

Sonderveröffentlichungen der Staatlichen Archive Bayerns
Nr. 2

Virtuelle Welten im Magazin

Aussonderung, Aufbewahrung, Sicherung und Nutzung

Vorträge der 5. Tagung des Arbeitskreises
„Archivierung von Unterlagen aus digitalen Systemen“
in München, 5. und 6. März 2001

Herausgegeben von
Karl-Ernst Lupprian

München 2003

Sonderveröffentlichungen der Staatlichen Archive Bayerns

(ISSN 1618-0739)

hrsg. von der Generaldirektion der Staatlichen Archive Bayerns.
Schriftleitung: Albrecht Liess

- Nr. 1: **Codex Laureshamensis.** Urkundenbuch der ehemaligen Fürstabtei Lorsch. Faksimileausgabe der Handschrift im Staatsarchiv Würzburg, hrsg. von der Generaldirektion der Staatlichen Archive Bayerns und dem Kuratorium Weltkulturdenkmal Kloster Lorsch e.V. Neustadt a. d. Aisch 2002, 2 Bände im Format 45 cm (h) x 32 cm (b).
Bd. 1: Faksimile mit 464 faksimilierten Farbseiten, Bd. 2: Aufsätze und Register (ISBN 3-7686-5185-1). 599,00
- Bestellungen über den Verlag Degener & Co., Postfach 1360, 91403 Neustadt a.d. Aisch, Tel. 09161/886039
- Nr. 2: **Virtuelle Welten im Magazin.** Aussonderung, Aufbewahrung, Sicherung und Nutzung. Vorträge der 5. Tagung des Arbeitskreises „Archivierung von Unterlagen aus digitalen Systemen“ in München, 5. und 6. März 2001, hrsg. von Karl-Ernst Lupprian (ISBN 3-921635-71-3). im Druck

Virtuelle Welten im Magazin

Aussonderung, Aufbewahrung, Sicherung und Nutzung

(vakat)

Sonderveröffentlichungen der Staatlichen Archive Bayerns
Nr. 2

Virtuelle Welten im Magazin

Aussonderung, Aufbewahrung, Sicherung und Nutzung

Vorträge der 5. Tagung des Arbeitskreises
„Archivierung von Unterlagen aus digitalen Systemen“
in München, 5. und 6. März 2001

Herausgegeben von
Karl-Ernst Lupprian

München 2003

Sonderveröffentlichungen der Staatlichen Archive Bayerns
herausgegeben von der Generaldirektion der Staatlichen Archive Bayerns

Schriftleitung: Albrecht Liess

Nr. 2: Virtuelle Welten im Magazin. Aussonderung, Aufbewahrung, Sicherung und Nutzung.
Vorträge der 5. Tagung des Arbeitskreises „Archivierung von Unterlagen aus digitalen Systemen“ in
München, 5. und 6. März 2001, herausgegeben von Karl-Ernst Lupprian

© Generaldirektion der Staatlichen Archive Bayerns, München

Satz und Gestaltung: Karin Werth

Gesamtherstellung: Danuvia Druckhaus Neuburg GmbH, 86633 Neuburg a.d. Donau

ISSN: 1618-0739

ISBN: 3-921635-71-3

Inhaltsverzeichnis

KARL-ERNST LUPPRIAN, Zur Einführung	6
HERMANN RUMSCHÖTTEL, Grußwort	7
JÜRGEN EHRMANN, Technisch-organisatorische Anforderungen an den elektronischen Rechtsverkehr	8
KLAUS-PETER VOLKMANN, Einführung und Einsatz von VIS im Geschäftsbereich des Bayerischen Staatsministeriums für Landesentwicklung und Umweltfragen	21
THOMAS SCHÄRLI, Das Projekt ELGAR zur elektronischen Unterstützung von Geschäfts- und Aktenregistrierungsprozessen im Kanton Basel-Stadt	29
THEKLA KLUTTIG, Zum DOMEA-Konzept und seiner Umsetzung im Freistaat Sachsen – ein Werkstattbericht	41
MICHAEL WETTENGEL, Anwendungsspezifische Fachkonzepte zur Umsetzung archivischer Anforderungen für die Aussonderung elektronischer Akten: Das Beispiel des Fachkonzepts FAVORIT-Office-Flow	48
CHRISTIAN KEITEL, Die archivische Bewertung elektronischer Statistiken	63
BETTINA MARTIN-WEBER, Erschließung und Nutzbarmachung digitaler Unterlagen im Bundesarchiv	69
NILS BRÜBACH, Der Einsatz von Metadaten-Sets in Schriftgutverwaltung und Archiv	77
FRANK M. BISCHOFF, Empfehlungen für die elektronische Archivierung. Ein Normentwurf der <i>Association Française de Normalisation</i>	89
MARIANNE DÖRR, Archivierung von digitalen Dokumenten in Bibliotheken	99
BJÖRN DEHMS, ANDREAS ENGEL und ULRICH MEYER, XML – ein Format zur dauerhaften Aufbewahrung und Nutzung einfacher, relationaler Datenbanken?	104
KARL-ERNST LUPPRIAN, Rematerialisierung digitaler Unterlagen – ein Weg zur Langzeitarchivierung? Das Projekt HD-Rosetta	118
Anhang:	
Programm	122
Teilnehmerverzeichnis	123
Inhaltsübersicht über die bisher erschienenen Tagungsbände des Arbeitskreises „Archivierung von Unterlagen aus digitalen Systemen“	125

Zur Einführung

Die 5. Tagung des „Arbeitskreises für die Archivierung von Unterlagen aus digitalen Systemen“ war mit 39 Teilnehmern aus Archiven in Deutschland, Österreich und der Schweiz, Bibliotheken, Verwaltung, Justiz, Forschung und Wirtschaft ähnlich gut besucht wie die vorangegangenen Veranstaltungen. Der Tagungsband, der nun vorgelegt werden kann, ist ungeachtet seines späten Erscheinens nach wie vor aktuell.

Eine gewisse Uneinheitlichkeit des Druckbilds ließ sich nicht vermeiden, da einige Beiträge nur als Powerpoint-Präsentationen bzw. PDF-Dateien mit eingebetteten Präsentationen vorlagen, deren Nachbearbeitung für den Druck nur eingeschränkt möglich ist. Angesichts der geringen Zahl dieser Beiträge erwies sich eine hybride Publikation (Buch und CD-ROM) wie beim Mannheimer Tagungsband als nicht opportun.

Die Reihenfolge der Beiträge versucht den Stoff in drei thematische Blöcke zu gliedern. Ein erster Block mit den Beiträgen von Jürgen Ehrmann, Klaus-Peter Volkmann und Thomas Schärli geht die Problematik aus der Sicht der Registraturbildner an.

Der zweite Block mit den Beiträgen von Thekla Kluttig, Michael Wettengel und Christian Keitel beschäftigt sich mit Bewertungs- und Aussonderungsfragen. Die Beiträge von Bettina Martin-Weber, Nils Brübach, Frank M. Bischoff und Marianne Dörr lassen sich unter den Schlagworten Metadaten, Erschließung, Benützung und Sicherung fassen.

Ein letzter Block mit den Beiträgen von Björn Dehms, Andreas Engel, Ulrich Meyer sowie Karl-Ernst Lupprian befasst sich mit technischen Problemen der dauerhaften Aufbewahrung elektronischer Unterlagen.

Eine Inhaltsübersicht der bislang erschienen vier Tagungsbände, das Tagungsprogramm und ein Teilnehmerverzeichnis beschließen den Band.

Den Autoren der Beiträge, die auch bei der Entstehung des Bandes für Rückfragen jederzeit hilfreich zur Verfügung standen, sei nochmals herzlich gedankt, besonders aber meiner Kollegin Karin Werth, die für die Gestaltung verantwortlich zeichnet und eine Reihe technischer Probleme zu lösen wusste.

München, im Februar 2003

Karl-Ernst Lupprian

Grußwort

Die diesjährige Tagung des Arbeitskreises hat den Charakter eines kleinen Jubiläums. Der institutionell nirgendwo eingebundene, völlig unabhängige Arbeitskreis hat es nunmehr ein halbes Jahrzehnt lang mit Erfolg verstanden, kompetente Gesprächspartner aus Archiven, Forschung und Verwaltung an einen Tisch zu bringen. Wenn man die Referate der bisherigen vier Tagungen in Münster, Ludwigsburg, Koblenz und Mannheim Revue passieren lässt, kann man nicht nur den dornigen Weg zu allmählich Gestalt annehmenden praktischen Lösungen nachvollziehen, sondern auch feststellen, dass sich Archive und Verwaltung ihrer gemeinsamen Verantwortung für die Bewahrung der behördlichen Überlieferung bewusst geworden sind.

So ist es nur folgerichtig, dass heute von den knapp 40 Teilnehmern nicht weniger als acht aus Verwaltung und Justiz in Baden-Württemberg und Bayern kommen. Ich freue mich, Herrn Ehrmann vom Justizministerium Baden-Württemberg, Herrn Volkmann vom Bayerischen Staatsministerium für Landesentwicklung und Umweltfragen, Herrn Fritsch vom Bayerischen Staatsministerium des Innern, Herrn Schmidbauer von der Regierung von Schwaben, Herrn Meier vom Oberlandesgericht München sowie die Herren Blabsreiter, Handl und Steidl vom Bayerischen Verwaltungsgerichtshof begrüßen zu dürfen. Mit den drei letztgenannten verbindet uns die Partnerschaft in einem von der Deutschen Forschungsgemeinschaft geförderten Forschungsprojekt zur Langzeitarchivierung.

Die Forschung ist durch Herrn Ulrich Meyer vom Institut für Wirtschafts- und Verwaltungsinformatik der Universität Koblenz-Landau vertreten, der anstelle des leider verhinderten Herrn Engel zu uns gekommen ist. Sein Institut war unter anderem an der Entwicklung des DOMEA-Konzepts beteiligt und bemüht sich um Lösungen für die Archivierung von digitalen Akten und Unterlagen aus Datenbanken.

Wenn man Lösungen realisieren und in praxistaugliche Produkte umsetzen will, ist die Kooperation mit der Wirtschaft unerlässlich. Ich begrüße Herrn Hofmaier von der Firma MFM in München, mit der uns ein langjähriges Geschäftsverhältnis verbindet, und Herrn Stubenvoll von der Firma Infowerk AG in Nürnberg. Beide haben zum Zustandekommen dieser Tagung mit Sachspenden beigetragen, wofür ihnen unser aller Dank gebührt.

Die Zeit erlaubt es mir leider nicht, alle Kolleginnen und Kollegen aus Archiven und Bibliotheken in Deutschland, Österreich und der Schweiz namentlich zu begrüßen. Ich heiße Sie bei uns herzlich willkommen.

Gestatten Sie mir zum Schluss einen knappen Überblick über die Situation in Bayern. Im Rahmen der Einführung der neuen Allgemeinen Geschäftsordnung, die dem Einsatz der IuK-Technik für die Vorgangsbearbeitung höchste Priorität zuweist, hat der Bayerische Ministerrat einen Beschluss gefasst, der die Beteiligung der Staatlichen Archive Bayerns bei der Einführung entsprechender Systeme festschreibt.

Die Generaldirektion der Staatlichen Archive Bayerns hat im Jahre 1998 eine Arbeitsgruppe eingesetzt, die Empfehlungen für Behörden und Gerichte des Freistaats Bayern zum Umgang mit digitalen Unterlagen von deren Entstehung bis zur Archivierung erarbeitet hat. Diese heuer publizierten Empfehlungen, die Sie auch in ihren Tagungsunterlagen finden, sind erfreulich positiv angenommen worden; der fachliche Dialog zwischen Archiven und Verwaltung ist angelaufen.

Von Ihrer Tagung erhoffen wir uns neue Erkenntnisse und Anregungen für die Bewältigung einer der wohl anspruchsvollsten Anforderungen, denen wir uns am Beginn des digitalen Zeitalters gegenüber sehen. Ich wünsche Ihnen fruchtbare Gespräche und einen angenehmen Aufenthalt.

Prof. Dr. Hermann Rumschöttel
Generaldirektor der Staatlichen Archive Bayerns

Technisch-organisatorische Anforderungen an den elektronischen Rechtsverkehr

Von

Jürgen Ehrmann

Zusammenfassung

Die weltweit im Zuge des WWW forcierten Arbeiten für einen sicheren eCommerce führen unausweichlich dazu, dass auch die Forderung nach elektronischer Antragsstellung bei Verwaltungsdienstleistungen aufkommt und auch „bedient“ werden kann. Die damit vorhandene Akzeptanz, vom Papier auf den elektronischen Austausch umzusteigen und die im Bereich eCommerce entwickelten und eingeführten Technologien auch für den elektronischen Rechtsverkehr einzusetzen, soll auch genutzt werden, um über Standardisierungen Justizabläufe wirtschaftlicher zu gestalten. Die in den letzten Jahren vorangetriebene Ausstattung der Justiz mit luK-Technik ermöglicht die Nutzung des elektronischen Datenaustauschs mit den damit verbundenen Vorteilen und Risiken. Die Konferenz der Justizministerinnen und -minister vom 7. bis 9. Juni 1999 hat die Bund-Länder-Kommission für Datenverarbeitung und Rationalisierung in der Justiz (BLK) beauftragt, Vorschläge für die Einführung des elektronischen Rechtsverkehrs vorzulegen. Neben den rechtlichen Fragestellungen sind grundsätzliche technische Rahmenbedingungen festzulegen, um eine einheitliche, wirtschaftliche und einfache Nutzung zu ermöglichen. Die praktische Umsetzbarkeit erfordert Festlegungen auf Standards und interoperable Produkte.

Im Rahmen der BLK-Arbeitsgruppe „Elektronischer Rechtsverkehr“ hat der Arbeitskreis „Querschnittsfragen Technik“ (AK-QUER) den Auftrag, die technischen und IT-organisatorischen Rahmenbedingungen einer elektronischen Kommunikation unter den Justizbehörden und mit den Verfahrensbeteiligten sowie der elektronischen Aktenführung zu analysieren und hierüber einen Bericht zu erstellen. Der AK-QUER hat sich bei der Erarbeitung seines Berichts mit vielen Kommunikationspartnern der Justiz, wie z.B. auch dem Arbeitskreis der Referenten der Landesarchivdirektionen, abgestimmt.

In dem Vortrag werden die konkreten organisatorisch-technischen Anforderungen der Justiz und die Empfehlungen des AK-QUER für die Bereiche

- Verschlüsselung und Elektronische Signatur
- Austauschformate und Austausch von Verfahrensdaten
- Elektronische Akte
- Archivierung (in den Justizbehörden und zur Übergabe an die Archive)

aufsetzend auf den technologischen und marktspezifischen Rahmenbedingungen dargestellt. Kurz angerissen werden auch die technischen Grundlagen symmetrischer und asymmetrischer Verschlüsselungsverfahren sowie des XML-Standards für die Datendarstellung und den Datenaustausch.

Foliensatz:

**BADEN-
WÜRTTEMBERG**



Technisch-organisatorische Anforderungen an den elektronischen Rechtsverkehr

5. Tagung des Arbeitskreises „Archivierung und Unterlagen
aus digitalen Systemen“
München, 5. März 2001



**JUSTIZ-
MINISTERIUM**

© 2000 Justizministerium Baden-Württemberg
RD Ehrmann, 5. März 2001

**BADEN-
WÜRTTEMBERG**

Elektronischer Rechtsverkehr



eCommerce
Online Banking
WWW

Mahnverf.
EGB
EHR

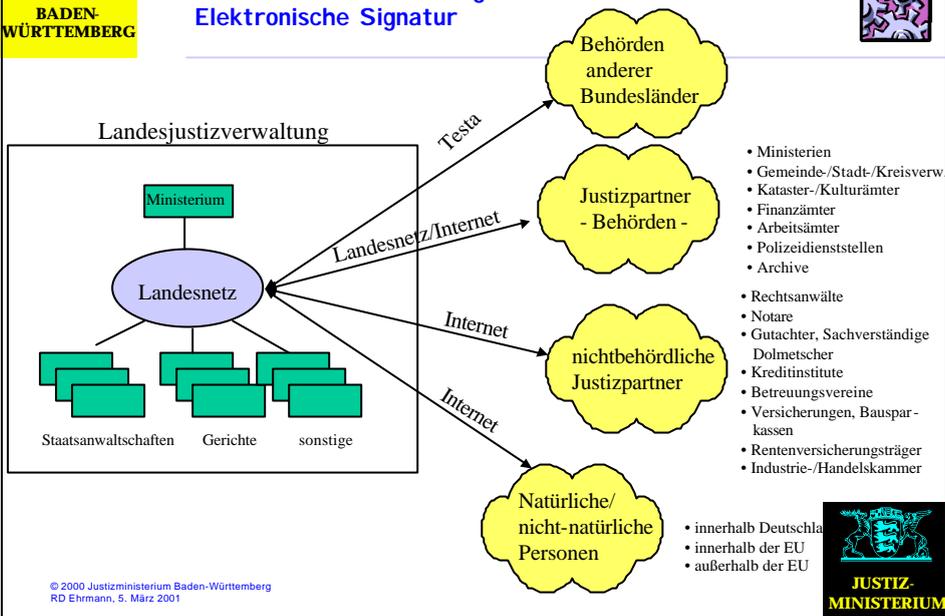
Justiz-EDV
Justiz-Netze



**JUSTIZ-
MINISTERIUM**

© 2000 Justizministerium Baden-Württemberg
RD Ehrmann, 5. März 2001

Kommunikationsbeziehungen der Justiz Elektronische Signatur



Auftrag des AK „Querschnittsfragen Technik im elektronischen Rechtsverkehr“



Verschlüsselung und elektronische Signatur

Austauschformate und Austausch von Verfahrensdaten

Elektronische Akte

Archivierung

Der Arbeitskreis „Querschnittsfragen Technik“ hat den Auftrag, die technischen und die IT-organisatorischen Rahmenbedingungen einer elektronischen Kommunikation unter den Justizbehörden und mit den Verfahrensbeteiligten sowie der elektronischen Aktenführung zu analysieren und hierüber einen Bericht zu erstellen. Der Bericht soll sich auch auf die Notwendigkeit der Festlegung von Standards und Formaten erstrecken und die Möglichkeiten für die Ausgestaltung der Schnittstellen zwischen Justizbehörden und anderen Stellen sowie Verfahrensbeteiligten aufzeigen.



Abstimmung des AK-QUER mit:

- ReferentenAK der LAD
- Bundesrechtsanwaltskammer
- Bundesnotarkammer
- KOOPA ADV (Innenverwaltungen)
- Landessystemkonzepten
- Bundesamt f. Anerkennung ausl. Flüchtl.
- TeleSec, SignTrust, Datev, BOS-Bremen
- Piloten der Justiz



Grundanforderungen an die Sicherheit

Integrität
Authentizität
Vertraulichkeit
Verbindlichkeit

State of the art

Benutzer-Passwort
Biometrische Verfahren
Aufbau geschlossener Benutzergruppen
Verschlüsselung
Elektronische (digitale) Signatur
Versiegelung von übertragenen Daten

Verschlüsselung und elektronische Signatur Kryptographie - Symmetrische Verfahren



Symmetrische Verschlüsselungsverfahren



schnell!
sicherer Schlüsselaustausch??
Authentisierung??

DES (Data Encryption Standard), DES-3 oder
IDEA (International Data Encryption Algorithm)

Verschlüsselung und elektronische Signatur Kryptographie - Public Key Verfahren

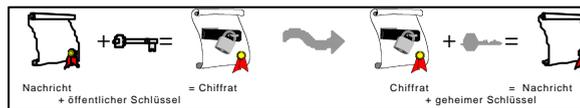


Asymmetrische Verschlüsselung, Public Key, RSA

Schlüsselpaar:

- Öffentlicher Schlüssel
- Geheimer Schlüssel

- a. Der Absender will sicherstellen, daß ausschließlich der Empfänger die Nachricht lesen kann:



- b. Der Empfänger will Gewißheit haben, daß die Nachricht vom Absender stammt:

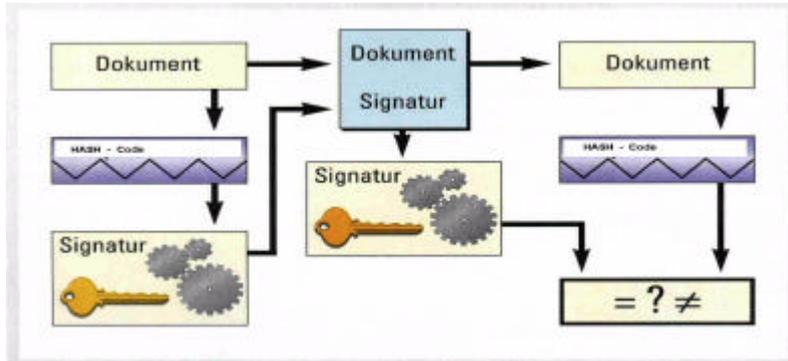


langsam, lange Schlüssel (Diskette oder Chipkarte) ---> Kombinierte Verfahren



Elektronische Signatur/Unterschrift:

Asymmetrisches Verschlüsselungsverfahren
Öffentlicher Schlüssel
Geheimer Schlüssel
+ Hashfunktion



© 2000 Justizministerium Baden-Württemberg
RD Ehrmann, 5. März 2001



Praktische Einsatzvoraussetzung:
Interoperabilität von Produkten

Test interoperabler Produkte (mail-orientiert)
im Sphinx-Projekt des BSI:

- Mailtrust-Standard 2.0
- Utimaco und CoConet
- Festlegung der öffentlichen Verwaltung (KOOPA ADV) auf Mailtrust 2.0

Online-Transaktionen (interoperable Produkte entstehen erst):

- BW-Card in den eBürgerdiensten
- Bremen online, Internetantragstellung im Mahnverfahren (OSCI)

Der AK-QUER empfiehlt:

- auf Konzepten und Standards von MailTrust aufzusetzen (X.501, X.509, PKCS#7, PKCS#11)
- bei mailorientiertem Austausch: MailTrust 2.0 u.höher
- bei online-Transaktionen: SSL 3.0, S-HTTP
- What You See Is What You Sign zu gewährleisten
- Piloten – auch abweichende – sind erlaubt

© 2000 Justizministerium Baden-Württemberg
RD Ehrmann, 5. März 2001



Elektronische Signatur Empfehlungen des AK-QUER (TrustCenter)



TrustCenter

- generieren sichere Schlüsselpaare
- ordnen öffentlichen Schlüssel Personen (oder Institutionen) zu
- dokumentieren dies in Zertifikaten
- führen in den Zertifikaten Zusatzinformationen (z.B. Attribute, Gültigkeitsdauern, Vertretungsberechtigungen)
- verwalten die Zertifikate (z.B. RevocationList)
- sperren Zertifikate

Nach (altem, strengen, teuren) dt. SigG zertifiziert

- TeleTrust
- SignTrust/BNotK

SigG:

(Die EU-SIG-RILi ist im Februar in das deutsche SigG umgesetzt worden)

Signatur: dient Authentifizierung

Qualifizierte Signatur: sicher erzeugt, sicher zugeordnet
und mit Zertifikat verknüpft

TrustCenter: Betrieb ist genehmigungsfrei und anzeigepflichtig

Akkreditiertes TrustCenter: Akkreditierung und Gütesiegel



Elektronische Signatur Empfehlungen des AK-QUER (TrustCenter)



Interoperabilität zwischen TrustCentern (Zertifikaten):

- X.509
- I S I S (Industrie Signature Interoperability Specification)
(derzeit nur SignTrust/BNotK, TeleSec strebt an)
- EESSI (European Electronic Signature Standardisation Initiative)

Der AK-QUER empfiehlt:

- Signaturen I S I S-konformer TrustCenter einzusetzen
- oder von TrustCenter-Anbietern, die sich auf zukünftige I S I S-Konformität verpflichtet haben
- der absehbare Wandel in den Formaten der Zertifikate (z.B. Stellenzahl zur Beschreibung der Attribute) in Projektplanung aufzunehmen und deshalb
- einzelne Zertifikatsteile zunächst restriktiv zu nutzen (Attributfeld nicht zu nutzen bzw. nur in wenigen Fällen)



Austauschformat für Schriftsätze:

Word? (2.0, 6.0, 8.0, Word2000)
RTF?
Wordperfect?
StarWriter?
....

Layouttreue?What-You-See-Is-What-You-Sign?Strukturierte Daten im Text?

Neuer Industrie-Standard: XML

**XML - Vorteile**

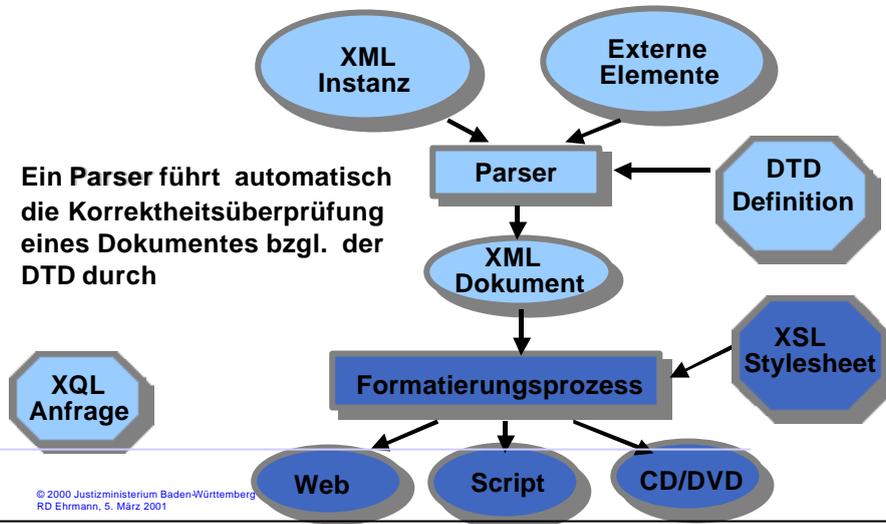
- **Trennung von Inhalt (XML), Präsentation (XSL) und Logik/Struktur (DTD)**
- **HTML-Nachfolger (Internet)**
- **Metasprachenkonzept**
- **Lesbar für Mensch und Maschine**
- **Einfach**
- **Offener Standard**
- **Erweiterbar (Justizdatensatz)**
- **Unterstützt UniCode (ASCII-Nachfolger)**
- **Hohe Stabilität (SGML-Subset)**
- **Layouttreue, Drucker oder CD**
- **Breite Herstellerunterstützung**
(zB Microsoft Word2000, IBM, SAG, Oracle, etc)

XSL, XQL, DTD HTML

```
<?xml version=1.0!>
<Antragsteller>
  <nname>Maier</nname>
  <vname>Franz</vname>
  <titel>Dr</titel>
  <wohntort>
    <stadt>Stuttgart</stadt>
    <plz>70000</plz>
    <strasse>Schillerplatz</strasse>
  </wohntort>
</Antragsteller>
```



Ein Parser führt automatisch die Korrektheitsüberprüfung eines Dokumentes bzgl. der DTD durch



```

<-- Beispiel fuer DTD, Datenstrukturen -->
<!-- ===== -->
<!DOCTYPE Stammdaten [
  <!ELEMENT Antragsteller (nname, vname, titel, wohnort)>

  <!ELEMENT nname (#PCDATA)>
  <!ELEMENT vname (#PCDATA)>
  <!ELEMENT titel (#PCDATA)>

  <!ELEMENT wohnort (stadt, plz, strasse)>
  <!ELEMENT stadt (#PCDATA)>
  <!ELEMENT plz (#PCDATA)>
  <!ELEMENT strasse (#PCDATA)>

]>
  
```





Der AK-QUER empfiehlt:

- **kurzfristig** verfügbar und damit sachgerecht sind
 - das RTF -Format (im Klartext lesbar und von vielen Editoren unterstützt)
 - das PDF -Format (weitverbreitet, layouttreu, komprimiert, aktive Inhalte sind ausschließbar)
- **kurz/mittelfristig** sind XML -Dokumente anzustreben, dabei sind
 - XML + XSLT zu signieren oder die jeweils gültige XSL vorzuhalten
 - DTD justizspezifisch zu definieren



Bildformate: TIFF, GIF, JPG, etc.....

Der AK-QUER empfiehlt:

- zum Austausch reiner Bildformate soll grundsätzlich das TIFF -Format verwendet werden
- nach der Abstimmung mit dem ReferentenAK der LAD wird die Übergabe von in TIFF -Format umgewandelten Schriftstücken akzeptiert, um Konvertierungsprobleme zu vermeiden
- elektronische Signaturen werden vor Übergabe aufgelöst und im Klartext übergeben (Gewährleistung der Richtigkeit „von Amts wegen“)
- wg. der weiten Verbreitung und der komprimierten Daten sieht der AK-QUER im PDF -Format eine mögliche Alternative



Papierakte ist geltende Vorschrift

Elektronischer Schriftverkehr forciert elektronische Aktenführung

Aktenteile liegen elektronisch vor

- Aus Geschäftsstellenautomation (SI JUS, HADES, JUSTUS, EUREKA, BajTech)
- Im Elektronischen Grundbuch
- Geschäftsberichte im Elektronischen Handelsregister

Gemeinsame Bearbeitung setzt elektronische Akte voraus
Übergabe an andere erfordert einheitliche Schnittstelle
und Suchstrategien

DOMEA – Konzept ist für Verwaltungsakten Vorgabe



Akte

Nur Papier

- +Handzeichen, Unterschriften im Original
- +Vertrautes Medium
- elektronische Teile ausdrucken
- Platzbedarf
- Suche schwierig
- paralleles Arbeiten (Ri und SE) schwierig

Hybridakte

- +passend für Einstieg
- organisatorisch schwierig
- wenig Vorteile

Nur elektronisch

- +Suche
- +wenig Platzbedarf
- +paralleles Arbeiten leicht
- Unterschriften in Klarauflösung
- am Bildschirm lesen schwierig
- Einscannen von Papierdokumenten
- Handling Altbestand





Der Ausbau des elektronischen Rechtsverkehrs wird dazu führen, dass die rein elektronisch vorliegenden Aktenbestandteile anwachsen werden. Damit gewinnt die Einführung einer elektronischen Akte an Bedeutung.

Der AK-QUER empfiehlt:

- Öffnungsklauseln in den Vorschriften der Länder zu schaffen, um praktische Erfahrungen durch Pilotierungen, die in konzeptionelle und organisatorische Vorgaben münden, zu ermöglichen
- das DOMEA-Konzept für Verwaltungsakten auf übertragbare Teile zu überprüfen
- entgegen der DOMEA-Konzeption für die Übergabe von Daten als Schnittstelle XML zu nutzen
- fachspezifisch (in den Justiz-Fachverfahren) zu prüfen, inwieweit aufsetzend auf der elektronischen Akte Workflow-Systeme einsetzbar sind
- fachspezifisch zu prüfen, ob die Übertragung von Aktenteilen über eine - Schnittstelle oder ob - der Versand der gesamten elektronischen Akte sachgerecht und wirtschaftlich ist



Daten aus Justizverfahren werden in den Justizbehörden selbst archiviert (ggf. nach Rumpfdatenbildung) und nach fachspezifischen Fristen an die Archive abgegeben.

Elektronische Daten veralten/verschwinden, weil

Datenträger nicht mehr lesbar sind:
(Lebensdauer CD ca. 15 Jahre).

Computersysteme und Lesegeräte hardwaremäßig nicht mehr verfügbar sind.

Die entsprechenden Programme und Programmsysteme (z.B. Datenmischung im EGB) nicht mehr verfügbar sind.

Visualisierungstechnologien (Editoren) nicht mehr verfügbar sind (Vorteil von Papier: keine Visualisierungstechnik nötig) und Technologien für die Prüfung von elektronischen Signaturen nicht mehr verfügbar sind.

