

Fachliche Grundsätze bei der Erschließung, Retrokonversion und Digitalisierung von analogem Archivgut am Beispiel des Landesarchivs Nordrhein-Westfalen

Von Frank M. Bischoff

Einleitung

Bei der Vorbereitung meines Vortrags im Kontext des Programms dieser Tagung ist mir klar geworden, dass eine Beschränkung auf die Darstellung von fachlichen oder technischen Grundsätzen bei der Digitalisierung von analogem Archivgut zu kurz greifen würde und eine Reihe von Voraussetzungen und Vorarbeiten deutscher Archive in den vergangenen Jahren ausblenden würde. Eine Digitalisierung von analogem Archivgut baut bei vielen öffentlichen Archiven auf jahrelange Bemühungen auf den Gebieten der Erschließung und der Verfügbarmachung dieser Erschließung im Internet auf und wäre ohne diese nicht sinnvoll umsetzbar. Während den Insidern diese Zusammenhänge offensichtlich sein mögen, bietet das heutige Referat doch Gelegenheit, die Etappen in ihren Zusammenhängen nochmals gerafft Revue passieren zu lassen. Die Überlegungen zur Erschließung von Archivgut, zur Retrokonversion von analogen Findmitteln sowie die im Zuge eines Aufbaus von Internet-Portalen zu berücksichtigenden Anforderungen sollen daher einer Reflexion über fachliche Grundsätze bei der Digitalisierung von analogem Archivgut vorangestellt werden.

Eine zweite Vorbemerkung bezieht sich auf den dehnbaren Begriff der „fachlichen Grundsätze“, der letztlich eine definierbare und verlässliche Qualität für die Herstellung eines Produkts suggeriert. Dabei ist es durchaus von Belang, ob die fachlichen Grundsätze im Sinne von Hausregeln nur für eine Einrichtung gelten oder für mehrere. Faktisch ist es nämlich in der globalisierten Welt des Internet von zentraler Bedeutung, dass sich viele Einrichtungen auf dieselben fachlichen Grundsätze berufen und nach diesen arbeiten. Wir befinden uns dann auf dem Terrain der Standards, die für bestimmte Industriezweige gelten können oder auch übergreifend respektiert werden und die über Normen national – z. B. im Kontext des Deutschen Instituts für Normung als DIN-Norm – oder international – z. B. im Kontext der International Standard Organization als ISO-Norm – nochmals einen quasi gesicherten Status erhalten. Unabhängig von der Frage, auf welchem Level die fachlichen Grundsätze angesiedelt sind, sei doch bereits hier unterstrichen, dass wir im Hinblick auf die Gewährleistung von eindeutigen Kommunikationsmöglichkeiten auf solche Grundsätze zunehmend angewiesen sind. Insofern ist es hilfreich, dass die DFG bei ihrer Förderung auf der Einhaltung von fachlichen Grundsätzen, Standards oder Normen im Interesse der Nachhaltigkeit der Ergebnisse besteht.

Erschließung von Archivgut in deutschen Archiven in der zweiten Hälfte des 20. Jahrhunderts

Archivische Erschließung umfasst nach gängiger archivwissenschaftlicher Auffassung zwei Arbeitsprozesse, nämlich Ordnung und Verzeichnung von Archivgut. Seit den 1960er Jahren wurden im westdeutschen Archivwesen gewisse inhaltliche Standards für die archivische Erschließung festgeschrieben, die meist mit einer 1967 erschienen Veröffentlichung zur archivischen Titelaufnahme von Johannes Papritz verbunden werden.¹ In Ostdeutschland fanden die schon früher publizierten, aber ähnlich angelegten Ordnungs- und Verzeichnungsgrundsätze (OVG)² für die staatlichen Archive der DDR von 1964 Anwendung.

In beiden Fällen handelt es sich eher um inhaltliche Standards, wobei die OVG einen etwas systematischeren Ansatz aufweisen. Zu den Elementen einer Verzeichnungseinheit zählen der Titel der Archivalie, ggf. Enthält-Vermerke, die Signatur, die Laufzeit, Bandzählung, Provenienzangaben, der Umfang und ggf. Verweise auf Aktenzeichen oder ältere Signaturen oder Registraturzeichen. Beiden deutschen Standards ist gemeinsam, dass sie die Titelaufnahme für eine Archivalie als Basis der Verzeichnung ansehen. Zugleich muss die Archivalie in ihren Entstehungskontext eingereiht werden, damit auch für den Benutzer die organisatorische und rechtliche Grundlage für das Handeln eines Registraturbildners und der Entstehungszweck einer Archivalie deutlich wird. Eine Archivalie wird folglich nicht als isoliertes Objekt in einer alphabetischen Ordnung abgebildet. Vielmehr müssen die Struktur eines Archivbestands und der Kontext einer Archivalie transparent bleiben oder gegebenenfalls auch transparent gemacht werden, um so der historischen Forschung alle Möglichkeiten einer quellenkritischen Untersuchung und Beurteilung zu erhalten. Die Archivalien eines Bestandes werden hierarchisch strukturiert und innerhalb einer Klassifikation abgebildet. Die Klassifikation versucht die Aufgabenstruktur oder Aktionsfelder des Registraturbildners zu gliedern und wiederzugeben oder zu rekonstruieren. Bei Organisationen mit einem funktionierenden Aktenplan wird die Klassifikation mit diesem identisch sein. Bei privaten Nachlässen, aber auch bei ungeordneten Beständen müssen Struktur und Kontext rekonstruiert werden.

