisher gab es „museum-digital“ in verschiedenen Versionen, jeder teilnehmende Museumsverband hatte seine eigene. Ein Artikel in Heise-Online (der zu sehr vielen Besuchen führte) zeigte auf, dass dies für das normale Publikum nicht unbedingt verständlich war. Warum sollte man, wenn man etwas suchte, zuerst in museum-digital:sachsen-anhalt, dann in museum-digital:rheinland-pfalz, dann in ... suchen? Mit Hochdruck wurde daher an einer übergeordneten Plattform gearbeitet. Diese ist jetzt ganz einfach unter www.museum-digital.de erreichbar. Diese Plattform basiert auf einer eigenen Datenbank, in welche die Informationen aus den verschiedenen regionalen Plattformen per Knopfdruck in ihrer jeweils aktuellen Form übernommen werden. Es kann also niemand (außer den Administratoren) direkt in diese nationale Datenbank hineinschreiben oder Änderungen an ihrem Inhalt vornehmen.

Der Nachteil dieser Lösung besteht darin, dass die Angaben in der nationalen und der regionalen Version nicht in jedem Augenblick gleich sein müssen. Wurde etwas in der regionalen Version aktualisiert, ist es erst nach einem Neuimport in der nationalen Version verfügbar. Diese Aktualitätsdifferenz beträgt zumeist einen halben Tag. Der Neuimport in die nationale Version geschieht noch per Knopfdruck, soll später aber vollautomatisch ablaufen.

Die großen Vorteile dieser Lösung liegen aber auf der Hand: An den bestehenden Strukturen musste nichts geändert werden,

d.h. die Museen in Sachsen-Anhalt, können sich auch weiter mit dem von ihnen mitgestalteten museum-digital:sachsen-anhalt identifizieren. Das Gleiche gilt für die Museen in Ostwestfalen-Lippe und andernorts. Der zweite Vorteil besteht darin, dass jenes Werkzeug, mit dem per Knopfdruck die nationale Datenbank erstellt wird, so gestaltet wurde, dass mit ihm jede beliebige Anzahl von Museen zu jeweils einem eigenen Portal zusammengefasst werden können. Es lässt sich also mit sehr wenig Aufwand ein Portal der Technikhistorischen Museen, eines der Rotkreuz-Museen, aller Museen der Stadt Halle (Saale), usw. erstellen. Die Technik für solche Objekt-Portale ist vorhanden – bisher hat sich der Bedarf aber noch nicht geäußert.

Mit dem Erstellen der nationalen Version wurden auch die kontrollierten Begriffe (Angaben zu Zeiten, Orten, Akteuren sowie Schlagworte) zusammengefasst. Jede Version von „museum-digital“ arbeitet seitdem mit zwei Datenbanken, einer für die üblichen Objektinformationen und einer gemeinsam genutzten für kontrolliertes Vokabular. Hierzu später mehr.

„Themator“

Wenn die Objekte schon einmal online verfügbar sind, dann müsste – so ein Gedanke, der in der AG Digitalisierung aufkeimte – es doch leicht möglich sein, die Objekte neben der Präsentation in der museum-digital-Datenbank auch in Form von digitalen Ausstellungen verfügbar zu machen. Dieser Gedanke war Ansporn den „Themator“ zu programmieren – ein Zusatzprogramm zum Erstellen von kleinen Themenportalen. Eines wurde ziemlich schnell klar: Eine digitale Ausstellung (oder ein „Thema“) stellt die Objekte in einen jeweils eigenen Kontext. Aus den üblichen (allgemeinen)

Objektbeschreibungen geht indes nur selten eindeutig hervor, worin die Beziehung des einzelnen Objektes zum jeweiligen Thema besteht. Es musste also möglich werden, für die Objektdarstellung innerhalb eines „Themas“ angepasste Objektbeschreibungen zu speichern. Ein einfaches zusammen Gruppieren der Objekte aus der Datenbank, versehen mit einem einleitenden Text, hätte niemals zu befriedigenden Darstellungen eines Themas geführt.

Gleichzeitig wurde deutlich, dass die Begrenzung auf Objekte, die bei „museum-digital“ selbst publiziert worden sind, eine Limitierung darstellt. Es wurde also ermöglicht auch „fremde“ Objekte in ein solches Thema einzubinden. Schließlich wurde deutlich, dass die vom Themator vorgegebene Form der Präsentation variabel sein muss. Eine digitale Ausstellung („Thema“) ist etwas Individuelles und sollte nicht genauso aussehen wie 125 andere auch. Im Themator wurden aus diesem Grund einige frei kombinierbare Darstellungsoptionen eingebaut.

Der Themator selbst ist so gestaltet, dass auch Außenstehende, beispielsweise Lehrer, sehr leicht ein Themenportal als Unterrichtsmaterial erstellen können. In der Bedienung ist der Themator an das Eingabe- und Bearbeitungswerkzeug von „museum-digital“ angelehnt – letztlich aber kann er auch ganz unabhängig von „museum-digital“ verwendet werden. Der Themator wird kontinuierlich weiterentwickelt. Die aktuelle Themenpräsentation ist als Begleitmaterial zur Brandenburgischen Landesausstellung 2014 konzipiert und behandelt das Thema der historischen Beziehungen zwischen Brandenburg und Sachsen.

LinkedData

Parallel zum Themator wurde ein anderer Gedanke aus der AG Digitalisierung in die Tat umgesetzt. Immer häufiger fiel im allgemeinen Diskurs zur Digitalisierung der Begriff „LinkedData“. Die Idee ist ganz einfach, statt für ein Objekt zu speichern, dass es in „Essen“ hergestellt wurde, speichert man als Herstellungsort die Angabe, „Essen“ und zudem, dass es sich um jenen Ort handelt, der in der online frei abrufbaren Datenbank „geonames“ die Nummer „2628810“ trägt. Diese Speicherung von Verweisen auf Normdaten sind natürlich noch keine LinkedData. Bei „museum-digital“ wurde eine Routine eingefügt, die beim Aufruf der entsprechenden Seiten an den passenden Stellen vollautomatisch weitere Daten vom geonames-Server einfließen lässt. Ein Klick auf den „Info“-Button hinter einem Ortsnamen lässt ein kleines Fensterchen aufgehen, in welchem eine OpenStreetMap-Karte mit der Lage des Ortes angezeigt wird, ebenfalls aber auch die Einwohnerzahl der entsprechenden Ortschaft. Die Einwohnerzahl ändert sich und wird bei „museum-digital“ gar nicht erst gespeichert, sie wird „on-the-fly“ vom geonames-Server geholt.

In gleicher Weise sind auch Personen und Institutionen über die Gemeinsame Normdatei der Bibliotheken (GND) und über Wikipedia verlinkt. Ruft jemand – durch Klick auf den Info-Button hinter einem Personennamen – weitere Informationen ab, so erscheinen alle Schreibweisen des Namens (aus der GND), ein Bild der Person (von Wikimedia) und in bis zu 8 Sprachen eine Kurzbeschreibung der Person (aus verschiedenen Wikipedia-Versionen über dbpedia). Auch diese Daten sind nicht bei „museum-digital“ gespeichert, sondern fließen stets aktuell ein.

In manchen W3C-Ausschüssen wurde die museum-digital-Anwendung von LinkedData als beispielhaft gesehen und in Protokollen vermerkt. Das hatte und hat internationale Nachfragen zur Folge.

NODAC mit Zauberstab

Alle kontrollierten Begriffe³ waren ja bereits mit der Erstellung der nationalen Version (s.o.) in einer separaten Datenbank zusammengefasst worden. Die LinkedData-Funktion von „museum-digital“ greift auf diese Datenbank zu. Im Ergebnis muss jede Person, jeder Ort, jede Zeit, jedes Schlagwort nur ein einziges Mal eingegeben werden. An der gespeicherten Stelle kann es mit Normdaten aus fremden Vokabularen zentral angereichert werden.

