

Die Archivierung digitaler Unterlagen der Verwaltung: Ein Projekt der Staatlichen Archive Bayerns, gefördert von der Deutschen Forschungsgemeinschaft

Karl-Ernst Lupprian

In den 90er Jahren des letzten Jahrhunderts intensivierte sich die Diskussion über die Archivierung elektronisch „geborener“ Unterlagen. Ein Meilenstein war der vom Committee on Electronic Records des Internationalen Archivrats bearbeitete und im Februar 1997 herausgegebene „Führer für das Verwalten elektronischer Unterlagen aus archivischer Sicht“.¹ Bereits im April des gleichen Jahres erschien ein kommentierter Literaturbericht zu diesem Thema aus der Feder von Alf Erlandsson.²

1998 bildete sich bei der Generaldirektion der Staatlichen Archive Bayerns die – bis heute bestehende – Arbeitsgruppe „Archivierung digitaler Unterlagen der Verwaltung“ (AG AdUV), deren Mitglieder sich zum Teil auf informationstechnische Fragen, zum Teil aber auf die Aktenaussonderung spezialisiert hatten. Das erste greifbare Ergebnis der AG war eine 2001 veröffentlichte

¹ Guide for managing electronic records from an archival perspective, Paris 1997 (International Council on Archives, Studies 8).

² Alf Erlandsson, Electronic records management: a literature review, Paris 1997 (International Council on Archives, Studies 10).

Broschüre mit Empfehlungen für die Behörden beim Umgang mit elektronischen Unterlagen.³

Im Frühjahr 1999 hatte die Generaldirektion einen Antrag an die Deutsche Forschungsgemeinschaft gerichtet, der den „Entwurf eines organisatorisch-technischen Konzepts für die Archivierung digitaler Unterlagen der Verwaltung zur Sicherung von Quellengrundlagen für die künftige historisch ausgerichtete Forschung“ als Thema benannte. Im April 2000 bewilligte die DFG die beantragte BAT-IIa-Stelle für 24 Monate und 15.000 DM für Geräte. Auf die daraufhin durchgeführte Stellenausschreibung, die einen Projektbeginn für Oktober 2000 vorsah, meldete sich jedoch kein geeigneter Bewerber. Erst nach einer Reduzierung des Projektumfangs und dadurch auch der Anforderungen an die Stelle gab es eine nennenswerte Anzahl von Bewerbungen. Zum 1. Januar 2002 konnte das Projekt mit Dr. Rodrigo Readı Nasser als wissenschaftlichem Mitarbeiter beginnen. Da er auf eigenen Wunsch halbtags beschäftigt war, lief das Projekt bis zum 31.12.2005.

Ziel des Projekts war die Entwicklung eines Modells für die Aussonderung und Übernahme elektronischer Akten (auf diese Teilmenge digitaler Unterlagen hatten wir uns beschränkt), ihre dauerhafte Speicherung (im Sinne der Langzeitarchivierung), ihre nach Möglichkeit automatisierte Erschließung und ihre Nutzung vor Ort im Archiv und online.

Dementsprechend sah das Arbeitsprogramm ursprünglich folgende Teile vor:

1. Eine Bestandsaufnahme der Forschung
2. Die Ermittlung der konstitutiven (essentiellen) Merkmale eines digitalen Akts
3. Die Ermittlung der notwendigen Metadaten
4. Die Festlegung eines Übernahmeverfahrens und der zulässigen Dateiformate
5. Die Konvertierung digitaler Akten auf Mikrofilm als langzeitstabilem Träger und die Rückdigitalisierung vom Film zum Zweck der Benützung
6. Die Nutzung der Metadaten für die Erschließung
7. Den Aufbau einer Testumgebung

Organisatorisch wurde das Projekt vom Autor geleitet und von der oben genannten AG AdUV begleitet. Wir gingen damals davon aus, dass die Mitglieder der AG hinreichend Arbeitszeit in das Projekt investieren könnten. Dies war jedoch ein Trugschluss, denn schon nach kurzer Zeit waren alle von ihren Tagesgeschäften so in Anspruch genommen, dass eine Zuarbeit für das Projekt nur in relativ geringem Umfang geleistet werden konnte.⁴ Dieser Umstand wirkte sich insbesondere negativ auf die Projektdokumentation und die Öffentlichkeitsarbeit aus. Der Zeitaufwand hierfür darf nicht gering angeschlagen werden; es wäre durchaus förderlich gewesen, eine zweite Stelle (Archivar des gehobenen Dienstes) zu beantragen.

