

ARCHIVALISCHE ZEITSCHRIFT

88. Band

FESTSCHRIFT
HERMANN RUMSCHÖTTEL
ZUM 65. GEBURTSTAG

Herausgegeben von
Gerhard Hetzer und Bodo Uhl

Sonderdruck
im Buchhandel nicht erhältlich



2006

BÖHLAU VERLAG KÖLN WEIMAR WIEN

Wissensräume im Archiv. Überlegungen zur Zukunft archiverischer Erschließung

VON
GERHART MARCKHGOTT

Den meisten Archiven steht mit Beginn dieses Jahrhunderts eine Metamorphose ins Haus: der Übergang ins Digitale Zeitalter. Die Anforderungen sind vielfältig. Kaum eine Verwaltung, privat oder öffentlich, kommt mehr ohne digitale Aufzeichnungen und Dokumente aus, und allmählich werden die daraus resultierenden Herausforderungen deutlich. Die langfristige Erhaltung digitalen Verwaltungsmaterials, der ‚elektronischen Akten‘, ist weltweites Thema. Archivarinnen und Archivare müssen mit Technikern, Juristen und Betriebswirten in deren Terminologie zu kommunizieren lernen. Man hat sich zu verständigen über ungewohnte Differenzierungen – z.B. zwischen technischem Archiv und Langzeitarchiv –, über scheinbare Selbstverständlichkeiten – z.B. die Definition und Unterscheidung zwischen Original und Kopie – und über völlig neue Fragen – z.B. Manipulationssicherheit, Berechtigungsverwaltungen und spezielle Aspekte des Datenschutzes.

Auch die Zeiten technischer Autarkie der Archive sind vorbei. Sie brauchen Partner, Dienstleister, die nicht nur über Knowhow und Ressourcen zur langfristigen Speicherung digitalen Archivgutes verfügen, sondern auch bereit sind, Verständnis für die gemeinsame Aufgabenstellung zu entwickeln. Solche Leistungen Dritter – auch wenn diese demselben Verwaltungsapparat angehören – kosten Geld, und so entstehen den Archiven früher oder später finanzielle Lasten, die schnell den Rahmen normaler Budgets sprengen. Neue Fragen der Wirtschaftlichkeit, bisher höchstens bei Bauvorhaben virulent, stellen sich und müssen überzeugend beantwortet werden.

Könnte aus den bisherigen Ausführungen der Eindruck entstehen, dass nur ‚lebende‘, direkt mit Aktenproduzenten verbundene Archive diese Veränderungen zu spüren bekommen, so ist diese Hoffnung (?) trügerisch. Denn die Digitalisierung ist auch

im Umgang mit konventionellem, physischem Archivgut in den letzten Jahren unentbehrlich geworden, und verschiedenste, von einander isolierte Werkzeuge und Anwendungen verlangen nun nach System, nach längerfristiger Planung und umfassenden Konzepten. Mit der unvermeidlichen Umstellung von analogen auf digitale Reproduktionsmethoden ändern sich viele Rahmenbedingungen archivarischer Tätigkeit. Es ist der Ausgleich zu suchen zwischen den evidenten Vorteilen digitaler Bereitstellung (zeit- und ortsungebunden, beliebig und verlustfrei reproduzierbar, ein für allemal herzustellen und dadurch ebenso schonend für die Originale wie wirtschaftlich interessant für das Archiv) und den latenten Gefahren, die sich aus der Maschinengebundenheit ergeben (kurze Innovationszyklen, unkalkulierbare Fehlfunktionen in undurchschaubaren Systemen und – nicht zuletzt – sekundenschnell exekutierte Löschbefehle).

Je selbstverständlicher der freie Zugang zu öffentlichen Archiven wird – in Österreich erst eine Errungenschaft des späten 20. Jahrhunderts – und je höher in der Folge die Benutzerzahlen steigen, umso dringender und unausweichlicher wird die digitale Bereitstellung zum Schutz häufig nachgefragter Archivaliengruppen. Digitale Bereitstellung setzt – nicht notwendig, aber doch vernünftiger Weise – digitale Erschließung voraus, die Transformation der altbewährten Verzeichnisse, Findbücher, Repertorien, Karteien und sonstigen Erschließungsmittel in digitale Formen höherer Ordnung. Kaum ein Archiv wird sich auf Dauer damit begnügen können, Images (Fotos) der überlieferten Verzeichnisse anzubieten. Vielmehr müssen diese maschinenlesbar gemacht, strukturiert und nach Möglichkeit in Datenbanken ‚übersetzt‘ werden, um tatsächlich jenen Mehrnutzen zu erbringen, den heutige Benutzer/innen legitimer Weise erwarten.