Der Vorteil dieser Lösung liegt auf der Hand: Es muss nicht jedes Museum die Normdaten herausuchen. Tatsächlich gibt es für die Normdatenanreicherung eine Redakteurin, die, von den Museen in Sachsen-Anhalt und vom Land bezahlt, die Kontrolle behält. Weitere ehrenamtlich tätige „Normdaten-Anreicherer“ sind ebenfalls aktiv.

Damit die Normdatenredaktion schnell von der Hand geht, wurde ein spezielles Werkzeug zur Terminologie-Kontrolle programmiert, welches durchaus mit teuren Programmen aus diesem Bereich mithalten kann. Besonders hilfreich ist in diesem Werkzeug der „Zauberstab“ – ein Button, der per Klick, soweit möglich, die bei „museum-digital“ zu speichernden Informationen aus den entsprechenden Normdaten-Vokabularen in aller Welt herausholt und in ein Formular

³ Begriffe: Personen- oder Institutionennamen, Ortsbezeichnungen, Zeitangaben und Schlagworte.

füllt, welches sich nach gründlicher Kontrolle mit einem weiteren Klick speichern lässt.

Dieses Normdatenwerkzeug („NODAC“) lässt sich auch verwenden, um kontrollierte Begriffe zueinander in Beziehung zu setzen. Hiermit lässt sich z.B. Hierarchisieren und Synonymisieren. Endlich kann man (i.e. der Normdatenredakteur) dem System „beibringen“, dass „Breslau“ eine ältere Bezeichnung von „Wroclaw“ ist, dass „Friedrich der Große“ auch „Friedrich II“ genannt wurde und dass „Schaukelpferd“ ein „Spielzeug“ ist. Das wiederum eröffnet ganz neue Trefferanzeigen! Wer in der nationalen Version (www.museum-digital.de) nach „Spielzeug“ sucht, dem wird auch das „Schaukelpferd“ als Treffer gegeben – unabhängig davon, ob im „Schaukelpferd“-Datensatz das Wort „Spielzeug“ auftaucht oder nicht.

Gemeinsam mit dem Schlagwort-Suchbegriff wird auf den Trefferseiten auch die hierarchische Einbindung der Begriffe ausgegeben. Ein Katalog sozusagen, der wiederum navigierbar ist und zum Sprung auf weitere Trefferseiten einlädt.

Der Schlagwort-Normdatenkatalog von „museum-digital“ ist „handgeprüft“, er lehnt sich im Wesentlichen an die Hierarchisierungen der GND an, hält sich in Zweifelsfragen aber an den „common sense“ – er ist ja erstellt für den Kontext einer (potentiell jedenfalls) alle Menschen erreichenden Internetpräsentation, Abweichungen von der GND werden in Kauf genommen.

Nutzbarmachung von Thesauri

Für die Arbeit in Museen, ob involviert in „museum-digital“ oder nicht wurden verschiedene museumsspezifische Thesauri, die sich

meistens auf die Objektklassifikation beziehen, als Nachschlagewerk veröffentlicht. Diese Thesauri sind damit erstmals in Gänze über einen ganz normalen Browser frei navigier- und nutzbar. Nach und nach werden weitere Thesauri an dieser Stelle veröffentlicht. Das Gleiche gilt für den museum-digital-eigenen Schlagwortkatalog, der bald öffentlich zugänglich und über eine API referenzierbar sein wird. Dann wird es auch von außen möglich sein, diesen Schlagwort-Katalog zu verwenden. Die, aus „museum-digital“ in Portale wie Europeana oder die sogenannte Deutsche Digitale Bibliothek exportierten, Datensätze enthalten alles Nötige, damit der museum-digital-Schlagwortkatalog leicht als LinkedData von anderen genutzt werden kann. Denn wer etwas nimmt (von geonames, von der GND) der sollte auch etwas geben.

Die eigene Seite

Das Ziel von „museum-digital“ war und ist die Zusammenführung von Objektinformationen aus vielen Museen. Dennoch war von Anfang an der Suchschlitz umschaltbar auf „Suchen in einem bestimmten Museum“ oder „Suchen in einer bestimmten Museumssammlung“. Als die Idee aufkam, dass jedes Museum doch auch die eigenen Objekte auf der eigenen Seite zeigen können sollte, waren nur einige Feinanpassungen nötig um dieses Ziel zu erreichen. Mittlerweile haben etliche Museen diese Funktion genutzt:

Bei „museum-digital“ haben Sie ihren Objekten ein „Zuhause im Internet“ gegeben - von hier aus werden die Daten in die eigenen Seiten eingespielt oder an Portale exportiert. Der große Vorteil dieser Lösung besteht darin, dass Objektinformationen nur an einer einzigen Stelle aktualisiert werden müssen. Andere Ausspielungen der

Informationen können von hier aus automatisch, halb-automatisch oder manuell aktualisiert werden.⁴

Selbst exportieren

Um den Austausch mit Portalen aller Art (Fachportale, Europeana, ddb) zu erleichtern wurden die Exportmöglichkeiten erweitert, indem neben einem museum-digital.xml und einem museumdat.xml jetzt auch ein LIDO.xml exportierbar ist. Die Daten lassen sich also in verschiedenem Umfang und verschiedener Darstellungsform aus- und eben auch weitergeben. Insbesondere die Möglichkeit im LIDO-Format, das sich zum gängigen Standard der Museumswelt entwickelte und das museumdat-Format ablöste, zu exportieren, erleichtert die Daten-Kommunikation mit Portalen enorm.

Das Exportmodul von „museum-digital“ wurde mit vielfältigen Einstellmöglichkeiten neu gestaltet. Bis auf Feldebene herab ist frei bestimmbar, was in den Export aufgenommen werden soll und was nicht. Diese eigentlich recht komplexe Materie (welche Informationen zu welchen Objekten exportieren) ist programmtechnisch so umgesetzt, dass die Bedienung sehr leicht fällt. Allen Museen ist es nunmehr möglich einen:

- jeweils eigens bestimmten Export im jeweils gewählten Format oder

⁴ Einige Beispiele: `<http://www.museum-digital.de/san/index.php?sv=+&done=yes&suin=2&extern=yes&exWho=2>` [12.04.2014]; `<http://www.haendelhaus.de/de/sammlungen/digitalisate.html>` [12.04.2014]; `<http://www.landesmuseum-stuttgart.de/sammlungen/digitaler-katalog/>` [12.04.2014]; `<http://www.wernigerode.de/de/grafiksammlung.html>` [12.04.2014].

- einen Standard museum-digital.xml-Export (der alle Informationen zu den Objekten enthält) oder
- einen Standard LIDO-Export (der nur Publikationsinformationen enthält)

selbst zu erhalten. Einige wenige Klicks genügen und der Server schickt dem Museum die entsprechenden Daten in einer ZIP-Datei, welche das Museum selbst weitergeben oder zur Datensicherung speichern kann. Für Museen wird es auf diese Weise einfacher, sich an (Portal-) Projekten zu beteiligen, die Administration von „museum-digital“ muss nicht mehr aktiv werden.