Das grundlegende Prinzip der Erschließung ist die Provenienz, also die Einordnung von Archivalien in ihren Entstehungs- oder Überlieferungszusammenhang. Ohne die Kenntnis um die Herkunft einer Archivalie lässt sich ihr Entstehungszweck nicht, nur unzulänglich oder nur mit hohem Rekonstruktionsaufwand nachvollziehen. Wenn in einem Schriftstück dargestellt wird, welche körperlichen oder psychischen Zwangs-

¹ Johannes Papritz, Die archivische Titelaufnahme bei Sachakten, Marburg 1967 (= Veröffentlichungen der Archivschule Marburg, 4)

² Ordnungs- und Verzeichnungsgrundsätze für die staatlichen Archive der Deutschen Demokratischen Republik, hg. von der Staatlichen Archivverwaltung im Ministerium des Innern der Deutschen Demokratischen Republik, Berlin 1964

maßnahmen eine Person im Nationalsozialismus erlitten hat, ist es durchaus relevant zu wissen, ob sich diese Angaben in einer privaten Familienkorrespondenz befinden, ob sie aus der Patientenakte eines Sanatoriums entstammen oder ob sie Bestandteil einer Entschädigungsakte nach Bundesentschädigungsgesetz sind, die angelegt wurde, weil die betreffende Person eine wirtschaftliche Entschädigung für erlittenes Unrecht im Nationalsozialismus beantragt hat.

In deutschen Archiven ist demnach eine Erschließungsform vorherrschend, die die einzelne Archivalie nach inhaltlichen Standards verzeichnet und innerhalb einer Aufgaben- oder Sachgruppenhierarchie abbildet. Die Summe der Verzeichnungseinheiten, die zu dem zu erschließenden Bestand oder Teilbestand gehören, wird zusammen mit einer Einleitung, die oft eine knappe Registraturbildner- und Bestandsgeschichte umfasst, ggf. ergänzt um Verweise oder diverse Indices zu einem Findbuch zusammengefasst. Selbst in elektronischen Systemumgebungen spricht man im deutschen Archivwesen heute immer noch von Findbüchern, im Internet oft von Online-Findbüchern.

Neben den Findbüchern wurde schon im selben Jahrzehnt des Erscheinens der OVG und der Papritz'schen Veröffentlichung ein weiteres Hilfsmittel entwickelt, das den Benutzern den Überblick über die Bestände eines Archivs erleichtern und Aufbau und Struktur der Bestände transparent machen soll. Im Staatsarchiv Detmold erschien 1970 die erste gedruckte Beständeübersicht.³ Das Hauptstaatsarchiv in Düsseldorf folgte 1974⁴ und das Münstersche Staatsarchiv schließlich 1979 aufbauend auf Vorarbeiten in den frühen 1960er Jahren.⁵ Die Beschreibung eines Bestands ist weniger stark genormt, als die der Titelaufnahme, weist aber doch zumeist gewisse Grundelemente auf, wie die knappe Erläuterung der Geschichte und der Aufgaben des Registraturbildners, die in dem Bestand vorkommenden Betreffe, die Laufzeit und der Umfang. Die Beständeübersicht bietet einen ersten Überblick über den gesamten Inhalt eines Archivs und spiegelt dessen Gliederung in einzelne Bestände und Beständegruppen wider. Ist die einzelne Titelaufnahme im Findbuch innerhalb einer Klassifikation eingereiht, so findet sich der einzelne Bestand innerhalb des Gesamtaufbaus der Bestände des Archivs abgebildet, also eingereiht in die Tektonik. Gemäß dem in deutschen Archiven seit rund 100 Jahren geltenden Provenienzprinzip erfolgt die Einteilung in Bestände nicht auf inhaltlicher oder sachthematischer Grundlage, sondern nach der Herkunft der Unterlagen. Hintergrund dieses Ord-

³ Die Bestände des Staatsarchivs und Personenstandsarchivs Detmold, hg. v. Nordrhein-Westfälischen Staatsarchiv Detmold, Detmold 1970 (= Veröffentlichungen der staatlichen Archive des Landes Nordrhein-Westfalen, Reihe B, H. 3)

⁴ Die Bestände des Hauptstaatsarchivs Düsseldorf: Kurzübersicht, hg. vom Hauptstaatsarchiv Düsseldorf, Düsseldorf 1974 (= Veröffentlichungen der staatlichen Archive des Landes Nordrhein-Westfalen, Reihe B, H. 4)

⁵ Die Bestände des Nordrhein-Westfälischen Staatsarchivs Münster: Kurzübersicht, hg. v. Nordrhein-Westfälischen Staatsarchiv Münster, Münster 1979 (= Veröffentlichungen der staatlichen Archive des Landes Nordrhein-Westfalen, Reihe B, H. 1); eine erste Fassung bereits maschinenschriftlich vervielfältigt unter dem Titel: Die Bestände des nordrhein-westfälischen Staatsarchivs Münster: Kurzübersicht, Münster 1962.

nungsprinzips ist die Erkenntnis, dass man einzelne Sachverhalte systematisch nach Aufgaben und Zuständigkeiten bzw. nach der Herkunft von Archivalien recherchieren kann und diese Sachverhalte zugleich den Entstehungskontext zu ihrer Interpretation benötigen.

Neue Anforderungen an die archivische Erschließung und die Findmittel seit den 1990er Jahren

Seit den 1990er Jahren erlebt die archivische Erschließung in Deutschland in mehrfacher Hinsicht eine Neuausrichtung, die vor allem mit der Entwicklung einer internationalen Standardisierung, mit dem Zwang zur Rationalisierung und der Notwendigkeit zu einem erhöhten Benutzerservice auch jenseits der Lesesäle ursächlich zusammenhängt.

Internationale Standardisierung

Auf der Basis von Vorarbeiten kanadischer Archivare entwickelte eine beim Internationalen Archivrat angesiedelte Arbeitsgruppe einen Erschließungsstandard, die *General International Standard Archival Description*, kurz ISAD(G), die in ihrer ersten Fassung 1993 veröffentlicht wurde.⁶ Zwar macht auch dieser Standard inhaltliche Vorgaben und gibt insgesamt 26 Elemente an, die eine Archivalienverzeichnung umfassen kann. Viel wichtiger und im Vergleich zu deutschen Standards hervorzuheben ist, dass es sich bei ISAD(G) um einen Strukturstandard handelt. ISAD(G) geht von einer mehrstufigen Verzeichnung aus. Es wird nicht festgelegt, bis zu welchem Level die Verzeichnung erfolgen muss. Verzeichnung ist auf der Ebene des Bestands ebenso möglich, wie auf der Ebene eines Teilbestands, einer Aktengruppe einer Akte oder eines Einzelblattes. ISAD(G) verlangt aber eine hierarchische Verzeichnung, die jeweils angibt, auf welcher Stufe sie sich bewegt. Darüber hinaus setzt ISAD(G) voraus, dass die Verzeichnung vom Allgemeinen zum Besonderen geht, dass die hierarchischen Beziehungen klar dokumentiert und jederzeit nachvollziehbar sind und dass eine Redundanz von Angaben möglichst vermieden werden sollte.