Lösung: regelmäßiger Technologietransfer mit Bestätigung der Richtigkeit von Amts wegen



Der AK-QUER empfiehlt:

- eine einheitliche Vorgehensweise für die Langzeitarchivierung für die
 - Archivierung in den Justizbehörden und
 - für die Übergabe an die Landesarchive
- die durch den technologischen Fortschritt erforderlichen regelmäßigen Konvertierungen sind konzeptionell und wirtschaftlich in den Justizprojekten zu berücksichtigen
- für die Langzeitarchivierung (in den Justizbehörden z.B. nach Verfahrensabschluss und anschließender Rumpfdatenbildung) sollten die Daten in TIF F umgewandelt werden, da die Gefahr des Datenverlustes hier gering ist
- nach der Abstimmung mit dem ReferentenAK der LAD wird die Übergabe von in TIF F-Format umgewandelten Schriftstücken akzeptiert
- elektronische Signaturen werden vor der Archivierung aufgelöst und im Klartext abgelegt (Gewährleistung der Richtigkeit „von Amts wegen“)
- wg. der weiten Verbreitung und der komprimierten Daten sieht der AK-QUER im PDF -Format eine mögliche Alternative
- andere Standards (mittelfristig sollte XML angestrebt werden) für die Langzeitarchivierung sind - abgestimmt über die BLK - in Abstimmung mit den LAD zu konzipieren



Verschlüsselung und Elektronische Signatur

Austauschformate und Austausch von Verfahrensdaten

Elektronische Akte

Archivierung

Vielen Dank für Ihre Aufmerksamkeit!

Einführung und Einsatz von VIS im Geschäftsbereich des Bayerischen Staatsministeriums für Landesentwicklung und Umweltfragen

Von Klaus-Peter Volkmann

**AK "Archivierung von Unterlagen
aus digitalen Systemen" - 05. März 2001**



Einführung und Einsatz von VIS im Geschäftsbereich des Bayerischen Staatsministeriums für Landesentwicklung und Umweltfragen

MR Klaus-Peter Volkmann

Archiv 2001 - 1

StMLU-Volkmann-02/01

Gliederung



- Papier - oder elektronisches Dokument?**
- Projekthistorie und -ziele im BStMLU**
- Kriterien und Strategien**
- Erfahrungen des BStMLU**
- Aussonderung und Archivierung**
- Stand und Perspektiven in den Verwaltungen**

Archiv 2001 - 2

StMLU-Volkmann-02/01

Papier als Informationsträger



Dokumente im Geschäftsprozess

Registratur

Rechtsverbindlichkeit!

Karteiverwaltung

**Wissen?
- wie?
- für wen?**

Sicherheit heute?

Archiv

Archiv 2001 - 3

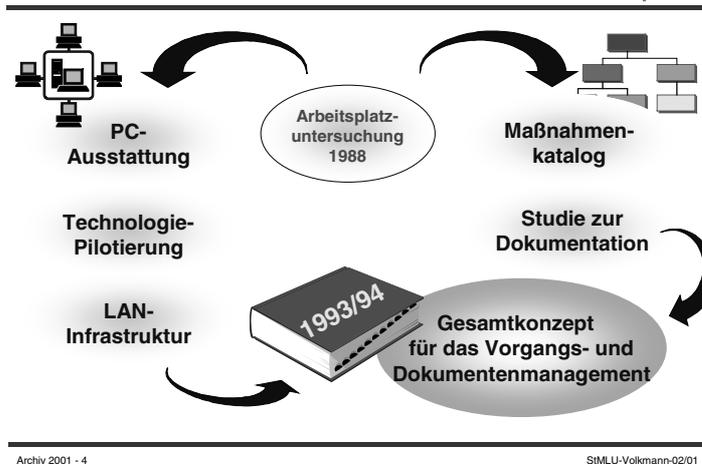
StMLU-Volkmann-02/01

Die Papierform der klassischen Verwaltungsarbeit ist der Garant für Rechtssicherheit im Verwaltungshandeln gegenüber Bürger und Unternehmen wie auch intern. Sie ist primär darauf ausgelegt, Entscheidungen und die damit verbundenen Arbeitsabläufe inhaltlich zu dokumentieren und Verbindlichkeit zu gewährleisten. Eine Nutzung von Inhalten aktueller Vorgänge – z.B. durch paral-

lelen Zugriff zur Beschleunigung von Abläufen oder als Referenz für wiederkehrende Fälle – ist nur begrenzt möglich. Auch steht in Papierform dokumentiertes Wissen nur demjenigen zur Verfügung, der Zugriff auf die jeweiligen Dokumente hat.

Elektronisch gespeicherte Dokumente erlauben eine quasi unbegrenzte inhaltliche Erschließung und Nutzung für viele Personen bzw. Stellen in der Verwaltung und eröffnen damit völlig neue Optionen der Arbeitsweise wie auch qualitative Verbesserungen im Verwaltungshandeln, setzen dabei aber voraus, dass Vorgänge und Dokumente nicht als „Besitzstand“ damit befasster Stellen betrachtet werden und die gleichen Sicherheiten bezüglich Authentizität, Integrität und Vertraulichkeit realisiert werden können, wie sie bei der Papierform gegeben sind.

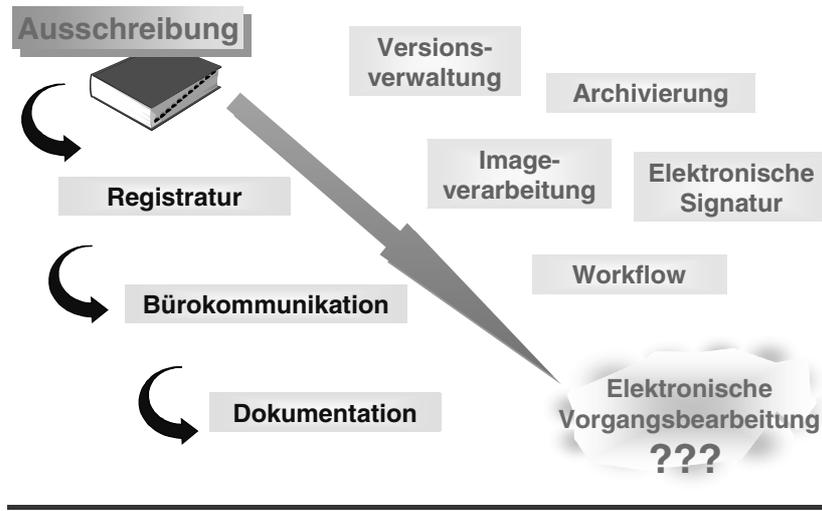
Vorgeschichte des Projektes



Das StMLU befasst sich seit Anfang der 90er-Jahre mit der Frage der elektronischen Form der Vorgangsbearbeitung. Ausgangspunkt war eine umfassende Arbeitsplatzuntersuchung bezüglich der Arbeitsweise, Ablaufprozesse und des Informationsbedarfs im Ministerium, sowie der technischen Optionen für ein elektronisches Dokumentenmanagement. Das Ergebnis dieser Analyse führte zunächst zum Aufbau einer modernen technischen Infrastruktur für alle Arbeitsplätze auf PC-Basis mit Vernetzung,

sowie zu einem Maßnahmenkatalog für ein stufenweises Vorgehen in organisatorischer, personeller und technischer Hinsicht. Aspekte und Anforderungen der elektronischen Dokumentation wurden dabei gesondert untersucht und bildeten die Grundlage für ein Gesamtkonzept des Vorgangs- und Dokumentenmanagements.

Ausschreibung einer Vision

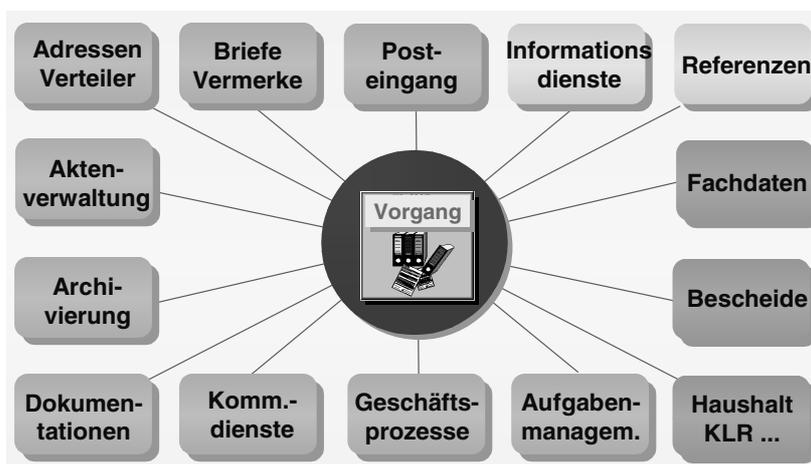


Archiv 2001 - 5

SIMLU-Volkmann-02/01

Das vorgenannte Konzept wurde 1994 im Sinne einer „Vision“ zum Gegenstand einer Ausschreibung gemacht, verbunden mit der konkreten Realisierung einer ersten Stufe der Umsetzung, nämlich der Installation eines Registraturverfahrens, das geeignet sein sollte, Vorgänge in Form elektronischer Objekte zu bilden und zu verwalten – als Kern, auf dem stufenweise weitere Realisierungsschritte zukünftig aufsetzen konnten.

Aufgaben und Anforderungen



Archiv 2001 - 6

SIMLU-Volkmann-02/01

Die konzeptionellen Anforderungen an ein System des Vorgangs- und

Dokumentenmanagements sind vielfältig. Sie erfordern insbesondere ein hohes Maß an Integrationsfähigkeit verschiedenster Funktionsbausteine, die alle miteinander in Beziehung stehen – z.B. Registratur, Dokumentenerstellung, Archivierung, Kommunikationsdienste,

Workflow-Verfahren etc. Zusätzliche Anforderungen ergeben sich aus der erforderlichen

Einbeziehung von Dokumenten und Vorgängen, die in fachspezifischen DV-Anwendungen oder Verwaltungsanwendungen (z.B. Haushaltsvollzug) entstehen. Außerdem müssen die Inhalte von Vorgängen und Dokumenten erschlossen und möglichst unmittelbar und auf einfache Weise an Informationssysteme zur Bereitstellung als Informationsgrundlagen nach außen übergeben werden können.

Ziele und Nutzen



- Beschleunigung von Arbeitsabläufen durch Parallelzugriff auf Dokumente**
- unmittelbare Transportwege und asynchrone Kommunikation**
- einfache Sachstandsverfolgung durch Transparenz der Bearbeitung intern und ggf. für externe Partner**
- Vorgänge und Dokumente als dauerhafte Informationsquelle für alle - jederzeit zugänglich**
- Verbesserte Suchmöglichkeiten**
- Hohe Verfügbarkeit und Personenunabhängigkeit**
- Einsparungen / Kostenreduzierung (?) - ja, aber ...**

Archiv 2001 - 7

StMLU-Volkmann-02/01

Der mit der Einführung eines Dokumentenmanagementsystems verbundene Nutzen ist primär qualitativer Art und nur begrenzt monetär messbar. Der Aufwand für Realisierung, Einführung und den dauerhaften Betrieb darf nicht unterschätzt werden. Einsparungen an einer Stelle stehen neue Aufwände an anderer Stelle gegenüber. Die Frage nach dem „ob“ stellt sich jedoch nicht – nur die Frage nach dem „wie“. Der

Erfolg ist sicher, wenn die richtigen Strategien gewählt und die mit der Realisierung verbundenen Notwendigkeiten beachtet werden.

Strategien und Kriterien



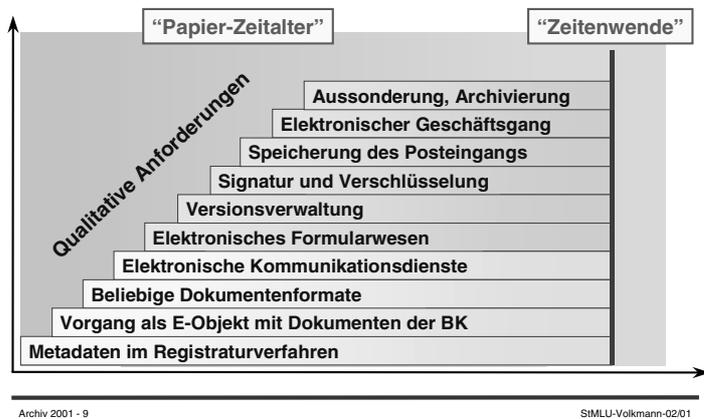
- Einheitliche, flächendeckende Nutzung**
- Stufenkonzept für den Aufbau**
- Technologische Modularität**
- Freiheitsgrade bezüglich unterschiedlicher Nutzungen und Anwendungsbereiche**
- Integrationsfähigkeit mit Fachverfahren und Integrierten Lösungen (z.B. SAP)**
- Offenheit bezüglich Plattformen und Datenbanken**
- Standardlösung mit beeinflussbarer und berechenbarer Zukunft**
- langfristig verlässlicher externer Partner als Generalunternehmer mit Verwaltungskompetenz**

Archiv 2001 - 8

StMLU-Volkmann-02/01

Die bei der Einführung eines Dokumentenmanagementsystems zu beachtenden Kriterien sind vielfältig und bedürfen umfassender konzeptioneller Vorarbeiten sowie einer konsequenten Berücksichtigung im Projektverlauf.

Stufenkonzept

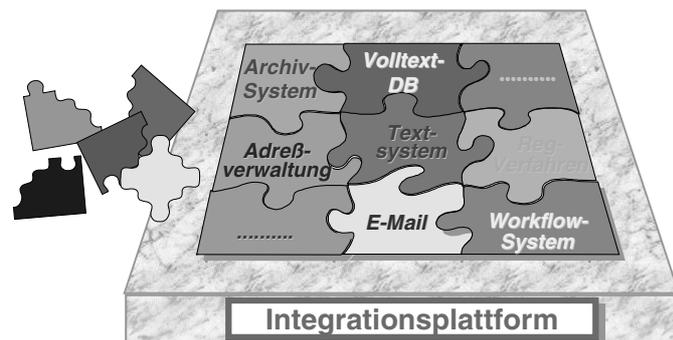


Angesichts des Umfangs des Projekts ist – aus Sicht des StMLU – ein verträgliches, von allen Betroffenen mit getragenes und langfristig angelegtes Stufenkonzept bei der Vorgehensweise unerlässlich. Die Abfolge der Stufen ist dabei durchaus variabel und je nach den jeweiligen Rahmenbedingungen begrenzt steuerbar. Eine Ablösung der rechtsverbindlichen Papierform durch elektronische Dokumente und Vorgänge ist erst dann möglich, wenn alle benötigten Funktionen technisch so realisiert sind, dass sie die gleichen

qualitativen Anforderungen erfüllen wie das Papier derzeit. Die rechtlichen Voraussetzungen für die Ablösung der Papierform durch elektronische Verfahren und Objekte sind entsprechend anzupassen, dieser Prozess ist derzeit im Gange.

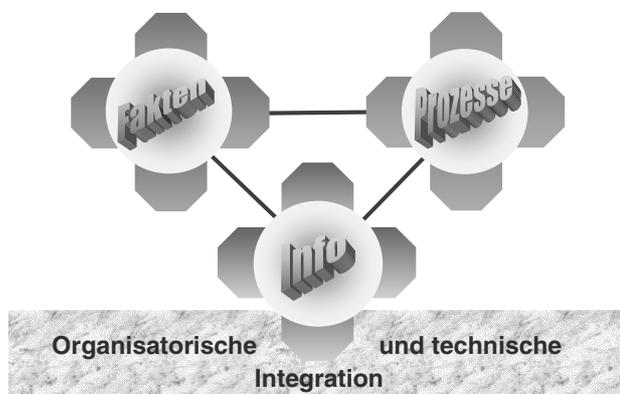
Die Vielfalt der im Laufe der Zeit benötigten technischen Funktionen erfordert ein technisches Integrationskonzept, das die Interoperabilität der verschiedenen Komponenten des Dokumentenmanagements gewährleistet. Erforderlich ist insbesondere eine Integrationsplattform zur redundanzfreien Realisierung aller funktionsübergreifenden Administrationsaufgaben (z.B. Benutzerverwaltung, Rechteverwaltung, Organisations- und Prozessregeln, Objektverwaltung etc.).

Modularität und Integration

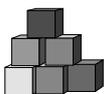


Das Projekt des StMLU sah von Beginn an vor, die Vorgangs- und Dokumentenbearbeitung in ein übergreifendes Informationsmanagement einzubeziehen. Insbesondere wurde gefordert, die Integrationsfähigkeit bzw. erforderliche Schnittstellen zu Faktendatenbanken und Fachanwendungen, Informationssystemen und allgemeinen Verwaltungsprozessen intern wie behördenübergreifend sicherzustellen – und zwar bezüglich der Arbeitsorganisation wie auch hinsichtlich der technischen Realisierung.

Strategisches Ziel



DMS-Projekt im BStMLU



- Gründe für die Entscheidung für VIS
- Office-Software mit E-Mail / Internet / VIS als Standard-Konfiguration
- flächendeckende Ausstattung aller Arbeitsplätze (ca. 500) einschließlich Leitungsebenen seit 1997/98
- Rolle von VIS im Verhältnis zum Intranet
- Sonderlösungen für spezielle Aufgaben auf VIS-Basis (z.B. Bewerberverwaltung, Redendokumentation, Bürgerreferat)
- Anpassung der Konfiguration als Reaktion auf Erfahrungen im Praxisbetrieb
- Anpassungen an den technologischen Fortschritt (z.B. Windows, Office)
- Pilotierung in nachgeordneten Behörden

Archiv 2001 - 12

StMLU-Volkmann-02/01

Im Rahmen der vorgenannten Ausschreibung wurde dem Produkt VIS der Zuschlag erteilt. Die Realisierung bezieht Standard-Office-Produkte mit ein. Ein wesentliches fachliches Argument für diese Entscheidung war u.a. die Modularität des Realisierungskonzepts wie auch die bereits damals erkennbare unmittelbare Ausrichtung des Produkts auf den spezifischen Bedarf der Verwaltung. Die Einführung hat gezeigt, dass ein Dokumentenmanagementsystem dieser Art als „Werkzeugkasten“ verstanden werden muss, mit dessen Hilfe unterschiedlicher Anwen-

dungsbedarf abgedeckt werden kann. Voraussetzung dafür sind entsprechende umfassende Organisationskonzepte.

VIS wird (in einer zwischenzeitlich überholten Version) im StMLU flächendeckend (an ca. 500 Arbeitsplätzen) eingesetzt. Im Zuge der Einführung einer neuen Version wird nunmehr die flächendeckend einheitliche Einführung in allen Behörden des Geschäftsbereichs angegangen. Entsprechende Pilotierungen laufen bereits.

Eines der wichtigsten Erfolgskriterien ist die Akzeptanz des Systems durch das Personal. Dabei darf der erforderliche Wandel der Arbeitsweisen und der damit einhergehende Unterstützungs- und Betreuungsbedarf nicht unterschätzt werden. Diese Aspekte haben ein deutlich höheres Gewicht als Details des konkret gewählten Produkts.

Erfahrungen - Personal



- Geringe Technik-Vorbehalte
- Erwartungsdruck versus Zweifel - Nutzen-Transparenz
- Erfordernis eines Bewusstseinswandels (Öffnung für vernetztes Denken und Arbeiten)
- hoher Informationsbedarf auf allen Ebenen
- Probleme der Qualifizierung
- Akzeptanz arbeitstechnischer und technologischer Standards
- Umstellung auf neue Arbeitsweisen - mit und ohne Papier

Archiv 2001 - 13

StMLU-Volkmann-02/01

Technische Systeme zur Unterstützung der Verwaltungsarbeit können nur dann erfolgreich und nutzbringend eingesetzt werden, wenn sie mit dem organisatorischen Rahmen engstens abgestimmt sind. Dies kann häufig nur erreicht werden, wenn individuelle Arbeits-

weisen durch einheitliche, ggf. neue Formen der Aufgabenabwicklung ersetzt werden, bei denen die Möglichkeiten – und Risiken – der technischen Lösung angemessen berücksichtigt werden. Hierzu bedarf es einer kreativen Auseinandersetzung mit den Nutzungspotenzialen, einer entsprechenden Folgenabschätzung und weitreichender Vorgaben in den verschiedensten Bereichen.

Erfahrungen - Organisation



- Vereinbarungen mit den Personalvertretungen
- Abstimmung Organisation / Technik
- Vereinheitlichung von Arbeitsweisen
- Aktuelle Reorganisationsmaßnahmen
- Prüfung der Ablauforganisation
- Kreative Erschließung von Nutzenpotenzialen
- neue Administrationsaufgaben
- neue Regularien (E-Mail, Internet, etc.)
- Koordinierung im Geschäftsbereich sowie ressortübergreifend

Archiv 2001 - 14

StMLU-Volkmann-02/01

Technologische Aspekte



- Standard-Software - nicht Eigenentwicklung**
- Funktionale Modularität der Lösung**
- Offenheit für Technologie-Alternativen**
- Datenschnittstellen für Import und Export**
- Offenheit für Org- / Rechte-Konfiguration (Umsetzbarkeit von Reorganisationsmaßnahmen)**
- Flexible Benutzerschnittstelle (Arbeitsplatztypen)**
- Administrationsaufwand**
- Generalverantwortung für die Lösung insgesamt**

Die Anforderungen an das Vorgangs- und Dokumentenmanagement in der Verwaltung sind zu komplex, um hier mit Eigenentwicklungen operieren zu können. Bei der Auswahl eines Produktes auf dem Markt ist daher entscheidend, in welchem Maße die Nutzung gestaltbar, der Einsatz konfigurierbar ist. Außerdem ist der Aspekt der Interoperabilität mit anderen Systemen zu beachten, damit die Durchgängigkeit elektronischer Prozesse nicht gefährdet wird. Schließlich muss be-

Archiv 2001 - 15

StMLU-Volkmann-02/01

rücksichtigt werden, dass im laufenden Betrieb Administrationsaufwand anfällt (Nutzerverwaltung, Zugriffsrechte, Organisationsanpassungen etc.).

Aussonderung und Archivierung



- Standardisierte / konfigurierbare Fristen je nach Dokumentenklasse bzw. Objekttyp**
- Auflisten von Objekten durch Vorab-Auswertung zur Kontrolle durch die Verantwortlichen**
- Unterscheidung zwischen Metadaten / Objekten**
- differenzierte Behandlung von Verweisen auf Dokumente und Vorgänge**
- Unterschiedliche Verantwortlichkeiten je nach Objektklasse**
- Exportschnittstelle und Kennzeichnung der Behandlung in den Metadaten**



Archiv 2001 - 16

StMLU-Volkmann-02/01

Im Hinblick darauf, dass die im behördeninternen Vorgangs- und Dokumentenmanagement anfallenden elektronischen Vorgangsobjekte später ggf. zentralen Archiven übergeben werden müssen, sind klare und detaillierte Vorgaben in dieser Hinsicht unverzichtbar, sowohl was die beschreibenden Merkmale von Vorgängen und Dokumenten anbelangt, wie auch die technische Form der Übergabe. Da die Archivverwaltungen wohl nicht behördenspezifische technische und organisatorische Schnittstellen bewältigen können, sind sie in besonderer Weise gefordert, diesbezüglich entsprechende Vorgaben zu machen.

DMS-Stand und Perspektiven



BStMLU

- Pilotierung von VIS^{kompakt} im nachgeordneten Behördenbereich
- Neue Version: VIS^{kompakt} 2.0
- Integration mit Windows 2000 (Active Directory)

Ressorts in Bayern

- Bewertung der PARO-Ergebnisse des BMI
- Beschluss der Organisationsreferenten
- Aktuelle Initiativen einzelner Behörden (auf Basis von CSE/DOMEA)
- "Schub" durch "E-Government"-Lösungen

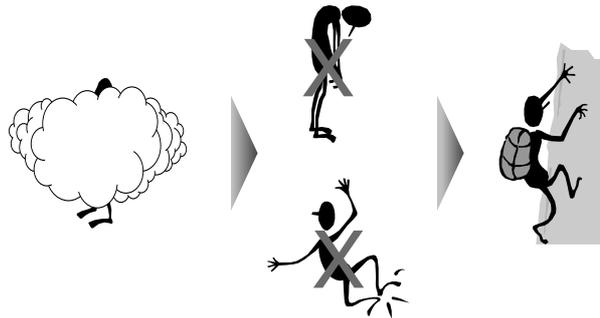
Das StMLU war das erste Ressort in Bayern, das erste Schritte zur Realisierung der elektronischen Vorgangsbearbeitung unternommen hat. Weitere Ressorts haben zwischenzeitlich ebenfalls begonnen. Eine einheitliche Vorgabe eines Produktes wurde bewusst vermieden, stattdessen werden die Vorarbeiten des Bundes als Entscheidungshilfe herangezogen. Grundlage bei jeder Produktwahl muss in jedem Falle das DOMEA-Fachkonzept sein.

Archiv 2001 - 17

StMLU-Volkmann-02/01

Eine Beschleunigung der Aktivitäten der Verwaltungen in diesem Bereich ist im Zuge der eGovernment-Initiativen zu erwarten, da nur dann eine durchgängige elektronische Bearbeitung von Verwaltungsprozessen zu erreichen ist, wenn dafür auch intern die entsprechenden Infrastrukturen zur Verfügung stehen.

Dokumentenmanagement - ein weiter Weg ...



Archiv 2001 - 18

StMLU-Volkmann-02/01

Die Aufgabe ist komplex, viele Details sind zu beachten, die Umsetzung ist ein langwieriger Prozess. Erfolg verspricht einzig ein Weg der kleinen, behutsamen Schritte!

Kontakt



MR Klaus-Peter Volkmann
 Leiter des Referats IuK-Betrieb
 und Bürokommunikation

Bayerisches Staatsministerium für
 Landesentwicklung und Umweltfragen
 Rosenkavalierplatz 2
 81925 München

Tel. 089 / 9214-3370
 Fax 089 / 9214-203370
 e-mail: klaus-peter.volkmann@stmlu.bayern.de

Archiv 2001 - 19

StMLU-Volkmann-02/01

Das Projekt ELGAR zur elektronischen Unterstützung von Geschäfts- und Aktenregistrierungsprozessen im Kanton Basel-Stadt

Von Thomas Schärli

Anlässlich der 3. Tagung des Arbeitskreises „Archivierung von Unterlagen aus digitalen Systemen“ (Koblenz 1999) wurde das Lebensphasen-Modell des Staatsarchivs Basel-Stadt (StABS) für elektronische Unterlagen vorgestellt.¹ Dieses beruht auf einem klassischen „Records Lifecycle“-Ansatz, mit strikter Rollentrennung zwischen Schriftgutverwaltung („Registratur“) und Archiv, jedoch mit intensivierter archivischer Einflussnahme auf den Unterlagenproduktionsprozess. Vorarchivische Interventionen erfolgen parallel auf den Handlungsebenen Recht, Organisation und Technik. Im Bereich der Technik besteht die Absicht, den Unterlagenbildnern, nebst anderen Dienstleistungen, geeignete Werkzeuge anzubieten. Von einem solchen Werkzeug soll nachstehend die Rede sein.

Ziele und Hintergründe

Das Projekt ELGAR („Elektronische Geschäfts- und Aktenregistrierung“) verdankt seinen Ursprung einem glücklichen Umstand: Bereits 1994–1998 hatte sich das Staatsarchiv Basel-Stadt (StABS) an einem Projekt zur elektronischen Unterstützung von Parlaments-, Regierungs- und Departementsgeschäften namens GEKOBAS beteiligt.² Der Schwerpunkt dieses Vorhabens lag auf der Kontrolle und Steuerung termingebundener Aufträge. Dank zahlreicher Meta-Daten auf Aktenplan- und Dossier- bzw. Aktenstufe kann das Archiv einen grösseren Reichtum an automatisch archivierbaren Findmittel-Informationen erhoffen. Allerdings eignet sich das realisierte System nicht für die Aggregation von Unterlagen aus beliebigen Geschäftsprozessen und fällt deshalb für einen verwaltungsweiten Einsatz ausser Betracht.³

Im Rahmen seines eigenen Systems PRISMA verfügt das StABS hingegen über ein Werkzeug zur Verwaltung von Geschäftskontakten (mit unterlagenbildenden Organen, Archivgut-Benützern und weiteren Geschäftspartnern).⁴ Darin können auch Dossiers gebildet und pro Dossier Ereignisse (mitsamt allfälligen Folgeaktivitäten) festgehalten sowie Dokumente zugeordnet werden. Seine Handhabung ist denkbar einfach. Einziger Nachteil: An zahlreichen Aktendossiers müssten mehrere Partner beteiligt werden können. Ferner fehlen gewisse Voraussetzungen, die ein „recordkeeping system“ erfüllen muss, um eine authentische und ver-

¹ Thomas Schärli, Das archivische „Lebensphasen“-Modell im Kontext elektronischer Informationssysteme; ein Werkstattbericht des Staatsarchivs Basel-Stadt als Beitrag zur Methodendiskussion im Bereich der Archivierung elektronischer Unterlagen. In: Digitale Herausforderungen für Archive (Materialien aus dem Bundesarchiv, Heft 7), Koblenz 1999, S. 51–58.

² Ebd. S. 56: Projekt GEKOBAS, Teilsystem ALLKOM.

³ Die drei im Rahmen von GEKOBAS realisierten Teilsysteme werden derzeit durch ein Nachfolgesystem namens KONSUL ersetzt. Dieses stammt von demselben Hersteller und soll Anfang 2002 produktiv aufgechaltet werden. Wegen verschiedener Mängel ist das Teilsystem ALLKOM (für die Geschäftsführung auf Departementsstufe) bisher auf zögerliche Akzeptanz gestossen. Als das Projekt ELGAR gestartet wurde, war von den KONSUL-Plänen noch nichts bekannt. Dieses Beispiel zeigt, wie unberechenbar sich die Umstände archivischer Einflussnahme auf den Design von Systemen präsentieren können.

⁴ Zum Projekt PRISMA vgl. Thomas Schärli, Das Informatiksystem PRISMA im Staatsarchiv Basel-Stadt: Ziele, Einsatz, Konsequenzen (Anhang zum Jahresbericht 1999 des Staatsarchivs Basel-Stadt, S. 40–64). Die im Rahmen von PRISMA realisierte Software wird inzwischen unter dem Namen scopeArchiv als Standardprodukt vertrieben.

lässliche Überlieferung sicherzustellen; für solche Zwecke war das betreffende Produkt begreiflicherweise nicht konzipiert.⁵

Zufällig erhielt nun die Herstellerfirma zu gleicher Zeit von zwei anderen kantonalen Dienststellen den Auftrag, eine Geschäftskontrolle mitsamt Dokumentenablage, die auch für Wissensmanagement-Zwecke nutzbar sein sollte, einzuführen. Sie entschloss sich deshalb, auf der Basis der bereits vorhandenen Systemkomponenten ein zusätzliches, auf recordkeeping-Bedürfnisse ausgerichtetes Produkt zu realisieren, und fragte das StABS an, ob es bereit wäre, an dessen Entwicklung in der Rolle eines fachlichen Beirats mitzuwirken.

Vorarbeiten

Seitens des StABS sind seit 1999 verschiedene weitere Meilensteine erreicht worden:⁶

- Das Projekt PRISMA wurde in seinen wesentlichen Teilen 1999 erfolgreich abgeschlossen.
- Eine Gesamtsicht auf verwaltungsweite Informations- und Überlieferungsprozesse ist im 1999 entstandenen *Rahmenkonzept „Integration des Unterlagenmanagements im Kanton Basel-Stadt“* festgehalten.
- Aus dem genannten Rahmenkonzept ging eine *Lösungsarchitektur für die gesamte Lebenszeit elektronischer Unterlagen* unterschiedlicher Herkunft hervor. Darin wird modellhaft versucht, einen standardisierten Transportfluss von Datenobjekten, die aus Dokumentenmanagement-, Datenbank-Umgebungen oder hybriden Anwendungssituationen stammen können, zu beschreiben (Abbildung 1).

Dank dieser Vorarbeiten wird das StABS heute von seinen massgeblichen verwaltungsinternen Partnern als „competent player at the electronic records management table“ wahrgenommen.

Problemstellung

Aus Archivsicht besteht das Hauptproblem darin, den neuen Herausforderungen mit minimalen Personalressourcen nachkommen zu müssen. Durch die intensivierete vorarchivische Einflussnahme entsteht eine personelle Zusatzbelastung. Nur wenn es gelingt, so viele Überlieferungsprozesse als möglich zu standardisieren, lässt sich der Mehraufwand auf ein vernünftiges Mass beschränken. Dies setzt allerdings einen vorübergehenden Sondereffort voraus.

Wie Abbildung 1 zeigt, erfüllt ein dem Archivsystem vorgelagertes Übernahme-Werkzeug diesen Zweck. Mit dessen Hilfe können heterogene Metadaten-Gebilde durch unterschiedliche Formulare abgebildet und in das Archivsystem integriert werden. Dadurch vermindert sich der Programmieraufwand für Datenübernahmen auf die Hälfte. Es verbleibt aber immer noch die Aufgabe, archivwürdige Daten aus einer Vielzahl von Systemen auf die Übernahme-Schnittstelle zu leiten.