Inventarisierungswerkzeug

Es war im Laufe der Arbeit sehr schnell deutlich geworden, dass Publizieren und Inventarisieren zwei grundlegend verschiedene – sich jedoch in Teilaspekten überschneidende Tätigkeiten sind, die sich wechselseitig bedingen und befördern können – aber nicht müssen. Der „Königsweg“ – so wurde ermittelt, liegt wohl darin, gleich bei der elektronischen Inventarisierung die Internet-Publikation mitzubedenken und zum Beispiel allgemeinverständliche, publizierbare Objektbeschreibungen zu verfassen. Diese Empfehlungen für das Inventarisieren sind natürlich nur dann weiterführend, wenn das entsprechende Instrumentarium zur Verfügung steht - wenn die Museen über (wenigstens eine) Datenbank verfügen. Viele Museen, so hatte es ja schon die Analyse am Beginn von „museum-digital“ gezeigt, verfügen jedoch nicht über die entsprechende Software, auch nicht über die Mittel, sich diese zu beschaffen.

Für diese Museen wurde in „museum-digital“ die Möglichkeit geschaffen direkt zu inventarisieren, d.h. neben den für die Publikation gedachten Informationen auch jene Angaben zu speichern, die für die Objektverwaltung (einschließlich Eigentumsnachweis, Standorthistorie, Ausstellungshistorie und Rechtsstatus) vonnöten sind. Die Idee dabei ist nicht, einen Ersatz für bestehende Inventarisierungssoftware zu schaffen und in Konkurrenz zu den entsprechenden Firmen zu treten, sondern vielmehr all jenen, die sich eine solche Software nicht leisten können (oder scheuen, sich in eine solche einzuarbeiten) etwas an die Hand zu geben. „Museum-digital“ ist und wird vor allem für die Publikation optimiert.

Nachdem der Anfang gemacht war und die ersten Museen aktiv die Inventarisierungsfunktionen von „museum-digital“ nutzten, kam schnell der Wunsch auf, auch diese zu verfeinern. Mittlerweile können bis zu 100 zusätzliche Felder ausgefüllt, eigene Hilfetexte für jedes Feld erstellt, Merklisten generiert und Feldeinträge im Batch-Modus bearbeitet werden. Listenausdrucke mit vielfältigen Sortier- und Darstellungsoptionen sind genauso möglich wie der selektive Export und die Sicherung aller Inventarisierungs- und Publikationsinformationen. Zur Datensicherung kann auch eine klassische Karteikarte gedruckt werden. Die Inventarisierungsfunktionen werden weiter ausgebaut, indem bald zu den Objekten auch Dokumente, Bilder, Videos und Ton-Dateien geladen werden können, die dann zwar im System vorhanden, aber nur nach Maßgabe des Museums auch der Öffentlichkeit präsentiert werden. So wird es möglich beispielsweise auch Restaurierungsberichte oder Scans der historischen Karteikarten zum Objekt zu speichern.

Museen sind im Allgemeinen vor allem voneinander verschieden. In einem kooperativen Vorhaben wie „museum-digital“ ist es deshalb nicht leicht, ein und dieselben Felder für alle Museen zu verwenden. Kunstgeschichtliche Museen benötigen möglicherweise ein Feld für „Stil“, ein technikhistorisches Museum benötigt ein solches nicht. Es hat sich jedoch gezeigt, dass auch bei gleicher Eingabemaske die Art der Eingabe (sowohl bei Publikations- als auch bei Inventarisierungsinformationen) immer auch vom Eingebenden selbst abhängt. Wenn zwei Archäologen das gleiche Objekt datieren, dann schreibt der eine einen Zeitraum mit Anfang und Ende, der andere aber den Namen einer „Kultur“ in die Datenbank. Bevor man hier (und es gibt viele ähnliche Stellen) die große Standardisierung für alle Museen erreichen kann, sollte – und die Erfahrung zeigt, dass dies möglich ist – wenigstens in den einzelnen Einrichtungen eine Standardisierung stattfinden.

Wenn es bei allen entsprechenden Objekten aus Museum A immer heißt „Spandau“, dann kann man das leicht automatisch in „Berlin-Spandau“ ändern. Wenn es aber manchmal „Spandau“, ein andermal „Spandau bei Berlin“, dann „Berlin, Spandau“, dann „Berlin, Ortsteil Spandau“, heißt, dann wird es irgendwann schwierig noch etwas zu finden, geschweige denn die Ortsangaben mit Normdaten anzureichern oder bei Exporten für externe Projekte und Portale verarbeitbar zu haben. Bei „museum-digital“ wurden einige technische Hilfsmittel eingesetzt, um auch die Inventarisierungsangaben gleichförmig zu machen – wenigstens innerhalb einer einzelnen Einrichtung.

Going Europe

Schon sehr früh in seiner Entwicklung war „museum-digital“ eng mit europäischen Projekten verwoben, die zumeist im Kontext der Europeanana standen (s.o.). Aus dieser gemeinsamen Arbeit entstand eine Reihe von Kontakten außerhalb Deutschlands. Weil die Ausgangslage in vielen Ländern ähnlich wie in Deutschland ist, kam die Idee auf, „museum-digital“ zu übersetzen. Mittlerweile gibt es eine eigene Version auf Ungarisch⁵ und die polnische Version ist in Arbeit⁶.

Die Zukunft wird zeigen, ob sich diese Versionen lohnen, aber auch da gilt es einen langen Atem zu haben. Die Zeichen stehen jedenfalls gut. Eine englischsprachige Version ist ebenfalls bereits verfügbar – die brauchen die polnischen Kollegen als Ausgangspunkt für die Übersetzung. Eine italienische Version ist bereits vorbereitet, braucht aber noch viel Arbeit. Im Rahmen seines Studiums hat ein Student eine vollständige Version (Eingabe- und Ausgabeseite) in Bahasa Indonesia erstellt – ungewiss, ob sich Museen finden, die damit arbeiten werden. Schließlich ist Inventarisieren (das, was man in Mitteleuropa darunter versteht) und digitales Publizieren (was wir dafür halten) wenigstens in seinen Feinheiten kulturell determiniert.

Nächste Schritte ?

„Museum-digital“ wächst in die Tiefe (Publizieren und Inventarisieren) und Breite (regional, national, international). Mittlerweile gibt es mehr als 20 einzelne Ausspielungen (nicht alle wirklich öffentlich),

⁵ Ungarische Version <www.museum-digital.de/hu/portal> [12.04.2014].

⁶ Polnische Version <www.museum-digital.de/pl/portal> [12.04.2014].

die auf Wunsch vollkommen unabhängig voneinander betrieben werden können. Dies macht die Pflege der Software, vor allem aber das Einführen neuer Funktionen schwierig. Dennoch stehen zwei Erweiterungen aktuell zur Diskussion, die von vielen Museen gewünscht werden: Ein Modul für die Verwaltung von Fundzusammenhängen (Archäologie-Modul) und ein Modul für Museumsbibliotheken nach bibliothekarischen Standards (womit auch eine Erweiterung der Exportmöglichkeiten auf andere Exportformate verbunden ist).

In gewisser Weise vermindert die entstandene Komplexität die Flexibilität. Genau diese Flexibilität aber ist eines der wesentlichen Kennzeichen von „museum-digital“. Es muss weiterhin möglich sein schnell und unbürokratisch auf neue Ideen, Wünsche und Notwendigkeiten aus den Museen und aus der allgemeinen Entwicklung der Technik und des Konsumtionsverhaltens zu reagieren. Bevor also weitere Module entwickelt werden, steht eine Revision der Software und mögliche schlankere Version des Bestehenden auf der Tagesordnung.

Der Aufwand lohnt

Die Internetpräsenz von „museum-digital“ hat mittlerweile jeden Tag zwischen 1.000 und 1.500 Besucher, im Monatsschnitt ca. 30.000 bis 40.000. Das ist nicht gerade wenig dafür, dass auch erst 40.000 Objekte online verfügbar sind. Die Besucher bleiben durchschnittlich 3 Minuten auf den Seiten von „museum-digital“ und besuchen dabei 5,7 einzelne Webseiten. Die Seiten sind so programmiert, dass die Besucher leicht von einer zu einer weiteren Seite geführt werden.