Während also die deutsche Verzeichnungstradition die Beschreibung des Archivguts auf den beiden Stufen Bestand und Archivalie vorsieht, stellt ISAD(G) das frei und lässt Zwischenstufen zu. Gemeinsam ist beiden die hierarchische Strukturierung der Erschließung, in der deutschen Tradition explizit mit den Tektonikstufen in der Beständeübersicht resp. der Klassifikation im Findbuch, in ISAD(G) demgegenüber sehr offen und ohne weitere Festlegungen definiert. Das Provenienzprinzip ist in ISAD(G) ebenso grundlegend, wie es in der deutschen Verzeichnungstradition verankert ist.

⁶ Internationale Grundsätze für die archivische Verzeichnung, übersetzt u. bearb. v. Rainer Brüning u. Werner Hegewaldt, Marburg 1994 (= Veröffentlichungen der Archivschule Marburg, 23)

Während ISAD(G) in anderen, auch deutschsprachigen europäischen Ländern in den letzten Jahren immer stärker rezipiert wurde, bleibt die Resonanz in Deutschland aufgrund der eigenen Standards bis zum heutigen Tag eher verhalten. Wenn sich Prinzipien von ISAD(G) auf kurz oder lang doch durchsetzen, dann einerseits aufgrund von Rationalisierungszwängen und andererseits wegen der Notwendigkeit, digitale Erschließungsinformationen kommunizierbar und austauschbar zu machen, was wiederum einen entsprechenden, an ISAD(G) angelehnten Auszeichnungsstandard erfordert, auf den ich noch eingehe.

Rationalisierung

Im Zuge der knapper werdenden öffentlichen Kassen gerieten die Archive seit den 1990er Jahren in den Blick der Haushälter.⁷ Zum Teil wurden Unternehmensberatungen mit der Analyse der Archive beauftragt, zum Teil nahmen sich die Archive selbst unter die Lupe und analysierten ihre Situation. Dabei stellte sich in den meisten staatlichen Archivverwaltungen heraus, dass die Schere zwischen den Archivalienzugängen und den Erschließungsdefiziten immer größer wird, dass sich die Erschließungsrückstände für die zurückliegenden Archivalienzugänge in vielen Bundesländern bereits auf mehrere hundert Personenjahre addiert und dass die Erschließung der Neuzugänge mit dem verfügbaren Personal kaum zu leisten ist.

Die Schlussfolgerung aus diesem Dilemma führte in Nordrhein-Westfalen zu einer stärkeren Konzentration auf die Fachaufgabe der Erschließung. Die Erschließungstiefe ist in vielen Landesarchiven inzwischen an ein realistisches Maß angepasst worden. Sicherlich greift die Formel

tiefe Erschließung = gut und teuer
flache Erschließung = schlecht und billig

in vielen Fällen zu kurz. Einerseits kann eine tiefe Erschließung im Einzelfall Folgeaufwände für das Magazin-, Lesesaal- und Restaurierungspersonal mindern, andererseits kann auch eine flache Erschließung den notwendigen Informationsbedarf der Forschung oftmals hinreichend befriedigen bzw. überhaupt erst ermöglichen, dass Archivgut zugänglich wird. Fest steht in jedem Fall, dass nicht erschlossenes Archivgut brach liegendes historisches Erbe ist, das Kosten verursacht, ohne Nutzen und Erkenntnis stiften zu können. Vor diesem Hintergrund ist Tiefenerschließung fast nur noch bei ausgewählten Beständen und mit Drittmittelförderung möglich, während flache Erschließung inzwischen zur Norm geworden ist. Noch ist eine generelle Wendung zur Stufenerschließung nach ISAD(G) nicht erkennbar, doch wird diese Möglichkeit zunehmend diskutiert und in Einzelfällen auch angewendet.⁸

⁷ Vgl. zu dieser Thematik auch die Beiträge in dem Band Benutzerfreundlich – rationell – standardisiert. Aktuelle Anforderungen an archivische Erschließung und Findmittel. Beiträge zum 11. Archivwissenschaftlichen Kolloquium der Archivschule Marburg, hg. v. Frank M. Bischoff, Marburg 2007 (= Veröffentlichungen der Archivschule Marburg, 46)

⁸ Edgar Kutzner, Paradigmenwechsel: Von Papritz zu ISAD(G) und EAD, in: Archivar 61 (2008) S. 268-275.

Eine Maßnahme im Rahmen der Reduzierung der Erschließungsaufwände stellt die Festlegung von Standardzeitsätzen für die Verzeichnung dar. In vielen staatlichen Archiven werden solche Normzeiten zur Steuerung des Aufwands genutzt. Da mit diesen Maßnahmen die erheblichen Erschließungsrückstände nicht zeitnah bewältigt werden können, wurden mancherorts Priorisierungskriterien entwickelt, die die Reihenfolge für die Rückstandsbearbeitung festlegen. Die Qualität der vorhandenen Findmittel, noch bestehende rechtliche Beschränkungen einer freien Nutzungen, die Nachfrage der historischen Forschung, die Nutzungshäufigkeit und konservatorische Belange spielen dabei meist die Hauptrolle.