Eine zusätzliche Vereinheitlichung ergibt sich aus dem Grundsatz, dass primär administrative Prozesse archivisch bewertet und die aus ihnen erwachsenden Unterlagen (Aufzeichnungen) identifiziert sowie systematisch ausgesondert werden sollten. Solche Prozesse werden in Basel-Stadt mehrheitlich durch Datenbank-Applikationen unterstützt.⁷ Nicht alle Inhalte eines

⁵ Die Partnerverwaltung von PRISMA beruht auf dem Standardprodukt scopePartner von scope solutions ag, das 1998 ohne Anpassungen beschafft wurde.

⁶ Vgl. Schärli (wie Anm. 1) S. 53 f.

⁷ Workflow-Technologien werden in Basel-Stadt verhältnismässig selten eingesetzt. Grund dafür sind die Heterogenität der zugrundeliegenden Middleware-Plattformen sowie eine eher kleine Zahl an gleichförmigen

Datenbanksystems sind archivisch gleich interessant. Für statistische Auswertungszwecke genügen unter Umständen gewisse Zeitschnitte, um die Kontinuität entsprechender Archivzeugnisse, wie Amtsbuch-Serien, Register, Karteien usw. sicherzustellen.

„Elektronische Akten“ hingegen sind mehrheitlich Dokumente (Dateien) in Verbindung mit struktur-, inhalts-, kontext- und darstellungsbezogenen Zusatzinformationen (Metadaten), die nicht ohne explizite Zusatzmassnahmen in einem einzigen System verfügbar sind. Vor allem Kontext-Metadaten müssen zusätzlich manuell erfasst werden, sofern sie nicht von den Systemtransaktionen ab, die sie erzeugten, abgeleitet werden können.

Die Heterogenität der eingesetzten Systeme und System-Plattformen macht diese Aufgabe nicht einfacher. Viele Unterlagen, die aus der Sicht der zugrundeliegenden administrativen Aufgaben und Geschäftsprozesse gemeinsam abgelegt werden müssten, sind heute auf eine Vielzahl logischer Speicher-Orte (File Server, Dokumentenmanagementsysteme usw.) verteilt.

Diese Ausgangslage bewog das Staatsarchiv und die Basler Firma scope solutions ag, gemeinsam ein *spezialisiertes Werkzeug für den Querschnittprozess „Aktenbildung“* („record-keeping“) zu entwickeln, und zwar so, dass sich dieses in beliebige technische Umgebungen möglichst reibungslos integrieren lassen sollte.

Lösungsarchitektur

Vor dem Hintergrund der skizzierten Problemstellung mag die bereits erwähnte Lösungs-Architektur veranschaulichen, welcher systematische Rahmen für das konzeptionelle Vorgehen massgeblich war:

- In diesem Schema hat das Archivsystem **PRISMA** (Mitte unten) die Funktion eines „Rückhaltebeckens“ für Metadaten (Archivverzeichnisse unterschiedlicher Form) sowie Verknüpfungen zu Unterlagen, die auf Papier oder elektronisch archiviert werden können.⁸
- Dem System PRISMA vorgelagert ist ein **Übernahme-Werkzeug**, mit dessen Hilfe beliebige (Meta-)Daten-Aggregationen aus Ursprungssystemen durch Formulare abgebildet und an PRISMA übergeben werden können.⁹ Damit verbleibt aber immer noch die Aufgabe, in den Ursprungssystemen festzulegen, welche Daten oder Dateien dem Übernahme-Werkzeug zugeleitet werden sollen, und die dafür notwendigen Funktionen zu realisieren. Dies kann sehr zeitraubend und kostspielig werden, sofern es nicht gelingt, die aktenbildenden Funktionen selbst zu standardisieren.
- In der Datenbank-Welt (rechts) besitzt der Kanton Basel-Stadt eine Datenlogistik-Strategie, die auch für Archivierungszwecke genutzt werden kann. Über einen zentralen **Datenmarkt** lassen sich operationelle Daten zwischen verschiedenen Applikationen austauschen sowie historisierte Daten zu Auswertungszwecken an ein Data Warehouse („**Datenlager**“) übergeben. An das letztere wird auch das StABS angeschlossen werden, so dass der Archivierungsprozess auf einheitliche Weise erfolgen kann.¹⁰ In diesem Rahmen sollte die

Geschäftsprozessen (vermutlich auch einer Verwaltungskultur, die der systematischen Gestaltung von Büroprozessen bisher keine besondere Bedeutung beigemessen hat).

⁸ Das eingesetzte Produkt scopeArchiv kann je nach Grundkonfiguration auch digitale Datenobjekte oder Dateien speichern. PRISMA (als Bezeichnung eines Projekts bzw. Subsystems) und scopeArchiv sind im Rahmen dieses Schemas nicht deckungsgleich.

⁹ Als Datenaustausch-Format werden XML-Schemata verwendet.

¹⁰ Das Projekt „Datenlager“ befindet sich noch in einer Vorkonzeptphase. Eine Machbarkeitsstudie wurde im Frühjahr 2001 abgeschlossen. Sie enthält konkrete Anforderungen und Lösungsansätze für Archivierungsprozesse. Aus Kapazitätsgründen (Die verfügbaren Personalressourcen sind auf absehbare Zeit durch das Projekt

Kontinuität der bereits erwähnten Karteien, Register und Amtsbuch-Serien in einem grösseren Umfang gesichert werden können.

- Auf der Seite der **Bürokommunikations- und Dokumentenmanagementsysteme** ist die Ausgangslage weniger komfortabel. Mehrere eMail- und Middleware-Plattformen konkurrenzieren sich; der gegenseitige Datenaustausch ist mit zahlreichen Erschwernissen verbunden. Die massgebliche Einheit bildet hier nicht der Datensatz, sondern das Dokument (die Datei). Kontext-Metadaten sind, wenn überhaupt, nur lückenhaft in diesen Systemen mitgespeichert. Das Problem eines dauerhaft haltbaren Dateiformats soll im Rahmen eines besonderen Projekts (**BEDA**: „Basler elektronisches Dokumentenarchiv“) gelöst werden. Dabei wird nach Möglichkeit auf die Ergebnisse grösserer Referenzprojekte wie ELAN oder ARELDA abgestellt.¹¹

Das derzeit am schlechtesten gelöste Problem betrifft, wie bereits erwähnt, die gegenseitige Integration der **Datenbank- und Dokumentenmanagement-Welten**. An diesem Punkt setzt das Projekt **ELGAR** an. Dessen Ziel besteht darin, aus Dateien und zugehörigen Transaktionsaufzeichnungen aus Datenbank-Applikationen oder Workflows elektronische „Akten“ („records“) zu bilden und in standardisierten Kontext-Strukturen („archival bonds“ gemäss L. Duranti) darzustellen.¹²

Um den Analogievergleich mit dem Entwässerungssystem nochmals aufzunehmen: Bildet PRISMA das „Rückhaltebecken“, so wäre ELGAR als „Vorfluter“ zu betrachten.

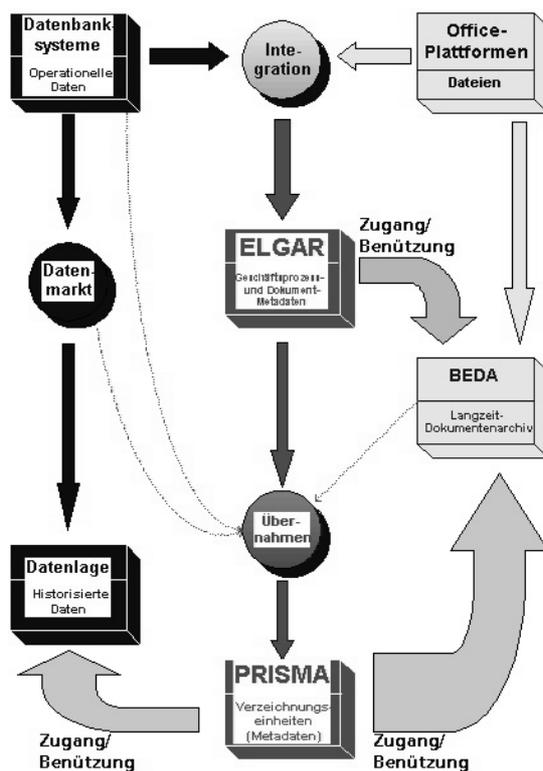


Abb. 1:
Lösungsarchitektur zur
Archivierung
elektronischer
Unterlagen im Kanton
Basel-Stadt

„Datenmarkt“ vollständig ausgelastet) wird mit der nächsten Projektphase frühestens in ein oder zwei Jahren begonnen werden können.

¹¹ Zu ELAN vgl. Andreas Engel, Elektronische Archivierung und Nutzung (ELAN): Eine Projektidee zur Entwicklung von Verfahren für die dauerhafte Archivierung und Nutzung von elektronischen Geschäftsunterlagen aus Behörden und Betrieben. In: Auf der Suche nach archivischen Lösungsstrategien im digitalen Zeitalter. Beiträge zur 4. Jahrestagung des Arbeitskreises „Archivierung von Unterlagen aus digitalen Systemen“ im Stadtarchiv Mannheim, 10.–11.4.2000, hrsg. von Ulrich Nieß (Sonderveröffentlichungen des Stadtarchivs Mannheim 26), Mannheim 2001, S. 73–84. ARELDA („Archivierung elektronischer Datenbanken und Dateien“) ist ein mehrjähriges Projekt des Schweizerischen Bundesarchivs, das mittels experimentellen Prototyping die dauerhafte Archivierbarkeit unterschiedlicher Datenbank-, Daten- und Dateitypen untersucht.

Funktionale Erfordernisse

Soweit die Vorgaben aus Archivsicht. Für den Software-Partner standen andere Aspekte im Vordergrund: Das beabsichtigte Produkt musste in eine Angebots-Nische vorstossen und unmittelbare Kundenbedürfnisse abdecken können – im Klartext: sich rasch und in grösserer Zahl verkaufen lassen. Nicht der flächendeckende Einsatz in grossen Organisationen, sondern die unkomplizierte Einführung in überschaubaren Betriebseinheiten bildete für ihn den Ausgangspunkt. Das Werkzeug sollte handlich, möglichst selbsterklärend und in die jeweilige gewohnte Arbeitsumgebung mühelos integrierbar sein. Den Fokus bildete die effiziente, transparente Unterstützung der täglichen Büroarbeit in Teams oder beliebig strukturierten Personengruppen, nicht zuletzt auch in virtuellen Organisationen. Mit einem Minimum von Bearbeitungsschritten und Eingabeaufwand sollte ein optimaler Output (in Form von übersichtlichen, informativen Bildschirmanzeigen oder Berichten sowie an den zutreffenden Ort weitergeleiteten Arbeitsergebnissen) erzielt werden können.¹³

Zwischen den Ansprüchen eines ergonomisch hochstehenden, effizient einsetzbaren Bürowerkzeugs und denjenigen einer mit geschäftsprozess-orientierter Kontext-Information gut dotierten, glaubwürdigen und verlässlich gebildeten Wissensbasis galt es die bestmögliche *Synthese* zu finden. Eine solche musste mehr als einen Kompromiss darstellen.

Letztlich ist aber auch das Archiv auf vorarchivische Aktenverwaltungssysteme von bester Qualität angewiesen: Da am täglichen Geschäftsgang die gesamte Bandbreite einer Belegschaft, mit unterschiedlichster Prägekraft und technischer Versiertheit, beteiligt ist, spielt die Akzeptanz des Personals für solche Instrumentarien eine Schlüsselrolle. Werden diese unzutreffend genutzt oder gar umgangen, ist das Kriterium einer vollständigen Nutzung nicht erfüllbar. Ohne Vollständigkeit ist auch die Verlässlichkeit des benützten Systems problematisch. Dadurch sinken die Beweiskraft („evidence“) der dem Archiv angebotenen Daten bzw. die Nachweisbarkeit stattgefundenen Vorkommnisse und die Nachvollziehbarkeit von Geschäftsprozessen.¹⁴

Das ELGAR-Geschäftsobjekt-Modell

Der archivische Anteil am konzeptionellen Design des neuen Produkts bestand hauptsächlich darin, die ihm bekannten Erkenntnisse aus internationalen Projekten im Bereich des „electronic records management“ einzubringen und dafür zu sorgen, dass wegweisende oder aussichtsreiche Standards durch das System unterstützt werden können.

Aufmerksam verfolgt wurden in diesem Rahmen vor allem die Arbeiten am internationalen Records Management Standard (ISO 15489) und dessen Vorbildern (dem australischen Standard AS 4390), einschliesslich daran anknüpfenden Verfeinerungen, namentlich den „Australian Recordkeeping Metadata Standards“ sowie „Recordkeeping Metadata Schemata“ (RKMS), ferner dem europäischen Projekt MoReq.¹⁵

¹² Luciana Duranti, The Archival Bond. In: Archives and Museum Informatics 11, 3–4 (1997) S. 213–218.

¹³ Das Nahziel des Produkts scopeDossier bildet dessen Nutzung als spezialisierte Fachapplikation, mit der die täglich benötigten Daten über Geschäftspartner, stattgefundenen Kontakte, geplante Termine und Aufträge, konsultierten, bearbeiteten oder weitergeleiteten Dokumente usw. im jeweiligen Geschäftszusammenhang verwaltet werden. Erweiterungen, bis hin zu einem „enabling system“, das im Hintergrund zu Handen anderer Applikationen die Geschäftsaufzeichnungen und Dokumente richtig „versorgt“ und bereitstellt, wurden aber konzeptionell mitberücksichtigt. Vgl. die Ausführungen im Kapitel „Geplante Erweiterungen“.

¹⁴ Luciana Duranti, The Concepts of Reliability and Authenticity and Their Implications. In: Archivaria 39 (1995) S. 5–10, sowie weitere Publikationen derselben Autorin (vgl. <http://www.slais.ubc.ca/people/faculty/duranti/duranti.htm>).

¹⁵ Zur Metadaten-Problematik vgl. die Ergebnisse des internationalen Recordkeeping Metadata Workshop der Archiefschool Amsterdam, 5.–8. Juni 2000 (publiziert auf <http://www.archiefschool.nl/amf/sem2000.htm>) Ferner

Abbildung 2 repräsentiert das „high level model“ der wichtigsten für ELGAR definierten Geschäftsobjekte. Verwandtschaften mit dem Recordkeeping Metadata Schema sind unverkennbar: Geschäftsaufzeichnungen, Partnern, Ordnungssysteme (Aktenpläne) und Akten lassen sich jeweils in unbeschränkt viele hierarchische Stufen gliedern und können zueinander in beliebigen Beziehungen stehen.¹⁶

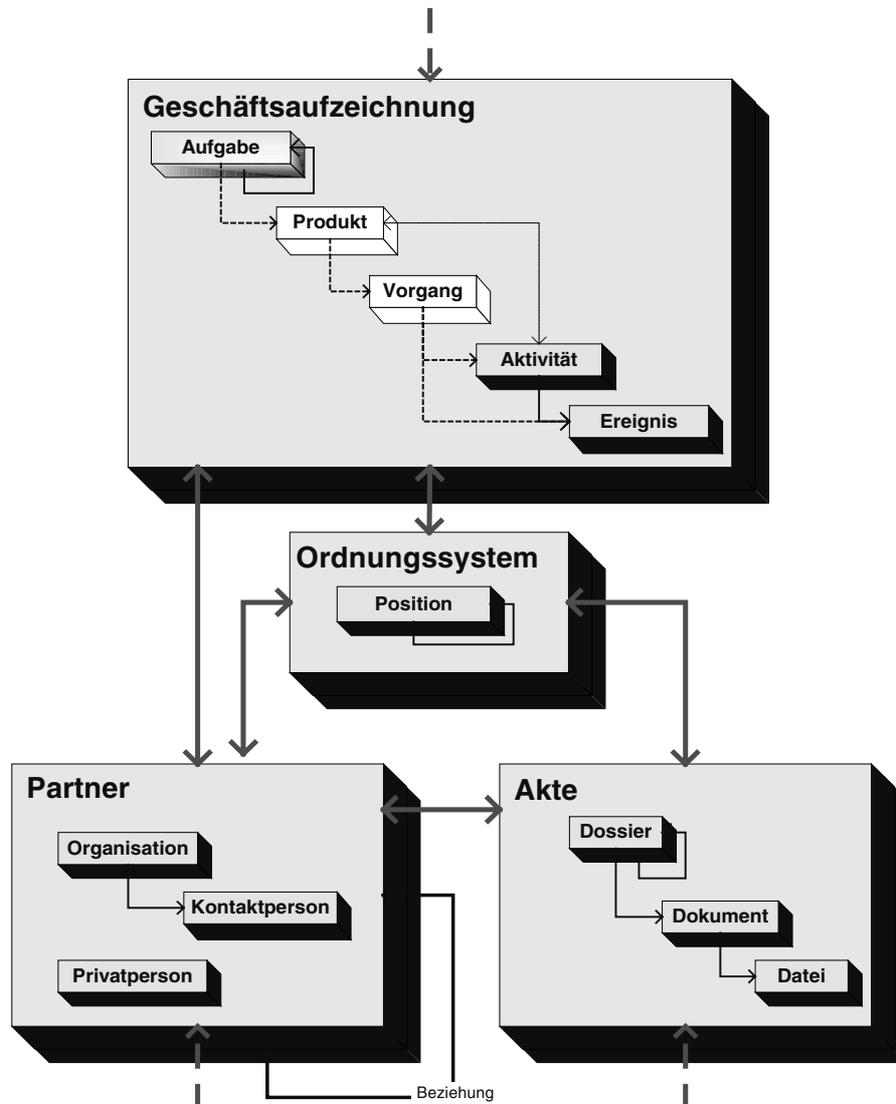


Abb. 2: Geschäftsobjekt-Modell ELGAR¹⁷

Michael Wettengel, Australische Metadaten-Standards: ein Thema für Archive? In: Digitale Herausforderungen für Archive (wie Anm. 1) S. 95–103. Zu ISO 15489 und MoReq: Nils Brübach, ebd. S. 105–112.

¹⁶ Der RKMS eignet sich für das Meta-Modell vor allem wegen seines hohen, an der australischen Verzeichnungs-tradition („series principle“) ausgerichteten Abstraktionsgrades. Als heuristisches Raster wurden auch die Functional Requirements für Evidence in Recordkeeping des Pittsburgh-Projekts beigezogen (vgl. <http://www.sis.pitt.edu/~nhprc/>). Die Daten- und Prozessmodelle des UBC-Projekts „The Preservation of the Integrity of Electronic Records“ (vgl. <http://www.interpares.org/UBCProject/>) wurden vor allem für den konkreten Design beigezogen.

¹⁷ Alle Entitäten bis auf Aufgaben, Produkte und Vorgänge sind heute im Produkt scopeDossier abgebildet. Auch die Entität „Aufgaben“ ist, da bereits für PRISMA entwickelt, verfügbar und kann mit wenig Aufwand eingebunden werden. Die Ergänzung durch zusätzliche Entitäten für Produkte und Vorgänge – das heisst Workflows, welche Aktivitäts- und Ereignisketten zusammenfassen können – ist als Bestandteil einer weiteren Ausbaustufe vorgesehen. Auf diese Weise soll ELGAR in Zielsetzungen der wirkungsorientierten Verwaltungsführung und des e-Government eingebunden werden können. Mit den unterbrochenen Pfeilen ist sodann angedeutet, dass Daten aus Drittsystemen übernommen und gemäss ELGAR-Standard abgelegt werden können. Bei Partner-

In der Realität werden freilich nicht alle möglichen Beziehungen implementiert. Der einfacheren Bedienbarkeit halber muss das theoretische Modell beim praktischen Design auf konkrete Erfordernisse abgestimmt werden können.

Systemdesign

Zu den schwierigsten Schritten im Software Engineering gehört das Umsetzen von Konzepten in lauffähige Programme. An diesem Punkt trennt sich die Spreu vom Weizen. Einen besonders kritischen Erfolgsfaktor stellt die Kommunikation zwischen Auftraggeberin und Designer dar. Im Falle von ELGAR brachten beide Partner bereits eine mehrjährige, erfolgreiche Erfahrung aus dem Projekt PRISMA mit. So konnte mit minimalem Zeitaufwand eine beidseits befriedigende Lösung für unmittelbare Bedürfnisse als auch für weiterreichende Perspektiven gefunden werden.

Den Angelpunkt des Produkts *scopeDossier* bildet die Akte (in der Schweiz: das Dossier). Eine solche muss konkret ausgeführte Geschäftsvorgänge dokumentieren können. Sie kann mit unterschiedlich umfangreichen Geschäften oder Projekten, aber auch mit konkreten Geschäftsfällen (Personen: Kunden, Partnern usw., Sachobjekten oder anderen Identifikationsmerkmalen) zusammenfallen.

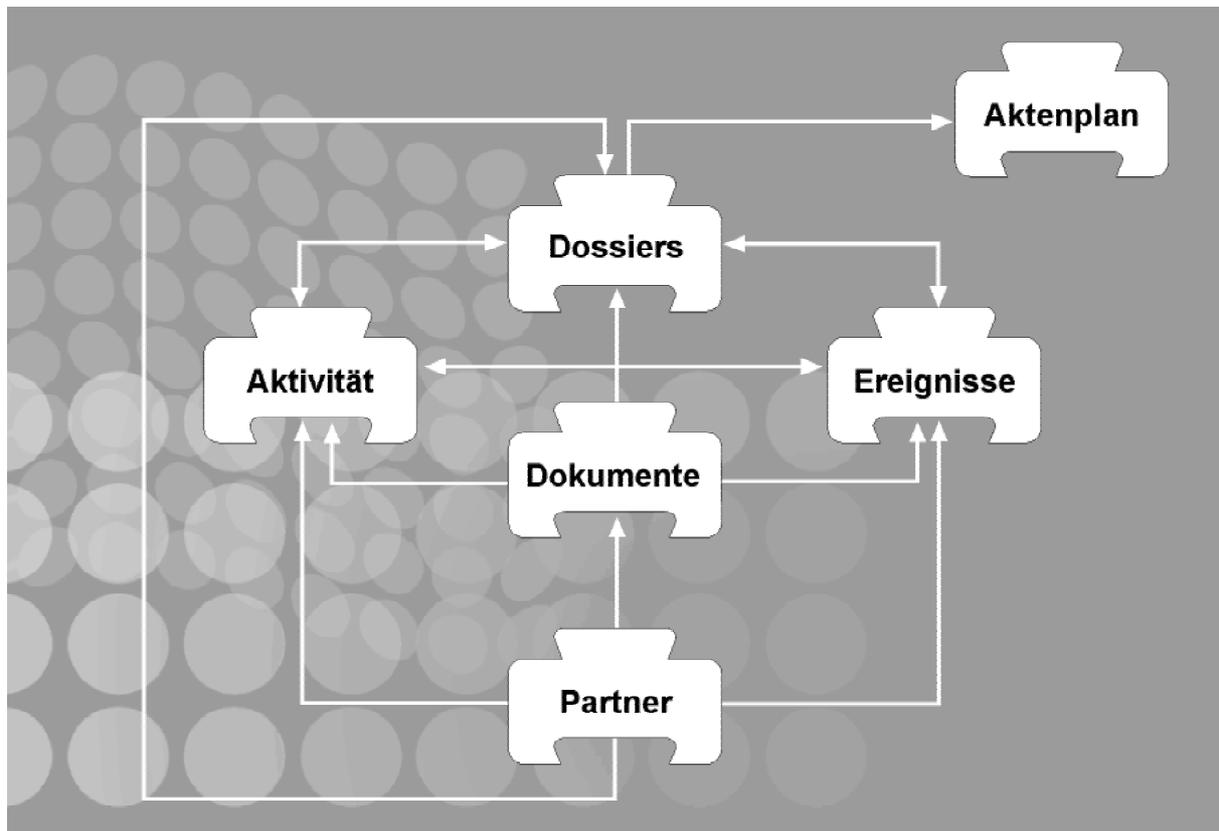


Abb. 3: Implementierte Geschäftsobjekte und deren Beziehungen im Produkt *scopeDossier*

Jedes geschäftsrelevante *Dossier* muss in ein Ordnungssystem bzw. *Aktenplan* eingeordnet werden können. Sodann sind mit ihm beliebig viele *Partner* (in unterschiedlichen Rollen), *Dokumente* (in evt. mehreren Versionen), beiläufig oder nachträglich aufgezeichnete *Ereignisse* sowie gezielt geplante *Aktivitäten* verknüpfbar. Aus erledigten *Aktivitäten* werden eben-

Daten ist dies bereits heute der Fall, indem diese via einen Service des kantonalen Datenmarktes anhand der zentralen Personendatenbestände übernommen oder aktualisiert werden können. Geschäftsprozess-abhängige Varianten von Metadata Sets können zu jeder Entität dank frei definierbarer Formulare in beliebiger Zahl verwendet werden.

falls *Ereignisse*, die pro Dossier in chronologischer Reihenfolge angezeigt werden können, so dass der Geschäftsgang jederzeit vollständig ersichtlich ist. Dokumente, die einer Aktivität oder einem Ereignis zugeordnet werden, sind automatisch auch mit dem entsprechenden Dossier und der massgeblichen Aktenplan-Position verknüpft.

Die abgebildeten Geschäftsobjekte gleichen den Seiten eines Würfels. Woher immer der Blick fällt: Es braucht höchstens zwei bis drei Mausklicks, um alle verknüpften übrigen Objekte – zum Beispiel sämtliche Dossiers, hängige Aktivitäten und Partner, denen ein bestimmtes Dokument zugeordnet wurde – vollständig und in der gewünschten Reihenfolge sichtbar zu machen.

Dieses Grundkonzept ist erweiterbar: Aktivitäten und Ereignisse lassen sich mittels Workflow-Technologien zu vordefinierten Vorgängen zusammenfassen und automatisieren. Die abgearbeiteten Schritte werden durch das System selbst dokumentiert. Unterschiedliche Dokumentenablagensysteme („document stores“) können integriert werden. Via Dokumentvorlagen werden Metadaten aus scopeDossier automatisch in erzeugte Dateien übertragen und umgekehrt.

Dank eines zugrundeliegenden Komponenten-Frameworks ist es ferner möglich, beliebige weitere Module, die zum Teil bereits aus dem Projekt PRISMA verfügbar sind, in die ELGAR-Strukturen einzubinden.¹⁸

Stand des Projekts

Wie bereits angeführt, erfolgte der Start zum Projekt ELGAR spontan, wenn auch durchaus auf der Grundlage eines breit angelegten Vorwissens und solider konzeptioneller Vorarbeiten. Inzwischen ist die erste Version des Produkts scopeDossier mit den in Abbildung 3 dargestellten Modulen betriebsbereit und wird in sechs Dienststellen aus fünf von insgesamt sieben Departementen eingeführt. Weitere Dienststellen stehen kurz vor einem Beschaffungsentcheid.¹⁹

Aus der Sicht des StABS ergibt sich die Chance, selbst als „code of best practice“ aufzutreten und exemplarisch vorzuführen, wie ein zukunftsorientiertes records management aussehen könnte. Davon verspricht es sich einen nachhaltigen Effekt, der im Verlauf der nächsten Jahre – und vor dem Hintergrund anderer tiefgreifender Veränderungsprozesse wie „new public management“, „electronic government“ usw. – weitere Kreise ziehen wird. Ein leistungsfähiges records management wird freilich nicht durch geeignete technische Werkzeuge allein in Gang gesetzt, sondern bedarf ergänzender organisatorischer Massnahmen, die über den Zuständigkeitsbereich des StABS hinausgehen. Entsprechende Schritte wurden eingeleitet.

¹⁸ Beispielsweise wurde für das Projekt PRISMA ein Aufgabenmodul entwickelt, das durch zusätzliche Geschäftsobjekte für Produktgruppen und Produkte bzw. Leistungen zur Unterstützung einer New Public Management-Wissensbasis ergänzt und allen ELGAR-BenutzerInnen zur Verfügung gestellt werden kann.

¹⁹ Unter den aktuellen Kunden finden sich Stellen mit sehr unterschiedlichen Geschäftsprozessen und Nutzungsbedürfnissen wie zum Beispiel das Hochbau- und Planungsamt, die Denkmalpflege, die zentralen Einwohnerdienste (mit Anbindung an den Datenmarkt) oder die kantonale Pensionskasse (mit ausgeprägten Falldossiers).

Mittelfristig muss ein gesamtkantonaler ELGAR-Standard definiert werden, der die wesentlichen Anforderungen an ein taugliches records management in einer software-unabhängigen Form verankert. Letztlich gilt es sodann, diesen Standard so umfassend als möglich in den gesamtkantonalen IT-Infrastrukturen zu verankern.²⁰

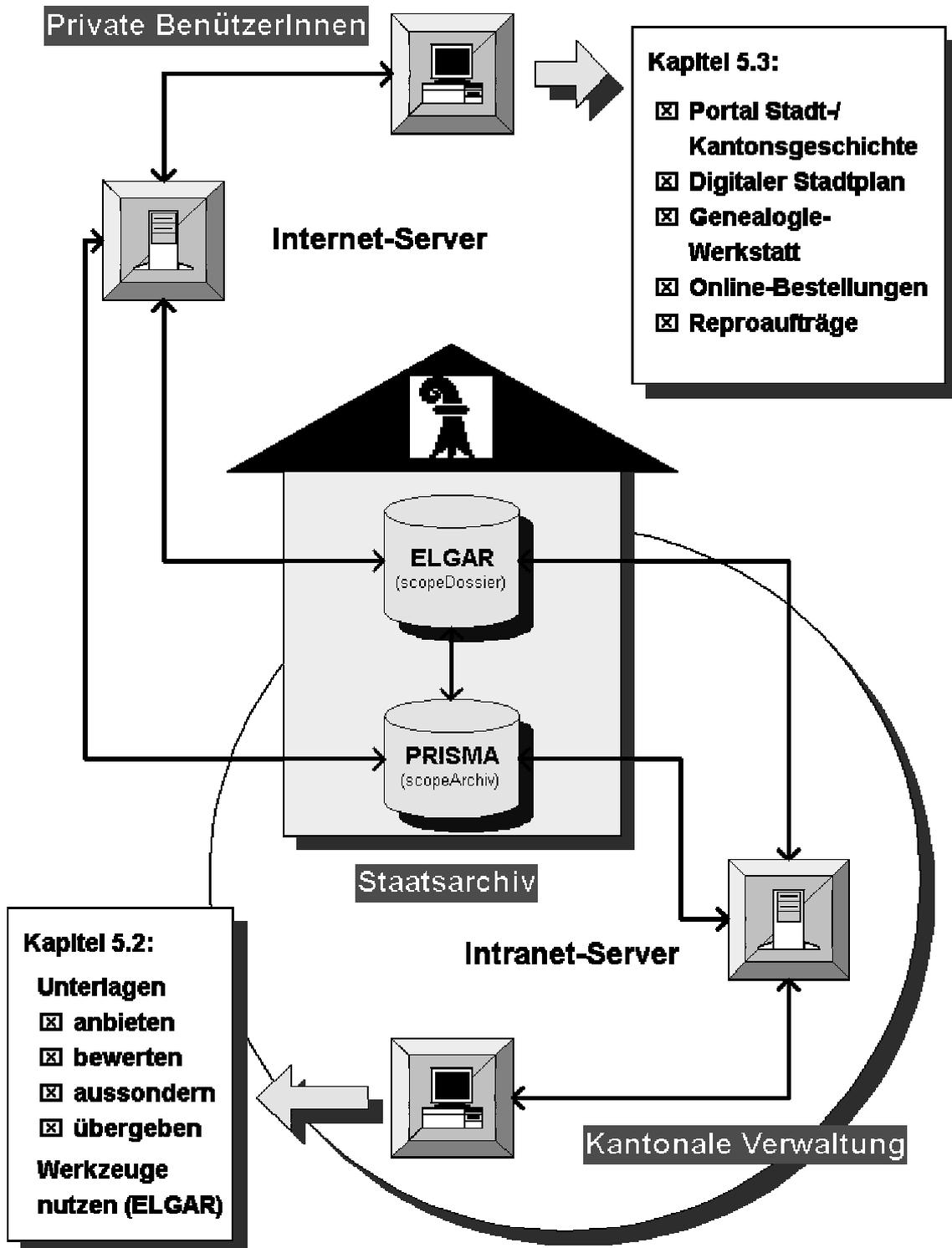


Abb. 4: Das Staatsarchiv als eGovernment-Vorreiter

²⁰ Zu den Bestandteilen dieses ELGAR-Standards gehören primär, anknüpfend an den ISO 15489-Standard, Organisationsvorschriften mit Regelwerken und Merkblättern.