Dass sich der Aufwand für die Museen lohnt, zeigt sich aber vielleicht noch deutlicher im Feedback, in den Fragen aus dem Publikum

an die Museen, in den Informationen, die fachkundige Sammler und andere zur Ergänzung schicken. Es zeigt sich generell darin, dass die einzelnen Institutionen mit den (potentiellen) Besuchern auf einer neuen Ebene in Kontakt treten und auch darin, dass ein verstärkter Austausch untereinander stattfindet.

„Museum-digital“ ist letztlich das Produkt der Ideen aus den Museen selbst und dieses Produkt ist erfolgreich: Viele Besucher, immer mehr Museen, die mitmachen möchten, immer mehr nationale und internationale Versionen. Darauf kann man stolz sein – und auch deshalb lohnt es sich.

Mit Google umgehen

Die meisten der Besucher der Internetseiten von „museum-digital“ kommen über Google. Die museum-digital-Software ist von Beginn an dem Gedanken gefolgt, dass immer neue Portale zu errichten das Publikum nur verwirren kann. Die Mehrheit der Nutzer kennt nur eine Suchmaschine und ist nicht daran gewöhnt zu überlegen, „ich habe Frage A“ und muss deshalb wohl zu „Portal B“ gehen. Die Mehrheit der Internetnutzer folgt dem Prinzip der Serendipität. Sie können ihre Fragen nicht genau stellen und erst beim Recherchieren und den ersten im Netz angebotenen Antwortvorschlägen wird ihnen die eigene Fragestellung klarer. Dieses Prinzip bedient Google in hervorragender Weise und wahrscheinlich ist es (auch) deshalb so erfolgreich.

„Museum-digital“ setzt also darauf, dass die Nutzer bei Google sind und an Google gewöhnt sind. Ergo ist es das Ziel bei möglichst vielen Suchbegriffen im Google-Ranking einen Platz auf der ersten Seite zu erlangen. Zu diesem Zweck wird die Software von

„museum-digital“ im Hintergrund ständig überarbeitet und auf diese Absicht hin optimiert. Sicherlich könnte „museum-digital“ im Design gefälliger oder unterhaltsamer daherkommen, auch daran wird gearbeitet – aber: Das Design steht immer erst an dritter Stelle. Davor stehen die Findbarkeit (durch und bei Google) und die Nutzbarkeit (Übersichtlichkeit).

Dieses Konzept scheint aufzugehen. Wenn die Nutzer bei Google sind – dann müssen auch die Museen dort vertreten sein. Dennoch: Im Laufe der Zeit und mit wachsender Anzahl von verfügbaren Objekten wird „museum-digital“ für immer mehr Leute, insbesondere aber für Museumskollegen, selbst zu einem Portal. Nutzer, die genau wissen wonach sie suchen und die genau wissen, dass es eine Chance gibt, es bei „museum-digital“ zu finden.

Warum konnte „museum-digital“ erfolgreich werden

In erster Linie ist „museum-digital“ ein Vorhaben der Museen selbst. Das schafft Identifikation und ermöglicht Freude am Mitmachen. Jedes Museum hat – und dies unterscheidet „museum-digital“ sehr von Portalen wie etwa der Europeana oder der sogenannten Deutschen Digitalen Bibliothek – jederzeit die volle Kontrolle über seine Daten, kann stets Aktualisierungen vornehmen oder entscheiden, das Vorhaben zu verlassen. Dann ist „museum-digital“ extrem günstig, die Teilnahme ist kostenfrei. Jeder Grad von Involviertheit ist möglich. Es ist keine „intelligente Linksammlung“ wie die meisten Portale, sondern der Ort, an dem die Museen ihre Publikationsinformationen selbst speichern können. Es ist in dieser Hinsicht das Gegenstück zu Linksystemen wie der ddb – bei denen vorausgesetzt wird, dass die Objektinformationen bereits online verfügbar sind.

Jeder, auch das ist wichtig, wird in gleicher Weise ernst genommen. Nicht jeder Vorschlag lässt sich umsetzen – aber es ist ganz egal, woher dieser Vorschlag kommt, ob aus einem kleinen oder großen Museum. „Museum-digital“ ist flexibel, in gewissem Maße unkontrolliert und frei von unnötiger Bürokratie – in einem Wort: lebendig. Das ist ganz gewiss ein Schlüssel seines Erfolges. „Erfolgreich sein“ ist ohnehin etwas Relatives und lässt sich nur an den Zielen messen.

Das grundsätzliche Ziel von „museum-digital“ ist es allen Museen zu ermöglichen, sich selbst über ihre Objekte im Internet auf möglichst einfache Weise möglichst effizient bekannt zu machen. Diese Möglichkeit geschaffen zu haben, ist der eigentliche Erfolg des Projektes. Es liegt in der Natur einer solchen Zielstellung, dass stets Weiterentwicklungen und Verbesserungen möglich sind. Solche Verbesserungen werden bei „museum-digital“ weiterhin umgesetzt, die Arbeit für die Museen wird weiter reduziert und die Effektivität der Publikation (bei „museum-digital“ z.B.: beim Export an Portale, beim Google-Ranking) weiter gesteigert. Das System ist in einem beständigen Prozess der Optimierung und diesen Prozess gestalten alle mit.

Ein wenig Struktur

Spontan und ohne „Masterplan“ ist „museum-digital“ in den letzten 5 Jahren auf mehr als 20 Versionen gewachsen, an denen sich mehr als 320 Museen mit mehr als 40.000 Objekten beteiligen. Doch „museum-digital“ ist Ehrenamt, ist persönlicher Einsatz - ist Wollen, Wirken und Werden – wie schon eingangs gesagt. Natürlich geht

das nicht wirklich ohne „Struktur“ – nur, dass diese Struktur bei „museum-digital“ gewachsen und nicht geplant ist.

Im Grunde ist jede eigene Version von „museum-digital“ ihr eigener Kosmos. Es gibt stets einen oder mehrere Koordinatoren, die entweder aus Museen eines regionalen Verbandes, aus der Geschäftsstelle eines Museumsverbandes oder aus den Reihen einer regionalen Museumsbetreuungsstelle stammen. Diese Koordinatoren organisieren zusammen mit ihren Trägereinrichtungen Treffen für Museen. Dabei erklären sie den Gebrauch der „museum-digital“-Software und achten auf die Einhaltung der Mindestanforderungen der Publikationswürdigkeit der Informationen. Diese regionalen Träger haben (und nutzen) die Chance, Projekte einzureichen und mit Museen aus ihrer Region durchzuführen. Bisweilen kann ein Teil der Mittel auch genutzt werden, um die Software (die ja alle nutzen) weiter zu entwickeln.

Auf deutschlandweiter Ebene gibt es jedoch bislang kaum Koordination und vor allem keine Stelle, die als juristische Person fungieren könnte. Hier werden die Ideen aus den Museen gesammelt und die Software entwickelt, Qualitätsmaßstäbe festgeklopft und Normdaten angereichert. All dies geschieht fast ausschließlich im Ehrenamt. Für Bodenhaftung sorgt der Kontakt zu den Museen, zu den regionalen Trägern und die stete Rücksprache mit der AG Digitalisierung des Museumsverbandes Sachsen-Anhalt.