Benutzerservice

Vor etwa 15 Jahren haben sich die ersten Archive dem Internet geöffnet und Informationen ins World Wide Web eingestellt. Die Bedeutung dieses Mediums ist seither in atemberaubender Geschwindigkeit gewachsen. Es geht längst nicht mehr darum, dass Archive ihre Informationen im Internet als Add-On zu ihrem Standardservice im Lesesaal bereitstellen. Vielmehr werden archivische Bestände und Benutzungsangebote nur noch wahrgenommen, wenn sie auch im Internet verfügbar sind.

In Nordrhein-Westfalen wurde vor 14 Jahren mit Unterstützung der Deutschen Forschungsgemeinschaft (DFG) die Arbeit an einem archivischen Internetportal aufgenommen. Heute ist das nordrhein-westfälische Archivportal mit 30 Mio. Zugriffen pro Jahr vermutlich eine der am häufigsten frequentierten archivischen Websites in Deutschland. Wenn wir den Erfolg dieser Website fortschreiben wollen, müssen wir zunächst ermitteln, was ihn ausgemacht haben könnte. Ich rekurriere dafür gern auf eine im Jahr 2008 abgefasste Examensarbeit zur Frage der Integration von Beratungsangeboten in Online-Findmitteln eines Marburger Archivreferendars, der das Phänomen sehr pointiert zusammengefasst hat. Auf der Basis einer Auswertung des Internet-Nutzerverhaltens stellt er fest: „Internetbesucher von archivischen Homepages fahnden nach jenen exklusiven Informationen, die nur von den Archiven bereitgehalten werden. Ihre Interessenlage entspricht damit sehr genau jener des klassischen Archivbenutzers. Sie suchen Inhalte und kein kulturelles Entertainment.“⁹

Das nordrhein-westfälische Archivportal wollte von Anfang an vor allem die Beständeübersichten der Archive bieten. Die Findbücher sollten zu einem späteren Zeitpunkt folgen. Heute sind in dem Portal *archive.nrw.de* rund 480 Archive aus NRW vertreten, die meisten mit Kurzübersichten über ihre Bestände. Immerhin findet der Benutzer aber auch schon ein paar Hundert Findbücher. Ungeachtet der Erfolge des Portals lassen sich doch verschiedene Problemlagen erkennen, die dazu führen, dass das weitere Wachstum langsamer voranschreiten wird.

⁹ Christoph Volkmar, Service für den virtuellen Nutzer. Vorschläge zur Integration von Beratung in Online-Findmitteln, Marburg 2008, S. 16, <http://www.landesarchiv-bw.de/sixcms/media.php/25/Transferarbeit_Volkmar.pdf.> eingesehen am 18.08.2008.

Knappe Kurzübersichten kleiner Archive zählen oft nur wenige Seiten. Aber selbst die ausführlichen Beständeübersichten großer Archive passen zwischen zwei Buchdeckel. Der Aufwand, die Beständeübersichten von über 400 Archiven in kurzer Zeit internetfähig aufzubereiten ließ sich folglich in einer konzertierten Aktion schultern und noch ohne XML-ausgezeichnete Dateien – XML gab es zu der Zeit noch nicht – realisieren. Bei Findbüchern stellt sich das Problem bereits deutlich komplexer dar:

1. Die Zahl der Findbücher in nordrhein-westfälischen Archiven dürfte in die zehntausende gehen. Das Volumen ist also nicht mit den Beständeübersichten vergleichbar.
2. Die Findbücher lagen vor 15 Jahren noch im Wesentlichen in einer analogen Papierfassung vor. Um sie in ein Internetportal zu überführen, müssen sie aber in digitaler Form verfügbar sein.

Will man die Findbücher von mehr als vierhundert Archiven in einem Internetportal strukturiert abbilden und recherchierbar machen, dann müssen diese Findbücher gewissen Mindestanforderungen an die Einheitlichkeit ihrer Struktur genügen, die durch eine Software nach einem einheitlichen Muster verarbeitet werden können. Möchte man noch einen Schritt weiter gehen und auch an die Bereitstellung von digitalisiertem Archivgut denken, erhöhen sich diese Anforderungen zusätzlich.

Vor dem Hintergrund dieser Überlegungen schien die Ende der 1990er Jahre von dem gerade scheidenden Präsidenten des Bundesarchivs ausgegebene Formel plausibel zu sein, dass die Archive 100 Prozent ihrer Beständeübersichten, 10 Prozent ihrer Findmittel und 1 Prozent ihrer Archivalien online zugänglich machen sollen.¹⁰ Nachdem in vielen Archiven die erheblichen Erschließungsrückstände und der Aufwand der nachträglichen Digitalisierung ins Bewusstsein gelangt waren, schien eine andere Erwartungshaltung illusorisch.

Retrokonversion in deutschen Archiven

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

¹⁰ Hartmut Weber, Digitale Repertorien, virtueller Lesesaal und Praktikum im WWW – neue Dienstleistungsangebote der Archive an die Forschung, in: Fundus – Forum für Geschichte und ihre Quellen 4 (1999) S. 212.

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

Vor dem Hintergrund dieser Einsicht beantragte das Hauptstaatsarchiv in Düsseldorf im Jahr 2000 bei der DFG Mittel für ein Projekt zur „Entwicklung von Werkzeugen zur Retrokonversion archivischer Findmittel“.¹¹ Das im Rahmen dieses Projekts entwickelte Tool ermöglichte, aufgrund von inhaltlichen und Layout-Merkmalen die in maschinengeschriebenen oder gedruckten Findmitteln enthaltenen Informationen zu identifizieren, zu zerlegen und strukturiert in einem XML-Dokument abzulegen. Allerdings mussten die Parameter für die diversen Merkmale im Detail angepasst werden. Wegen der Unterschiedlichkeit der Findbücher war der Anpassungsaufwand hoch, da nahezu für jedes Findbuch individuell definiert werden mussten.