Nächste Schritte

Im ersten Quartal des Jahres 2001 erarbeitete das StABS einen Ausgabenbericht zu Händen der Regierung und des Parlaments mit Bezeichnung „Zweite Informatisierungsetappe“. Unter dem Motto „Vom Aufbau der Instrumentarien zur Kundenorientierung“ strebt dieser ein attraktiveres Angebot gegenüber der Öffentlichkeit (mittels eines verbesserten Internet-Auftritts und online-Zugang zu digitalisierten Archivfindmitteln und rege nachgefragten Bild-dokumenten) wie auch gegenüber der Verwaltung an. Im letzteren Fall sollen vorarchivische Prozesse neu gestaltet und via Intranet abgewickelt werden. Dabei dient ELGAR zusammen mit PRISMA als Datenbank für operationelle Geschäftsdaten (vgl. Abbildung 4). Das Produkt scopeDossier soll in diesem Rahmen durch folgende **Zusatzfunktionalitäten** ergänzt werden:

- ein **Datenaustausch-Modul** zwischen verschiedenen ELGAR-Installationen und weiteren Geschäftspartnern
- ein „**records capturing assistent**“, der sich in beliebige Datenbank-Applikationen integrieren lässt, um die Ergebnisse bestimmter Transaktionen direkt in ELGAR (als Ereignisse, evt. Dokument-Metadaten mitsamt zugehörigen Dateien) abzulegen
- ein Geschäftsobjekt **Vorgang**, das mehrere Aktivitäten und/oder Ereignisse als Workflow zusammenfasst²¹
- einer Schnittstelle mit vordefinierbaren Aktualisierungsregeln für die **Publikation** von Dokumenten oder ganzen Dossiers **im Internet**.²²
- Module zur **Unterstützung von New Public Management**-Vorgaben und deren Umsetzung in der Verwaltungspraxis.²³

Aus der verwaltungsinternen Vernehmlassung dieses Ausgabenberichts ergaben sich verschiedene überaus nützliche **Seiteneffekte**. So wurde der Vertreter des StABS in eine Arbeitsgruppe, die sich mit der Ablösung der aktuellen Bürokommunikationsplattformen befasst, berufen, um darin die Schnittstellen zu ELGAR sicherzustellen. Gleichzeitig laufen Bestrebungen im Hinblick auf einen Ausbau der heutigen Datenlogistik zu einer umfassenden Geschäftsprozess- und Dokumentenlogistik sowie eines verwaltungsweiten Wissensmanagements. Diese Ziele wurden sowohl in der kantonalen eGovernment-Strategie als auch im vierjährigen Politikplan der Regierung verankert, wobei dem StABS und dem Statistischen Amt eine federführende Rolle zukommen soll.

Beim Ausbau von ELGAR wird davon ausgegangen, dass noch während mehrerer Jahre Papierablagen vorherrschen dürften. Der Übergang zu einem elektronischen records management wird sich im besten Fall über mehrere Zwischenphasen erstrecken. Angelehnt an den DOMEA- oder den norwegischen NOARK-Standard sind zwei Vor- und vier Ausbaustufen vorgesehen.

Zu den Vorstufen zählen der Ist-Zustand (Ablieferung von Abgabeverzeichnissen als Word- oder Excel-Dokumente bzw. vereinzelt Access- oder Filemaker-Datenbanktabellen). Dieser soll so rasch als möglich durch browserfähige und direkt in PRISMA integrierbare Formulare ersetzt werden.

²¹ Ein solcher Vorgang entspricht wohl in der Praxis ziemlich gut dem Vorgang, wie er in DOMEA gehandhabt wird.

²² Vgl. Recordkeeping Metadata Standard for Commonwealth Agencies. Version 1.0, Canberra 1999, S. 43–45.

²³ Vgl. Anm. 18. Die neuen Verwaltungsmodelle übertragen der Peripherie eine grössere Autonomie beim Vollzug ihrer Aufgaben. Als Kompensation muss die Transparenz dieses Aufgabenvollzugs für die jeweiligen Auftraggeber (Parlament, Regierung, Departement) verbessert werden können.

Die ELGAR-*Ausbaustufen* gliedern sich in:

- die Installation des *Aktenplan-Moduls* auf stand alone-Arbeitsplätzen (für das Registratur-Personal, mit Lesezugriff der übrigen MitarbeiterInnen via Browser)
- die Verwaltung von *Papierdossiers* (Aktenplan- und Dossier-Metadaten sowie weitere all-fällige Aufzeichnungen werden elektronisch, die Akten auf Papier archiviert)
- den Übergang zur *vollelektronischen Aktenführung* (mit elektronischer Archivierung von Metadaten und Dateien bei Einhaltung der ELGAR-Anwendungsstandards) sowie, als Fernziel
- den Vollausbau zu einem zentralen „business knowledge repository“.

Begleitend sind organisatorische Vorkehren zu treffen. Nicht zuletzt gilt es, die Qualität des Records Managements zu verbessern und die noch ungelöste Frage der Zuständigkeiten in diesem Aufgabenfeld zu klären.

Vorläufiges Fazit und Ausblick

Zusammenfassend ist eine geradlinige Kontinuität von der eingangs erwähnten Strategie zu den im Rahmen von ELGAR erfolgten sowie geplanten Umsetzungsschritten feststellbar. Die Idee, eine gesamt-kantonale Informatikstrategie gewissermassen von hinten (vom Archiv aus) aufzurollen, erschien anfänglich mehr als wilde Utopie, hat sich inzwischen aber in den Köpfen der verwaltungsinternen Partner bereits fest etabliert. Obwohl dies alles mit äusserst knappen Personalressourcen bewältigt werden musste, wurden die wesentlichen Ziele ausnahmslos erreicht.

Das Ressourcenproblem wird freilich auch in den nächsten Jahren das grösste Risiko darstellen. Durch den grossen Aufwand an konzeptionellen Vorarbeiten, der während Jahren geleistet werden musste, blieben zwangsläufig andere Aufgaben liegen. Die Erschliessungsrückstände sind in den letzten Jahren weiter angewachsen. Nach wie vor bleiben ferner grössere Unterlagenbestände aus dem Papierzeitalter, vereinzelt über das Jahr 1970 zurück, zu bewerten; das Risiko, dass partielle Überlieferungslücken entstehen, wurde angesichts der nur zurückhaltend wahrnehmbaren vorarchivischen Präsenz nicht kleiner.

In den nächsten Jahren steht deshalb die Strategie des StABS im Zeichen verstärkter Massnahmen zur Minimierung von Überlieferungsverlusten. Die Pionierphase in der Domäne der Archivinformatik gilt als abgeschlossen.

Dies bedeutet, dass andere kantonale Stellen fortan mehr Verantwortung übernehmen müssen. Andernfalls wäre das Gelingen des Gesamtvorhabens noch nicht ausreichend abgesichert.

Ein anderes Risiko stellt das gewählte Vorgehen dar. Üblicherweise werden in einem Projekt zuerst die fachlichen Anforderungen definiert und dann mittels Ausschreibung der geeignetste IT-Partner ausgewählt. Im vorliegenden Fall verlief es umgekehrt: Zuerst fanden sich zwei Partner mit konvergenten Ideen, um die Machbarkeit einer Idee anhand eines konkreten Produkts zu demonstrieren. Anschliessend sollte ein software-unabhängiger Standard erarbeitet werden.

Wegen der angetönten Ressourcenprobleme wird das StABS allerdings in nächster Frist keine Zeit für einen solchen Standard finden. Wie weit sich das vorhandene Produkt scopeDossier in der kantonalen Verwaltung ausbreiten kann, ist vorläufig ungewiss. Im allgemeinen überwiegt heute das Bedürfnis, Aktenführungs-Funktionen in weitflächig eingeführten Middleware-Systemen grosser Lieferanten einzubetten. Welcher der beiden Lösungsansätze sich in der Praxis durchsetzen wird, muss die Zukunft erweisen.

Ein Vorteil des in Basel gewählten Lösungsansatzes besteht zweifellos in der aussergewöhnlichen Kohärenz von recordkeeping und Archivierung. Dank Produkten, die vom gleichen

Hersteller stammen und auf identischen Konstruktionsprinzipien beruhen, wird das Risiko des Informationsverlustes beim Datentransfer von ELGAR nach PRISMA marginalisiert. Zweitens erlaubt die Idee des „records capturing“, neue, auf die Beschaffenheit elektronischer Unterlagen angepasste Bewertungsverfahren zu entwickeln.²⁴ Entsprechende Anpassungen und System-Ausbauten beinhalten in dieser Phase (wo noch keine marktfähigen Standardlösungen erhältlich sind) unter eingespielten Partnern das kleinere Risiko als in ambitionösen Grossprojekten. Drittens bietet das abgestufte Vorgehen Chancen, aus Misserfolgen oder Fehlern zu lernen und das eingesetzte Produkt sukzessive zu verbessern.

Um nochmals den Bogen zur eingangs erwähnten Strategie des StABS zu schlagen: Mit dem beschriebenen Instrumentarium ist es unerheblich, ob ein „records lifecycle“-Konzept oder ein „records continuum“-Ansatz verfolgt wird: Die Zuständigkeiten für Schriftgutverwaltung und Archivierung mögen zusammenfallen oder abgetrennt sein; das gedankliche Konstrukt des Überlieferungsprozesses bildet ein Ganzes.

²⁴ Statt nachträglich Samples zu bilden oder eine qualitative Auslese zu treffen, wird entscheidender werden, welche Systeme archivrelevante Geschäftsprozesse unterstützen und welche Arbeitsergebnisse (Aufzeichnungen und bearbeitete Dokumente) aus solchen Geschäftsprozessen heraus definiert und in ein recordkeeping system geleitet werden sollen. Vgl. dazu Margaret Hedstrom und David Bearman, Reinventing Archives for Electronic Records. In: Electronic Records Management Strategies, hrsg. von M. Hedstrom, 1995, S. 97 ff., ferner Niklaus Bütikofer, Bewertung als Voraussetzung für die elektronische Archivierung. In: Arbido 16 (2001) Nr. 4, S. 10–12.

Zum DOMEA-Konzept und seiner Umsetzung im Freistaat Sachsen – ein Werkstattbericht¹

Von Thekla Kluttig

Die sächsische staatliche Archivverwaltung unternahm 1999 einen ersten konkreten Schritt in Richtung Archivierung elektronischer Unterlagen mit der Beschaffung eines Rechners mit Sonderausstattung für das Sächsische Hauptstaatsarchiv Dresden, der zur Sichtung und Bewertung von elektronischen Unterlagen dienen sollte und bereits gedient hat. Parallel wurde die vom Hauptstaatsarchiv 1995/96 begonnene Erfassung von IT-Anwendungen in den anbieterpflichtigen Stellen fortgeführt.²

Doch je konkreter die Befassung mit IT-Anwendungen in Verwaltung und Justiz wurde, um so deutlicher wurde auch, wie begrenzt die Einflussmöglichkeiten des einzelnen Archivs bzw. Archivars auf die Einführung oder die Entwicklung einer IT-Anwendung sind. Als Beispiel möge nur das in mehreren Bundesländern eingesetzte Programm zur Geschäftsstellenautomation in Staatsanwaltschaften dienen, das auf der Tagung dieses Arbeitskreises im vorigen Jahr vorgestellt wurde.³

In einer Zwischenbilanz stellte die sächsische Archivverwaltung zu Beginn des vorigen Jahres fest, dass durch intensive und aktive Kontaktaufnahme mit anbieterpflichtigen Stellen zwar eine gewisse Sensibilisierung von Mitarbeitern aus Justiz und Verwaltung für die Problematik zu erzielen war (und uns die Dringlichkeit der Beschäftigung damit immer wieder vor Augen geführt wurde). Solange aber über die Vorgehensweise bei der Archivierung elektronischer Unterlagen nicht grundsätzlich entschieden war,⁴ stießen wir schnell an unsere Grenzen: Der Feststellung archivwürdiger elektronischer Unterlagen in einer anbieterpflichtigen Stelle folgten bei dort vorhandener Kooperationsbereitschaft zwei Fragen: Wie sollten die Daten abgegeben werden und wohin sollten sie abgegeben werden? Unser Ziel war daher ein zweifaches: Erstens mussten wir versuchen, bei der Einflussnahme auf die Verwaltung „weiter oben einzusteigen“, ggfs. auch Bund-Länder-übergreifend; zweitens mussten wir eine Entscheidung vorbereiten, wie die Archivierung elektronischer Unterlagen zukünftig erfolgen sollte.

Wie sind unsere organisatorischen Rahmenbedingungen? Die sächsische Staatsregierung unterhält das Hauptstaatsarchiv Dresden mit seinen Außenstellen, dem Staatsarchiv Chemnitz und dem Bergarchiv Freiberg, sowie das Staatsarchiv Leipzig. Die Dienst- und Fachaufsicht liegt beim Referat 44 (Archivwesen) im Staatsministerium des Innern, das gleichzeitig die Funktion einer Landesarchivdirektion wahrnimmt. Die Zuständigkeit für den Bereich der elektronischen Unterlagen ist hier Bestandteil der Zuständigkeit für das Aufgabengebiet „Überlieferungsbildung“: Diese Einbindung der elektronischen Unterlagen in den allgemei-

¹ Der folgende Beitrag wurde für die Drucklegung geringfügig überarbeitet, T.K.

² Vgl. Angela Ullmann, Vorsorge gegen das Datenchaos. In: Sächsisches Archivblatt 2 (1999) S. 11. – Dies., Maschinenlesbare Daten in staatlichen Behörden und Archiven in Sachsen. Vortrag auf dem 3. Brandenburgischen Archivtag am 25.03.1999. In: Brandenburgische Archive 13 (1999) S. 4–6, sowie: Elektronisches Nachweissystem „Anbieterpflichtige Stellen“ im Sächsischen Hauptstaatsarchiv Dresden. In: Sächsisches Archivblatt 1 (2000) S. 19–21.

³ Thekla Kluttig – Angela Ullmann, Digitale Registerdaten als Anbietersliste – Neue Möglichkeiten bei der Bewertung von Verfahrensakten der Staatsanwaltschaften. In: Auf der Suche nach archivischen Lösungsstrategien im digitalen Zeitalter. Beiträge zur 4. Jahrestagung des Arbeitskreises „Archivierung von Unterlagen aus digitalen Systemen“ im Stadtarchiv Mannheim, 10.–11.4.2000, hrsg. von Ulrich Nieß (Sonderveröffentlichungen des Stadtarchivs Mannheim 26), Mannheim 2001, S. 25–30.

⁴ Als Stichwort sei nur genannt: Eigenarchivierung oder Übertragung eines Teils der Archivierungsaufgaben an ein Landesrechenzentrum? Diese Frage wurde von deutschen Archivverwaltungen durchaus unterschiedlich beantwortet.

nen Kontext der Überlieferungsbildung und die Bündelung der entsprechenden Verantwortlichkeit erscheinen uns sehr wichtig, um eine sachlich nicht gerechtfertigte Isolierung des Aufgabengebiets zu verhindern. Der zuständige Referent im Referat Archivwesen hat als Angehöriger des Innenministeriums und als Fachvorgesetzter der Verantwortlichen in den vier Staatsarchiven gute Voraussetzungen, um die Kooperation mit den Fachverwaltungen und Maßnahmen zur Durchsetzung archivischer Anforderungen zu bündeln.

Ein zentraler Ansprechpartner hinsichtlich des Einsatzes von Informationstechnik ist die bei der Staatskanzlei angebundene KoBIT, die Koordinierungs- und Beratungsstelle für Informationstechnik in der sächsischen Landesverwaltung.⁵ Die KoBIT steht wiederum in enger Verbindung mit dem Arbeitskreis Informationstechnik (AK IT), in dem die Leiter der IT-Referate der Ministerien vertreten sind und u.a. ressortübergreifende Standards festlegen. So empfahl der AK IT 1997, Ausschreibungen, die Leistungen im Zusammenhang mit der Vorgangsbearbeitung und/oder dem Dokumentenmanagement beinhalten, auf der Grundlage der Anforderungen und der Gliederungssystematik des Projektes „DOMEA“ der Koordinierungs- und Beratungsstelle der Bundesregierung für Informationstechnik in der Bundesverwaltung (KBSt) vorzunehmen.⁶ Die Orientierung am DOMEA-Konzept wurde damit für die Staatsverwaltung verbindlich.

Warum und in welcher Form befasst sich die sächsische Archivverwaltung mit dem DOMEA-Konzept? In der sächsischen Verwaltung wird keineswegs schon eifrig elektronische Vorgangsbearbeitung betrieben – elektronische Vorgangsbearbeitung und die Führung elektronischer Akten sind bisher nur pilotiert.⁷ Seit Jahren im Echtbetrieb sind dagegen verschiedene Systeme zur Schriftgutverwaltung, die die Metadaten zum analogen Schriftgut enthalten. Der Befund ist einfach: In zahlreichen anbieterpflichtigen Stellen des Freistaates Sachsen, so in allen Staatsministerien, wird mit verschiedenen Registraturprogrammen gearbeitet. Alle Ministerien müssen dem Hauptstaatsarchiv Dresden ihre Unterlagen anbieten. Das Hauptstaatsarchiv wird einen Teil dieser Unterlagen übernehmen – und die einzig sinnvolle Lösung kann nur sein, die in der Verwaltung vorhandenen Metadaten als Bewertungsgrundlage zu nutzen und als Erschließungsdaten mit zu übernehmen. Das heißt, dass die verschiedenen Registraturprogramme in der Lage sein müssen, vom Archiv definierte, einheitliche Standards einzuhalten und ein einziges im Archiv eingesetztes Erschließungsprogramm mit Metadaten zu beliefern. Und damit befinden wir uns in der Stufe 1 des DOMEA-Konzepts.

Das Referat Archivwesen sah es auch daher als große Chance an, in einer Arbeitsgruppe unter Federführung der KoBIT mitzuwirken, die im Sommer 2000 damit beauftragt wurde, Empfehlungen zur „Archivierung elektronischer Akten“ zu erarbeiten. Das Innenministerium ist darin mit dem IT-Referat und dem Referat Archivwesen vertreten. Über einige Ergebnisse der Arbeitsgruppentätigkeit, die sich auf das DOMEA-Konzept und seine Umsetzung in Sachsen beziehen, soll im folgenden berichtet werden.⁸

⁵ Verwaltungsvorschrift der Sächsischen Staatsregierung zur Koordinierung der Planung von Vorhaben und des Einsatzes von Haushaltsmitteln bei Vorhaben im Bereich der Informationstechnik (IT) in der Landesverwaltung des Freistaates Sachsen vom 16. November 1998. In: Sächsisches Amtsblatt 1998, Bl. 846.

⁶ Koordinierungs- und Beratungsstelle der Bundesregierung für Informationstechnik in der Bundesverwaltung (KBSt) (Hrsg.), DOMEA. Aufbau eines Pilotsystems für Dokumentenmanagement und elektronische Archivierung im IT-gestützten Geschäftsgang, Bonn 1997.

⁷ Auch deshalb ist eine Konzentration auf diese Verfahren unter Vernachlässigung von IT-Fachverfahren und Informationssystemen etc. nicht zu rechtfertigen.

⁸ Die Protokolle der Sitzungen sind auf dem Kommunikations-Server für die IT-Experten der sächsischen Landesverwaltung (Kositex) einsehbar. Dieser von der KoBIT betriebene Server dient als wesentliches Kommunikations- und Informationsmittel hinsichtlich aller Aspekte des IT-Einsatzes in der Landesverwaltung.

Das DOMEA-Konzept dürfte in diesem Kreis hinlänglich bekannt sein – trotzdem will ich drei wesentliche Merkmale kurz in Erinnerung rufen:⁹

1. Es unterscheidet zwischen drei Formen der IT-unterstützten Vorgangsbearbeitung: Systemen zur Schriftgutverwaltung, Systemen zur elektronischen Aktenablage und Systemen zur elektronischen Vorgangsbearbeitung. Das vom Bundesarchiv mitentwickelte Stufenkonzept der Archivierung und Aussonderung soll bei allen drei Nutzungsformen praktiziert werden.
2. Es entwickelt ein vierstufiges Aussonderungsverfahren: Zunächst können in einem Aussonderungskatalog unbefristete Vernichtungsgenehmigungen auf Aktenplanebene erteilt werden. Dann werden anbiutungspflichtige Akten (bzw. Aktenschnitte) in einem elektronischen Anbieterverzeichnis mit Datensatzbeschreibungen zu drei Ebenen aufgelistet und so dem Bundesarchiv angeboten. Daraufhin entscheidet das Bundesarchiv auf der Grundlage des Anbieterverzeichnisses auf Aktenebene über die Archivwürdigkeit und übersendet das entsprechend ergänzte Anbietersverzeichnis als Bewertungsverzeichnis zurück an die aktenführende Stelle. Schließlich übersendet die aktenführende Stelle das um die kassablen Positionen bereinigte Bewertungsverzeichnis als Abgabeverzeichnis – zusammen mit der eigentlichen Aussonderungsdatei – an das Bundesarchiv. Aussonderungseinheiten sind allerdings nicht elektronische Akten, sondern sog. Aktenschnitte, die die aussonderungsreifen Vorgänge der Akte umfassen. Wird DOMEA auf papierbasiertes Schriftgut angewandt, bezieht sich die Aussonderung statt auf Vorgänge auf einzelne Bände der Akte.¹⁰
3. Das elektronische Anbieterverzeichnis enthält Datensatzbeschreibungen zu drei Ebenen: zu der (abgeleiteten) *Betreffseinheit* (d.h. der untersten Ebene des Aktenplans), zu der Akte und zu den Vorgängen in der Akte. Zu den Datensatzbeschreibungen später mehr.

DOMEA geht also zunächst von einer Dreistufigkeit des physisch bzw. digital vorliegenden Schriftguts aus: Dokumente werden zu Vorgängen formiert, diese zu Akten genommen, die wiederum zu bestimmten *Betreffseinheiten* angelegt wurden. Der Aktenplan bildet „nur“ den Ordnungsrahmen; der tatsächliche Aktenbestand wäre in einem Aktenverzeichnis nachzuweisen. Zur Verdeutlichung ein Beispiel:

<i>Gruppe</i>	031	<i>Beamte</i>
<i>Untergruppe:</i>	0313	<i>Ausbildung und Prüfung von Beamten</i>
<i>Betreffseinheit:</i>	0313.80	<i>Ausbildungs- und Prüfungsordnungen des Bundes und der anderen Bundesländer</i>
<i>Sachakte:</i>	0313.80/0	<i>Sammelsachakte</i>
<i>Sachakte:</i>	0313.80/1	<i>APOs gehobener Archivdienst des Bundes und ...</i>
<i>Sachakte:</i>	0313.80/2	<i>APOs höherer Archivdienst des Bundes und ...</i>
<i>Sachakte:</i>	0313.80/3	<i>APOs gehobener Verwaltungsdienst des Bundes ...</i>
		<i>etc.</i>

⁹ Vgl. zum folgenden: KBSt (Hrsg.): Konzept zur Aussonderung elektronischer Akten (Schriftenreihe der KBSt 40), Köln 1998 sowie KBSt (Hrsg.), Konzept Papierarmes Büro (DOMEA-Konzept). Dokumentenmanagement und elektronische Archivierung im IT-gestützten Geschäftsgang (Schriftenreihe der KBSt 45), Köln 1999. Vgl. auch die aktuellen Informationen unter www.kbst.bund.de. Einen knappen, aber präzisen Überblick über das Pilotprojekt DOMEA und weitere Projekte zur IT-gestützten Vorgangsbearbeitung bietet Michael Wettengel, *Verwaltungsmodernisierung und IT-gestützte Vorgangsbearbeitung in der Bundesverwaltung*. In: *Auf der Suche nach archivischen Lösungsstrategien im digitalen Zeitalter* (wie Anm. 3) S. 35–46.

¹⁰ Einen völlig anderen Weg geht das Dänische Reichsarchiv: Daten aus elektronischen Registraturprogrammen werden in Fünfjahresblöcken übernommen. Die Behörde schließt also alle fünf Jahre die Datenbank, Aktenvermerk über die Ergebnisse eines Besuchs im Dänischen Reichsarchiv im April 2001, Registratur des Sächsischen Staatsministeriums des Innern, Az.: 44-7510.12/1; s.a. Kirsten Villadsen Kristmar, *Neue Vorschriften des Dänischen Staatsarchivs für die elektronische Datenablage*. In: *INSAR. Europäische Archivnachrichten* 8 (2001) S. 3 u. 7.

Die Sachakte 0313.80/1 umfasst z.B. die Vorgänge zu den einzelnen APOs des gehobenen Archivdienstes des Bundes und der Länder; diese Vorgänge umfassen wiederum eine verschiedene Zahl von Dokumenten.

Bei der Diskussion der Übertragbarkeit von DOMEA stieß nun unsere Arbeitsgruppe schnell auf ein grundsätzliches Problem: In den Behörden des Freistaates Sachsen werden in der Regel keine Akten im oben dargestellten Sinne gebildet.¹¹ Die Registraturordnung in ihrer novellierten Fassung vom Februar 2000 legt nur fest: „Unterlagen, die sachlich zusammengehören, sollen vorgangsweise geordnet werden.“¹² Und der Landeseinheitliche Aktenplan (LEAP) bezeichnet als „Akte“ die kleinste Aktenplaneinheit, in der Terminologie von DOMEA also die (ggfs. abgeleitete) *Betreffseinheit*. Zu dieser Akte werden in sächsischen Behörden Vorgänge angelegt. Die Metadaten zu diesen Vorgängen und zu den einzelnen Schriftstücken in den Vorgängen werden im Registraturprogramm registriert. Die Akten-, besser: Vorgangsbildung stellt sich also wie folgt dar:

<i>Gruppe</i>	<i>031</i>	<i>Beamte</i>
<i>Untergruppe:</i>	<i>0313</i>	<i>Ausbildung und Prüfung von Beamten</i>
<i>Akte:</i>	<i>0313.80</i>	<i>Ausbildungs- und Prüfungsordnungen des Bundes und der anderen Bundesländer</i>
<i>Vorgang:</i>	<i>0313.80/1</i>	<i>APO gehobener Verwaltungsdienst in Bayern</i>
<i>Vorgang:</i>	<i>0313.80/2</i>	<i>APO höherer Archivdienst in Nordrhein-Westfalen</i>
<i>Vorgang:</i>	<i>0313.80/3</i>	<i>APO gehobener Archivdienst in Schleswig-Holstein</i>
<i>Vorgang:</i>	<i>0313.80/4</i>	<i>APO höherer Archivdienst in Baden-Württemberg</i>
<i>Vorgang:</i>	<i>0313.80/5</i>	<i>APO gehobener Verwaltungsdienst in Hessen</i>
<i>etc.</i>		

Die so genannte Akte „Ausbildungs- und Prüfungsordnungen des Bundes und der anderen Bundesländer“ ist virtuell; das physisch vorliegende Schriftgut ist dagegen zweistufig aufgebaut (Vorgang / Dokumente). In der Praxis sieht das dann so aus:¹³

Aktenzeichen	Titel	Stand	Bearbeiter	Dok.
<u>64-0313.80-07/10</u>	Staatsarchivreferendare höherer Dienst Nordrhein-Westfalen	zda		1
<u>64-0313.80-07/11</u>	Ausbildungs- und Prüfungsordnung für den gehobenen und höheren Archivdienst in Schleswig-Holstein	Bearb	Ref. 64	1
<u>64-0313.80-07/6</u>	Ausbildungs- und Prüfungsordnung für den höheren Archivdienst im Land Niedersachsen	Bearb	Kluttig	1
<u>64-0313.80-07/7</u>	Entwurf einer Verordnung über die Ausbildung und Prüfung für die Laufbahn des höheren u. gehobenen Archivdienstes des Landes Mecklenburg-Vorpommern (Ausbildungs- und Prüfungsordnung höherer Archivdienst-APOhAD M-V)	zda		4
<u>64-0313.80-07/8</u>	Entwurf einer Landesverordnung über die Ausbildung und Prüfung für die Laufbahnen des gehobenen u. höheren Archivdienstes des Landes Schleswig-Holstein (APOArchD)	zda		1
<u>64-0313.80-07/9</u>	Ausbildung mittlerer Archivdienst Bayern	Bearb	Kluttig	1

¹¹ Vgl. hierzu Thekla Kluttig, Akten, Vorgänge, Dokumente – Tendenzen in der behördlichen Schriftgutverwaltung. In: Der Archivar 52 (2000) S. 22–26.

¹² Gemeinsame Verwaltungsvorschrift der Sächsischen Staatskanzlei, des Sächsischen Staatsministerium des Innern u.a. über die Verwaltung von Unterlagen (VwV Registraturordnung – VwVRegO) vom 2.2.2000. In: Sächsisches Amtsblatt Nr. 8 vom 24.2.2000, S. 158–161.

¹³ Screen-Shot einer Maske des im Sächsischen Staatsministerium des Innern eingesetzten Registraturprogramms, die ein Rechercheergebnis zeigt.

Eine wichtige Frage war daher: Sieht DOMEA eine dreistufige Schriftgutverwaltung zwingend vor? Ist die Bildung von physischen Akten und Vorgängen also eine Voraussetzung für die Orientierung an DOMEA? Wir haben das Konzept noch einmal intensiv studiert und festgestellt, dass dem nicht so ist. So ist z.B. mit Bezug auf die Anbietung auf den drei Ebenen *Betreffseinheit*, *Akte* und *Vorgang* festgeschrieben: „Wird auf eine Vorgangsbildung verzichtet, so wird statt der Angaben zu den Vorgängen ein entsprechendes Verzeichnis zu den Dokumenten in einer Aussonderungsportion geliefert.“¹⁴ Die Bildung von Vorgängen innerhalb von Akten ist bei DOMEA lediglich optional: Daher bleibt letztlich vage, ob die elektronische Akte zweistufig (*Dokumente*, *Akte*) oder dreistufig (*Dokumente*, *Vorgänge*, *Akte*) gebildet wird. Die Möglichkeit einer zweistufigen Schriftgutverwaltung wird also eingeräumt.

Welche Konsequenzen hat nun dieser Befund für die Umsetzbarkeit der ersten Stufe von DOMEA in den sächsischen Behörden einschließlich der Archivverwaltung – und warum beschäftigen wir uns so intensiv mit dieser Frage?

- Wir müssen entscheiden, auf welcher Ebene wir unsere Bewertungsentscheidung treffen wollen und welche Informationen uns dazu zur Verfügung stehen sollen.
- Wir müssen in diesem Zusammenhang prüfen, ob die von DOMEA vorgeschlagenen Datensatzbeschreibungen bei Anbietungs-, Bewertungs- und Abgabeverzeichnissen unseren Bedürfnissen entsprechen.
- **Wir müssen entscheiden, auf welcher Ebene wir die „Archivnummer“ (in der Terminologie von DOMEA), also die Archivalien- und Lagerungssignatur, vergeben und was damit zukünftig unsere physische Archivalieneinheit im Magazin sein wird.**
- Wir müssen entscheiden, welche Metadaten zukünftig als Erschließungsdaten in das Archiv übernommen werden – denn die traditionelle Erschließung durch Verzeichnung von Akten wird es zukünftig nicht mehr geben.
- Mit den Entscheidungen, die wir jetzt treffen, werden die Weichen auch für die zukünftige Anbietung und Bewertung elektronischer Akten gestellt.

Auf die ersten beiden Punkte – und damit unsere derzeitigen Ansätze für eine modifizierte Umsetzung des DOMEA-Konzepts – möchte ich nun eingehen. Wie bereits dargestellt, sieht das Aussonderungsverfahren nach DOMEA vor, dass das Anbieterverzeichnis Datensatzbeschreibungen zu drei Ebenen enthält: der (abgeleiteten) *Betreffseinheit*, der *Akte* und den *Vorgängen*. Bewertet wird ausschließlich auf *Aktenebene*, und nur die Datensatzbeschreibungen dieser Ebene werden im Bewertungsverzeichnis zurück an die Behörde gesandt.