Jeder einzelne dieser Punkte stellt eine Herausforderung dar und ergibt in der Regel ein Arbeitsprogramm für Jahre. Der Bogen der Maßnahmen reicht von vergleichsweise trivialen Aufgaben wie der Organisation von Massenscans und der Beschaffung von Massenspeichern bis zur Entwicklung von Gesamtstrategien und Visionen.

Der vorliegende Beitrag versucht sich in letzterer Kategorie. Vordergründig könnte man die folgenden Überlegungen auch

jener Art von Problemen zuordnen, die wir gar nicht hätten, wenn es keinen Computer gäbe. Aber eben nur vordergründig: In Wahrheit wird durch die EDV nur ein Grundproblem archivischer Erschließung wieder deutlich, an das wir so gewöhnt sind, dass wir es üblicher Weise nicht als Problem wahrnehmen bzw. nicht für lösbar halten: die Vieldimensionalität menschlichen Wissens.

Es wird nach Überzeugung des Autors schon in wenigen Jahren eine der wichtigsten Aufgaben intelligenter Archivierung sein, zur traditionellen Kernkompetenz der Erhaltung und Erschließung physischer Dokumente eine neue Kernkompetenz der Erschließung und Vermittlung von Wissen zu entwickeln: und zwar Wissen nicht im Sinne linearer, durch Vorträge oder Publikationen vermittelter Erkenntnisse, sondern dynamisches, interaktiv entstehendes, individuelles Wissen künftiger Benutzer/innen und Interessent/innen.¹ Die Archive werden sich auf diesem Gebiet gegen zahlreiche und (über?)mächtige Konkurrenten (Bibliotheken, Verlage und andere, wissensorientierte Unternehmen völlig neuer Art) behaupten müssen. Die folgenden Ausführungen sollen nicht nur diese Überzeugung näher erklären und untermauern, sondern den Leser/innen auch Gelegenheit bieten, die eigene Position bzw. Situation in diesem Szenario zu verorten und vielleicht auch schon konkrete Schlüsse oder Erkenntnisse daraus zu ziehen.

Faktum 1:

Jedes Archiv verfügt über Findmittel unterschiedlicher Form und Qualität. Wenn auch vielerorts schon einiges für deren Digitalisierung geschehen ist und manche Neuerfassungen in Datenbanken erfolgten, dominiert doch noch das gedruckte und ge-

¹ Die Begriffe Daten – Information – Wissen werden meist intuitiv verwendet und inhaltlich kaum unterschieden. Dieser Beitrag folgt der Mehrheit der Definitionen, die semantisch interpretierte Daten als Information bezeichnen; zu Wissen werden solche Informationen, wenn Menschen sie wahrnehmen, verarbeiten und reproduzieren oder nutzen können. (Vorsicht: andere Definitionen verwenden die Wörter „Information“ und „Wissen“ genau umgekehrt: Information als situationsbezogenes Wissen). Siehe dazu auch die Definition in: Rainer KUHLEN – Thomas SEEGER – Dietmar STRAUCH (Hrsg.), Grundlagen der praktischen Information und Dokumentation, Bd. 2, 5. Aufl. München 2004, S. 130 f.

bundene „Verzeichnis“. Dieser Zustand wird zunehmend als unbefriedigend empfunden. So läuft derzeit ein DFG-gefördertes Vorprojekt zur Retrokonversion solcher Verzeichnisse, um Entscheidungsgrundlagen für eine größere Initiative zur Konversion von archivischen Findmitteln zu schaffen. Angesichts der Fülle von ungelösten Problemen – Schrifterkennung, maschinengestützte Analyse bzw. Strukturerkennung und Annotation, Normalisierung etc. – ist mit einer flächendeckenden Retrokonversion in größeren Archiven wohl erst gegen Ende dieses Jahrzehnts zu rechnen. Dennoch wächst jetzt schon die Quantität maschinenlesbarer Findmittel von Jahr zu Jahr, und eines Tages wird dieser Konversionsvorgang abgeschlossen sein. Die heute noch papiergebundenen Informationen werden maschinenlesbar sein. Damit eröffnen sich völlig neue Möglichkeiten, die schon jetzt, im Vorfeld, bedacht und diskutiert werden sollten, um sie dann optimal nutzen zu können.