Um von solchen punktuellen Ansätzen weg zu einer flächendeckenden Retrokonversion archivischer Findmittel und zugleich zu einer Anschubförderung durch die Deutsche Forschungsgemeinschaft zu gelangen, mussten zunächst das Mengengerüst ermittelt und Umsetzungsszenarien entwickelt werden. Das Landesarchiv Nordrhein-Westfalen führte 2005 mit Unterstützung der DFG eine Vorstudie zur Retrokonversion archivischer Findmittel durch. Der Abschlussbericht vom Juni 2006 geht von mehr als 60 Millionen zu digitalisierenden Titelaufnahmen in den öffentlichen Archiven Deutschlands aus. Die Vorstudie schlug eine Anschubfinanzierung durch die DFG vor, bei der ein Antragsteller die Hälfte der DFG-Fördersumme auch als Eigenanteil beisteuern sollte. Außerdem empfahl die Vorstudie die Einrichtung einer Geschäftsstelle, um die deutschen Archive zu beraten, die Retrokonversionsarbeiten zu versteigern und die Abwicklung der Retrokonversionsaufträge durch Dienstleister zu begleiten. Seit September 2007 ist die mit DFG-Förderung eingerichtete Koordinierungsstelle Retrokonversion an der Archivschule Marburg aktiv und bündelt die bundesweiten Retrokonversionsmaßnahmen. Zwischen 2008 und 2010 dürften im Rahmen der DFG-Förderlinie einschließlich des Eigenanteils der beteiligten Archive annähernd 2½ Mio. Verzeichnungseinheiten retrokonvertiert worden sein.

¹¹ Vgl. hier und im Folgenden Frank M. Bischoff, Aufgaben und Erfahrungen der DFG-geförderten Koordinierungsstelle Retrokonversion an der Archivschule Marburg, in: Retrokonversion, Austauschformate und Archivgutdigitalisierung. Beiträge zum Kolloquium aus Anlass des 60-jährigen Bestehens der Archivschule Marburg, zugleich 14. Archivwissenschaftliches Kolloquium der Archivschule Marburg am 1. und 2. Dezember 2009, Marburg 2010 (=Veröffentlichungen der Archivschule Marburg, 51), S. 15-47, bes. S. 19ff.

Standardisierte Austauschformate für Erschließungsinformationen

Als Zielsysteme, in die die retrokonvertierten Verzeichnungseinheiten eingespeist werden, müssen die Programme der Archive angesehen werden. Das ist allein schon deshalb notwendig, weil viele archivrechtliche Normen der Länder Findmittel unter einen ähnlich hohen Schutz stellen, wie das Archivgut selbst, und weil diese digitalen Findmittel in den Archiven auch überarbeitet, ergänzt und angereichert werden. Dort, wo keine daten- bzw. persönlichkeitschutzrechtlichen Belange betroffen sind, dürfen und wollen die Archive Findmittel aber öffentlich zugänglich machen, auch in Fremd-datenbanken und Archivportalen.

Die Deutsche Forschungsgemeinschaft hat als Bedingung für eine Retrokonversionsförderung den teilnehmenden Archiven auferlegt, ihre Findmittel nicht nur in eigenen Datenbanken, auf eigenen Homepages oder in regionalen Archivportalen zugänglich zu machen, sondern auch über ein zentrales Archivportal, das eine Bündelungsfunktion für das deutsche Archivwesen erfüllen soll. Hintergrund dieser Forderung ist die Orientierung an den Interessen der Forschung. Man geht von der Philosophie eines One-Stop-Shop aus, in der ein zentrales Informationsportal die Wünsche der Forscher nach einem Maximum an Information am ehesten befriedigen kann. Dezentrale Informationen werden an einem Ort fachlich angemessen zusammengeführt, um den Forscher nach erfolgreicher Recherche von dort aus dann zielgerichtet an die eigentliche Provenienzadresse, also an die Informationssysteme der Archive zu verweisen.

Dieses Portal existiert noch nicht und es hat sich in den letzten Jahren herausgestellt, dass es unter den Bedingungen föderaler Strukturen nicht trivial ist, ein zentrales Archivportal zu errichten und zu finanzieren, das allen Archivsparten zugänglich ist, den staatlichen und kommunalen ebenso, wie den kirchlichen und den privaten Archiven, unter letzteren z. B. bedeutende Adels- und Unternehmensarchive. Derzeit zielen die Bemühungen darauf, im Kontext der Deutschen Digitalen Bibliothek (DDB) ein deutsches Archivportal zu errichten.

Vor diesem Hintergrund setzt die Digitalisierung von Findmitteln in einer heterogenen Archivlandschaft voraus, dass die in ein digitales Format konvertierten Findbücher mit relativ geringem Aufwand in verschiedene Zielsysteme portiert werden können. Will man die Kommunikation von Erschließungsinformationen zwischen unterschiedlichen Systemen sicherstellen, benötigt man dazu einen gemeinsamen Sprachcode, ein standardisiertes Austauschformat für Erschließungsinformationen.¹²

Das Zielformat muss verschiedene Eigenschaften aufweisen. Zu den formalen Anforderungen an ein Austauschformat zählt, dass es anerkannt, standardisiert, verbreitet, mit einfachen Mitteln verarbeitbar und nicht proprietär ist. Die strukturellen Anforderungen verlangen, dass die Findbuchstrukturen und die Kontextinformationen zu den einzelnen Findbucheinträgen verlustfrei abgebildet werden können. Die inhaltli-

¹² Vgl. ebd., S. 27ff.

chen Anforderungen schließlich entsprechen den inhaltlichen Anforderungen an herkömmliche, analoge Findmittel und bedingen, dass alle gebräuchlichen Elemente in Findmitteln adäquat dargestellt werden können.

Im Rahmen des Düsseldorfer Projekts zur Entwicklung von Retrokonversionswerkzeugen wurde das sog. Standardaustauschformat (Saft-XML) entwickelt, das sich an der deutschen Erschließungstradition inhaltlich und terminologisch orientiert und einen verlustarmen Austausch von Erschließungsinformationen zwischen verschiedenen Systemen gewährleisten soll. SAFT-XML wird als Austauschformat für Findbücher in dem nordrhein-westfälischen Archivportal benutzt.