In Sachsen sind nun die Ebenen „(abgeleitete) *Betreffseinheit*“ und „*Akte*“ identisch. Wir mussten also in einem ersten Schritt überlegen, ob wir DOMEA entsprechend *eine* Ebene festlegen können, auf der die Bewertung stattfindet. Dabei stellten wir fest, dass es zwei Varianten für die Umsetzung von DOMEA auf unsere Verhältnisse gibt:

1. Variante

DOMEA	Sachsen
(abgeleitete) <i>Betreffseinheit</i>	(abgeleitete) <i>Betreffseinheit</i>
<i>Akte</i>	<i>Vorgang</i>
<i>Vorgang</i>	–
<i>Dokument</i>	<i>Dokument</i>

¹⁴ Konzept zur Aussonderung elektronischer Akten (wie Anm. 9) S. 26 und 42, s.a. Thekla Kluttig, *Akten, Vorgänge, Dokumente* (wie Anm. 11) S. 26.

2. Variante

DOMEA	Sachsen
(abgeleitete) Betroffenseinheit (Ebene Aktenplan)	Betroffenseinheit (sechsstellig) (LEAP)
Akte	(abgeleitete) Betroffenseinheit (Fachaktenpläne der Ressorts)
Vorgang	Vorgang
Dokument	Dokument

Bei der Variante 1 würde auf Vorgangsebene bewertet. Das würde eine sehr differenzierte Bewertungsentscheidung ermöglichen. Aber: Allein im Staatsministerium des Innern werden jährlich durchschnittlich rd. 12.900 Vorgänge zu rd. 1980 Akten (Betroffenseinheiten) angelegt. Diese Zahl von Vorgängen müssten wir also – jeweils etwa zehn Jahre später – jährlich bewerten.

Variante 2 würde hingegen die Zahl der Einheiten, die zu bewerten sind, deutlich reduzieren. Im Falle von Behörden, die den LEAP nicht durch differenzierte Fachaktenpläne erweitern, sondern stattdessen eine hohe Zahl von einzelnen Vorgängen zu einzelnen Aktenplanpositionen anlegen, würden wir uns aber mit der Entscheidung für Variante 2 auf eine ausschließliche Bewertung auf Aktenplanebene verpflichten.

Wir haben beide Varianten diskutiert und festgestellt, dass wir uns nicht auf eine Bewertungsebene beschränken möchten. Stattdessen werden wir versuchen, folgendes Verfahren durchzusetzen: Für die Mehrheit der Fälle wird schon auf der Ebene der (abgeleiteten) Betroffenseinheit bewertet. In speziellen Fällen, z.B. bei bestimmten Betroffenseinheiten und abhängig von der anbietenden Behörde, werden die Vorgänge genauer betrachtet. Falls sinnvoll, wird eine Bewertungsentscheidung auf dieser Ebene gefällt. Das bedeutet, dass wir uns für die Variante 2 entscheiden, weil sie auf den Normalfall besser zugeschnitten ist, die Ausnahme aber ebenfalls zulässt.¹⁵

Damit komme ich zum zweiten Punkt: Wir prüfen in diesem Zusammenhang, ob die von DOMEA vorgeschlagenen Datensatzbeschreibungen bei Anbiere-, Bewertungs- und Abgabeverzeichnissen unseren Bedürfnissen entsprechen. Wir sind hier noch mitten in der internen Diskussion, deshalb steht das folgende unter einem gewissen Vorbehalt. Die Arbeitsgruppe hat festgestellt, dass wir uns in vielen Punkten an den Datensatzbeschreibungen von DOMEA orientieren können. Diese Datensatzbeschreibung – hier am Beispiel des Bewertungsverzeichnisses – ist im wesentlichen wie folgt:¹⁶

Lfd. Nr.	Feldname	Datentyp	Bedeutung
1	beh_kurz	Text (10)	Amtliche Kurzbezeichnung der abgebenden Behörde
2	akt_stelle	Text (30)	Bezeichnung der aktenführenden Stelle oder Organisationseinheit
3	dat_uhrzeit	Datum / Uhrzeit	Erstellungsdatum und Uhrzeit des Abgabeverzeichnisses
4	satzart	Text (1)	Satzart, hier: A = Akte
5	lfd_nr	Num	Laufende Nummer der Aufbewahrungseinheit
6	aktenzeichen	Text (30)	Aktenzeichen
7	inhalt	Text (250)	Inhaltsangabe
8	vg_num	Text (30)	Vorgangsnummern von bis; bei Papierakten: Bandnummern von bis
9	laufzeit	Text (30)	Laufzeit von bis

¹⁵ Dieser Bewertungsschritt muss natürlich eingebunden sein in einen analytischen Prozess, s. hierzu Robert Kretschmar, Die „neue archivistische Bewertungsdiskussion“ und ihre Fußnoten. In: Archivalische Zeitschrift 82 (1999) S. 36 f.

¹⁶ Datensatzbeschreibung zu einer Akte im Bewertungsverzeichnis. In: Konzept zur Aussonderung elektronischer Akten (wie Anm. 9) S. 37.

Lfd. Nr.	Feldname	Datentyp	Bedeutung
10	aufbew_dauer	Num	Aufbewahrungsdauer (nur für Papierakten)
11	aufbew_end	Num	Endjahr der Aufbewahrungsfrist (nur für Papierakten)
12	medium	Text (1)	Papierform oder elektronische Form
13	archiv_num	Num	Archivnummer des Bundesarchivs
14	erg_buchst	Text (3)	Ergänzungsbuchstabe der Archivnummer
15	bew_status	Text (1)	Bewertungsstatus (Bewertungsergebnis des Bundesarchivs)

Wir halten aber einige Modifizierungen für sinnvoll. Abweichend von DOMEA sehen wir z.B. nicht vor, dass jeder Datensatz die Angaben zur abgebenden Behörde und zu Datum und Uhrzeit der Erstellung des Verzeichnisses enthält. Die Redundanz lässt sich vermeiden, indem dem Verzeichnis ein Vorsatz nach folgendem Muster vorangestellt wird:

Vorsatz

Feldname	Länge	Bedeutung
Satzart	1	Header
VerzArt	3	Art des Verzeichnisses (AUS, BEW oder ABG)
Ident	10	Identifikator des Verzeichnisses (pro Behörde eindeutig)
Behörde	100	Kurzbezeichnung der abgebenden Behörde
Herkunft	255	abweichende Herkunft der Unterlagen, wenn nicht bei der abgebenden Behörde entstanden
ErstDatum	10	Erstellungsdatum im Format TT.MM.JJJJ
ErstUhrz	5	Erstellungsuhrzeit im Format HH:MM

Ein Nachsatz nach dem Verzeichnis wiederholt diese Angaben und enthält darüber hinaus die Gesamtzahl der Sätze. Er ermöglicht damit eine Kontrolle der Vollständigkeit des Verzeichnisses. Auch die Satzbeschreibungen sind aus unserer Sicht noch in weiteren Punkten zu modifizieren:

- Das Feld Vorgangsnummer von–bis halten wir für nicht nötig. Wir würden daher auf dieses Feld verzichten.
- Die Angabe „von–bis“ in einem Textfeld erschwert die maschinelle Verarbeitung. Das Feld Laufzeit von-bis wird daher in zwei Felder getrennt.
- Es gibt kein Feld, in das die abgebende Stelle eine interne Nummer eintragen könnte, die für das Archiv irrelevant ist, die aber beim Rücklesen des Bewertungsverzeichnisses ins System die Zuordnung von Akten und Vorgängen im Registraturprogramm wesentlich vereinfachen würde. Wir würden den Satz um ein solches Feld ergänzen.
- **Es gibt kein Feld, in das die Behörde einen Bewertungsvorschlag eintragen könnte – eine Regelung, die unsere Registraturordnung vorsieht. Wir würden den Satz ebenfalls um ein solches Feld ergänzen.**
- Es gibt kein Feld, in das das Ende der Aufbewahrungsfrist eingetragen werden könnte. Da das Sächsische Archivgesetz und die Registraturordnung die vorfristige Abgabe von Schriftgut als Möglichkeit vorsehen, ist ein solches Feld aus unserer Sicht erforderlich.
- Es gibt kein Feld, in dem ein Verweis auf Personendatenschutz erfolgen könnte.

Als derzeitiger Diskussionsstand in der von der KoBIT geleiteten Arbeitsgruppe „Archivierung elektronischer Unterlagen“ ergeben sich folgende Schlussfolgerungen: Auch wenn die sächsischen Behörden keine Akten im Sinne von DOMEA bilden, ist die Aussonderung behördlichen Schriftguts in Anlehnung an DOMEA möglich. Die geplante Empfehlung zur Archivierung elektronischer Unterlagen kann allerdings nicht lediglich auf DOMEA verweisen. Sie muss vielmehr eine Begriffsklärung voranstellen, das Verfahren zur Transformation darlegen und sachsenspezifische Regelungen treffen und erläutern.

Anwendungsspezifische Fachkonzepte zur Umsetzung archivischer Anforderungen für die Aussonderung elektronischer Akten:

Das Beispiel des Fachkonzepts Favorit[®]-OfficeFlow[®]

Von Michael Wettengel

1. Die „DOMEA[®] Empfehlungen“ für die Aussonderung elektronischer Akten als Ausgangspunkt

Auf den letzten beiden Tagungen des Arbeitskreises „Archivierung von Unterlagen aus digitalen Systemen“ wurde bereits über das „Konzept zur Aussonderung elektronischer Akten“ berichtet, das 1998 im Rahmen des Projektes DOMEA[®] erarbeitet wurde.¹ Es enthält „Empfehlungen“ des Bundesarchivs für die Bundesverwaltung, wie elektronische Akten aus IT-gestützten Vorgangsbearbeitungssystemen ausgesondert, bewertet und archiviert werden können.² Diese „Empfehlungen“, die vom Interministeriellen Koordinierungsausschuss gebilligt wurden, sind Bestandteil des „Konzepts Papierarmes Büro (DOMEA[®]-Konzept)“, das 1999 veröffentlicht wurde; sie gelten unabhängig von bestimmten Produkten oder spezifischen Anwendungen.³

Die fortgeschriebenen Anforderungen des DOMEA[®]-Konzepts bildeten die Grundlage für eine Ausschreibung, die Ende 1999 vom Beschaffungsamt des Bundesministeriums des Innern durchgeführt wurde, um Produkte auf Übereinstimmung mit dem Konzept Papierarmes Büro zu überprüfen.⁴ Dies bedeutet, dass die „Empfehlungen“ des Bundesarchivs auch ein Bestandteil der Anforderungen für die Evaluierung waren. Die Ausschreibung mündete im Abschluss von Rahmenverträgen für konzeptkonforme Produkte. Das Ausschreibungsver-

¹ Zu den Berichten auf den Tagungen des Arbeitskreises „Archivierung von Unterlagen aus digitalen Systemen“ 1999 und 2000 vgl. Andreas Engel – Andrea Kern, DOMEA. Pilotprojekt der Bundesverwaltung zum Dokumentenmanagement und zur elektronischen Archivierung im IT-gestützten Geschäftsgang. In: Michael Wettengel (Hrsg.), Digitale Herausforderungen für Archive (Materialien aus dem Bundesarchiv 7), Koblenz 1999, S. 79–88; Michael Wettengel, Aussonderung und Archivierung elektronischer Akten. In: Ebd. S. 89–96; ders., Verwaltungsmodernisierung und IT-gestützte Vorgangsbearbeitung in der Bundesverwaltung. In: Auf der Suche nach archivischen Lösungsstrategien im digitalen Zeitalter. Beiträge zur 4. Jahrestagung des Arbeitskreises „Archivierung von Unterlagen aus digitalen Systemen“ im Stadtarchiv Mannheim, 10.–11.4.2000, hrsg. von Ulrich Nieß (Sonderveröffentlichungen des Stadtarchivs Mannheim 26), Mannheim 2001, S. 35–46.

² Andreas Engel – Andrea Kern – Michael Wettengel, Empfehlung des Bundesarchivs zur Aussonderung elektronischer Akten. In: Koordinierungs- und Beratungsstelle der Bundesregierung für Informationstechnik in der Bundesverwaltung (KBSt) (Hrsg.), Konzept zur Aussonderung elektronischer Akten (Schriftenreihe der KBSt 40), Köln 1998, S. 13–44.

³ KBSt (Hrsg.), Konzept Papierarmes Büro (DOMEA[®]-Konzept). Dokumentenmanagement und elektronische Archivierung im IT-gestützten Geschäftsgang (Schriftenreihe der KBSt 45), Köln 1999; vgl. auch KBSt (Hrsg.), Abschlußbericht zum Projekt DOMEA[®]. Dokumentenmanagement und elektronische Archivierung im IT-gestützten Geschäftsgang (Schriftenreihe der KBSt 41), Köln 1999.

⁴ Vgl. Bundesministerium des Innern (Hrsg.), DOMEA[®]-Telegramm Nr. 6, Evaluierung der Konformität von Vorgangsbearbeitungssystemen mit dem „Konzept Papierarmes Büro (DOMEA[®]-Konzept)“ und Abschluss von Rahmenverträgen – Bericht zu den Ergebnissen der Ausschreibung – (KBSt-Brief Nr. 3/00), Berlin, Juni 2000; KBSt (Hrsg.), Ergebnisse der Evaluierung. Konformität von Vorgangsbearbeitungssystemen mit dem Konzept Papierarmes Büro (DOMEA[®]-Konzept). Dokumentenmanagement und elektronische Archivierung im IT-gestützten Geschäftsgang (Schriftenreihe der KBSt 51), Köln 2000. URL: http://www.kbst.bund.de/bei_uns/hoppel/domea/evaluierung/.

fahren stellte daher einen wichtigen Anreiz für Behörden und Anbieter dar, sich auch frühzeitig mit der Frage der Aussonderung elektronischer Akten zu befassen.⁵

Die Frage, wie sich das Aussonderungskonzept in Zukunft bewähren wird, hängt jedoch eng damit zusammen, ob Anbieter bereit sein werden, die Anforderungen der „Empfehlungen“ in ihre Produkte als Systemmodule zu integrieren. Nur wenn die Aussonderung elektronischer Akten als funktionaler Bestandteil in Softwareprodukten enthalten ist, bestehen gute Chancen, dass das Aussondungsverfahren bei den Anwendern erfolgreich verlaufen wird.

2. Das Projekt „Aussonderung“ des Anwender-Arbeitskreises (AK) Favorit

Es war daher ein Glücksfall, dass sich der AK Favorit im Frühjahr 2000 auf Anregung der Forschungsstelle für Verwaltungsinformatik der Universität Koblenz mit der Frage der Aussonderung befasste. Die Integration der Aussonderungsfunktionalität in Favorit[®]-OfficeFlow[®] bildete nach der Sitzung des AK Favorit am 16. Februar 2000 den Arbeitsschwerpunkt des folgenden halben Jahres. Beteiligt waren neben der Firma Debis und der Forschungsstelle für Verwaltungsinformatik von Anwenderseite das Bundesverwaltungsamt, das Gemeinsame Gebietsrechenzentrum Köln und das Nordrhein-Westfälische Landesamt für Datenverarbeitung und Statistik sowie von archivischer Seite das Bundesarchiv und das Nordrhein-Westfälische Hauptstaatsarchiv Düsseldorf.⁶

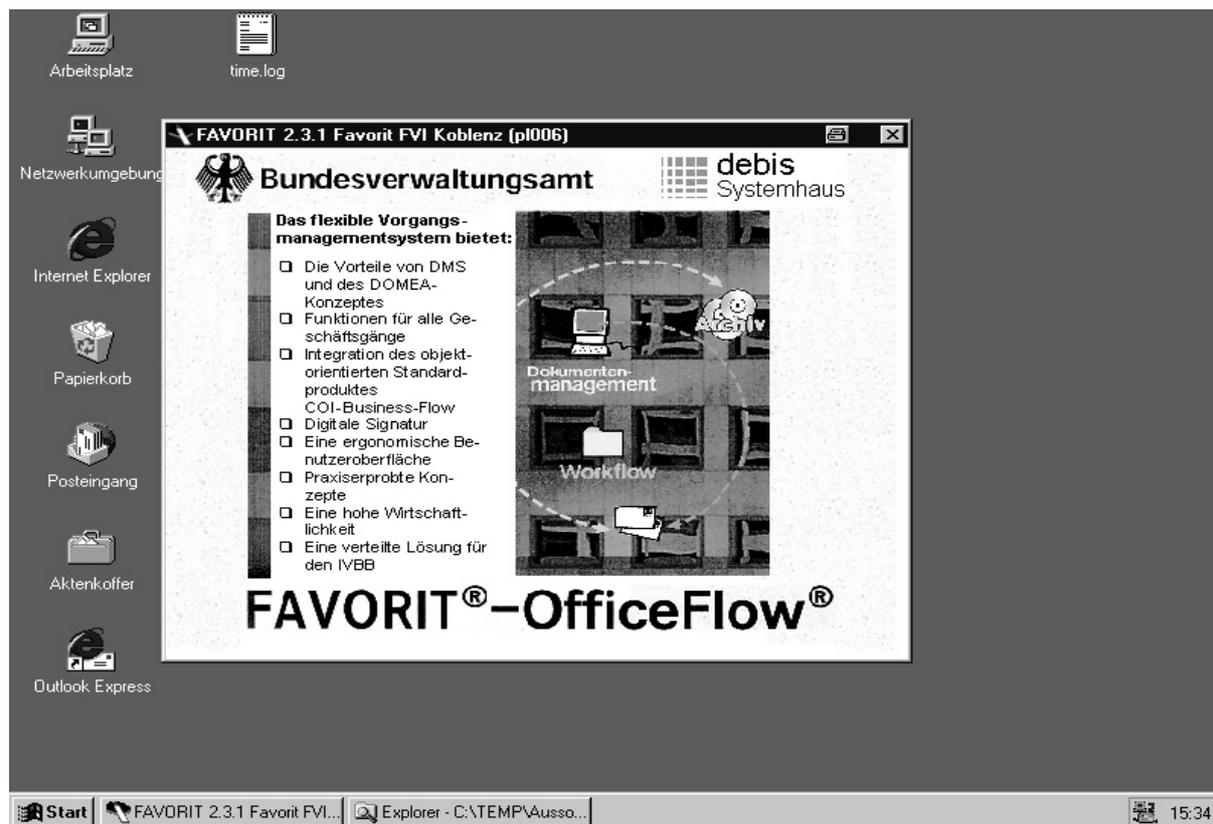
Favorit[®]-OfficeFlow[®] ist ein Vorgangsbearbeitungssystem, das derzeit im Bundesverwaltungsamt im Wirkbetrieb für die Verwaltung und Wiedereinziehung von BAföG-Darlehen Anwendung findet, darüber hinaus aber auch schon in anderen Abteilungen dieser Behörde eingesetzt wird. In der Nordrhein-Westfälischen Ministerialverwaltung findet derzeit ein Piloteinsatz von Favorit[®] statt. Sowohl in der auf der Grundlage der fortgeschriebenen Anforderungen des DOMEA[®]-Konzepts Ende 1999 durchgeführten Ausschreibung des Beschaffungsamts als auch an dem Evaluationsverfahren im Bundesministerium des Innern, das im selben Jahr stattfand, hat Favorit[®] zusammen mit zwei bzw. vier anderen Produkten erfolgreich abgeschlossen.⁷ Favorit[®] gehört damit zu den wichtigsten im öffentlichen Sektor eingesetzten Produkten im Bereich der IT-gestützten Vorgangsbearbeitung. Wie das DOMEA-System der Firma CSE sieht Favorit[®] verschiedene Nutzungsstufen vor, als Schriftgutverwaltungssystem, als elektronische Aktenablage oder als Vorgangsbearbeitungssystem.

⁵ Offenbar trifft dies nicht nur auf Bundesbehörden zu, bei denen die Notwendigkeit geeigneter Schnittstellen zunehmend beachtet wird; auch für Landesbehörden soll DOMEA-Konformität ein Qualitätskriterium bilden, wie ein Erfahrungsbericht für Sachsen feststellt. Thekla Kluttig, Strategien und Spielräume archivischer Behördenberatung. In: Nils Brübach (Hrsg.), Der Zugang zu Verwaltungsinformationen – Transparenz als archivische Dienstleistung. Beiträge des 5. Archivwissenschaftlichen Kolloquiums der Archivschule Marburg (Veröffentlichungen der Archivschule Marburg 33), Marburg 2000, S. 147–154, hier S. 149.

⁶ Vgl. Michael Wettengel, Verwaltungsmodernisierung und IT-gestützte Vorgangsbearbeitung (wie Anm. 1) S. 38.

⁷ DOMEA[®]-Telegramm Nr. 6, Evaluierung der Konformität von Vorgangsbearbeitungssystemen mit dem „Konzept Papierarmes Büro (DOMEA[®]-Konzept)“ und Abschluss von Rahmenverträgen, S. 8 f.; Vgl. zu Favorit[®] Johannes Keusekotten – Gabriele Sauermann, Erfolgreiche Verwaltung mit Favorit[®]-Office Flow[®]. In: Auf der Suche nach archivischen Lösungsstrategien im digitalen Zeitalter (wie Anm. 1) S. 69–72.

Abbildung 1: Favorit®-OfficeFlow® 2.3.1



Als Grundlage für das Aussonderungsverfahren in Favorit®-OfficeFlow® einigte sich der Arbeitskreis auf die im DOMEA®-Aussonderungskonzept enthaltenen „Empfehlungen“ des Bundesarchivs. Die dort vorgesehenen Funktionalitäten für die Aussonderung, insbesondere das Bewertungsverfahren nach dem Stufenkonzept sowie die Abgabe einer Aussonderungsdatei mit den im Konzept vorgesehenen Formaten und Strukturen sollten in Favorit®-OfficeFlow® implementiert werden. Einige Modifikationen und Präzisierungen waren jedoch erforderlich. Sie betrafen insbesondere die folgenden Bereiche:

Präzisierungen, die über die im Aussonderungskonzept vorgesehenen „Empfehlungen“ hinausgehen;

Modifikationen oder Ergänzungen der „Empfehlungen“ aufgrund systemspezifischer Besonderheiten von Favorit®-OfficeFlow®;

Modifikationen oder Ergänzungen der „Empfehlungen“ aufgrund besonderer Anforderungen von Archivverwaltungen.

Sie werden in den folgenden Kapiteln näher beschrieben.

2.1. Präzisierungen der „Empfehlungen“

Die über die „Empfehlungen“ hinausgehenden erforderlichen Präzisierungen ergaben sich aus den Diskussionen des Arbeitskreises. Sie resultierten aus der Notwendigkeit, die Anforderungen der „Empfehlungen“ in Favorit® zu implementieren und Verfahren für die Umsetzung in der praktischen Anwendung festzulegen. Sie umfassten vor allem folgende Punkte:

Zunächst war eine Entscheidung zu fällen, welche der in dem Registratursystem gespeicherten Angaben an die Archive übergeben werden sollen.

Ausgesonderte elektronische Akten sind in der Behörde zuverlässig zu vernichten; das Verfahren dafür war festzulegen, wenn die Daten – wie im Bundesverwaltungsamt – auf CDs gespeichert sind.

Die Metadaten der ausgesonderten Akten sollen in der abgebenden Behörde recherchierbar bleiben; dies soll aber außerhalb der Registraturkomponente von Favorit[®] erfolgen.

Das Verfahren der automatischen Übernahme des Aussonderungsstatus (anbieten/übernehmen/vernichten) von einer höheren Ebene (z.B. Aktenplangruppe oder Betroffenseinheit) auf die einzelnen Akten war festzulegen.

Bei der Einrichtung eines Web-Zugangs für die zuständigen Archive auf die Aussonderungsportion innerhalb der abgebenden Behörde sind die Zugriffsrechte zu definieren.

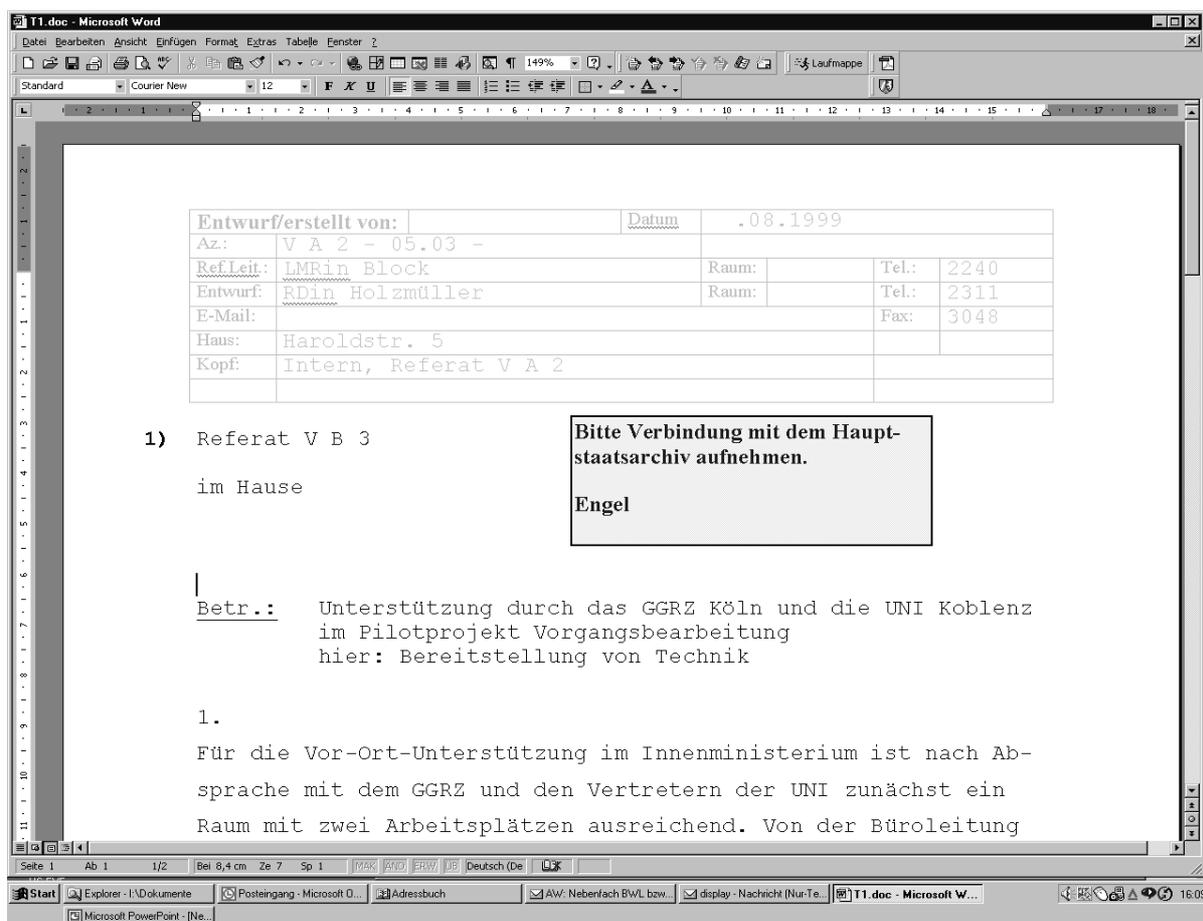
2.2. Änderungen aufgrund von Besonderheiten von Favorit[®]

Den wichtigsten Bereich bildeten die Modifikationen oder Ergänzungen aufgrund der spezifischen Besonderheiten und Funktionen von Favorit[®]-OfficeFlow[®]. Zwar bietet das DOMEA[®]-Konzept prinzipiell produktneutrale „Empfehlungen“ für die Aussonderung, als Modellfall und Erfahrungshintergrund für deren Erarbeitung hatte jedoch das in der Koordinierungs- und Beratungsstelle der Bundesregierung für Informationstechnik in der Bundesverwaltung (KBSSt) eingesetzte System DOMEA der Firma CSE gedient. Anpassungen an Spezifika von anderen Systemen und Anwendungen sind daher erforderlich. Bei Favorit[®] waren drei Aspekte von besonderer Bedeutung:

Erstens: In Favorit[®] können Dokumente mit Eingangsstempeln und Geschäftsgangvermerken sowie mit einer Art elektronischer Notizzettel (sogenannte „gelbe Zettel“) versehen werden. Diese werden nicht zusammen mit dem Dokument, sondern als zusätzliche Folien (als Layers) gespeichert, die während der Anzeige über das Originaldokument gelegt werden. Wenn das Dokument beim Erstellen der Aussonderungsdatei mit den Layers in einem TIFF-Bild abgespeichert wird,⁸ kann der Text des Dokuments unlesbar werden. Das Dokument soll daher einmal ohne Layers und ein zweites Mal mit Layers ausgesondert werden, d.h. dass Dokumente in vielen Fällen zweimal hintereinander abgelegt werden. Die „gelben Zettel“ sollen grundsätzlich entfernt werden, wenn ein Dokument zu den Akten genommen wird. Durch organisatorische Regelungen muss festgelegt werden, dass derartige „gelbe Zettel“ keine aktenrelevanten Informationen enthalten dürfen.

⁸ Zur Frage der Formate vgl. Raphael Ostermann, Potentielle Dateiformate zur Langzeitarchivierung von Dokumenten unter Berücksichtigung von Primär- und Metainformationen. In: Michael Wettengel (Hrsg.), Digitale Herausforderungen für Archive (wie Anm. 1) S. 25–35.

Abbildung 2: „Gelbe Zettel“ in Favorit®-OfficeFlow®



Zweitens: In den mit Favorit® gebildeten elektronischen Akten sind häufige elektronische Verweise (Links) auf Dokumente und Vorgänge in anderen Akten typisch. Hier besteht die Gefahr, dass sich nach der Bewertung in einer archivwürdigen Akte Verweise auf Dokumente und Vorgänge in kassierten Akten befinden. Aus diesem Grunde wurde beschlossen, spätestens bei der Aussonderung Verweise auf Dokumente in anderen Akten in „echte Kopien“, d.h. in elektronische Dokumente, an den Verweisstellen umzuwandeln. Diese „Kopien“ sind mit der Akte, in der sich der Verweis auf sie befand, auszusondern, auch wenn sie nicht das gleiche Geschäftszeichen tragen.

Drittens: Anders als im DOMEA®-Konzept vorgesehen ist bei Favorit® die Vorgangsebene nicht Träger der Kontext- und Bearbeitungsinformationen (Geschäftsgangvermerke, Verfügungen, Zeichnungen, Laufweginformationen etc.). Favorit® bildet Informationen zum Laufweg und zur Bearbeitung in sogenannten „Umlaufmappen“ ab. Zum Zwecke der Aussonderung werden diese Informationen – wie im DOMEA®-Konzept beschrieben – in Form eines sogenannten „Mitzeichnungsblatts“ aufbereitet, exportiert und in das TIFF-Format umgewandelt.

Es erwies sich als günstig, dass Vertreter der Herstellerfirma und verschiedener Anwender von Favorit® an dem Projekt beteiligt waren, da auf diese Weise verschiedene Sichtweisen eingebracht und auf Spezifika aufmerksam gemacht werden konnte, die sonst möglicherweise nicht beachtet worden wären.

2.3. Änderungen aufgrund zusätzlicher Anforderungen von Archiven

Als weiteres Archiv, das für Behörden zuständig ist, die Favorit[®] anwenden, war das Nordrhein-Westfälische Hauptstaatsarchiv zusammen mit dem Bundesarchiv im AK Favorit vertreten und hatte eine Reihe ergänzender Anforderungen eingebracht. Weitere Archive hatten ebenfalls die Gelegenheit wahrgenommen, zu den Vorentwürfen des Aussonderungskonzepts Stellung zu nehmen.

Bereits in der Eröffnungssitzung des Arbeitskreises wurden eine Reihe von zusätzlichen Anforderungen vorgebracht; sie betrafen insbesondere die folgenden Punkte:

Auch bei generellen Kassationsentscheidungen für eine Aktenplangruppe oder Betreffseinheit sollte es in Ausnahmefällen möglich sein, einzelne Akten als archivwürdig einzustufen, ohne dass dies eine Rückwirkung auf die anderen Akten der Aktenplangruppe oder Betreffseinheit haben darf. Die aus der höheren Ebene übernommenen Voreinstellungen sollen daher prinzipiell manuell änderbar sein. Aus den Registraturdaten muss dabei hervorgehen, daß eine manuelle Anpassung vorgenommen wurde.

Die Namen der Einsender von Dokumenten sollten recherchierbar bleiben.

Aus Gründen der Nachvollziehbarkeit von Bewertungs-Entscheidungen sollen dem zuständigen Archiv ein reduzierter Metadaten-Satz der in der Behörde vernichteten Akten sowie eine statistische Auszählung der vernichteten Akten, Vorgänge und Dokumente geliefert werden.