Dennoch, um zugleich wirkungsvoll und effektiv zentral zu agieren und auch, um sich an nationalen und internationalen Projekten und Ausschreibungen (zentral) beteiligen zu können, ist recht bald eine Organisationsform auf der deutschlandweiten Ebene zu

etablieren. Dazu gibt es bereits Ideen – die aber erst noch umzusetzen sind. Wichtig ist, dass diese Zentrale kein „Wasserkopf“ wird und keine „Kontrollinstanz“ - nichts, das – wie zu oft in vergleichbaren Projekten - die Absicherung der eigenen Existenz vor die gesetzte Aufgabe stellt. Sie muss von den Beteiligten auf allen Ebenen gemeinsam getragen werden.

Kurzes Fazit

„Museum-digital“ ist Bewegung von unten (aus den Museen selbst), die jetzt langsam oben (auf deutschlandweiter Ebene) angekommen ist. Es unterscheidet sich darin wesentlich von Projekten wie der Deutschen Digitalen Bibliothek oder Europeana, die „von oben herab“ entwickelt wurden und denen, die teilnehmen wollen Vorgaben machen – die oft nur von den großen Einrichtungen befriedigend zu erfüllen sind.

Der Ansatz von „museum-digital“ ermöglicht es allen Museen mitzugestalten. Sie sind keine bloßen „Datenlieferanten“: Es sind ihre eigenen Daten, die sie auf ihrer eigenen Plattform nach von ihnen mitentwickelten Vorgaben in von ihnen mitbestimmter Weise präsentieren.

Der Ansatz „von oben“ erleichtert es eine finanzielle Absicherung auch auf Dauer zu erreichen. Bei „museum-digital“ wird stattdessen versucht möglichst ohne finanzieller Unterstützung auszukommen. Dennoch, wenn das Vorhaben weiterhin so wächst wie bisher, wird in absehbarer Zeit mehr als Ehrenamt vonnöten sein. „Museum-digital“ wird auch damit umgehen können – solange es seinem Grundgedanken, ein für alle offenes, in hohem Maße flexibles Vorhaben von Museen für Museen zu sein, treu bleibt.

Dieser selbst gesetzte Anspruch lässt „museum-digital“ seiner eigenen Geschwindigkeit folgen. Diese ist schwer vorhersehbar und ist ebenso bedingt vom Einfallsreichtum und den Möglichkeiten der Beteiligten wie von der sich rasant entwickelnden Wirklichkeit des Internets. Die vergangenen Jahre haben gezeigt, dass man bei „museum-digital“ am besten mit langem Atem zu gutem Ende kommt. Wobei ein Ende nicht vorgesehen ist.

Die Podiumsdiskussion des 18. Archivwissenschaftlichen Kolloquiums – eine Zusammenfassung

von Stephanie Oertel

Abgerundet wurde das 18. Archivwissenschaftliche Kolloquium zum Ende des zweiten Tages mit einer Podiumsdiskussion. Innerhalb dieser zusammenfassenden und abschließenden Diskussion rückten Fragen beider Kolloquiumstage noch einmal in den Mittelpunkt. Gerald Maier, Landesarchiv Baden-Württemberg, moderierte die Podiumsdiskussion und ließ neben den Vertretern der Kultureinrichtungen auch das Publikum zu Wort kommen. Andreas Berger nahm als Vertreter der Archive teil. Rüdiger Hohls vertrat die Sicht der Wissenschaft und Stefan Rohde-Enslin die Sicht der Museen. Alle Diskussteilnehmer waren bereits als Referenten zu Wort gekommen. Mit der Beantwortung der Podiumsfragen konkretisierten sie ihre Sichtweise zur Digitalisierungsoffensive in den Kultureinrichtungen und gingen auf Forderungen und Probleme ein. Die Podiumsdiskussion bündelte damit die Themen des Kolloquiums und förderte die offene Diskussion aller Teilnehmer.

Welche Bedeutung hat die Digitalisierungsoffensive in den Archiven für die Forschung? Mit dieser Frage startete die Podiumsdiskussion und hinterfragte damit die wissenschaftlichen Ansprüche an die digitalen Quellen. Bezogen auf einen langen Zeitraum wird damit vermutlich ein großer Teil des analogen Archivguts in die digitale Form überführt und verfügbar, schlussfolgerte Rüdiger Hohls und sieht das deutlich als Vorteil für die Wissenschaftler, die immer mehr unter einem Rationalisierungsdruck stehen. Mit der digitalen

Bereitstellung erhalten sie auf unkompliziertem Weg Zugang zu Archivgut. Der Vorlauf der Bibliotheken und die damit einhergehenden Digitalisierungserfahrungen sind gewinnbringend einzubinden beim Aufbau eigener archivischer Digitalisierungskompetenzen. Die Archive treten dabei als Informationsdienstleister in Erscheinung, ganz im herkömmlichen Berufsbild verbleibend, indem sie digitale Recherchematerialien bereitstellen. Eine Diskrepanz stellt sich lediglich in der Abbildung der Tektonik und damit der Einbindung der Kontextinformationen an das Digitalisat dar, denen nicht von allen Wissenschaftlern derselbe Stellenwert beigemessen wird wie innerhalb der Archivwelt. Andreas Berger betonte, dass mit der reinen Onlinestellung der Images, d.h. ohne Einbindung von Angaben zur Archivtektonik und den deskriptiven, strukturellen, administrativen und technischen Metadaten aus archivischer Sicht kein Mehrwert für die Forschung, sondern ein Verlust einhergeht. Das Auffinden der digitalen Quelle muss über eine eindeutige Adressierung garantiert werden. Die wissenschaftliche Einordnung der Digitalisate wird jedoch erst durch die Anreicherung von Kontextinformationen ermöglicht. Die Recherche und damit das Auffinden der digitalen Quellen erfolgt auf dreierlei Wegen, so Andreas Berger weiter, zum einen über die systematische Suche der archivischen Klassifikation, deren Abbildung über die Normdaten erfolgen könnte, zum anderen über die am häufigsten verbreitete Google-Schlitz-Suche, die auch als Eingangstor des Auffindens betrachtet wird und zu guter Letzt in der bildlichen Darstellung des Stöberns. Die übergreifende Verlinkung der Projektinformationen und -ergebnisse im Web wird im digitalen Zeitalter immer wichtiger, betonte die Forscherseite nochmals und erhielt von allen Teilnehmern Zustimmung.

Braucht das deutsche Archivwesen eine nationale interdisziplinäre Digitalisierungsstrategie? Die Antwort war ein klares Nein zu einer einheitlichen deutschen Strategie von den Vertretern der Gedächtnisinstitutionen und der Forschung. Die Diskussionsteilnehmer sprachen sich alle für die Etablierung von Digitalisierungsstrategien auf regionaler wie lokaler Ebene aus, in denen sich die Bestände und sammlungsbezogenen Schwerpunkte der Archive widerspiegeln. Mit der Festlegung der lokalen Digitalisierungsstrategien wird die neue Kernaufgabe „Archivgutdigitalisierung“ in den Archiven formuliert. Damit geht die Sensibilisierung für die benötigten personellen und finanziellen Ressourcen einher. In diesem Zusammenhang regte Rüdiger Hohls an, auch die Bereitstellung von Archivgut auf Anfragen (on demand Digitalisierung) zu berücksichtigen. Andreas Berger fasste zusammen, dass es zu Nutzeranfragen und der Bestellung von Einzelscans (on demand) im Archivwesen kein einheitliches Vorgehen gibt. Bislang finden diese Digitalisate keinen Eingang in die Archivdatenbank, sondern werden im Anschluss wieder gelöscht. Sinnvoll wäre hier die Digitalisierung des gesamten Bestandes über die angefragten Einzelscans hinaus, jedoch fehlen dazu die Ressourcen. Hingegen sind die Digitalisierungsprojekte darauf ausgelegt vollständige Bestände zu digitalisieren und diese als Master dauerhaft vorzuhalten. Die daraus generierten Derivate werden für die Benutzung bereitgestellt. In vielen meist größeren Archiven ist das bereits Alltag. Innerhalb der anderen Archive muss dieses Aufgabengebiet noch geplant und implementiert werden. Kopfzerbrechen bereitet dabei die Finanzierung, da auch im Archiv keine neuen Stellen geschaffen werden. Daraus resultieren meist Projekte, die über Zweit- oder Drittmittel finanziert und durch befristete Projektstellen bearbeitet