Faktum 2:

Wie schon in vergangenen Jahrhunderten leben auch heute noch die Archive vom Wissen ihrer Archivarinnen und Archivare. Dieses stellt neben (nicht nach!) den Archivalien selbst den wertvollsten Inhalt eines Archives dar. Ein Teil dieses Wissens sollte sich im Lauf der Tätigkeit in den Verzeichnissen manifestieren, die der/die Archivar/in erstellt. Ein weiterer Teil archivischen Wissens fließt oftmals in Publikationen ein, die – freilich nur in un- oder bestenfalls teilstrukturierter Form – zusätzliche Informationen, vertiefende Teilaspekte oder Verbindungen zu benachbarten Themen für spätere Leser festhalten. Vereinzelt wird auch ein Teil des Wissens durch Lehrtätigkeit weitergegeben. Aber ein beträchtlicher, vielleicht der größte Teil archivischen Wissens bleibt zeitlebens im Kopf, im individuellen Gedächtnis gefangen und wird allenfalls in Beratungen anlassbezogen und meist mündlich weitergegeben. So bedeutet auch heute noch in jedem Archiv der Abgang von Mitarbeiter/innen einen gravierenden, partiell unersetzlichen Wissensverlust.

Gemeinsamer Nenner dieser beiden Fakten ist das Verhältnis von Archiven zu Informationen und Wissen. Im Wortgebrauch dieses Beitrages enthalten Archivalien digital oder analog codierte Informationen über Objekte oder Ereignisse der Vergangenheit;

Verzeichnisse dagegen enthalten (Meta-)Informationen über den Inhalt von Archivalien oder über den Zugriff auf diese. Diese Informationskategorien bilden das Baumaterial für die Brücke vom forschenden Menschen der Gegenwart zum aufgezeichneten Geschehen der Vergangenheit. Aber Material alleine ergibt noch keine Brücke. Wie wenig hilfreich Informationen für sich alleine genommen im Archiv sind, lässt sich an unerfahrenen, nicht oder schlecht beratenen Archivbenutzer/innen beobachten, die hilflos in Verzeichnissen blättern oder sich tage-, manchmal wochenlang sinn- und erfolglos durch Papierberge wühlen.

Erst wenn es durch Anwendung fremden Wissens (= archivari-sche Beratung) oder eigenen Vorwissens (= Archiverfahrung) gelingt, die für die eigene Zielsetzung relevanten Informationen ausfindig zu machen bzw. herauszufiltern und aus deren Verknüpfung (im Fall der Verzeichnisse) handlungsleitendes Wissen oder (im Fall der Archivalien) erkenntnistiftendes Wissen zu machen, stellt sich Erfolg ein. Mit anderen Worten: *Erfolgskritischer Faktor für Archivrecherchen ist die Verknüpfung von Informationen.*

Wie wichtig die Beachtung dieser relativ trivialen Feststellung für die Entwicklung der Archive im digitalen Zeitalter sein wird, kann eine einfache Überlegung verdeutlichen. Stellen Sie sich vor, der gesamte Inhalt eines Archives, Archivalien wie Verzeichnisse, wäre maschinenlesbar erschlossen und ein ungeübter Benutzer müsste mit heutigen Methoden und Werkzeugen ein bestimmtes Thema oder Dokument recherchieren. Er würde in der Regel kläglich scheitern: Unmengen sinnloser ‚Treffer‘, nicht berücksichtigter Begriffsvarianten, veralteter Wörter und nicht erkannter Zusammenhänge würden zu einem Informationsgau führen, der jede vernünftige Arbeit unmöglich macht.

Die Bereitstellung digitalen oder digitalisierten Archivgutes bedeutet erst dann eine Qualitätsverbesserung, wenn adäquate Erschließungsmethoden zur Verfügung gestellt werden können.

Das Problem ist weder fachspezifisch noch neu. Schon seit Jahrzehnten wird interdisziplinär in Mathematik, Informatik, Sprachwissenschaft und verwandten Spezialdisziplinen an derar-

tigen Aufgabenstellungen gearbeitet.² Zwar sind Maschinen noch weit davon entfernt, menschliches Denkvermögen abbilden oder nachahmen zu können, aber die Methoden und Werkzeuge für die Generierung von und den Umgang mit Wissen gehen mittlerweile doch weit über das Niveau sequentieller Suchen und ‚einfacher‘ Suchmaschinen hinaus. Wenn man sich also der Einsicht nicht verschließt, dass maschinenlesbares Archivgut zwar eine Voraussetzung, aber bei weitem nicht das entscheidende Kriterium eines funktionsfähigen digitalen Archives darstellt, so ist die Auseinandersetzung mit neuen Methoden und Werkzeugen der Erschließung und Präsentation unverzichtbares Gebot zukunftsorientierter Archiventwicklung.