Ein später hinzugekommener Wunsch war die Ermöglichung der Festlegung der Aussonderungsattribute „Aussonderungsstatus“ und „Aufbewahrungsfrist“ auf Aktenplangruppen-Ebene (im Fachkonzept zur Aussonderung in Favorit[®]-OfficeFlow[®] „Betreffsgruppen“ genannt). Ein auf Aktenplangruppen-Ebene festgelegter Aussonderungsstatus muss ebenso wie die Aufbewahrungsfrist automatisch auf die dazugehörigen Betreffseinheiten und Akten übertragen werden. Der Aussonderungsstatus soll zusätzlich zu den Werten „Anbieten“ oder „Vernichten“ auch den Wert „Übergeben“ einnehmen können, wenn die betreffenden Akten ohne weiteres Bewertungsverfahren archivwürdig sein sollen.

3. Das „Fachkonzept zur Aussonderung in Favorit[®]-OfficeFlow[®]“ als Ergebnis

Als Ergebnis der Arbeit des AK Favorit wurde von der Firma Debis ein „Fachkonzept zur Aussonderung in Favorit[®]-OfficeFlow[®]“ erstellt, das jetzt in der Version 1.1 vom August 2000 vorliegt und diesem Beitrag im Anhang in leicht gekürzter Form beigelegt wurde.

In der „Überblick“ genannten Einleitung werden zunächst die Archivgesetze des Bundes und der Länder als rechtliche Grundlage für die Aussonderung hervorgehoben. Für das Verfahren der Aussonderung wird das „Konzept zur Aussonderung elektronischer Akten“ zugrunde gelegt. Im ersten Abschnitt über die Einbettung des Fachkonzepts in Favorit[®]-OfficeFlow[®] werden danach unter anderem die notwendigen Änderungen an der Favorit[®]-Oberfläche beschrieben. Diese umfassen zunächst die Ergänzung der Maske zu den Betreffseinheiten um die Attribute „Aussonderungsstatus“ und „Aufbewahrungsfrist“ sowie die Einrichtung einer zusätzlichen Registerkarte „Aussonderung“ für die Registriermaske. Diese enthält folgende Angaben:

Für Vorgänge:

Die Aussonderungsattribute: zdA-Datum, Aufbewahrungsfrist, Aufbewahrungsende und Aussonderungsstatus.

Die sogenannten Archivattribute: Archivnummer, Laufende Nummer, Ergänzungsbuchstabe und Bewertungsstatus.

Für Akten:

Die Aussonderungsattribute: Aufbewahrungsfrist und Aussonderungsstatus.

Die einzelnen Abläufe nach dem Stufenkonzept in den „Empfehlungen“ des Aussonderungskonzepts, wie beispielsweise die Erstellung eines Anbieterverzeichnisses und der Import des Bewertungsverzeichnisses, sollen über das Menü der Hauptmaske aufgerufen werden können.

Im Hauptteil des Fachkonzepts Favorit[®] werden die Anforderungen aus fachlicher Sicht in Form einer Ablaufbeschreibung dargestellt. Dabei umfasst die so genannte „erste Stufe“ des Fachkonzepts im wesentlichen die Stufen eins bis drei der „Empfehlungen“ des Bundesarchivs. Sie enthalten die folgenden Funktionen:

Erfassung und automatische Übertragung („Vererbung“) der Aussonderungsattribute: Dies bedeutet, dass die Attribute Aufbewahrungsfrist und Aussonderungsstatus zu Aktenplangruppen, betreffseinheiten und Akten erfasst und automatisch auf die dazugehörigen niedrigeren Ebenen (bei Aktenplangruppen: betreffseinheiten, Akten und Vorgänge) übertragen werden. Manuelle Änderungen auf einer unteren Ebene sind jedoch möglich.

Übersicht über Aussonderungsattribute

Erstellen des Anbieterverzeichnisses nach der Stufe 2 der „Empfehlungen“ des Bundesarchivs

Import des Bewertungsverzeichnisses in Favorit[®] nach der Stufe 3 der „Empfehlungen“ des Bundesarchivs

Ausbau der generischen Recherche in Favorit[®] nach den Aussonderungsattributen Aufbewahrungsfrist und Aussonderungsstatus

Einschränkung der Bearbeitungsfunktionalität für Vorgänge der Aussonderungsportion: Das heißt, dass die Vorgänge, die nach Ablauf der Aufbewahrungsfrist in das Anbieterverzeichnis für das zuständige Archiv geschrieben wurden, für die weitere Bearbeitung in Favorit[®] gesperrt sind. Es ist dann in der Behörde nicht mehr möglich, den Vorgang wiederaufleben zu lassen oder die Metadaten des Vorgangs oder dessen Dokumente zu verändern. Die schlussverfügbaren Dokumente der entsprechenden Vorgänge können nicht mehr in einem anderen Vorgang registriert oder übernommen werden. Auch Querverweise auf bzw. von den Vorgängen und Dokumenten der Aussonderungsportion sind prinzipiell nicht mehr zulässig. Auf diese Weise sollen nachträgliche Veränderungen bereits für die Aussonderung vorbereiteter Unterlagen unmöglich gemacht werden.

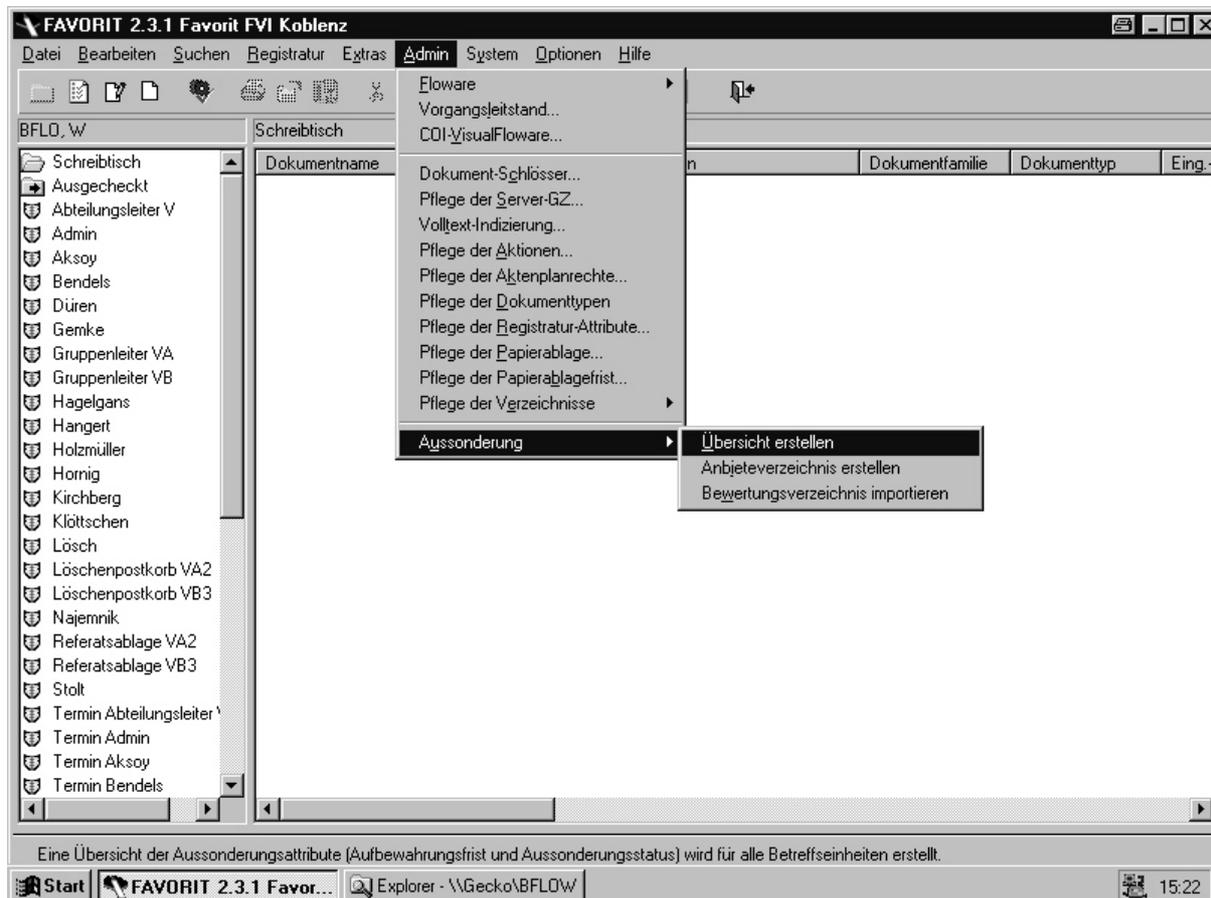
Die so genannte „zweite Stufe“ des Fachkonzepts Favorit[®] behandelt die Stufe vier des Stufenkonzepts der „Empfehlungen“ des Bundesarchivs und beschäftigt sich mit der Erstellung der Aussonderungsdatei und der Durchführung der Abgabe an das zuständige Archiv, der Löschung der ausgesonderten Dokumente und Metadaten in Favorit[®], der Einrichtung eines Web-Zugriffs für das zuständige Archiv sowie der automatischen Protokollierung der Aussonderungsaktivitäten. Diese Stufe der Aussonderung liegt in der vorliegenden Version des Fachkonzepts Favorit[®] noch als Grobkonzept vor. Den Abschluss des Fachkonzepts bilden die Datensatzbeschreibungen des Anbieterverzeichnisses und des Bewertungsverzeichnisses, die denjenigen der „Empfehlungen“ des Bundesarchivs entsprechen. Für die Datensatzbeschreibung des Aussonderungsverzeichnisses und die Struktur der Abgabedatei wurde auf die entsprechenden Ausführungen der „Empfehlungen“ verwiesen.

4. Die systemtechnische Umsetzung des „Fachkonzepts“ und Ausblick

In der inzwischen fertiggestellten Version 2.3.1 von Favorit[®]-OfficeFlow[®] wurden die Funktionalitäten der „ersten Stufe“ des Fachkonzepts (die Stufen eins bis drei der „Empfehlungen“) systemtechnisch umgesetzt. In einem weiteren Schritt soll jetzt die im gegenwärtigen Fachkonzept nur grob beschriebene „zweite Stufe“ näher spezifiziert und in die künftige Version 2.4 von Favorit[®] integriert werden. Die betroffenen Archive werden diese Entwick-

lung weiterhin begleiten müssen, um eine adäquate Umsetzung der fachlichen Anforderungen und eine befriedigende Weiterentwicklung des Fachkonzepts zu gewährleisten.

Abbildung 3: Aussonderung in Favorit®-OfficeFlow® 2.3.1



Ein wichtiges Thema wird dabei beispielsweise die Frage der bereits erwähnten „gelben Zettel“ darstellen. Die derzeitige Lösung, diese vor der Aussonderung von den Dokumenten zu entfernen und durch organisatorische Maßnahmen sicherzustellen, dass sie nicht zur Mitteilung aktenrelevanter Informationen verwandt werden, vermag aus fachlicher Sicht nicht zu befriedigen. Eine Ablage der Dokumente mit „gelben Zetteln“ ist jedoch nicht möglich, da sonst der Text des Dokuments verdeckt werden könnte. Eine Ablage der „gelben Zettel“ als eigene Dokumente wirft jedoch massive Mengenprobleme auf, die ohnedies Gegenstand der künftigen Gespräche sein werden. Schon jetzt wird die Tatsache, dass ein Dokument einmal ohne Layers und ein zweites Mal mit als Layers gespeicherten Eingangsstempeln, Geschäftsgangvermerken und Verfügungen abgelegt wird und zusätzlich noch alle im Geschäftsgang erzeugten und abgelegten Versionen eines Dokuments ausgesondert werden sollen, zu einem erheblichen Bedarf an Speicherkapazitäten führen. Hier sind gerade Archive gefragt, Lösungen zu finden; und eine fachliche Diskussion sollte vor diesen technischen Problemen nicht vorschnell aufgeben. Es gilt dabei zu bedenken, dass auch bisher schon vielfach dem Bearbeiter überlassen bleibt, was in die Akten gelangt.

Als Resümee bleibt festzuhalten, dass sich das 1998 erarbeitete „Konzept zur Aussonderung elektronischer Akten“ und die darin enthaltenen „Empfehlungen“ bislang bewährt haben. Ohne größere Modifikationen konnten bereits viele der archivischen Anforderungen an die IT-gestützte Vorgangsbearbeitung systemtechnisch implementiert werden. Die archivischen Anforderungen sind, wie dieses Beispiel zeigt, technisch umsetzbar und in Standard-Anwendungen integrierbar.

Es erwies sich als außerordentlich hilfreich für die Mitarbeit der Archive im AK Favorit, dass auf ein solches Konzept zurückgegriffen werden konnte. Auf diese Weise liegt ein Standard vor, der lediglich angepasst und in Systemanwendungen implementiert werden muss. Archive sollten daher solche Konzepte bereits fertig vorlegen können, bevor elektronische Systeme in Behörden ihres Zuständigkeitsbereichs eingeführt werden. Im Bereich der Vorgangsbearbeitungssysteme hat die Implementierung von Aussonderungsfunktionalitäten in Systemanwendungen begonnen, dies zeigt nicht nur das Beispiel von Favorit[®]. Auch für das im Bundesministerium für Verkehr, Bau- und Wohnungswesen im Einsatz befindliche Archivierungssystem DORIS (Dokument Retrieval und Informationssystem) der Firma Haessler⁹ soll von September bis Dezember 2001 ein Aussonderungsmodul erarbeitet werden. Da dieses System nicht nur in Bundesbehörden eingesetzt werden dürfte, befasst sich die Arbeitsgruppe (Spiegelgremium) „Elektronische Systeme in Justiz und Verwaltung“ der Archivreferentenkonferenz des Bundes und der Länder mit einem Aussonderungskonzept für dieses System. Diese neue ARK-Arbeitsgruppe wurde mit dem Ziel eingesetzt, „archivfachliche Anforderungen an elektronische Systeme von Justiz und Verwaltung einzubringen“.¹⁰

Die archivfachliche Verständigung über Standards bei der Archivierung von Unterlagen aus digitalen Systemen sollte daher eine zentrale Aufgabe von archivischen Arbeitskreisen und Arbeitsgruppen darstellen. Dabei sollten konkrete Vorhaben zur Festlegung von Schnittstellen stets in Projektgruppen unter Einbeziehung von Archiven, Anwendern, Herstellerfirmen und ggf. auch Dienstleistern stattfinden, wenn diese für eine der betroffenen Parteien Serviceleistungen durchführen. Auf diese Weise wird eine möglichst vollständige Berücksichtigung aller relevanten Aspekte erleichtert und eine problemlose Implementierung ermöglicht. Aus der Arbeit am Fachkonzept Favorit[®]-OfficeFlow[®] im AK Favorit liegen bereits Erfahrungen vor, aus denen ähnliche Projekte künftig profitieren können.

⁹ Vgl. Herbert Schweitzer, DORIS im Bundesministerium für Verkehr, Bau- und Wohnungswesen. Flexibel und mobil mit elektronischen Akten. In: BBB Informationen Heft 244, Köln 9/10 (2000) S. 8–10.

¹⁰ Protokoll der 91. Archivreferentenkonferenz des Bundes und der Länder in Nürnberg am 10.10.2000, TOP 5.4, S. 24.

Anhang

Flexibles Archivierungs- und Vorgangsbearbeitungssystem im IT-gestützten Geschäftsgang (FAVORIT®-OfficeFlow® 2.3)

Fachkonzept zur Aussonderung in FAVORIT®-OfficeFlow® Version 1.1 (gekürzte Fassung)

Auftraggeber: Bundesverwaltungsamt
Auftragnehmer: debis Systemhaus GEI, Bonn
Version: 1.1
Stand: 09.08.2000

- 1 Aussonderung in FAVORIT®-OfficeFlow®
 - 1.1 Überblick
 - 1.2 Einbettung in FAVORIT®-OfficeFlow®
 - 1.3 Anforderungen aus fachlicher Sicht (Ablaufbeschreibung)
 - 1.3.1 Erfassung und Vererbung der Aussonderungsattribute
 - 1.3.2 Übersicht über Aussonderungsstati und Aufbewahrungsfristen
 - 1.3.3 Erstellen des Anbieterverzeichnisses
 - 1.3.4 Import des Bewertungsverzeichnisses
 - 1.3.5 Ausbau der Generischen Recherche
 - 1.3.6 Einschränkung der Bearbeitungsfunktionalität
 - 1.3.7 Erstellen der Abgabedateien für das zuständige Archiv
 - 1.3.8 Löschfunktionalität
 - 1.3.9 WEB-Zugriff
 - 1.3.10 Protokollierung verschiedener Aussonderungsdaten
 - 1.4 Dateistrukturen
 - 1.4.1 Stuktur des Anbieterverzeichnisses (in dieser Veröffentlichung nicht enthalten)
 - 1.4.2 Struktur des Bewertungsverzeichnisses (in dieser Veröffentlichung nicht enthalten)
 - 1.4.3 Struktur der Abgabedateien (in dieser Veröffentlichung gekürzt)
- Aussonderung in FAVORIT®-OfficeFlow®

Überblick

Aufgrund der Archivgesetze des Bundes und der Länder sind die öffentlichen Stellen verpflichtet, Papier- und elektronisches Schriftgut nach Ablauf der Aufbewahrungsfristen dem zuständigen Archiv zur Übernahme anzubieten. Das Archiv entscheidet, welche angebotenen Vorgänge übernommen werden und welche vernichtet werden können.

Diese Gesetze dienen als Grundlage für das „Konzept zur Aussonderung elektronischer Akten“ (KBSt-Schriftenreihe, Band 40).

Auf Basis dieses Aussonderungskonzeptes wird in FAVORIT® ein Aussonderungsverfahren für Akten und Vorgänge umgesetzt, wodurch eine regelmäßige, fristengesteuerte Aussonderung ermöglicht wird. Auf diese Weise werden Vorgänge, deren Aufbewahrungsfrist nach der zdA-Verfügung abgelaufen ist, nach Zusammenarbeit mit dem zuständigen Archiv (Bundesarchiv bzw. Hauptstaatsarchiv des jeweiligen Bundeslandes) aus dem System entfernt.

Einbettung in FAVORIT®-OfficeFlow®

In der Version 2.2.3 von FAVORIT®-OfficeFlow® wurde bereits eine Zusatzkomponente mit ersten Funktionalitäten zur Aussonderung von Vorgängen realisiert. Für die Version 2.3.1 ist die Umsetzung der ersten Stufe der Aussonderung vorgesehen. Eine Erweiterung der Funktionalität erfolgt in der zweiten Stufe und wird in die Version 2.4 integriert. Die Schritte der zweiten Stufe sind in diesem Konzept nur beispielhaft in Form eines Grobkonzeptes beschrieben. Ein fachliches Feinkonzept wird mit der Version 2.4 erstellt.

An der FAVORIT®-Oberfläche werden die Masken „Betreffseinheit anlegen/ändern“ und die Registriermaske verändert. Die Maske zu den Betreffseinheiten wird um die beiden Aussonderungsattribute „Aussonderungsstatus“ und „Aufbewahrungsfrist“ ergänzt. Die Registriermaske erhält für Akten und Vorgänge eine zusätzliche Registerkarte „Aussonderung“, auf welcher folgende Informationen dargestellt werden:

- Für Vorgänge:
- Aussonderungsinformationen: zdA-Datum, Aufbewahrungsfrist, Aufbewahrungsende und Aussonderungsstatus
- Archivinformationen: Archivnummer, Laufende Nummer, Ergänzungsbuchstabe und Bewertungsstatus.
- Für Akten:
- Aussonderungsinformationen: Aufbewahrungsfrist und Aussonderungsstatus.

Die Aufbewahrungsfrist legt den Zeitpunkt der Aussonderung nach der zdA-Verfügung fest. Der Aussonderungsstatus kann die Werte „Anbieten“, „Vernichten“ und „Übergeben“ einnehmen. „Anbieten“ bedeutet, daß die auszusondernden Vorgänge dem zuständigen Archiv in einem Anbieterverzeichnis zur Verfügung gestellt werden. Das Archiv entscheidet dann, ob die angebotenen Aktenschnitte übergeben werden sollen. Steht der Aussonderungsstatus auf „Vernichten“ werden die entsprechenden Vorgänge nach Ablauf der Aufbewahrungsfrist im System gelöscht. Der Status „Übergeben“ sagt aus, daß diese Vorgänge dem zuständigen Archiv ohne weitere Bewertung nach Ablauf der Aufbewahrungsfrist übergeben werden.

Sämtliche Funktionalitäten, wie z.B. Erstellung des Anbieterverzeichnisses, Import des Bewertungsverzeichnisses oder die gesamte Löschfunktionalität werden über Menüpunkte der Hauptmaske aufgerufen.

Anforderungen aus fachlicher Sicht (Ablaufbeschreibung)

Erfassung und Vererbung der Aussonderungsattribute

Die Aussonderungsattribute „Aussonderungsstatus“ und „Aufbewahrungsfrist“ existieren für Betreffsgruppen, Betreffseinheiten, Akten und Vorgänge. In der Maske „Betreffseinheit anlegen/ändern“ können diese Attribute für die jeweilige Betreffseinheit bzw. die Betreffsgruppe verändert werden. In der Registriermaske auf der Registerkarte „Aussonderung“ besteht die Möglichkeit, die beiden Attribute für die entsprechende Akte zu verändern. Der Aussonderungsstatus bzw. die Aufbewahrungsfrist werden für den Vorgang in der Registrierung angezeigt, können jedoch auf dieser Ebene nicht mehr verändert werden.

Anlegen von Betreffseinheiten und Objekten

Beim Anlegen einer Betreffseinheit bzw. einer Betreffsgruppe werden der Aussonderungsstatus und die Aufbewahrungsfrist festgelegt. Der Aussonderungsstatus wird über einen Radiobutton, der den Wert „Anbieten“, „Vernichten“ oder „Übergeben“ einnehmen kann, eingestellt. Die Aufbewahrungsfrist wird in einer Combobox ausgewählt. Dort wird eine Liste

mit Jahreszahlen angeboten, die nicht veränderbar ist. Die Liste enthält die Werte „1“, „2“, „3“, „4“, „5“, „6“, „7“, „8“, „9“, „10“.

Die unter einer Betreffsgruppe angelegten Betreffseinheiten erben die Aussonderungsattribute der übergeordneten Einheit. Wird eine neue Betreffseinheit direkt unter der Teilaktenplanebene angelegt, so steht der Defaultwert des Aussonderungsstatus auf „Anbieten“ und der der Aufbewahrungsfrist auf „2 Jahre“.

Wird eine Akte angelegt, so erbt diese den Aussonderungsstatus und die Aufbewahrungsfrist der zugehörigen Betreffseinheit bzw. der Betreffsgruppe. Ebenso erhalten neu angelegte Vorgänge die Aussonderungsattribute der entsprechenden Akte.

Veränderung der Aussonderungsattribute

Der Aussonderungsstatus und die Aufbewahrungsfrist können auf Betreffsgruppen-, Betreffseinheiten- und auf Aktenebene verändert werden. Bei der Änderung der Aussonderungsattribute einer Betreffsgruppe werden die zugehörigen Akten mit den neuen Attributen versehen, wenn diese dieselben Werte der Betreffsgruppe aufweisen. Der geänderte Aussonderungsstatus bzw. die geänderte Aufbewahrungsfrist der Betreffsgruppe werden auf die darunterliegenden Betreffsgruppen bzw. Betreffseinheiten und die jeweils zugehörigen Akten übertragen, wenn die Attributwerte übereinstimmen.

Werden die Aussonderungsattribute für eine Betreffseinheit geändert, so erhalten alle Akten der Betreffseinheit, welche die Werte der entsprechenden Betreffseinheit aufweisen, den neuen Wert. Die Akten, die schon veränderte Attribute im Vergleich zur Betreffseinheit aufweisen, behalten die vorhandenen Werte.

Bei einer Veränderung der Aufbewahrungsfrist bzw. des Aussonderungsstatus einer Akte wird ein spezielles Recht in der Organisationsdatenbank benötigt. Diese manuelle Änderung wird im Protokoll zur Akte festgehalten. Alle Vorgänge, die sich in der entsprechenden Akte befinden, erhalten die neu festgelegten Attribute.

Auswirkungen der Geschäftszeichenänderung

Die Geschäftszeichenänderung bewirkt eine neue Zuordnung von Objekten im Aktenplan, wodurch auch eine Anpassung der Aussonderungsattribute durchgeführt wird. Eine Änderung von Geschäftszeichen für Akten ist zur Zeit in FAVORIT® nicht möglich. Wird das Geschäftszeichen eines Vorgangs geändert, so erhält dieser den Aussonderungsstatus bzw. die Aufbewahrungsfrist der neuen Akte.

Übersicht über Aussonderungsstatus und Aufbewahrungsfristen

Der Aussonderungsstatus soll auf Betreffsgruppen-, Betreffseinheiten- bzw. auf Aktenebene in Absprache mit dem zuständigen Archiv festgelegt werden. Über den Menüpunkt *Admin* → *Aussonderung* → *Übersicht* kann eine Übersicht über die Aussonderungsstatus und Aufbewahrungsfristen aller Betreffseinheiten aus FAVORIT® exportiert werden. Diese Übersicht wird im Dateiformat „csv“ (mit Kommata getrennte Variablen) abgespeichert und kann somit mit Microsoft Excel geöffnet werden. Zu den Betreffseinheiten werden das Aktenplankennzeichen, der Name der Betreffseinheit sowie die beiden Aussonderungsattribute aufgelistet.

Unter dem Menüpunkt *Admin* → *Aussonderung* sind auch weitere Funktionalitäten der Aussonderung, die weiter unten beschrieben werden, zu finden. Diese Funktionen können nur ausgeführt werden, wenn der Benutzer das ADMIN-Recht innehat.

Erstellen des Anbieterverzeichnisses

Nach Ablauf der Aufbewahrungsfristen können die Vorgänge, deren Aussonderungsstatus auf „Anbieten“ steht, dem zuständigen Archiv im Anbieterverzeichnis zur Verfügung gestellt werden. Diese Funktion wird über das Menü der Hauptmaske *Admin* → *Aussonderung* → *Anbieterverzeichnis erstellen* aufgerufen. Es wird eine ASCII-Datei mit verschiedenen Meta-informationen zu den anzubietenden Vorgängen mit den zugehörigen Akten und Betreffseinerheiten erstellt. Der Aufbau der ASCII-Datei richtet sich nach dem „Konzept zur Aussonderung elektronischer Akten“ (KBSt-Schriftenreihe, Band 40) und wird in Kapitel 2 dieses Konzeptes detailliert beschrieben.

Import des Bewertungsverzeichnisses

Auf der Grundlage des exportierten Anbieterverzeichnisses wird vom zuständigen Archiv ein Bewertungsverzeichnis erstellt. Dieses umfaßt die Datensätze der Akten des Anbieterverzeichnisses, die um die Archivnummer des zuständigen Archivs, den Bewertungsstatus und den Ergänzungsbuchstaben der Archivnummer ergänzt werden. In FAVORIT® wird die Möglichkeit geschaffen, diese ASCII-Datei zu importieren (Menüpunkt *Admin* → *Aussonderung* → *Bewertungsverzeichnis importieren*). Die Registraturattribute der Vorgänge der Aussonderungsportion werden um die oben genannten Archivinformationen ergänzt.

Die Archivnummer bzw. der Ergänzungsbuchstabe zu dieser kennzeichnen das zuständige Archiv. Der Bewertungsstatus „K“ bezeichnet die Vorgänge, die vernichtet werden sollen. Vorgänge, die den Status „A“, „S“ und „T“ erhalten, sollen dem Archiv übergeben werden.

Ausbau der Generischen Recherche

Die Suchkriterien der Generischen Recherche werden um die Aussonderungsattribute „Aussonderungsstatus“ und „Aufbewahrungsfrist“ erweitert. In der Trefferliste wird zu den gefundenen Objekten der Aussonderungsstatus und das Aufbewahrungsende zusätzlich angezeigt. Das Aufbewahrungsende wird aus dem zDA-Datum und der Aufbewahrungsfrist errechnet, wobei die Aufbewahrungsfrist mit Ablauf des Jahres beginnt, in welchem die zDA-Verfügung durchgeführt wurde.

Einschränkung der Bearbeitungsfunktionalität

Die Vorgänge, die nach Ablauf der Aufbewahrungsfrist in das Anbieterverzeichnis für das zuständige Archiv geschrieben wurden, sind für die weitere Bearbeitung in FAVORIT® gesperrt. Es ist nicht mehr möglich, den Vorgang wiederaufleben zu lassen oder die Metadaten zu verändern. Ebenso können die Metadaten der Dokumente, die dem Vorgang angehören, nicht mehr verändert werden. Bei Papierdokumenten wird die Änderung der Metadaten durch die Empfangsbestätigung verhindert, d.h. der Menüpunkt der Hauptmaske *Extras* → *Empfangsbestätigung* ist inaktiv.

Ein Registrieren von Dokumenten in die zur Aussonderung angebotenen Vorgänge kann nicht mehr durchgeführt werden. Die schlussverfügbaren Dokumente der entsprechenden Vorgänge können nicht mehr in einen neuen bzw. vorhandenen Vorgang registriert werden. Ebenso ist es nicht mehr möglich, von den schlussverfügbaren Dokumenten per Drag&Drop eine Referenz zu erzeugen oder diese in vorhandene Vorgänge zu übernehmen.

Es dürfen keine Querverweise auf bzw. von den angebotenen Vorgängen und den dort enthaltenen Dokumenten erstellt werden.

Erstellen der Abgabedateien für das zuständige Archiv

Vorgänge, die dem zuständigen Archiv angeboten wurden und den Bewertungsstatus „A“, „S“ oder „T“ erhalten haben, werden dem Archiv zur Verfügung gestellt. Ebenso werden die Vorgänge, die den Aussonderungsstatus „Übergeben“ aufweisen und deren Aufbewahrungsfrist abgelaufen ist, ausgesondert.

Diese Funktionalität wird in FAVORIT® über das Menü der Hauptmaske *Admin* → *Aussonderung* → *Abgabedateien erstellen* aufgerufen. Es werden unterschiedliche Informationen zusammengestellt, die in unterschiedliche Dateiformate konvertiert werden. Die Primär- und Bearbeitungsinformationen (z.B. Laufwegs- und Mitzeichnungsinformationen) werden im TIFF-Format übergeben, während die dazugehörigen Metainformationen im ASCII-Format dargestellt werden. Bild- und Layerinformationen von eingescannten Dokumenten werden in einem TIFF zusammengeführt. Das Dokument wird einmal ohne Layer und ein zweites Mal mit Layer ausgesondert, damit die Dokumentinformationen lesbar bleiben. Notizen in Form von gelben Zetteln werden von den Dokumenten entfernt. Zu Versionsdokumenten werden alle vorhandenen Versionen ausgesondert.

Befinden sich in den auszusondernden Vorgängen Referenzen auf Dokumente in anderen Akten, werden diese kopiert, und die Kopie wird ausgesondert. Umgekehrt müssen auch die Referenzen auf Dokumente in für die Aussonderung vorgesehenen Vorgängen aufgelöst werden.

Elektronisch verschlüsselte oder signierte Dokumente sind zum Zeitpunkt der Aussonderung geöffnet, damit die jeweiligen Dokumente ohne Einsatz spezifischer Schlüssel oder Software gelesen werden können.

Die Struktur der Aussonderungsdateien wird aus dem „Konzept zur Aussonderung elektronischer Akten“ übernommen. Die Akten der Aussonderungsportion erhalten ein Aktendeckblatt, welches den jeweiligen Datensatz des importierten Bewertungsverzeichnisses mit der Archivnummer beinhaltet. Ebenfalls werden alle Metainformationen der Akte ausgegeben. Zu den auszusondernden Vorgängen werden die Metainformationen, die Dokumentenliste und die Bearbeitungsinformationen ausgeliefert. Die Primärinformationen der Vorgangsdokumente werden dem Archiv mit ihren Metainformationen übergeben.

Mit den Aussonderungsdateien wird ein Abgabeverzeichnis erstellt, welches die Informationen des Bewertungsverzeichnisses ohne die zu vernichtenden Akten enthält. Zusätzlich wird ein Auszählprotokoll übergeben, welches die Anzahl der Datensätze der Aussonderungsportion zu Betreffseinheiten, Akten und Vorgängen sowie die Anzahl zu vernichtender und auszusondernder Vorgänge enthält. Die zu vernichtenden Vorgänge sind solche, die dem zuständigen Archiv gar nicht angeboten wurden, da deren Aussonderungsstatus auf „Vernichten“ steht und die Vorgänge, die vom Archiv als zu Vernichten gekennzeichnet wurden.