werden. Doch sind neben dem Projektpersonal, die für das Projektmanagement zuständig sind, weitere Archivmitarbeiter in den Ablauf des Digitalisierungsworkflows involviert. Das Archivgut wird bspw. von Magazinern ausgehoben. Restauratoren, wenn vorhanden, stellen die Bestandserhaltung mit konservatorischen und restauratorischen Maßnahmen sicher und machen Vorgaben für den Scanprozess (bspw. Öffnungswinkel, Verwendung von Auflicht- oder Einzugsscanner). Für die Digitalisierung mittels Dienstleister sind vergäberechtliche Aspekte zu prüfen und für Inhouse-Projekte ist eigenes Digitalisierungspersonal notwendig. In beiden Fällen ist das Verwaltungspersonal des Archivs hinzuzuziehen. Die Ergebnisse müssen anhand der Originale durch geschultes Fachpersonal auf ihre Vollständigkeit überprüft und so ihre Qualität sichergestellt werden. Und auch für die dauerhafte Speicherung und die digitale Bereitstellung der Ergebnisse über die Projektzeit hinaus sind personelle Ressourcen der internen IT-Abteilung einzubinden und ggf. Kooperationen mit Rechenzentren oder Dienstleistern einzugehen. Der Digitalisierungsworkflow umfasst alles in allem einen nicht zu unterschätzenden Kostenaufwand, fasste Gerald Maier zusammen, für dessen Etablierung im Archivwesen Anschubfinanzierungen bspw. über Förderprogramme notwendig sind.

„Wie lässt sich das Semantic Web und damit die maschinelle Verarbeitung von Informationen innerhalb des Digitalisierungsworkflows einbinden?“, lautete die dritte Frage der Podiumsdiskussion. Das Semantic Web ist ein Konzept zur Weiterentwicklung des Internets, in dem eine inhaltliche Bewertung von strukturierten Texten möglich ist. Von zentraler Bedeutung ist hier die hinreichende Vernetzung der digitalen Daten im Web, die u.a. mit der Verwendung

von Normdaten möglich scheint. Aktuell haben sich diese Normdaten aber noch nicht im Archivwesen etabliert. Für die zentrale Anreicherung der Digitalisate mit Informationen durch Experten ist außerdem eine zentrale Bereitstellung von Kontextinformationen bspw. über eine Plattform zielführend. Hierfür ist Vertrauen in die jeweiligen Kompetenzen notwendig, das noch aufgebaut werden muss. Denn wer lässt schon gern „Fremde“ an „seine“ Daten? Darüber hinaus wäre der Rücklauf der Forschungsdaten in die interne Archivdatenbank gewinnbringend und bedarf der technischen Realisierung. Das Kenntlichmachen der ursprünglichen und der neuen Informationen stellt dabei die geringste Hürde dar und wird u.a. aktuell in Crowdsourcing-Projekten des Landesarchivs Baden-Württemberg getestet, die die wissenschaftliche Community zur Informationsanreicherung aufruft.¹

Auf welchem Weg lassen sich die Forscherinteressen mit der Zugänglichkeit der digitalen Quellen verbinden? Diese vierte Frage sprach u.a. die standardisierte Erhebung von Erschließungsinformationen und deren Bereitstellung an. Zielführend für die digitale Bereitstellung ist die standardisierte Eingabe innerhalb der Erschließung bzw. Katalogisierung der Bestände. Spezial- oder Sonderfälle treten in allen drei Gedächtnisinstitutionen auf und müssen berücksichtigt werden, meinte Stefan Rohde-Enslin. Zu Schwierigkeiten kann es innerhalb von Datenbankfeldern kommen, die meist mit „Anmerkung“ oder „Sonstiges“ benannt werden und in denen eine freie

¹ Vgl. Erfassung von Kriegsgräberlisten im Crowdsourcing-Projekt des Landesarchivs Baden-Württemberg mit dem Verein für Computergenealogie <http://wiki-de.genealogy.net/Labw_kgl> [29.06.2014].

kung“ oder „Sonstiges“ benannt werden und in denen eine freie Texteingabe möglich ist. Beim Auslesen der Datenbank und der Einspielung in Portale werden diese Felder entweder nicht berücksichtigt oder falsch zugeordnet und führen damit zu Mehraufwand. Bereits bei der Eingabe können diese Fehler durch Ausfüllen über Auswahltext unterbunden werden. Zu berücksichtigen ist, dass die Anordnung der digitalen Daten in der eigenen Datenbank sich von der Wiedergabe im Portal unterscheidet. Denn mit dem Datentransfer findet eine Vereinheitlichung statt und damit kann eine Datenbeschneidung einhergehen, der man sich bewusst sein sollte. Andreas Berger informierte, dass sich die Kolleginnen und Kollegen aus den Archiven bislang noch nicht mehrheitlich mit dem Gedanken angefreundet haben, die Findmittel ausschließlich flach erschlossen oder das Archivale unerschlossen und eventuell auch mit Fehlern digital bereit zu stellen. Der damit einhergehende unfertige Charakter und die mangelnde Qualität halten viele Kollegen davon ab, da die Erschließung eine zentrale archivische Aufgabe ist und die Ergebnisse der Kritik der Fachwelt ausgesetzt sind. Nur einige wenige schlagen bislang diesen Weg ein. Das Internet überprüft sich selbst, führte Stefan Rohde-Enslin als Argument an und belegte dies am Bsp. von Wikipedia². Rüdiger Hohls forderte im Zeichen der Wissenschaft hingegen die möglichst rasche Zugänglichmachung vieler vollständiger Bestände und hält die Etablierung von Digitalisierungsstrategien in den jeweiligen Einrichtungen für zielführend. Die Diskussionen zur Erschließungstiefe unter den Archivaren sind spartenübergreifend

² Vgl. <<http://de.wikipedia.org/wiki/>> [29.06.2014].

auszuweiten, da sie sich den meisten Wissenschaftlern nicht erschließen.

Mit dem Thema der Sicherungsverfilmung von Archivalien beschäftigte sich die fünfte Podiumsfrage. Lassen sich Mikrofilme als Ausgangsmaterial für die Digitalisierung, auch im Hinblick auf die Bestandserhaltung des Archivguts heranziehen? Aus archivischer Sicht wird die Möglichkeit, die vom Bundesamt für Bevölkerungsschutz und Katastrophenhilfe erstellten Mikrofilme und davon angefertigten Nutzungsduplikate für die Digitalisierung zu nutzen, schon länger durchdacht und erprobt. Diese Digitalisierungsmethode verspricht schneller und günstiger zu sein als die Digitalisierung von Archivgut. Die Originale selbst erfahren eine Schonung und in kürzester Zeit könnte ein digitaler Basisbestand aufgebaut werden, der die Möglichkeit zur Anreicherung und Suchbarkeit zulässt. Jedoch ist die Archivcommunity im Hinblick auf die Qualität der Mikrofilme geteilter Meinung. Die Einschätzungen reichen von „die Qualität reicht aus“ bis hin zu „die Qualität ist zu schlecht und muss im Einzelfall geprüft werden“. Fakt ist: Fehler, die bereits während der Mikroverfilmung aufgetreten sind, lassen sich auch bei der Digitalisierung nicht oder nur mit großem Aufwand beheben. Qualitätsmängel können bspw. die fehlende Schärferegulierung für jedes einzelne Archival im Herstellungsprozess des Mikrofilms sein oder die falsche Reihenfolge der Aufnahme und damit der Verlust des Zusammenhangs des Bestandes. Den Vergleich zwischen der Kosteneffizienz der Digitalisierung vom Mikrofilm und der Digitalisierung vom Original widmet sich innerhalb des Produktivpiloten das Sächsische Staatsarchiv.