Auf internationaler Ebene findet die aus Amerika kommende *Encoded Archival Description* (EAD) als Standard zur Kodierung von Findbüchern inzwischen weltweit Beachtung. EAD ist weniger eng der deutschen Erschließungstradition verhaftet, eignet sich aber gleichermaßen als Format für den Transfer von Erschließungsinformationen. EAD ist sehr stark an ISAD(G) orientiert. Die Tag-Library von EAD umfasst im Anhang sog. *Crosswalks* zu ISAD(G), die eine eindeutige Zuordnung der Elemente gewährleistet, so dass über den Weg eines Findbuchauszeichnungsstandards resp. eines Austauschformats der Erschließungsstandard ISAD(G) auch in die deutsche Erschließungspraxis Einzug nehmen wird.

Beide Formate sind als Zielformate für die Retrokonversion zugelassen, im Falle von EAD sogar lediglich eine Untermenge der EAD-Tag-Library. Zugleich ist jede Abweichung davon ausgeschlossen. Für die beiden Formate wurde ein Mapping erstellt, das den kontrollierten und verlustfreien Wechsel zwischen SAFT und EAD gewährleistet. Die strikte Vorgabe der DFG, für die Teilnahme am Retrokonversionsförderprogramm eines der beiden Auszeichnungsformate anzuwenden, beinhaltet zugleich eine Chance für das deutsche Archivwesen. Hier wird ein nachhaltiger Standard für den Austausch und die Kommunikation von Erschließungsinformationen definiert, eine in der Praxis angewendete Norm geschaffen. Wenn Informationen vernetzt werden sollen, wird man von solchen Standards profitieren, die über das einzelne Archiv oder eine Archivsparte hinaus die technische Verständlichkeit der Erschließungsinformationen überhaupt erst gewährleisten. Es deutet sich bereits heute an, dass sich nach Abschluss der Retrokonversion der Findbücher deutscher Archive die Einführung von Standards für den Findmittelaustausch im Rückblick als wichtigste Weichenstellung darstellen könnte.

Mögliche Grundsätze und Perspektiven für eine Digitalisierung von archivalischen Quellen

Der Ende der 1990er Jahre erhobene Anspruch an die deutschen Archive, 100 Prozent der Beständeübersichten, 10 Prozent der Findbücher und 1 Prozent der Archivalien in digitaler Form ins Internet zu stellen, konnte immerhin über rund eine Dekade hinweg als Orientierung gelten. Heute ist er in mehrfacher Hinsicht überholt und

kann nicht mehr als Leitmotiv für eine mittel- bis langfristige Arbeitsplanung dienen. Einerseits sind die Archive in der Pflicht, den gesamten Bestand ihrer Findmittel der Forschung online zur Verfügung zu stellen, sobald diese keinen rechtlichen Beschränkungen mehr unterliegen. Andererseits ist das Online-Angebot von Archivgut digitalisaten in erster Linie durch die Knappheit der verfügbaren Mittel limitiert, so dass bei dem bewerteten und als wertvoll befundenen Schriftgut der Archive nicht ein willkürlicher Prozentsatz für die Digitalisierungsbemühungen angesetzt werden darf, sondern allenfalls Regeln für eine Priorisierung der Arbeiten ermittelt werden müssen. Und schließlich muss man den Gedanken einer gleichmäßigen und gleichzeitigen Erreichung derart postulierter Ziele fallen lassen. Eine annähernde Homogenität der öffentlichen Archive im Hinblick auf die Digitalisierung von Archivgut ist absehbar nicht zu erwarten. Insofern werden die Phasen der Zielerreichung von Archiv zu Archiv unterschiedlich ausfallen, was zur Folge hat, dass die Digitalisierung von archivalischen Quellen zu einem Zeitpunkt aufgenommen werden muss, zu dem viele Archive ihre Findmittel noch nicht oder nicht vollständig datenbankgestützt vorhalten.

Unbestreitbar richtig bleibt an der Prognose von Hartmut Weber die Reihenfolge, in der die Leistungen zu erbringen sind. Grundlage für die Digitalisierung von Archivgut ist die Verfügbarkeit von digitalen Findmitteln. Die Erschließungsinformationen lassen die Provenienz der Archivalien erkennen und bieten den organisatorischen und strukturellen Kontext, in dem Archivgut verständlich, auswertbar und referenzierbar wird. Insofern schaffen archivische Erschließungsinformationen Authentizität. Dort, wo Erschließungsinformationen noch nicht in digitaler Form vorliegen, müssen sie deshalb zunächst geschaffen oder konvertiert werden, bevor die Digitalisierung des Archivguts selbst in Angriff genommen wird. Damit ist zugleich ein wichtiger Grundsatz für die Digitalisierung und Online-Stellung von archivalischen Quellen benannt: Digitalisate von archivalischen Quellen müssen in den gleichen hierarchischen und kontextuellen Strukturen wiedergegeben werden, in der die Erschließungsinformationen selbst organisiert und präsentiert werden.

Mit einer großflächigen Digitalisierung analogen Archivguts und der Bereitstellung der Digitalisate in Lesesälen oder im Internet sind vier positive Effekte verknüpft:

1. Der Nutzer profitiert von einem schnellen Zugriff auf die Digitalisate unabhängig von Ausbezeiten und Personalkapazitäten der Archive. Kann er gar von seinem Arbeitsplatz zu Hause oder an der Hochschule auf die Digitalisate zugreifen, so spart er Zeit, Reisekosten und ggf. auch Reproduktionskosten. Darüber hinaus ist er frei von den Beschränkungen der Öffnungszeiten.
2. Die letzten Jahre haben gelehrt, dass insbesondere jüngere Menschen überwiegend solche Informationen rezipieren und verarbeiten, die im World Wide Web zugänglich sind. Archive, die nicht nur ihre Findmittel, sondern auch ihre archivalischen Quellen online stellen, erhöhen die Chance, dass diese wahrgenommen und erforscht werden und tragen damit zur Erfüllung ihrer gesetzlichen Aufgaben bei.