Als Projektpartner und künftige „Lieferanten“ elektronischer Akten konnten der Bayerische Verwaltungsgerichtshof und das Bayerische Staatsministerium der

³ Digitale Unterlagen. Entstehung – Pflege – Archivierung. Empfehlungen für die Behörden des Freistaates Bayern, hrsg. von der Generaldirektion der Staatlichen Archive Bayerns, München 2001 (online verfügbar unter <http://www.gda.bayern.de/digempfpdf>).

⁴ An dieser Stelle möchte ich meiner Kollegin Margit Ksoll-Marcon und meinen Kollegen Lothar Saupe, Bernhard Grau und Christoph Bachmann für die im Rahmen des Möglichen geleistete Mitarbeit, insbesondere für viele Stunden fruchtbarer Diskussion, herzlich danken.

Finanzen gewonnen werden. Im Lauf des Projekts wurden einige Kooperationspartner gefunden: Dr. Readi konnte die Erfahrungen des Kompetenzzentrums Germanistik an der Universität Trier (Prof. Gärtner) mit SGML nutzen. Das Institut für Softwaretechnologie der Universität der Bundeswehr München (Prof. Borghoff) unterstützte uns bei Problemen mit der Hardware und brachte Erfahrungen aus einem Archivierungsprojekt mit der Bayerischen Staatsbibliothek ein, zu deren Digitalisierungszentrum (Frau Dr. Dörr) ebenfalls fruchtbare Kontakte aufgebaut wurden. Der Seniorchef der Firma mfm Hofmaier in München, Dipl.-Ing. Dietrich Hofmaier, half uns mit wertvollen Fachinformationen, bestärkte unsere Sicht des Mikrofilms als langzeitstabilem Träger digitaler Information und vermittelte uns den Kontakt zur Firma Zeuschel, die ein Gerät für die Ausbelichtung elektronischer Daten auf Film auf den Markt gebracht hat und auch an der Weiterentwicklung eines Rückscanners mit OCR-Lesung der Metadaten Interesse zeigte.

Der Projektverlauf kann hier nicht im einzelnen dargestellt werden. Die Ergebnisse zeigen, dass nicht alle Ziele erreicht werden konnten. Erfolgreich war der Abschluss des Metadatenkatalogs für Sachakten der Verwaltung.⁵ Er konnte vollständig in die Anforderungen an das in Bayern in Einführung befindliche Dokumenten-Managementsystem (DMS) integriert werden und wird zum Teil auch in die aus X-DOMEA abgeleitete Aussonderungsschnittstelle X-Archiv einfließen. Des weiteren konnte das Modell eines digitalen Sachakts – unter Berücksichtigung hybrider Überlieferung – veröffentlicht werden.⁶

Wenig erfolgreich verlief dagegen die testweise Übernahme von digitalen Sachakten aus einer Behörde. Diese war sehr kooperationswillig, die das DMS betreuende Firma leider weniger. Es hat sich jedoch dabei gezeigt, dass unser Aktenmodell auch für sehr komplexe Schriftgutstrukturen geeignet ist. Die für die Darstellung von Struktur und Inhalt einer solchen Abgabe in einem Webbrowser notwendigen Programme wurden erstellt und an einem Dummy erfolgreich getestet.

Um einen durch die Verzögerungen beim oben genannten Übernahmetest drohenden Leerlauf zu vermeiden, wandte sich Dr. Readi einer Abgabe von Statistikdaten aus einer aufgelösten Grenzpolizeibehörde zu. Die Abgabe bestand aus Magnetbändern und Disketten im Format 5,25 Zoll. Die Bänder konnten gelesen werden, nachdem das Institut für Mathematik der Universität München ein passendes Laufwerk zur Verfügung gestellt hatte. Schwieriger war das Auslesen der Disketten; hier konnte uns das Institut Professor Borghoffs weiterhelfen, nur eine Diskette trotzte allen Versuchen. Problematisch war zunächst die Darstellung der etwa 10.000 Statistiken selbst, die mit dem Textprogramm HIT der Firma Siemens geschrieben waren. Da dessen Steuerzeichen nicht dokumentiert sind, half nur „trial and error“. Mittlerweile liegen die Statistiken als ASCII-Dateien auf CD-ROM vor.