Wie steht es um die Digitalisierungsinfrastruktur im deutschen Archivwesen? Diese sechste Frage hat sich der technischen Ausstattung und dem Digitalisierungsworkflow gewidmet, die bislang nur in den Bibliotheken weit verbreitet sind und für deren Etablierung die Archive und Museen noch Nachholbedarf sehen. So bemängelte Andreas Berger, dass die Nachvollziehbarkeit des Workflows bislang nur mit manuellen Laufzetteln hergestellt wird und sich dafür bislang noch kein einfaches Software-Tool durchgesetzt hat. Eine mögliche Abhilfe testet im Produktivpiloten das LWL-Archivamt für Westfalen mit dem Einsatz von Barcodes im Digitalisierungsprozess. Mit dem Einlesen des dem Archivale zugewiesenen Barcodes kann der jeweilige Arbeitsstand im Workflow abgerufen und damit der Prozess nachvollzogen werden.

Als siebte Frage formulierte Gerald Maier: Wie sollen die Digitalisate präsentiert werden? Von der Wissenschaft wird hier lediglich der formale Zugang der digitalen Quellen gefordert. Aus archivarischer Sicht ist für die Nachvollziehbarkeit der Entstehungsgenese des Archivals neben der Bereitstellung der digitalen Quellen auch die Anreicherung von Kontextinformationen unabdingbar. Eine spartenspezifische Ansicht im Viewer wird daher gewünscht. In bibliothekarischen Projekten, die eine DFG-Förderung erfahren haben, hat sich der DFG-Viewer durchgesetzt. Die Gründe liegen auf der Hand, denn die Einbindung des DFG-Viewers ist für DFG-geförderte Projekte verpflichtend. Da dieser Viewer lediglich auf die Repräsentation von Büchern ausgelegt ist, bedarf es einer Weiterentwicklung des DFG-Viewers. Qualitätsverbesserungen von Seiten der Bibliotheken sind ebenso zu erhoffen wie Anpassungen für die archivische Materialvielfalt und die Einbindung der Archivtektonik. In den Museen ist

der DFG-Viewer nicht existent, gab Stefan Rohde-Enslin zu bedenken. In den Bibliotheken und Archiven werden parallel zum DFG-Viewer eigene Viewer-Anwendungen genutzt und ausgebaut.

Zum Abschluss der Podiumsdiskussion beantworteten die Diskutanten die Frage, ob sich Web 2.0 Technologien im Archivwesen und in die Digitalisierungsprojekte einbinden lassen. Andreas Berger meinte dazu, dass die Archivwelt dem Thema noch kritisch gegenüber steht, aber auch ein Potenzial für die Zukunft sieht, die junge Generation einzubinden. So informieren einige wenige Archivkollegen regelmäßig über ihre Homepage oder Facebook zu Projekten und Aktionen im Archiv, twittern oder betreiben Blogs. Dieser Trend lässt sich auch in den Bibliotheken und Museen verfolgen.

Als Fazit der Podiumsdiskussion lässt sich festhalten, dass mit dem Austausch innerhalb des Archivwesens und über dieses hinaus mit den Wissenschaftlern aus Forschung und Lehre, den Bibliotheken und Museen als weiteren Gedächtnisinstitutionen sowie mit den Dienstleistern für das technische Knowhow viele Anregungen zum Thema Digitalisierung im Archivwesen und Impulse für den Produktivpiloten gegeben wurden. Das komplexe Thema lässt sich aber binnen zwei Tagen nicht abschließend diskutieren, sodass der Diskurs aktiv weitergeführt werden muss. Darüber hinaus sind Kooperationen zu bilden und die Synergien dazu zu nutzen, im Bereich der Digitalisierung die bereits etablierten Standards zu verbreiten. In den Archiven sind lokale Digitalisierungsstrategien zu formulieren, Workflows festzulegen und die technische Umsetzung und Ausstattung zu realisieren. Die Finanzierungsmöglichkeiten müssen für die Projektplanung eruiert werden. Abzuwägen sind hierfür die Akquise über

Zweit- und Drittmittel aber auch Stiftungsgelder und Fundraising. Archivische Digitalisierungskompetenzen sind weiter aufzubauen. Dafür wird kooperatives Vorgehen als wegweisend betrachtet, um das schriftliche kulturelle Erbe nachhaltig in digitalisierter Form im digitalen Zeitalter bereitzustellen.

Autorenverzeichnis

Reinhard Altenhöner

Deutsche Nationalbibliothek
Adickesallee 1
60322 Frankfurt/Main

Dr. Elke Bauer

Herder-Institut für historische
Ostmitteleuropaforschung –
Institut der Leibniz-Gemeinschaft
Gisonenweg 5-7
35037 Marburg

Dr. Irmgard Christa Becker

Archivschule Marburg
Bismarckstr. 32
35037 Marburg

Dr. Andreas Berger

Historisches Archiv der
Stadt Köln
Heumarkt 14
50667 Köln

Prof. Dr. Dr. Rainer Hering

Landesarchiv Schleswig-Holstein
Prinzenpalais
24837 Schleswig

Rolf Källman

National Archives of Sweden
Digisam, Riksarkivet
115 88 Stockholm
SCHWEDEN

Martin Liebetruth

Georg-August Universität Göttingen
Niedersächsische Staats- und
Universitätsbibliothek Göttingen
Papendiek 14 (Hist. Gebäude)
37070 Göttingen

Guenther Muehlberger

University of Innsbruck
Department for German Language
and Literature Digitisation and
Digital Preservation (DEA)
Innrain 52
6020 Innsbruck
ÖSTERREICH

Dr. Sylvia Necker

Institut für Zeitgeschichte
München-Berlin
Dokumentation Obersalzberg
Leonrodstr. 46 b
80636 München

Stephanie Oertel

Deutsches Kunstarchiv im
Germanischen Nationalmuseum
Kornmarkt 1
90402 Nürnberg

Barbara Reed

LI
15 Wentworth Ave
Sydney
2000 NSW
AUSTRALIEN

Stefan Rohde-Enslin

Staatliche Museen zu Berlin
Preußischer Kulturbesitz
Institut für Museumsforschung
In der Halde 1
14195 Berlin

Dr. Andrea H. Schneider

Gesellschaft für
Unternehmensgeschichte e.V.
Sophienstr. 44
60487 Frankfurt/Main

Dr. Susan Splinter

Neue Deutsche Biographie
Historische Kommission bei der
Bayerischen Akademie
der Wissenschaften
Alfons-Goppel-Str. 11 (Residenz)
80539 München

Dr. Peter Wörster

Herder-Institut für historische
Ostmitteleuropaforschung –
Institut der Leibniz-Gemeinschaft
Dokumentensammlung
Gisonenweg 5-7
35037 Marburg

Veröffentlichungen der Archivschule Marburg

Stand: Februar 2015

**Nr. 1: DÜLFER, K.; KORN, H.-E.:
Gebräuchliche Abkürzungen des 16.-
20. Jh.**, bearb. von K. Uhde, Nachdr. d.
9., überarb. Aufl., 2013, 93 S.
ISBN 978-3-923833-76-4 € 11,80