3. Für die Archive trägt die Digitalisierungsoffensive vielfach den Charakter einer Schutzreprografie, die es ermöglicht, dem Benutzer die gewünschten Informationen an die Hand zu geben, ohne den Erhaltungszustand der Originale zu gefährden. Insofern leistet die Schutzdigitalisierung einen Beitrag zur Reduzierung der Kosten der Bestandserhaltung, weil Originale nicht mehr oder zumindest seltener in die Benutzung gegeben werden. Der Einsturz des Stadtarchivs Köln hat derartigen Erwägungen eine zusätzliche Bedeutung verliehen.
4. Auf internen Servern oder auf Webservern bereit gestellte Archivgutdigitalisate helfen, die Arbeit der Archive rationeller zu gestalten. Bestellungen müssen nicht mehr entgegengenommen, bearbeitet und kontrolliert werden. Archivgut muss nicht mehr ausgehoben, transportiert und reponiert werden. Und schließlich müssen die bereits digitalisierten Bestände nicht nochmals im Interesse einzelner Benutzeranliegen reprografiert werden.

Das Landesarchiv Nordrhein-Westfalen (resp. seine Vorgängereinrichtungen) hat in den ausgehenden 1990er Jahren begonnen, analoges Archivgut zu digitalisieren. Das Technische Zentrum des Landesarchivs hostet gegenwärtig Archivgutdigitalisate im Umfang von rund 65 TB und macht diese zum Teil in den Lesesälen der verschiedenen Standorte zugänglich. So werden allein am Standort Brühl der Abteilung Rheinland jährlich mehr als 1,2 Mio. Archivalienseiten digital aufgerufen und eingesehen. Der erste online gestellte Band der Edition der nordrhein-westfälischen Kabinettsprotokolle für die 6. Legislaturperiode (1966-1970) hat bereits im ersten Jahr seiner Onlinestellung über 35.000 Zugriffe erzielt, während von der Printversion pro Jahr kaum mehr als eine Handvoll Exemplare verkauft werden.

Natürlich handelt es sich gegenwärtig nur um einen Bruchteil der rund 150 laufenden Kilometer Archivgut des Landesarchivs, der in die Digitalisierungsarbeiten einbezogen wird. So werden die wertvollen mittelalterlichen Urkundenbestände in Münster und in Düsseldorf Zug um Zug digitalisiert. Im vergangenen Jahr wurde die auf etwa 10-15 Jahre kalkulierte Entsäuerung und Digitalisierung von rund 1,2 Mio. Entnazifizierungsakten der Abt. Rheinland des Landesarchivs begonnen und die Digitalisierung der Gestapo-Akten auf der Basis von vorliegenden Sicherungsfilme fortgesetzt. Neben den Kirchenbüchern des alten Reichs und den Zivilstandsregistern des 19. Jahrhunderts wird die im Zusammenhang mit dem Einsturz des Kölner Stadtarchivs beeinträchtigte Kölner Personenstandsüberlieferung seit 2010 ebenfalls anhand der Zweitüberlieferung am Brühler Standort des Landesarchivs digitalisiert und soll dem Stadtarchiv Köln noch im Jahr 2011 zur Verfügung gestellt werden.

In den zitierten Beispielen wurde die Entscheidung für eine Digitalisierung aufgrund von Priorisierungskriterien im Landesarchiv getroffen. Hohes Alter und hoher Marktwert können dabei ebenso eine Rolle spielen wie großes Forschungsinteresse oder häufige Benutzernachfrage. Ein Sonderfall ist die Reaktion auf die Kölner Katastrophe, die zu dem Entschluss der Anfertigung einer digitalen Zweitüberlieferung führte, wobei man zum Verständnis ergänzen muss, dass Personenstandsregister immer

doppelt geführt werden und im Kölner Fall die Doppelüberlieferung durch den Einsturz des Stadtarchivs bis auf Weiteres nicht zugänglich und im schlimmsten Fall in Teilen zerstört ist.

Die Diskussion um die Vorgehensweise bei der Digitalisierung von archivalischen Quellen ist sicherlich in einzelnen Archiven, aber noch nicht in der Fläche des deutschen Archivwesens geführt worden. Auf die Frage, welche archivalischen Quellen vorrangig digitalisiert werden sollen und welche ggf. später zu behandeln sind, müssen Antworten gefunden werden. Dabei könnten die folgenden Merkmale für die Priorisierung eine Rolle spielen:

- Archivaliengattung
(z. B. Urkunden, Amtsbücher, Karten und Pläne, Akten, Bilder usw.)
- Historische Bedeutung
(z. B. herausragende Bedeutung im Kontext der Überlieferung eines Archivs)
- Nutzungsfrequenz
(z. B. häufig bestellte Bestände, häufig reproduzierte Bestände)
- Qualität des Bearbeitungsstands
(z. B. flach oder tief erschlossene Bestände, bereits über Online-Findmittel referenzierbare Bestände)
- Laufzeit
(z. B. bis 1918/nach 1918 oder bis 1945/nach 1945 – Grundvoraussetzung ist, dass keine rechtlichen Beschränkungen einer Veröffentlichung entgegen stehen)
- Geschlossenheit der Bestände
(z. B. nur ganze Bestände/Teilbestände oder auch verschiedene Bestandssplitter zu thematischen Gesichtspunkten)
- Bestandserhalterische Aspekte
(dürfte insbesondere dann von Bedeutung sein, wenn mit der Digitalisierung die Zugänglichkeit von archivalischen Quellen aufrecht erhalten werden kann, weil der Originalbestand zu stark geschädigt ist, um ihn weiterhin in der Benutzung zu belassen)
- Arbeitsorganisatorische Kriterien
(z. B. für eine Restaurierung/Entsäuerung vorgesehene Bestände, verfilmte/verfichte Bestände)

Eine Digitalisierung und Online-Bereitstellung archivalischer Quellen muss sich auch mit den erforderlichen technischen und fachlichen Standards auseinandersetzen. Hier gibt es einerseits Vorarbeiten bei den Archiven, andererseits aber auch bei der DFG (Praxisregeln Digitalisierung¹³). Im Einzelnen sind folgende Fragen zu klären:

¹³ <http://www.dfg.de/download/pdf/foerderung/programme/lis/praxisregeln_digitalisierung.pdf> in der Fassung vom April 2009.