Das Abgabeverzeichnis dient dem zuständigen Archiv zur Überprüfung der Vollständigkeit und der Authentizität der Übergabe. Die abgebende Behörde bestätigt dem Archiv bei der Übergabe die Authentizität der zu übergebenden elektronischen Akten. Das Archiv bestätigt den Erhalt der übergebenen Dateien.

Löschfunktionalität

In FAVORIT® wird im Zuge der Aussonderung die Möglichkeit geschaffen, die entsprechenden Vorgänge zu löschen. Diese Funktionalität wird über das Menü der Hauptmaske *Admin* → *Aussonderung* → *Löschen* für die Vorgänge aufgerufen, deren Aussonderungsstatus auf „Vernichten“ steht und auf deren Übernahme das Archiv damit verzichtet hat. Es werden auch die Vorgänge gelöscht, deren Bewertungsstatus vom Archiv auf „K“ gesetzt wurde. Ebenfalls werden die dem Archiv übergebenen Vorgänge vernichtet.

Die Vorgangsdokumente, für die in FAVORIT® Mehrfachreferenzen bestehen, dürfen nicht gelöscht werden.

Wurde FAVORIT® für eine Archivierung auf Jukebox konfiguriert, so wurden die zu löschenden Vorgänge bereits auf CD gebrannt. Aus diesem Grund wird ein Reorganisationslauf vorgesehen, bei dem die zu vernichtenden Objekte von der CD entfernt werden.

WEB-Zugriff

Auf der Datengrundlage des Anbieterverzeichnisses findet die Bewertung durch die Bearbeiter des zuständigen Archivs statt. Dafür wird dem zuständigen Archiv die Berechtigung zur Einsichtnahme in die in FAVORIT® befindlichen Akten, Vorgänge und Dokumente der Aussonderungsportion erteilt. Diese Einsichtnahme wird über einen Webzugang realisiert, wobei die Zugriffsrechte auf die jeweilige Aussonderungsportion eingeschränkt werden.

Nach Durchführung der Aussonderung wird vom zuständigen Archiv ein WEB-Zugriff auf die ausgesonderten Vorgänge gewährleistet.

Protokollierung verschiedener Aussonderungsdaten

Um die unterschiedlichen Schritte der Aussonderung nachvollziehen zu können, werden einige Daten protokolliert. Der jeweilige Benutzer und das Ausführungsdatum der folgenden Funktionen werden festgehalten:

- Erstellen des Anbieterverzeichnisses
- Import des Bewertungsverzeichnisses
- Erstellen der Abgabedateien
- Löschen von Vorgängen

Diese Informationen werden über eine Ansichtsmaske, die über den Menüpunkt *Admin* → *Aussonderung* → *Protokoll* aufgerufen wird, angezeigt.

Dateistrukturen

Im folgenden wird die Struktur des Anbieter- und Bewertungsverzeichnisses dargestellt. Diese beiden Dateien liegen im ASCII-Format vor. Die einzelnen Attribute werden durch das Pipe-Zeichen getrennt dargestellt. Es wird jeweils eine Zeile pro Datensatz vorgesehen.

Die Attribute für *Betreffseinheiten (BE)*, *Akten* und *Vorgänge* werden im folgenden mit ihrer Beschreibung aufgelistet.

[...]

Struktur der Abgabedateien

Die Festlegung der Struktur der verschiedenen Abgabedateien, in welchen die Primär-, Bearbeitungs- und Metainformationen der ausgesonderten Objekte zusammengestellt werden, erfolgt in Anlehnung an das „Konzept zur Aussonderung elektronischer Akten“ (Schriftenreihe der KBSt 40). Der Dateiaufbau wurde bereits in Kapitel 6 beschrieben. In dem fachlichen Feinkonzept, welches zur Version 2.4 erstellt wird, wird die Struktur dieser Dateien spezifiziert.

Die archivische Bewertung elektronischer Statistiken

Von Christian Keitel

Die Unterlagen des Statistischen Landesamts Baden-Württembergs müssen der Archivverwaltung angeboten und daher von dieser auch bewertet werden. Statistiken können heute als eine Sonderform von Datenbanken angesehen werden. In einem weiteren Sinne beschäftigt sich dieser Beitrag daher mit der Bewertung von Datenbanken. Einen ersten Bezugspunkt unserer Überlegungen stellen die sozialwissenschaftlichen Datenarchive dar.¹

1. Sozialwissenschaftliche Datenarchive

Sozialwissenschaftliche Datenarchive verwahren die im Rahmen von sozialwissenschaftlichen Untersuchungen entstandenen Datensammlungen. In Deutschland wird diese Funktion von dem *Zentralarchiv für Empirische Sozialforschung* an der Universität zu Köln (ZA) ausgeübt. Nach einer von der *International Federation of Data Organizations for the Social Sciences* (IFDO) publizierten Studie werden durch die Datenarchive die wissenschaftliche Öffentlichkeit gestärkt, neue Forschungsvorhaben vorangetrieben und deren Kosten durch die Zweitauswertung der archivierten Daten reduziert.² Der möglichst schnelle und umfassende Zugang zu den Daten bildet daher das zentrale und dominierende Interesse dieser Archive. Eine entsprechende Schwerpunktsetzung findet sich auch bei den nationalen Datenarchiven. Daraus ergeben sich in diesem Zusammenhang vor allem zwei Konsequenzen:

1. Die Datenarchive sammeln möglichst viele Daten, ohne sie einer inhaltlichen Bewertung zu unterziehen.³
2. Die Daten werden in den aktuell gängigen proprietären Formaten angenommen und z.T. in anderen softwareabhängigen Formaten verwahrt und ausgegeben. Probleme der Langzeitarchivierung spielen eine untergeordnete Rolle.⁴

Die Erfahrungen sozialwissenschaftlicher Datenarchive können daher weder zur Bewertung noch zur Langzeitarchivierung statistischer Daten herangezogen werden.

¹ Die elektronischen Statistiken wurden von dem Verfasser im Rahmen eines in der Landesarchivdirektion angesiedelten Assessorenprojekts und in Zusammenarbeit mit Herrn Dr. Häußermann vom StA Ludwigsburg bewertet, dem er an dieser Stelle herzlich danken möchte. Sein herzlicher Dank gilt auch Frau Dr. Nicole Bickhoff (Landesarchivdirektion), Herrn Dr. Udo Schäfer (bisher Landesarchivdirektion, jetzt StA Hamburg) und Herrn Dr. Wolfgang Zimmermann (Landesarchivdirektion) für die gewährte Unterstützung sowie die Diskussion der Ergebnisse.

² http://www.ifdo.org/archiving_distribution/datprep_archiving_bfr.htm (Abruf 8.2.2001).

³ <http://www.za.uni-koeln.de/consultation/data-service/beratung-aufbereitung.htm> (Abruf 8.2.2001): „Bei der Akquisition wird darauf geachtet, dass das Datenmaterial bestimmten methodisch-technischen Kriterien genügt. Eine darüberhinausgehende Bewertung findet durch das ZA aber nicht statt.“ Die IFDO-Studie „Data Archive Workflow“ formuliert: „A general guideline is whether or not the data are usable for future scientific research.“ http://www.ifdo.org/archiving_distribution/archive_sbs_bfr.htm (Abruf 8.2.2001).

⁴ Im ZA „werden die Daten in einem Datenformat aufbereitet, die dem Benutzer eine moderne statistische Analyse etwa mit SAS oder SPSS ermöglicht“, <http://www.za.uni-koeln.de/consultation/data-service/beratung-aufbereitung.htm> (Abruf 8.2.2001). Die in der University of Essex beheimateten Data Archive nennen als standard formats ausschließlich proprietäre Produkte, <http://www.data-archive.ac.uk/depositingData/howtoDeposit.asp> (Abruf 8.2.2001).

2. Fragestellung

In den letzten 40 Jahren haben sich mehrere Autoren mit der Bewertung von statistischen Unterlagen beschäftigt. Zu nennen sind in diesem Zusammenhang Lisa Kaiser, Walter Grube, Klaus Döll, Rainer Stahlschmidt und zuletzt Niklaus Bütikofer. Bis auf den Diplompolitologen Döll sind alle Verfasser Archivare. Da wir heute jedoch nicht mehr vor dem „Problem des Hollerithmaterials“ und ähnlichen Schwierigkeiten stehen, können die Ergebnisse dieser Autoren nicht ungeprüft übernommen werden. Andererseits hat Terry Eastwood darauf hingewiesen, dass die Literatur zur Bewertung elektronischer Unterlagen nur hinsichtlich der technischen Spezifikationen von der sonstigen Bewertungsliteratur abweiche, ansonsten aber keine inhaltlich gestützten eigenen Bewertungskriterien erarbeitet habe.⁵ Die Frage lautet daher, ob sich die von den eingangs genannten Autoren gefundenen Thesen und Lösungen auf die heutige Situation übertragen lassen. Am Rand werde ich dabei auch auf die impliziten Einstellungen unseres eher philologisch gebildeten Berufsstandes eingehen.

3. Archivischer Informationsträger

Sollen maschinenlesbare Unterlagen als solche archiviert werden? In der älteren Literatur wurde die Übernahme maschinenlesbaren Materials aufgrund seiner Fragilität und der hohen Kosten für die erforderlichen Apparate abgelehnt.⁶ Die sozialgeschichtliche Forschung muss sich, so Walter Grube in einem 1954 vor dem 33. Deutschen Archivtag in Goslar gehaltenen Referat, mit der Vernichtung des Hollerithmaterials, also der Lochkarten, abfinden.⁷ Bereits in den 60er-Jahren hat der Diplom-Politologe Klaus Döll in seinem Gutachten dieser Auffassung widersprochen. Auftraggeber war das Bundesarchiv, das 1961 bei dem Fachausschuss für Methodenfragen der Deutschen Gesellschaft für Soziologie angefragt hatte, welche Statistiken aus Sicht der Sozialwissenschaften zu archivieren seien. Die Sozialwissenschaften, so Döll, könnten sich keinesfalls mit der in Aussicht gestellten Vernichtung dieses Materials zufrieden geben. Dieser Überzeugung entsprach aber damals kein adäquates Archivierungskonzept. Zwar nennt Döll Magnetbänder, er beschäftigt sich jedoch nicht mit Fragen der Haltbarkeit des Materials bzw. der Zugänglichkeit zur Information. Heute ist die Archivierung maschinenlesbarer Daten bereits in etlichen Nationalarchiven Realität, und es ist nur eine Frage der Zeit, bis sie auch von den Archiven anderer Träger praktiziert werden muss. Zugleich wächst das Interesse auf der Nachfrageseite. So verabschiedete der Verband der Historiker Deutschlands im Oktober 1994 in Leipzig eine Resolution, maschinenlesbare Daten keineswegs in analoger Form sondern maschinenlesbar zu archivieren.⁸ Auch die 28. Aktenaussonderungsbesprechung der baden-württembergischen Archivverwaltung kam am 15. Juli 1997 zum Ergebnis, dass maschinenlesbare Unterlagen dann nicht analog archiviert werden sollten, wenn dadurch ihre Auswertung verhindert werde.⁹ Dies kann bei den Statistiken schon angesichts ihres Umfangs unterstellt werden. Es ist daher erforderlich, Statistiken als elektronische Unterlagen zu archivieren und sie zuvor entsprechend zu bewerten. Dabei muss zunächst gefragt werden, an welcher Stelle im Prozess statistischer Datenverarbeitung archivwürdige Daten entstehen.

⁵ Terry Eastwood, *Appraisal of electronic records: A review of the literature in english*, 30.5.2000, <http://www.interpares.org/documents/ERAppraisalLiteratureReview.pdf>. (Abruf 8.2.2001).

⁶ Lisa Kaiser, *Zur Frage der Archivierung von Aktenbeständen staatlicher statistischer Ämter im Bundesgebiet*. In: *Der Archivar* 6 (1953) Sp. 113–124, hier Sp. 117. Theodore R. Schellenberg, *Die Bewertung modernen Verwaltungsschriftguts* (Veröffentlichungen der Archivschule Marburg 17), Marburg 1990, S. 63.

⁷ Walter Grube, *Das Problem der Massenakten*. In: *Der Archivar* 7 (1954) Sp. 253–262, hier Sp. 258.

⁸ Landesarchivdirektion, Registratur, AZ 7511.0/15.

⁹ Landesarchivdirektion, Registratur, AZ 7511.0/15.

4. Primärerhebungen

Nach Lisa Kaiser, Archivarin am Statistischen Bundesamt, wurden bis zum ersten Weltkrieg die Bestände der staatlichen statistischen Ämter „von der mehr philologisch-juristisch ausgerichteten Archivverwaltung im allgemeinen nicht als archivwürdig angesehen“.¹⁰ Später sei das Verständnis für den bevölkerungsgeschichtlichen Wert des Materials gewachsen, und vor dem Massenproblem stelle sich nun die Frage nach sinnvollen Kassationsentscheidungen. Demnach können Statistiken prinzipiell archivwürdig sein, eine Bewertung mit philologischen Mitteln wird ihnen aber nicht gerecht. Kaiser unterscheidet zwischen Erhebungsmaterial, Auswertungsmaterial (auf Lochkarten) und Tabellen, die generell bei den statistischen Ämtern verbleiben. Erhebungsmaterial sei dem Auswertungsmaterial vorzuziehen, „denn den von den Staatsarchiven vertretenen, vor allem historischen, Disziplinen bleiben gewöhnlich aus dem Material der maschinellen, aber auch der manuellen Aufbereitung nur selten oder in beschränktem Umfang wissenschaftlich bedeutungsvolle Aufgaben zu lösen“.¹¹ Ihre eigenen Bewertungsentscheidungen richtete die Autorin an den zeitgenössischen Benutzern aus, sie orientierte sich damit selbst an philologischen Kriterien.

Auch Walter Grube legte 1954 besonderen Wert auf die Archivierung von statistischem Urmaterial. Möglicherweise könne aber durch eine verfeinerte maschinelle Auswertung der Tabellen jener Bereich, der ausschließlich aus dem Erhebungsmaterial ersichtlich sei, vermindert werden.¹²

Döll verwahrte sich dann entschieden gegen die Auffassungen der beiden genannten Archivare. Für die Sozialwissenschaften sei eine solche Preisgabe des Hollerithmaterials keineswegs hinzunehmen, da es sich am besten zur Sekundäranalyse eigne. Auch sei das Tabellenmaterial den Primärerhebungen vorzuziehen: „Dieser mit der Verdichtung des Materials verbundene Informationsverlust mag dem Historiker bedeutend erscheinen, für den Sozialwissenschaftler wiegt er nicht schwer.“¹³

5. Bereinigte Daten

Mit der Hinwendung der Geschichtswissenschaften zu den Sozialwissenschaften wurde dieser Ansatz auch theoretisch fundiert. Unter dem Konzept der Metaquelle verstand man nun eine aus einer oder mehreren Originalquellen gewonnene Ansammlung von Informationen. Rainer Stahlschmidt berücksichtigte dieses Konzept, als er Ende der 70er-Jahre eine Studie „Zur Archivierung des Datenmaterials der amtlichen Statistik in Nordrhein-Westfalen“ veröffentlichte.¹⁴ Vorausgegangen war eine umfassende systematische Bestandsaufnahme der vom Landesamt für Datenverarbeitung und Statistik Nordrhein-Westfalen durchgeführten Erhebungen. Nach Stahlschmidt entstehen durch die Abschrift und Speicherung originaler Erhebungsbögen Metaquellen. Sofern nun Primär- und Metaquelle weitgehend identisch seien, könne auf die Archivierung des Urmaterials verzichtet werden. Primäres Ziel sei es, die in der Primärquelle enthaltenen Daten so weit wie möglich zu erhalten und zugänglich zu machen.

Statistiken versuchen bekanntlich die Erscheinungen abzubilden. Dies gelingt aber bei der Erhebung der Daten nicht in vollem Umfang. Erhobene Daten können mit der abzubildenden Materie nicht übereinstimmen. Vor einer Auswertung dieser Daten müssen sie daher mehrere Plausibilisierungsgänge durchlaufen, deren Ziel es ist, die unzutreffenden Daten zu entdecken

¹⁰ Kaiser (wie Anm. 6) Sp. 113.

¹¹ Ebd. Sp. 116.

¹² Grube (wie Anm. 7) Sp. 258 f.

¹³ Klaus Döll, Die Aufbewahrung sozialwissenschaftlich wichtiger Massendaten, Mschr. Köln o.J. [1965], S. 34.

¹⁴ Rainer Stahlschmidt, Zur Archivierung des Datenmaterials der amtlichen Statistik in Nordrhein-Westfalen (Veröffentlichungen der staatlichen Archive des Landes Nordrhein-Westfalen Reihe E: Beiträge zur Archivpraxis 3), Siegburg 1980.

und durch stimmige Daten zu ersetzen. Diese bereinigten Basisdaten enthalten dann ein größeres Maß an Übereinstimmung mit den abzubildenden Phänomenen. Sie weichen jedoch z.T. erheblich von dem Erhebungsmaterial ab. Sind daher die Erhebungsdaten bei der archivischen Bewertung vorzuziehen? Sollten Archivare die ursprüngliche Erhebung oder ein möglichst genaues Abbild der zu beschreibenden Phänomene überliefern? Hinzu kommt, dass die Archivierung unbereinigter Rohdaten jeden künftigen Benutzer dazu zwingen würde, den Plausibilisierungsprozess in Eigenregie durchzuführen. Dafür fehlen ihm nicht nur die zeitgenössischen Informationsmöglichkeiten, es wird ihm aller Wahrscheinlichkeit nach auch an der hierfür nötigen Zeit fehlen.¹⁵ Bereinigte Basisdaten sind daher unabhängig vom Grad ihrer Übereinstimmung mit dem Erhebungsmaterial diesem vorzuziehen. Da hier die Phänomene und nicht der Vorgang der Erhebung interessieren, unterscheidet sich die Bewertung von Statistiken grundsätzlich von der Bewertung klassischer Sachakten.

6. Aggregierte Daten

Nach ihrer Bereinigung werden die einzelnen Daten in der Regel zu größeren Einheiten zusammengefasst. Mit jeder Aggregation gehen dabei Detailangaben verloren. Während die Ausgangsdaten auch später noch beliebig aggregiert werden können, kann aus den aggregierten Daten nicht mehr auf die ursprünglichen Daten geschlossen werden. Abgesehen vom Döllgutachten tendiert daher die deutschsprachige Bewertungsliteratur zur Übernahme von Statistiken mit möglichst geringem Aggregierungsgrad. Niklaus Bütikofer hat in diesem Zusammenhang auf einen weiteren Punkt hingewiesen. Während der Statistiker die Daten im Kontext seiner gegenwärtigen Realität betrachten kann, stehen dem Historiker diese Zusatzinformationen nicht mehr zur Verfügung, er muss sie den Quellen entnehmen. Die Statistiken müssen daher den Historikern mehr Zusatzinformationen anbieten als den Zeitgenossen. Dies ist am ehesten bei unaggregierten Daten gegeben.¹⁶

Andererseits sollte der Benutzer schnell auf die Daten zugreifen und sie verarbeiten können.¹⁷ Von zentraler Bedeutung sind dabei die Verknüpfungsmöglichkeiten mit anderen Daten und Statistiken. Charles Dollar befürwortet daher eine Archivierung der ursprünglichen Zählheiten. Die besten Verknüpfungsmöglichkeiten böten schließlich Parameter wie Eigenname und Sozialversicherungsnummer.¹⁸ Auch Luciana Duranti hebt die zentrale Bedeutung der Verknüpfungsmöglichkeiten hervor. Sie verwendet diese Forderung allerdings als Argument gegen eine Übernahme nichtaggregierter Daten. Die gute Verknüpfbarkeit maschinenlesbarer Daten würde deren Informationswert bedeutend erhöhen.¹⁹

7. Bewertung der statistischen Unterlagen

Im Statistischen Landesamt (kurz StaLa) existieren folgende Formen statistischen Materials: Das Erhebungsmaterial, die aus ihm gebildeten aggregierten Tabellen bzw. Fachstatistiken, das Landesinformationssystem sowie die Publikationen. Das Erhebungsmaterial wird nur kurze Zeit vorrätig gehalten. Da es zu vernichtende Hilfsmerkmale enthält, ist es der Archivverwaltung aus Gründen des Datenschutzes kaum zugänglich. Zudem ist das plausibilisierte Material der anderen Gruppen den Erhebungsunterlagen vorzuziehen.

¹⁵ Niklaus Bütikofer, Archivierung von statistischen Daten. In: Studien und Quellen 16/17 (1990/91), Bern 1991, S. 263–292, hier S. 275.

¹⁶ Ebd. S. 269–274.

¹⁷ Döll (wie Anm. 13) S. 33 nennt die Kriterien Raumbeanspruchung, Informationswert und Auswertungsschnelligkeit.

¹⁸ Charles Dollar, Appraising machine-readable records. In: American Archivist 41 (1978) S. 423–430, hier S. 425.

¹⁹ Luciana Duranti, The thinking on appraisal of electronic records: Its evolution, focuses, and future directions. In: Janus, 1997(2), S. 47–67, hier S. 48.

Viele der Statistiken sind auf der Ebene des Landes oder der Regierungsbezirke durch Publikationen gut dokumentiert. Auf Kreisebene ist das Angebot eingeschränkt, auf Gemeindeebene sehr gering. Im Vergleich mit den Fachstatistiken und dem Landesinformationssystem enthalten sie aufgrund der geringen Regionalisierung bedeutend weniger Informationen, die zudem nur schwer einer größeren Auswertung zuzuführen sind. Publikationen decken das Feld der Statistik daher nur unzureichend ab.

Die wichtigste und größte Datenbank des Landesinformationssystems ist die Struktur- und Regionaldatenbank (SRDB). Letztlich gehen auf sie nahezu alle numerischen Datenbanken zurück, die vom StaLa betrieben werden. Daneben stehen noch Nachweisdatenbanken, die allerdings nicht vom StaLa selbst betrieben werden und folglich auch bei den Bewertungskonzepten der Betreiber (z.B. Innenministerium) zu berücksichtigen sind. Die SRDB erscheint zunächst archivwürdig, da in ihr eine große Zahl zentraler Daten gespeichert ist, die jeweils für relativ kleine Räume (i.d.R. auf der Ebene der Gemeinde) und oft über längere Zeiträume erhoben wurden. Andererseits stammen viele dieser Informationen aus den Fachstatistiken. Es muss daher eine Abwägung zwischen den Fachstatistiken und der SRDB getroffen werden. Die Entscheidung fiel aus den folgenden Gründen zugunsten der SRDB aus:

Nur die SRDB ermöglicht eine nahezu beliebige Verknüpfung der einzelnen Daten zu größeren Auswertungen bei zugleich minimalem Zeitaufwand.²⁰

Die SRDB enthält durchgehend einen hohen Plausibilisierungsgrad. Die der SRDB fehlende Dokumentation der Plausibilisierungsgänge wird durch den Charakter als amtliche Veröffentlichung kompensiert. Zudem ist auch ein großer Teil der Fachstatistiken nur unzureichend dokumentiert.

Die einzelnen Merkmale wurden über einen längeren Zeitraum erhoben und ermöglichen die Erstellung von Zeitreihen.

Wenige große Dateien sind nach der bisherigen Literatur einfacher und damit günstiger zu archivieren als etliche kleine.²¹

Wenigstens in unserer Zeit werden die SRDB-Daten stärker als die Daten aus den nicht in die SRDB eingeflossenen Fachstatistiken nachgefragt.

Die SRDB besteht aus zwei Ausgangsdatenbanken und einer Wertdatenbank. In der einen Ausgangsdatenbank werden die Schlüsselnummern für die Orte, in der anderen diejenigen für die erhobenen Merkmale und den Erhebungszeitpunkt gespeichert. Theoretisch ließe sich die SRDB als eine einzige Tabelle mit 5000 Spalten (für die Orte) und 350.000 Datensätzen (für die jeweilige Erhebung) darstellen, bei der die Werte dann in die einzelnen Felder eingetragen sind. Für die Archivverwaltung stellte sich daher die Frage einer Auswahl von Daten aus der SRDB. Auch diese Frage wird regelmäßig in der Literatur diskutiert.

Walter Grube sah sich durch die von ihm vorgeschlagene Übernahme von Erhebungsmaterial erheblichen Raumproblemen gegenüber und schlug daher vor, nur eine Volkszählung pro Generation geschlossen zu archivieren. Dazwischenliegende Zählungen sollten durch eine

²⁰ Den Belangen des Datenschutzes ist nach den Ausführungen des Datenschutzbeauftragten von Baden-Württemberg vom 12. Mai 1998 unter Auslegung der einschlägigen Bestimmungen des Bundesstatistikgesetzes und des Bundesarchivgesetzes Rechnung zu tragen (Landesarchivdirektion, Registratur, AZ 751.0607/1). Allerdings ist darauf hinzuweisen, dass hier keinesfalls mehrere zuvor existierende Anwendungen durch die Archivverwaltung verknüpft werden. Vielmehr existiert diese Datenbank bereits als solche im StaLa.

²¹ Kevin Ashley, Digital archive costs: Facts and fallacies. In: Proceedings of the DLM-Forum on electronic records. European citizens and electronic information: the memory of the information society, Luxemburg 2000, S. 121–128, hier S. 126. Die Kostenfrage stellt sich sowohl bei einer Übernahme dieser Daten als auch bei einer Vereinbarung mit dem StaLa, die Daten dauerhaft zu pflegen.

repräsentative Auswahl dokumentiert werden. Für Umbruchszeiten müsse schließlich eine besondere Bewertung stattfinden.²²

Das Döll-Gutachten empfiehlt ebenfalls, die einzelnen, regelmäßig durchgeführten Erhebungen in einem festen Zyklus zu archivieren, also jede 3., 5., 7. Erhebung zu übernehmen, da so alle Benutzer über gleich günstige Auswertungsmöglichkeiten verfügen könnten.²³ Durch eine Stichprobenwahl würden dagegen die Ergebnisse derartig festgelegt, dass sie eine Auswertung für spätere Fragestellungen leicht verhinderten.²⁴ Auch Stahlschmidt befürwortet eine zeitliche Auswahl in festen Abständen.²⁵

Nur in Form eines festen Intervalls ist also eine zeitliche Auswahl denkbar. Ein festes Intervall für die gesamte SRDB wird aber den zum Teil sehr unterschiedlichen Merkmalen nicht gerecht. Für jedes Sachmerkmal muss daher ein eigenes Intervall definiert werden. Dann können diese Merkmale aber nicht mehr durchgehend miteinander verknüpft werden. Demographische Untersuchungen würden schließlich nach Zeitschnitten generell nicht mehr durchführbar, beispielsweise könnten die Wanderungsbewegungen im Zuge der Deutschen Einheit nicht mehr nachvollzogen werden. Eine zeitliche Auswahl erscheint daher nicht praktikabel. Eine sachliche Auswahl ist wegen des Charakters der SRDB als umfassende amtliche statistische Beschreibung Baden-Württembergs problematisch, eine räumliche Auswahl würde die tiefe Regionalisierung der SRDB und damit deren Hauptstärke beeinträchtigen. Aus archivarischer Sicht ist daher die dauerhafte Sicherung der gesamten SRDB wünschenswert.²⁶

Indem wir nun das unzureichend dokumentierte Endprodukt der statistischen Bemühungen für archivwürdig erklären, verabschieden wir uns zugleich auch ein wenig von der philologisch begründeten Wertschätzung gegenüber den ursprünglichen Quellen. Auch werden die zu Recht geforderten Kontextinformationen zu Art, Herkunft, Entstehung und Zugänglichkeit der Daten in der SRDB nicht vorgehalten.²⁷ Statt dessen beeinflusste das technische Argument der Verknüpfbarkeit zentral die Bewertungsentscheidung. Auch bei anderen Datenbanken dürften sich die geschilderten Verhältnisse häufig finden lassen. Die Bewertung dieser elektronischer Unterlagen wird daher erhebliche Auswirkungen auf unser archivarisches Selbstverständnis haben.²⁸

²² Grube (wie Anm. 7) Sp. 258 f.

²³ Döll (wie Anm. 13) S. 47.

²⁴ Ebd. S. 42–46. Der Autor führt als weiteres Argument an, dass Stichproben selten Daten über kleine Regionaleinheiten enthielten (ebd. S. 46). Meines Erachtens ist die Problematik einer mittels Stichproben durchgeführten Erhebung nicht auf die Auswahl der zu archivierenden Einheiten zu übertragen.

²⁵ Stahlschmidt (wie Anm. 13) S. 15. Die einzelnen Statistiken bewertet Stahlschmidt zuvor durch negative Wertauswahl. Kassationskriterien sind insbesondere die hinreichende Publikation der Daten, ihre geringe historische Relevanz und die bessere Abbildung der Daten in anderen Quellen, Stahlschmidt, S. 95. Eine ähnliche Liste auch bei Meyer H. Fishbein, *Reflections on appraising statistical records*. In: *American Archivist* 50 (1987) S. 226–234, hier S. 231. Vergleichbar geht auch die staatliche Archivverwaltung Finnlands vor, vgl. Raimo Pohjola, *Appraisal and disposal of electronic records*, <http://www.narc.fi/dlm/6.pdf>, S. 37 (Abruf 8.2.2001).

²⁶ Ob diese Daten übernommen oder aufgrund einer Vereinbarung der Archivverwaltung mit dem StaLa dauerhaft in der Behörde verwahrt werden, soll in diesem Zusammenhang nicht thematisiert werden.

²⁷ Vgl. z.B. Franz-Josef Ziwes, *Überlegungen zur Bewertung digitaler Unterlagen aus dem Umweltinformationssystem Baden-Württemberg*. In: Udo Schäfer – Nicole Bickhoff (Hrsg.), *Archivierung elektronischer Unterlagen* (Werkhefte der Staatlichen Archivverwaltung Baden-Württemberg, Serie A, Landesarchivdirektion 13), Stuttgart 1999, S. 151.

²⁸ Vgl. die Kritik von Mats Burell, *Appraisal and information theory*, <http://www.narc.fi/dlm/7.pdf>, S. 44 f. (Abruf 8.2.2001): „As I mentioned earlier one might very well look upon archival sciences of the 1990s as a truly multidisciplinary discipline. But my impression is that the influences from other disciplines have mostly come from other disciplines within the Humanities, and not so much from the other disciplines outside the Humanities belonging to the rapidly growing Information Society. I believe that these other disciplines should gain on learning more about fundamental archival theory as well and archival theory could be developed by implementing theories from other information sciences.”

Erschließung und Nutzbarmachung digitaler Unterlagen im Bundesarchiv

Von Bettina Martin-Weber

Das Bundesarchiv gehört zu den wenigen Archiven in der Bundesrepublik Deutschland, die über praktische Erfahrung bei der Erschließung und Nutzbarmachung digitaler Unterlagen verfügen. Im vorliegenden Werkstattbericht werden gegenwärtige Praxis, Neuansätze und Perspektiven beschrieben. Die Ausführungen konzentrieren sich auf die Erschließung von Daten aus IT-Fachanwendungen in Dienststellen des Bundes und der DDR. Hierbei handelt es sich um Daten aus File-Systemen, modernen Datenbanksystemen und um Statistikdaten.¹

1. Erschließung digitaler Archivobjekte

Alle relevanten technischen und inhaltlichen Informationen zu archivierten Daten aus einer IT-Fachanwendung einer Provenienzstelle werden im Bundesarchiv bislang in einem sog. Findbuch zusammengestellt. Das „Findbuch“ bezieht sich auf eine digitale Archivalieneinheit, die sich inhaltlich und materiell-physisch gegen andere Dateneinheiten abgrenzen lässt und im folgenden als digitales Archivobjekt bezeichnet wird.² Als Produkt der archivischen Erschließung unterscheidet sich dieses „Findbuch“ von herkömmlichen Findbüchern, die üblicherweise auf der Ebene der Bestände aufsetzen und die Summe der Archivalieneinheiten einer Provenienzstelle, zumindest aber einen umfänglicheren Teil derselben, strukturiert nachweisen. Es enthält zwar durchaus vertraute Elemente, z.B. einen Überblick über die Geschichte der Provenienzstelle, beschreibt aber dann die Entstehung der Daten, den Datenaufbau, die Dateien, die Datenformate, die Datensatzstrukturen, den Feldaufbau, die Kodierungen usw. und dokumentiert die technische Bearbeitung der Daten im Archiv.