**Nr. 2: DÜLFER, K.; KORN, H.-E.:
Schrifttafeln zur dt. Paläographie des
16.-20. Jh.**, bearb. von K. Uhde,
13. Aufl., 2013, 184 S., 50 Tafeln.
ISBN 978-3-923833-28-3 € 30,80

**Nr. 3: PAPRITZ, J.: Die Karten-
titelaufnahme im Archiv**, 6. unveränd.
Aufl., 1998, 87 S. und ein Maßstabs-
schlüssel.
ISBN 978-3-923833-02-3 € 6,10

**Nr. 7: DEMANDT, K.E.: Laterculus
notarum.** Lateinisch-deutsche Interpre-
tationshilfen für spätmittelalterliche
und frühneuzeitliche Archivalien, 9.
unveränd. Aufl., 2014, 332 S.
ISBN 978-3-923833-06-1 € 26,80

**Nr. 20: MENNE-HARITZ, A.: Schlüs-
selbegriffe der Archivterminologie.**
Lehrmaterialien für das Fach Archiv-
wissenschaft, Nachdr. d. 3., durch-
ges. Aufl., 2011, 126 S.

ISBN 978-3-923833-09-2 € 16,80

**Nr. 23: BRÜNING, R.; HEEGEWALDT,
W.; BRÜBACH, N. (Übers. & Hrsg.):
ISAD(G) – Internationale Grundsätze
für die archivische Verzeichnung,**
Durchgesehener Nachdruck der 2.,
überarb. Ausgabe, 2006, 60 S.
(vergriffen, PDF-Version online)
ISBN 978-3-923833-07-8 € 10,40

**Nr. 24: STEIN, W.H.: Französisches
Verwaltungsschriftgut in Deutschland.**
Die Departementalverwaltungen 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. 27: UHDE, K. (Hrsg.): Qualitäts-
sicherung und Rationalisierungspoten-
tiale in der Archivarbeit.** Beiträge des
2. Archivwissenschaftlichen Kolloqui-
ums der Archivschule Marburg, 1997,
246 S.
ISBN 978-3-923833-53-5 € 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 Aufzeichnungen. Bericht über einer Studienreise nach Nordamerika, 10.-21. Mai 1999, 2000, 85 S.
(vergriffen, PDF-Version online)
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, 89 S.
ISBN 978-3-923833-67-2 € 8,20

Nr. 35: MENNE-HARITZ, A. (Hrsg.): Online-Findbücher, Suchmaschinen und Portale. Beiträge des 6. Archivwissen-

schaftlichen Kolloquiums der Archivschule Marburg, 2002, 238 S.
(vergriffen, PDF-Version online)
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. (Hrsg.): 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. Archivwissenschaftlichen 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 15.11.2004, 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 € 26,80

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 € 30,60

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 € 30,60

Nr. 46: BISCHOFF, F.M. (Hrsg.): Benutzerfreundlich – rationell – standardisiert. 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 an der Archivschule Marburg, 2008, 276 S.
(vergriffen, PDF-Version online)
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 € 21,30

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.
(vergriffen, PDF-Version online)
ISBN 978-3-923833-36-8 € 31,80

Nr. 50: POLLEY, R. (Hrsg.): Anbietung von Unterlagen öffentlicher Stellen an die Archive: Rechtslagen, Probleme, Lösungswege. Beiträge zu einem Workshop am 27.11.2008 an der Archivschule Marburg, 2011, 180 S.
ISBN 978-3-923833-37-5 € 22,80

Nr. 51: WENZEL, K.; JÄCKEL, J. (Hrsg.): Retrokonversion, Austauschformate und Archivgutdigitalisierung. Beiträge zum Kolloquium aus Anlass des 60-jährigen Bestehens der Archivschule Marburg, zugleich 14. Archivwissenschaftliches Kolloquium der Archivschule Marburg, 2010, 385 S.
ISBN 978-3-923833-38-2 € 29,80

Nr. 52: HIRSCH, V. (Hrsg.): Golden die Praxis, hölzern die Theorie? Ausgewählte Transferarbeiten des 41. und 42. wissenschaftlichen Kurses an der Archivschule Marburg, 2011, 323 S.
ISBN 978-3-923833-39-9 € 34,90

Nr. 53: BECKER, I.Ch.; HIRSCH, V.; WENZ-HAUBFLEISCH, A. (Hrsg.): Neue Strukturen – bewährte Methoden? Was bleibt vom Archivwesen der DDR. Beiträge zum 15. Archivwissenschaftlichen Kolloquium der Archivschule Marburg, 2011, 382 S.
ISBN 978-3-923833-40-5 € 34,80

Nr. 54: BECKER, I.Ch. (Hrsg.): Schutzfristen – Festlegung und Verkürzung. Beiträge zum Workshop der Archivschule Marburg am 3. Mai 2011, 2012, 127 S.
ISBN 978-3-923833-41-2 € 14,80

Nr. 55: BECKER, I.Ch.; HAFFER, D.; UHDE, K. (Hrsg.): Digitale Registraturen – digitale Archivierung. Pragmatische Lösungen für kleinere und mittlere Archive? Beiträge zum 16. Archivwissenschaftlichen Kolloquium der Archivschule Marburg, 2. unveränderte Auflage, 2014, 175 S.
ISBN 978-3-923833-42-9 € 15,80

Nr. 56: UHDE, K. (Hrsg.): Von A(mtsdruckschriften) bis Z(eitgeschichtliche Sammlungen) – Vielfalt im Archiv. Ausgewählte Transferarbeiten des 43. und 44. wissenschaftlichen Kurses an der Archivschule Marburg, 2013, 360 S.
ISBN 978-3-923833-43-6 € 26,80

Nr. 57: ASPELMEIER, J. (Hrsg.): Transparenz für die Bürger? Perspektiven historischer Öffentlichkeits- und Bildungsarbeit in Archiven. Beiträge zum 17. Archivwissenschaftlichen Kolloquium der Archivschule Marburg, 2014, 298 S.
ISBN 978-3-923833-44-3 € 24,80

Nr. 58: HAFFER, D. (Hrsg.): „The Hitchhiker's Guide to the Archival World“. Räume und Grenzen der Archivwissenschaft. Ausgewählte Transferarbeiten des 45. und 46. wissenschaftlichen Lehrgangs an der Archivschule Marburg, 2014, 284 S.
ISBN 978-3-923833-45-0 € 20,80

Nr. 59: BECKER, I.Ch.; HAFFER, D.; HIRSCH, V.; UHDE, K. (Hrsg.): Archiv – Recht – Geschichte. Festschrift für Rainer Polley, Marburg 2014, 482 S.
ISBN 978-3-923833-46-7 € 35,80

Nr. 60: BECKER, I.Ch.; OERTEL, S. (Hrsg.): Digitalisierung im Archiv – Neue Wege der Bereitstellung des Archivguts. Beiträge zum 18. Archivwissenschaftlichen Kolloquium der Archivschule Marburg, Marburg 2015, 276 S.
ISBN: 978-3-923833-47-4 € 28,80

Midosa-XML:

XML-basiertes Werkzeug zur Erfassung und redaktionellen Bearbeitung von Findbüchern. Hrsg. von der Archivschule Marburg, Handbuch Angelika Menne-Haritz, Bundesarchiv.

€ 250,00

Weitere Publikationen der Archivschule sind auf unserer Homepage recherchierbar.

Tel. +49 - 06421 - 1 69 71-0
Fax: +49 - 06421 - 1 69 71-10
E-Mail: archivschule@staff.uni-marburg.de
Internet: www.archivschule.de