- Welche Qualitätsanforderungen sind an die Digitalisate zu stellen?
(Bittiefe, Farbqualität, verwendete Formate)
- Können bestehende Qualitätsstandards angewendet oder fortgeschrieben werden
(z. B. DFG Praxisregeln) oder müssen neue Standards formuliert werden ?
- Welche technischen Metadaten müssen zur Verwaltung der Digitalisate bereitgestellt werden ?

Für eine technische Frage deutet sich bereits eine Vorentscheidung im deutschen Archivwesen an. Das Bundesarchiv, das nordrhein-westfälische Landesarchiv und eine Reihe weiterer Archive in Deutschland verwenden bereits jetzt den XML-gestützten Metadata Encoding and Transmission Standard METS¹⁴ als Containerformat für die Beschreibung von digitalen Objekten. METS kann auch hierarchisch strukturierte Codierungen aufnehmen, so wie sie mit EAD oder SAFT gegeben sind. Der Standard METS wird ebenso wie EAD von der Library of Congress verwaltet.

Wenn eine großflächige Digitalisierungskampagne im deutschen Archivwesen gestartet werden soll, dann sind neben den fachlichen und technischen Fragen sicherlich auch noch eine Reihe von organisatorischen Fragen zu klären. Dazu könnten Folgende gehören:

- Welche Möglichkeiten der operativen Umsetzung einer großflächigen Digitalisierung müssen in den Blick genommen werden ?
(z. B. Inhouse-Lösungen, Outsourcing, Einrichtung von Digitalisierungszentren, Nutzung bestehender Digitalisierungszentren usw.)
- Welche Kriterien sind bei einem Preis-Leistungs-Vergleich zu berücksichtigen ?
(z. B. Verweilzeit der Archivalien außer Haus, Gefährdungspotential, Transportumfänge und Versicherungsvolumina, Verfügbarkeit von High-End-Technologie usw.)
- Welche fachliche und organisatorische Unterstützung benötigen die Archive für großangelegte Digitalisierungskampagnen, welche fachliche und organisatorische Unterstützung kann sinnvoll geleistet und wie kann diese realisiert werden ?
(z. B. Informationsarbeit über die Fördermöglichkeiten, Stimulierung der Archive zur aktiven Beteiligung, organisatorische oder operative Unterstützung, Beratung bei der Vorgehensweise mit oder ohne Dienstleister usw.)
- Wie kann ein Großprojekt sinnvoller Weise starten ?
(z. B. mit einer vorgeschalteten Projektstudie, die eine Projektfineplanung erarbeitet, oder mit einem Pilotprojekt, das die erprobten Arbeitsschritte und Verfahren dokumentiert und entsprechende Empfehlungen für die Umsetzung gibt)

¹⁴ Vgl. dazu die informativen Online-Informationen des Bundesarchivs, <<http://www.bundesarchiv.de/daofind/mets/>>.

Fazit

Es gibt sicherlich einige Lehren, die sich aus den bisherigen Erfahrungen der Archive mit der Digitalisierung und Onlinestellung ihres Materials ziehen lassen. Dazu gehört die Einsicht, dass die archivische Erschließung unbedingt angelehnt an bestehende und anerkannte Standards erfolgen sollte, die auf den Grundprinzipien von Provenienz und Kontext aufbauen. Ebenso sollte für die digitale Strukturierung der Erschließungsinformationen auf einen anerkannten Standard wie etwa EAD rekurriert werden. Nur so lässt sich die Kommunikation von Erschließungsinformationen und Austauschbarkeit zwischen verschiedenen Plattformen gewährleisten, was eine Mindestanforderung an die Zukunftsfähigkeit und Nachhaltigkeit solcher Informationen darstellt. Damit ist nicht gemeint, dass die digitalen Systeme und Datenbanken, mit denen man arbeitet, vollkommen identisch mit Auszeichnungsstandards wie SAFT-XML oder EAD sein müssten. Sie müssen aber unbedingt in der Lage sein, ein fehlerfreies, wohlgeformtes und im Wesentlichen verlustfreies Austauschformat generieren zu können, so dass über solche Schnittstellen jederzeit die Kommunizierbarkeit der Daten gewährleistet ist. Die Erfahrung zeigt, dass man gut daran tut, solche Standardschnittstellen nicht in die Zukunft zu projizieren, sondern Systeme zu wählen, die darüber bereits nachweislich verfügen. Für manche kommerziellen Softwareanbieter scheint es eine fast unüberwindliche Hürde darzustellen, die versprochenen Standardschnittstellen auch wirklich zu realisieren.

Schließlich sollte man sich dessen bewusst sein, dass die Erschließungsinformationen für alle weiteren Arbeiten an Archivbeständen eine unabdingbare Voraussetzung darstellen. Erschließungsinformationen sind Referenzen und verleihen den archivischen Quellen Authentizität. Sie bieten der Forschung die notwendigen Kontextinformationen, um die Quellen kritisch beurteilen und auswerten zu können. Und schließlich sind sie Voraussetzung und Basis für eine Digitalisierung und Onlinestellung der Archivalien selbst.

Für das Hauptanliegen meines Beitrags, nämlich eine großangelegte Digitalisierungskampagne für archivalische Quellen sind ungeachtet der nicht unerheblichen Vorleistungen in den deutschen Archiven derzeit noch eine Reihe von Fragen zu beantworten. Manche Fragen werden sich durch eingehendere Analysen im Vorfeld klären lassen, doch wird es auch solche geben, die sich erst im operativen Betrieb stellen oder zumindest erst dann beantworten lassen. Sicher ist aber bereits jetzt, dass Standards zugrunde gelegt werden müssen, die eine Nachhaltigkeit großangelegter Digitalisierungskampagnen gewährleisten.

Ich danke Ihnen für Ihre Aufmerksamkeit !