Das weite Gebiet intelligenter Erschließung digitalen Archivgutes hat grundsätzlich zwei Zugänge: einerseits die kontextbezogene Speicherung von Informationen, das heißt die Erhaltung von möglichst vielen Verknüpfungen und Zusammenhängen zur einzelnen Informationseinheit, und andererseits die situationsbezogene Bereitstellung und benutzergerechte Darstellung von Suchergebnissen. Auch diese beiden Aspekte, Seiten derselben Medaille, sind nicht neu, wurden bisher aber – zumindest in Österreich – selten im archivischen Kontext diskutiert.

Übersetzen wir die abstrakten Formulierungen in den Archivalltag: Ein Bestand wird geordnet entweder nach vorhandenen, von der Provenienz mitgegebenen Strukturen (Aktenplan etc.) oder – in Ermangelung solcher – nach archivischen oder benutzungsorientierten Kriterien (Chronologie, Alphabet, Pertinenz etc.). Definierte Einheiten (Schachteln, Bündel, Akten etc.) werden mit formalen und inhaltlichen Metadaten beschrieben: Signatur, Datum, Ort, Personen, Inhalt, Laufzeit, Umfang ... In der Regel korrespondiert die Primärordnung eines Verzeichnisses mit der Aufstellung der Archivalieneinheiten im Speicher (Akte 1 ff. in Schachtel 1 ff.). Um nun diese praktische, aber wenig kunden-

² Statt einer subjektiven Auswahl von Zitaten sei auf eine deutschsprachige ‚summa‘ der Informationswissenschaft verwiesen, den 1. Band (Handbuch zur Einführung in die Informationswissenschaft und -praxis) des Werkes von KUHLEN – SEEGER – STRAUCH (wie Anm. 1) insbesondere auf den Abschnitt B „Methoden“ (S. 125–376) und die umfangreichen weiterführenden Literaturhinweise der einzelnen Beiträge.

freundliche Anordnung inhaltlich leichter zugänglich zu machen, wird das Verzeichnis durch einen oder mehrere Indices erschlossen (Orte, Personen, Materien etc.), die jeweils zusätzliche, manchmal sogar hierarchisch strukturierte Erschließungsdimensionen repräsentieren. Ist später eine Fragestellung „richtig“, d. h. korrespondiert sie mit den Erschließungskriterien des Bestandes, so sind rasche und zufriedenstellende Rechercheergebnisse zu erwarten. Ist die Frage dagegen „falsch“ im Sinn der Erschließung, so sind Erfolge nur durch Knochenarbeit zu erzielen, die sowohl Forscher/innen als auch Archivgut und -personal strapaziert. Ein praktisches Beispiel erlebten viele österreichische Archive anlässlich der Aufarbeitung nationalsozialistischen Unrechtes durch Historikerteams. Es liegt auf der Hand, dass viele durchaus sinnvolle, ja wichtige Forschungen solcherart be- oder gar verhindert werden.

Schon seit Jahrzehnten versuchen Archive, die klassische, papiergebundene Erschließungslogik von Verzeichnissen (Listen) und Karteien durch den Einsatz maschinengestützter Tabellen zu verbessern. Datenbanken haben den Vorteil, dass jede Spalte (jedes Datenfeld) eine Erschließungsdimension abbildet. Sekunden-schnell kann die Sortierung = Ordnung eines Bestandes der aktuellen Fragestellung angepasst werden, wenn ein geeignetes Datenfeld zur Verfügung steht: nach Orten, Personen, Begriffen etc. Hier wird erstmals das Problem geeigneter Präsentationsformen akut, im Fachjargon die „Benutzeroberfläche“. Um die Bedienung möglichst einfach zu halten, werden oft die Suchergebnisse nicht als Listen, sondern in Formularform dargestellt. Damit geht aber Wesentliches verloren: es ist nämlich nicht zu erkennen, welche Informationen (Archivalien, Bezüge, Begriffe etc.) ‚rundherum‘ noch vorhanden sind, die sowohl zur Einordnung des eigentlichen Treffers als auch für weiterführende Erkenntnisse relevant wären. Es ist deshalb immer wieder zu beobachten, dass Benutzer/innen trotz vermeintlich höheren Suchkomforts in Datenbanken nach statischen Darstellungen = gedruckten Verzeichnissen verlangen, um darin zu blättern und sich so ein Bild von der Informationssituation im konkreten Bestand zu machen. Klassische Verzeichnisse bieten eben einen wichtigen Kontext, der in reinen Trefferlisten fehlt.