Im November 2004 wurde ein interdisziplinäres Fachkolloquium veranstaltet, an dem knapp 40 Personen aus Archiven, Forschung und Verwaltung teilnahmen.

⁵ Digitale Unterlagen. Nr. 2: Metadaten für die Aussonderung und Archivierung digitaler Sachakten. Empfehlungen für die Behörden des Freistaates Bayern, hrsg. von der Generaldirektion der Staatlichen Archive Bayerns, München 2004.

⁶ Rodrigo Readi Nasser und Karl-Ernst Lupprian, Ein Modell für die elektronische Darstellung hybrider Verwaltungsakten in Archiven. In: Archivalische Zeitschrift 85 (2003), S. 147-173.- Erweiterte Fassung: Rodrigo Readi Nasser, Karl-Ernst Lupprian und Lothar Saupe, The Key Role of Metadata for Permanent Preservation of Digital Records in the Archival Environment – A Message from the Far Future. In: Metadata in Preservation. Selected Papers from an ERPANET Seminar at the Archives School Marburg, 3-5 September 2003, hrsg. von Frank M. Bischoff, Hans Hofman, Seamus Ross (Veröffentlichungen der Archivschule Marburg – Institut für Archivwissenschaft – Nr. 40), Marburg 2004, S. 185-203.

Für die Langzeitarchivierung elektronischer Unterlagen sollte die Eignung des Mikrofilms als Medium geprüft werden. Nach dem im Projekt entwickelten Modell besteht eine Aktenabgabe aus einer geordneten Reihe von Dateien, nämlich einer Metadatendatei der Abgabe (ASCII), Metadatendateien für jeden Aktenschnitt (ASCII) und Primärdateien (Bildformat), auf die von den jeweiligen Metadatendateien verwiesen wird. Aus dieser hierarchischen Ordnung wird eine Serie von Dateien gebildet (hierfür ist die Namengebung der Dateien wesentlich). Diese Serie wird auf Mikrofilm ausbelichtet. Dabei empfiehlt es sich, den für die Metadatendateien verwendeten Font vorher in einen OCR-Font (z.B. OCR-A oder OCR-B) genügender Größe zu wandeln, damit für den späteren Rückscan eine fehlerfreie Digitalisierung gewährleistet ist. Dieses Verfahren soll in Kooperation mit der Firma Zeutschel, die sowohl über einen Ausbelichter als auch einen schnellen Rückscanner verfügt, entwickelt werden.⁷

Die Erschließung digitaler Akten soll in Findmittel münden, die aber kein Sonderdasein neben denen für die Papierunterlagen führen dürfen. Es ist vielmehr eine integrierte Lösung anzustreben, die online zugänglich sein muss. Zum Abschluss unseres Projekts konnten die Grundlagen dafür gelegt werden,⁸ ein Prototyp ist in Arbeit.

⁷ Dabei sind Erfahrungen aus dem InnoNet-Projekt ARCHE zu berücksichtigen. Problematisch können die Mengen an Film werden, die – auch wenn man 16 oder mehr Aufnahmen auf einem Filmframe unterbringt – ein Archiv bei der Wahl dieser Lösung zu verwalten haben wird. Ein Ausweg wäre die Ausbelichtung des Bitstroms selbst auf Farbmikrofilm, wie im Projekt PEVIAR (www.peviar.ch) vorgeschlagen. Dabei ließe sich eine Speicherkapazität von ca. 700 MB auf einem Mikrofiche erreichen, und auch audiovisuelle Unterlagen könnten langzeitstabil gesichert werden. Dieser Weg bedingt jedoch eine sorgfältige Dokumentation zumindest der Algorithmen sowie der Dateiformate, sonst kann der Bitstrom nicht interpretiert werden.

⁸ Rodrigo Readi Nasser, Findmitteldatenbanken. Ein Beitrag zur Erschließung hybrider Akten. In: Archivalische Zeitschrift 87 (2005), S. 235-248.