Die gebundene Form des Findbuchs setzte der Einbindung von Dokumentationsunterlagen bislang Grenzen. Aktuelle Überlegungen zielen darauf, inhaltliche Erschließung, technische Beschreibung und Dokumentationsunterlagen zu digitalen Archivobjekten in strukturierten, auf Fortschreibung und Zuwachs hin offenen Dossiers³ zu organisieren. Die offene Form bietet sich auch an, da eine Publikation der Dossiers im allgemeinen nicht zweckmäßig ist, da sie nicht für die Recherche, sondern erst in einem fortgeschrittenen Stadium der Benutzung tatsächlich herangezogen werden. Gleichwohl steht einer online-Version der Textteile der Dossiers nichts entgegen.

Bislang liegen Findbücher für eine Vielzahl, aber keineswegs für alle digitalen Archivobjekte vor; sie repräsentieren jeweils vorläufig die vollständige Erschließung eines Objektes. Ihre Aufbereitung ist arbeitsintensiv und wird nie abgeschlossen sein, da die technische Bearbeitung (z.B. Migrationen) hierin nachzuhalten ist.

Neben diesen Findbüchern – bzw. künftig Dossiers – werden in verschiedenen Tabellen alle Bearbeitungen, Um- und Sicherungskopierungen sowie Konvertierungen und auch Anonymi-

¹ Das Bundesarchiv hat bisher keine elektronischen Akten übernommen.

² Der Begriff lehnt sich an den im Referenzmodell OAIIS (Open Archival Information System) gewählten Terminus des „Digital Information Object“ (vgl. Consultative Committee for Space Data Systems: Reference Model for an Open Archival Information System, CCSDS 650.0-W-5.0 White Book, 1999) an und meint die übernommenen primären Daten ohne weitere Darstellungs-, Meta- oder Erschließungsinformationen. Es kann sich hierbei um eine u.U. sehr umfangreiche Datei, aber auch um mehrere Dateien handeln.

³ Zu dieser Organisationsform und zum Begriff regte ein Besuch im Schweizer Bundesarchiv im Sommer 2000 an.

sierungen dokumentiert. Diese technische Dokumentation dient sowohl dem Nachweis der Bearbeitung als auch der internen Arbeitsplanung.

In der Archivdatenbank BASYS (Bundesarchiv – System), die v.a. als zentrales Erschließungsinstrument für Schriftgut seit den frühen 90er-Jahren im Bundesarchiv aufgebaut wurde, wird digitales Archivgut bislang nur in den sog. „Stammdatens“, in Informationen auf Bestandsebene und zu den Provenienzstellen nachgewiesen.⁴ Digitale Archivobjekte werden in BASYS bislang nicht erschlossen.

Eine Beständeübersicht über die maschinenlesbaren Daten im Bundesarchiv bietet der sog. Datenbestandskatalog (DBK), der die digitale Überlieferung der Provenienzstellen auf der Ebene der digitalen Archivobjekte nachweist. Sie ist im Intranet des Bundesarchivs für den internen Gebrauch einsehbar und wird Benutzern derzeit als Ausdruck oder Word-Datei zur Verfügung gestellt. Im Internet sind die maschinenlesbaren Bestände des Bundesarchivs im Beständebaum mit kurzen Beschreibungen auf der Ebene der digitalen Archivobjekte vollständig nachgewiesen.

2. Funktionen der Erschließung

Archivische Erschließung hat den Zweck, Benutzern das Auffinden gesuchter Archivalien zu ermöglichen und strukturelle Bezüge zu verdeutlichen. Die Erschließung digitaler Archivobjekte hat weitergehende Funktionen. Während sich konventionelles Archivgut im allgemeinen dem (kundigen) Benutzer sehr weitgehend selbst erschließt, da es eine Fülle relevanter Kontextinformationen in sich trägt, sind digitale Archivobjekte nur benutzbar, interpretierbar und analysierbar, wenn dem Benutzer vorab differenzierte Kontextinformationen und technische Daten an die Hand gegeben werden, die sich aus dem digitalen Archivobjekt selbst nicht erschließen lassen. Erst die Erschließung der Objekte stellt ihre Benutzbarkeit her.

Darüber hinaus hat Erschließung die Funktion, die Authentizität der digitalen Überlieferungsform nachzuweisen. Die Authentizität digitaler Objekte lässt sich nicht durch Ab- oder Vergleich mit anderen Originalen, Kopien, paralleler Überlieferung o.Ä. verifizieren oder falsifizieren. Allenfalls mangelnde Plausibilität der Daten lässt an ihrer „Echtheit“ zweifeln. Es müssen daher andere Wege des Nachweises der Authentizität der Unterlagen gefunden werden.

Dass unter den Bedingungen der Migration keine unbedingte Authentizität – im Sinne der Wahrung von Identität – möglich ist, liegt auf der Hand. Erst in der archivischen Praxis aber wird deutlich, in welchem Umfang Modifikationen an übernommenen Daten notwendig sein können, sollen sie in für die Langzeitarchivierung und Nutzung nach heutigem Stand geeignete Formate und Strukturen konvertiert werden. Betroffen sind nicht nur die Datenformate. Auch Eingriffe bei Dateistrukturen, Datensatzstrukturen und dem Aufbau von Datenfeldern können erforderlich werden. Dies gilt insbesondere für Daten aus älteren DV-Anwendungen, die häufig nicht heute üblichen Normalformen entsprechen.

Bei der Langzeitarchivierung von Daten aus IT-Fachanwendungen durch Migration – elektronische Akten sind bewusst ausgenommen – geht es nicht um die Bewahrung des „look and feel“, sondern um die Sicherung des Informationsgehaltes der Daten und den Erhalt der Auswertungsoptionen. Dies entspricht einer Form der Authentizität, die ich als „informationelle Authentizität“ charakterisieren möchte. Eine lückenlose, präzise Dokumentation der technischen Bearbeitung der Daten ist daher als Nachweis ihres in den Konvertierungs- und Kopierprozeduren unveränderten Informationsgehaltes unverzichtbar.

⁴ Die maschinenlesbaren Daten einer Provenienzstelle bilden hierbei einen eigenen Bestand, der sich durch den Zusatz „MD“ von der für den Schriftgutbestand gebildeten Archivsignatur unterscheidet (DA 5 Staatsrat der DDR, DA 5 MD Staatsrat der DDR – maschinenlesbare Daten).

Die vorhandenen Erschließungsinstrumente, vor allem Datenbanken, sind bislang nicht auf die besonderen Erfordernisse der Erschließung digitaler Überlieferung abgestimmt. Im Bundesarchiv erfolgte die Erschließung digitaler Archivobjekte in der beschriebenen Form bislang getrennt von der übrigen schriftlichen Überlieferung einer Provenienzstelle. Diese Entwicklung wurde durch organisatorische Rahmenbedingungen verstärkt.⁵ Es wird nun aber darauf ankommen, den Erschließungszusammenhang von konventioneller und digitaler Überlieferung einer Provenienzstelle herzustellen. Es werden Modelle zu entwickeln sein, mittels derer der Nachweis und die Erschließung von Daten aus IT-Fachanwendungen in Beziehung zur Erschließung der Schriftgutüberlieferung einer Provenienzstelle erfolgen kann. Für den Nachweis elektronischer Akten wird dies in Zukunft ohnehin unabdingbar sein.

Für die interne Informationspolitik im Bundesarchiv ist der Nachweis der digitalen Überlieferung in einer zentralen und allgemein zugänglichen Erschließungsumgebung aufgrund der Aufbauorganisation und der dislozierten Struktur von ganz besonderer Bedeutung.

Aber auch für den externen online-Zugriff auf Erschließungsinformationen, sei es in einer Datenbank-Replikation im Internet oder in datenbankgestützten Online-Findbüchern, ist der Nachweis digitaler Archivobjekte in BASYS wünschenswert.

3. Erschließungsinformationen im Benutzungsprozess

Umfängliche Erschließungsinformationen zu digitalen Archivobjekten sind notwendig, aber nicht alle Informationen sind in den unterschiedlichen Stadien eines Benutzungsprozesses von Relevanz.

Der Benutzungsprozess lässt sich in drei Phasen unterscheiden:

Erste Phase: Die Recherche in einem zentralen, allgemein zugänglichen Findhilfsmittel (Beständeübersicht, Datenbestandskatalog, Datenbank) führt den Benutzer zu einem digitalen Archivobjekt. Hierbei sollte auch die Struktur der parallelen Schriftgutüberlieferung einer Provenienzstelle wahrzunehmen sein.

Zweite Phase: Der Benutzer informiert sich in einem auf das einzelne Objekt bezogenen Dossier detailliert über Inhalt, Struktur, IT-technische und rechtliche Nutzungsbedingungen, Auswertungsoptionen etc. Vor allem die zweite Phase unterscheidet derzeit die Benutzung digitaler Überlieferung von einer gewöhnlichen Schriftgutbenutzung.

Dritte Phase: Der wissenschaftliche Benutzer wertet die Daten aus und verarbeitet sie weiter. In der Regel erfolgt dies nicht im Archiv, sondern mit Benutzungskopien am heimischen Rechner. Hiervon zu unterscheiden sind Anfragen nach Einzelinformationen, häufig privater Natur, die im allgemeinen nur durch gezielte Recherchen durch Archivpersonal zu beantworten sind.

In der ersten Phase reichen zum Nachweis der digitalen Überlieferung wenige Metadaten aus. Zu erfassen sind die Provenienzstelle und die Archivsignatur, die sprechende Bezeichnung der IT-Fachanwendung, die Laufzeit der Daten bzw. der Entstehungszeitpunkt der Daten, ggf. die Bezeichnung einzelner Dateien.

Nutzbare oder für die Aufnahme vergleichbarer Informationen bestimmte Felder existieren in der Erschließungsdatenbank im Bundesarchiv. Die im allgemeinen aussagekräftige Bezeichnung einer IT-Fachanwendung kann z.B. als Titel erfasst werden (z.B. *Kriminalstatistik, Zentraler Kaderdatenspeicher, Todesursachenstatistik*). Die Signatur eines digitalen Archivobjektes und die Laufzeit einer zugrundeliegenden IT-Fachanwendung sind problemlos zu

⁵ Die Zuständigkeit des Referates für maschinenlesbare Dateien erstreckt sich auf drei, an verschiedenen Orten untergebrachte Abteilungen des Bundesarchivs – DDR in Berlin, Bundesrepublik in Koblenz, Militärarchiv in Freiburg – und wird derzeit aus einer vierten Abteilung LS heraus in Koblenz wahrgenommen.

erfassen. „Sprechende“ Bezeichnungen einzelner, zu einem Archivobjekt gehörende Dateien können z.B. in den Enthält-Vermerken nachgewiesen werden.

Im Zuge der Entwicklung eines Datenmodells für eine Erschließung der digitalen Überlieferung im Kontext der Schriftgutüberlieferung werden die in BASYS vorhandenen Erschließungsansätze zu prüfen sein. Dabei wird kritisch zu hinterfragen sein, welche Informationen hierin zweckmäßigerweise aufzunehmen sind.

Erst in der zweiten Phase der Benutzung werden die übrigen Metadaten und Kontextinformationen dem Benutzer zur Verfügung gestellt. Sie sollen den Entstehungszusammenhang der Daten evident halten, ihren Aufbau und Inhalt beschreiben, die technische Bearbeitung dokumentieren sowie die Benutzungsmodalitäten festhalten und sind häufig um Dokumentationsunterlagen, Codepläne u.ä. angereichert. Es handelt sich hierbei im wesentlichen um die oben beschriebenen Findbücher bzw. Dossiers.

Erschließungsinformationen

<p><input type="checkbox"/> Metadaten</p> <ul style="list-style-type: none"> • Provenienzstelle • Archivsignatur • Bezeichnung der IT-Fachanwendung • Laufzeit • Aufbau der Daten: Dateiübersicht, Datensatz-, Datenfeldbeschreibung, Schlüssel 	<p><input type="checkbox"/> Kontextinformationen</p> <ul style="list-style-type: none"> • Aufgaben der Provenienzstelle • Beschreibung der IT-Fachanwendung • Genese der Daten • Dokumentationsunterlagen • Codelisten • Technische Bearbeitung <ul style="list-style-type: none"> – Konvertierung, Formate – Speichermedien • Benutzungsmodalitäten
---	---

Bettina Martin-Weber, Bundesarchiv

4. „Erschließungslandschaft“

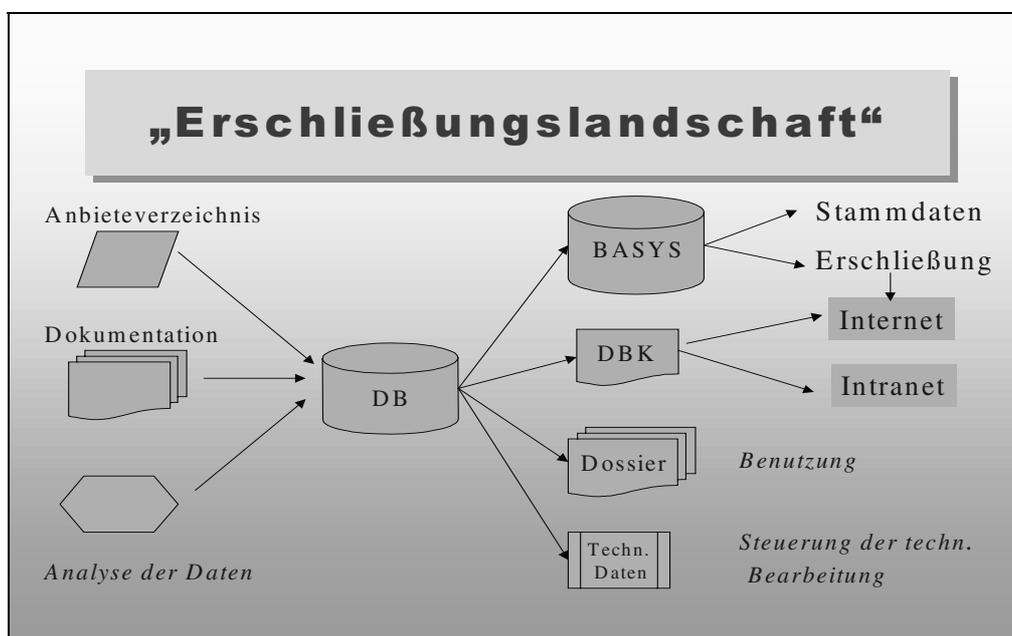
Es ist sinnlos, Daten zu archivieren, zu denen von den abgebenden Stellen keine aussagekräftigen Dokumentationsunterlagen geliefert werden, weil die Benutzbarkeit der Daten im allgemeinen nicht herzustellen ist. Daten tragen im Unterschied zu anderen Unterlagen meist nur wenige Metadaten und keine Kontextinformationen in sich. Die Archive sind darauf angewiesen, dass ihnen von den datenabgebenden Stellen alle relevanten Informationen zur Verfügung gestellt werden. Dies sollte in elektronischer Form in einem standardisierten Anbieterverzeichnis erfolgen, das Grundlage der Bewertung ist und für die spätere Erschließung durch Dokumentationsunterlagen zu ergänzen ist.⁶

Künftig sollen im Bundesarchiv die Informationen aus Anbieterverzeichnis, Auswertung der Dokumentationsunterlagen, Analyse der Daten und Protokollierung der technischen Bearbeitung in einer Datenbank zusammengeführt werden, die zunächst als „Insellösung“ in der zu-

⁶ Das Bundesarchiv führt derzeit in einem Pilotprojekt im Geschäftsbereich des Bundesministeriums für Verteidigung eine Erhebung zu IT-Fachanwendungen durch, die zunächst als Bewertungsgrundlage und dann als Basis eines differenzierteren Anbieterverzeichnisses für archivwürdige Anwendungen dienen soll. Um Aufwände bei den Behörden zu minimieren, wurden in der ersten Phase keine technischen Details erhoben. Diese sollen nach erfolgter Bewertung nur für die archivwürdigen Anwendungen erfasst werden.

ständigen Organisationseinheit aufgebaut wird, dann aber als Grundlage eines Datenmodells fungieren soll, das perspektivisch in BASYS zu realisieren ist. Aus der Datenbank als Informationspool sollen die jeweils relevanten Informationen in die verschiedenen Erschließungsinstrumente und Präsentationsformen fließen und die Steuerung der technischen Bearbeitung erfolgen.

Metadaten werden nach BASYS sowohl auf der Ebene der „Stammdaten“ der Bestände und Provenienzstellen als auch auf der Ebene der Archivalieneinheiten kopiert. Die Erschließungsdatenbank soll im Internet gespiegelt werden. Aus der Datenbank wird die Beständeübersicht „Datenbestandskatalog“ erstellt, die ihrerseits Grundlage der Präsentation von Bestandsinformationen im Intranet und Internet sein kann. Metadaten, Kontextinformationen und technische Beschreibung bilden die Grundlage des Dossiers. Die technische Beschreibung dient zur Steuerung der technischen Bearbeitung, also vor allem der Konvertierung und Sicherungskopierung.



5. Archivierungsstrategie und technisches Konzept

Im Bundesarchiv werden digitale Archivobjekte in nicht-proprietären Formaten als ASCII-Flat Files mit Feld- und Satztrennzeichen (csv-Dateien) gespeichert, derzeit als reine Primärdaten ohne weitere Metadaten und Kontextinformationen.

Da es sich bei den bisher archivierten Daten ganz überwiegend um DDR-Daten handelt, die nicht im Standard-EBCDIC, sondern zum Teil komprimiert und binär gepackt vorliegen, kann die Konvertierung nicht mit Standard-Tools durchgeführt werden. Vielmehr sind spezielle Konvertierungsprogramme erforderlich. Diese werden von externen Programmierern auf der Basis von Werkverträgen erstellt. Zum Auftragsvolumen gehört häufig die gleichzeitige Erstellung von Programmen zur Aufbereitung anonymisierter Benutzungsversionen personenbezogener Datenbestände. Die Konvertierung der Daten selbst und die Sicherung bzw. Umkopierung der Daten erfolgt im Bundesarchiv im Fachreferat.⁷

⁷ Für die Sicherungsroutinen ließe sich ein Outsourcing denken, allerdings sind hier auch noch nicht die Quantitäten angefallen, die die Vergabe zur kostengünstigeren Variante werden ließen.

Die Daten werden derzeit auf CD-R und DAT gesichert und können auf diesen Medien auch Benutzern zur Verfügung gestellt werden.⁸

Die Schutzfristen für DDR-Bestände werden bei wissenschaftlichen Benutzungsvorhaben regelmäßig verkürzt. Damit ist auch die Benutzung von Datenbeständen im Bundesarchiv bereits Realität. Allerdings besteht vor allem hinsichtlich der noch keineswegs abgeschlossenen Erschließung und technischen Bearbeitung der DDR-Daten ein drängender Handlungsbedarf.⁹

Bei der Übernahme von Daten aus der Bundesverwaltung steht das Bundesarchiv am Anfang und muss feststellen, daß sich die für die Übernahme der Akten eingeübten Routinen keineswegs auf die Übernahme von Daten übertragen lassen, sondern in Einzelfällen auch aufgrund der mit der Datenübergabe verbundenen Aufwände mit nicht unerheblichen Widerständen bei den anbietepflichtigen Stellen zu rechnen ist.

Angesichts der Rückstände bei der Bearbeitung bereits übernommener und benutzbarer Daten und der gewaltigen Aufgaben, die vor uns stehen, ist kritisch zu fragen, was in der Nutzbarmachung digitaler Überlieferung geleistet werden muss – und was geleistet werden kann.

Auf die Bedeutung einer sorgfältigen und vielschichtigen Erschließung wurde hingewiesen. Festgestellt habe ich in diesem Zusammenhang, dass die komplexe und zeitintensive Benutzung der Daten in aller Regel nicht im Archiv erfolgt. Hinzu kommt, dass die Benutzbarkeit der DDR-Daten eine Ausnahme darstellt, für bundesrepublikanische Daten aber i.d.R. Schutzfristen einer heutigen Nutzung entgegenstehen. Es macht daher derzeit keinen Sinn, archivierte Daten in einem ständigen Zugriff auf schnellen Speichermedien, auf Jukeboxen oder in Datenbanken zu halten, wie sie die „Archivierungskonzepte“ kommerzieller Anbieter für die kurz- oder mittelfristige Sicherung und Bereitstellung der Daten anbieten. Das Verfahren, die Daten auf Speichermedien vergleichbar der konventionellen Überlieferung im Magazin zu lagern und bei Bedarf auszuheben und zu kopieren, hat sich im Bundesarchiv bewährt. Dies gilt sowohl für die Nutzung als auch für die technische Bearbeitung im Rahmen der Sicherungsstrategie.

Auch dem Anspruch an online-Verfügbarkeit der Daten kann mit skeptischer Zurückhaltung begegnet werden. Der Aufwand, den die Präsentation sehr großer Datenbestände im Internet bedeutet, sowohl in der Aufbereitung als auch in der Datenpflege, dürfte in keinem Verhältnis zum Nutzen stehen.

Eine andere Frage ist die Aufbereitung der Benutzungsdateien. Sicher ist es wünschenswert, Daten komfortabler in einer Datenbankumgebung oder mit Statistikprogrammen aufbereitet zu nutzen. Vor allem die Anschaulichkeit der Daten im Archiv wäre so zu verbessern, auch interne Recherchen und Analysen würden erleichtert. Unseren Erfahrungen nach sind wissenschaftliche Benutzer aber in der Regel an der gesamten Datenbasis interessiert, die in längerfristig angelegten Projekten Grundlage umfangreicher Auswertungen und Bearbeitungen ist, die nicht im Archiv stattfinden. Hierfür sind die Daten in einer Form zur Verfügung zu stellen, die ihre unproblematische Weiterverarbeitung in Datenbanken oder mit Statistikprogrammen gewährleistet. Mehr muss nicht geleistet werden. Die Recherchen nach Einzelinformationen aber, wie sie bei den personenbezogenen DDR-Dateien zu erwarten, bisher allerdings

⁸ Für die Bereitstellung der Daten fallen an Auslagen 80 DM für DAT und CD und 11 DM für Disketten an. Vgl. die Verordnung über Kosten beim Bundesarchiv (Bundesarchiv-Kostenverordnung) vom 29. September 1997 (BGBl I 2380), geändert durch die 1. Verordnung zur Änderung der Bundesarchiv-Kostenverordnung vom 7. Nov. 2000 (BGBl I 1495).

⁹ Die Zeit drängt, weil häufig nur durch die Nachfrage bei Personen, die mit der Entstehung der Daten befasst waren, die erforderlichen Informationen für das technische und inhaltliche Verständnis der Daten zu beschaffen sind. Auch zeigen jüngste Erfahrungen, dass keineswegs alle vorhandenen Unterlagen tatsächlich erfasst und gesichert werden konnten, sondern dass auch heute noch Dokumentationsunterlagen zu DDR-Daten in Bundesbehörden liegen.

eher selten aufgetaucht sind, können bei einer schlichten Bereitstellung der Daten z.B. als Flat Files dem Benutzer kaum zugemutet werden, weil in jedem Fall Rechner, Software und Know-how vorhanden sein müssten. Diese Recherchen sind im Archiv mit eigenen Kräften durchzuführen.

Wenn ich im Falle der Aufbereitung der Daten auch angesichts begrenzter Ressourcen für Zurückhaltung plädiere, so scheinen mir auf dem Gebiet der Qualifizierung der Archivare Investitionen erforderlich zu sein. Ich möchte diese Forderung mit dem Begriff Methodenkompetenz beschreiben. Sowohl für eine weitsichtige, inhaltlich und technisch wohl begründete Bearbeitung der Daten als auch für die Erschließung und die Benutzerberatung halte ich profunde IT-Kenntnisse der archivischen Mitarbeiter und entsprechende Schulungsangebote für unverzichtbar.

XML als Perspektive

Als Vorstudie zu dem im vergangenen Jahr in Mannheim vorgestellten Projekt zur „Elektronischen Archivierung und Nutzung (ELAN)“ hat das Bundesarchiv gemeinsam mit der Forschungsstelle für Verwaltungsinformatik der Universität Koblenz im Rahmen einer Diplomarbeit die Beschreibungs- und Auszeichnungssprache XML auf ihre Anwendbarkeit für die Langzeitarchivierung einerseits und die Webpräsentation und Online-Recherche andererseits hin untersucht.¹⁰ Ausgewählt wurde hierfür der Archivdatenbestand „Datei Grenzzwischenfälle“ (DVH 32 MD/1). Es handelt sich um eine Datei, in der sämtliche Vorkommnisse an den Grenzen der DDR zur Bundesrepublik, zu West-Berlin, Polen und zur Tschechoslowakei zwischen dem 1. Dezember 1971 und dem 30. September 1990 erfasst sind. Pro Jahr wurden durchschnittlich 35.000 Meldungen der Grenzposten durch das Kommando der Grenztruppen der DDR aufgenommen. Insgesamt umfasst der Datenbestand 650.000 Datensätze, in ASCII-csv-Dateien ca. 180 MB.¹¹ Die „Datei Grenzzwischenfälle“ bot sich als Objekt an, weil sie bereits konvertiert als Flat File im Standard-ASCII-Format vorlag. Der Datenbestand wurde auch ausgewählt, weil er im Unterschied zu vielen anderen DDR-Datenbeständen des Bundesarchivs keine personenbezogenen Daten enthält und auch ohne zusätzliche Bearbeitung (Anonymisierung) keine besonderen Benutzungsprobleme aufwirft. Außerdem konnte unterstellt werden, daß gerade an diesen Daten neben dem wissenschaftlichen auch ein allgemeines öffentliches Interesse – publizistisch, lokalgeschichtlich, persönlich motiviert – besteht, eine Online-Nachfrage also wahrscheinlich ist.

Die prinzipiellen Vorteile von XML wurden bereits in verschiedenen Vorträgen anlässlich dieser Tagung thematisiert.

Ich meine, dass XML nicht nur neue Perspektiven für den systemunabhängigen Datenaustausch bei archivischen Findmitteln aufzeigt, sondern auch neue Möglichkeiten für die Langzeitarchivierung bietet, da durch die Einbettung von Informationen in ein XML-Dokument Informationsverluste begrenzt und gleichzeitig die Lesbarkeit und Verständlichkeit der archivierten Daten selbst erhöht werden kann. Einem Defizit der derzeitigen, oben beschriebenen Archivierungsstrategie, der Verlust struktureller und funktionaler Informationen bei komplexen Datenobjekten, der nur durch die Bereitstellung entsprechend aussagekräftiger Dokumentationsunterlagen und Beschreibungen zu kompensieren ist, kann mit XML wirksam begegnet werden.

¹⁰ Die Diplomarbeit von Björn Dehms, Langzeitarchivierung einfacher, relationaler Datenbanken. Entwicklung eines Prototypen zur Migration nach XML, wurde im Dezember 2000 vom Fachbereich Informatik der Universität Koblenz-Landau, Abteilung Koblenz angenommen.

¹¹ Im Rahmen der Tagung stellte Ulrich Meyer von der Forschungsstelle für Verwaltungsinformatik das Konvertierungsverfahren ausführlich da. Vgl. den Beitrag von Dehms, Engel und Meyer in diesem Band.

Mittels XML lassen sich semantische und strukturelle Informationen mit den Primärdaten verbinden und in einem Objekt darstellen und Metadaten, Informationen zur Provenienzstelle, Dateibezeichnungen, Angaben zur Laufzeit, zur Datenstruktur usw. im datentechnischen Verbund mit den Primärdaten in ein XML-Dokument einbinden¹². Bislang sind die archivierten Daten nur aufgrund der auf dem Speichermedium aufgetragenen Signatur und der Dateibezeichnung identifizierbar.

Es ist allerdings einschränkend festzustellen, dass die Präsentation von und die Recherche in großen Datenbeständen mit XML-eigenen Werkzeugen ohne Datenbankunterstützung nicht zu realisieren ist. Die Aufgabenstellung der Diplomarbeit ging von der Überlegung aus, dass das XML-Dokument, das für die Webpräsentation aufbereitet wurde, gleichzeitig den Anforderungen für die Langzeitarchivierung genügen sollte, also softwareunabhängig realisiert werden sollte. Auf den Einsatz eines proprietären Datenbanksystems wurde daher verzichtet. Der Versuch, die Präsentation der Daten in einer Webversion ohne Unterstützung durch ein Datenbankprogramm zu verwirklichen, schlug fehl. Mit akzeptabler Performance ließ sich nur ein sehr kleiner Ausschnitt der Daten präsentieren. Die für eine derartige Webpräsentation – die für den Benutzer nur exemplarischen Wert hat – erforderlichen Aufwände sind angesichts des zu erwartenden Nutzens nicht vertretbar.¹³

Das Bundesarchiv wird XML als Werkzeug zur Langzeitarchivierung im Auge behalten. Allerdings ist vor allem die Entwicklung leistungsfähigerer Tools zur Präsentation und Migration der Daten aus XML-Dokumenten abzuwarten. Vor allem aber dürfte die Entwicklung nativer XML-Datenbanken ganz neue Möglichkeiten eröffnen.

Das im Rahmen der Diplomarbeit erstellte Konvertierungsprogramm wurde inzwischen auf der Grundlage eines Werkvertrags leicht modifiziert und soll an weiteren Beständen des Bundesarchivs getestet werden.

Fazit

Erst die Erschließung digitaler Archivobjekte stellt deren Benutzbarkeit sicher. Vor allem in dieser Hinsicht unterscheidet sie sich von traditioneller Schriftguterschließung. Für die Bereitstellung und Verbreitung der Erschließungsinformationen sind geeignete Formen zu entwickeln. Sie sollten sich an der Funktion von Information in veränderten Benutzungsverfahren orientieren und die Zusammenführung aller Erschließungsinformationen zur gesamten Überlieferung einer Provenienzstelle zum Ziel haben. Dies ist die vordringliche Aufgabe. Langzeitarchivierung in XML würde die Voraussetzungen für die Erschließung erneut drastisch verändern, weil Metadaten, Struktur- und Kontextinformationen nicht in Findmitteln bereitgestellt, sondern weitgehend in die digitalen Archivobjekte selbst eingebunden würden. XML wird in Archiven derzeit vor allem für den Austausch archivischer Erschließungsinformationen getestet und eingesetzt. Ob es auch für die Langzeitarchivierung taugt, wird die Zukunft zeigen.

¹² Damit würden die archivierten Objekte sich den in OAIS als Archival Information Packages (AIP) beschriebenen Informationspaketen annähern.

¹³ Die im Rahmen der Diplomarbeit konzipierte Online-Auskunft war kurzzeitig auf der Webseite der Universität Koblenz einsehbar. Sie basiert auf Windows NT und ist auf dem Linux-Webserver des Bundesarchivs derzeit nicht lauffähig. Eine Anpassung der Daten steht noch aus.

Der Einsatz von Metadaten-Sets in Schriftgutverwaltung und Archiv

von Nils Brübach



Der Einsatz von Metadaten-Sets in
Schriftgutverwaltung und Archiv

Nils Brübach

Inhalt dieser Präsentation:

- Kurze Vorstellung folgender Ansätze:
 - ◆ Dublin-Core
 - ◆ Pittsburgh-Projekt
 - ◆ NARA
 - ◆ Australische Ansätze
 - ◆ Aktuelle europäische Entwicklungen
- Fragestellungen:
 - ◆ Metadaten: Technisches Hilfsmittel oder funktionales Instrument ?
 - ◆ Metadaten-Sets in Schriftgutverwaltung und Archiven – Einheitlichkeit oder Unterscheidung?

Aktuelle Ansätze: -Definitionen-

- „Metadata is one of those words that means different things to different people“
- „ Strukturierte Daten über Daten“
- Ausgangspunkt: Bibliothekarische Diskussion: Dublin Core Metadata Initiative (1992-1996) Metadata Element Set 1996
- Instrument zum Zugang elektronischer Publikationen im Internet

....Dublin Core:

- Unterscheidung von Inhaltsabhängigen, Inhaltsbeschreibenden und Inhaltsunabhängigen Elementen
- Metadaten-Elemente:
 - ◆ DC.Title: Titel der zugänglich zu machenden Quelle
 - ◆ DC.Creator: Verfasser/Urheber
 - ◆ DC.Subject: Thema, als Schlagwort z.B.
 - ◆ DC.Description: Inhaltsbeschreibung
 - ◆ DC.Publisher: Herausgeber, Verleger
 - ◆ DC.Date: Zeitliches, z.B. Publikationsdatum, Zeitangaben zu Veränderungen, Versionen
 - ◆ DC.Type: Dokumententyp der Resource

....Dublin Core

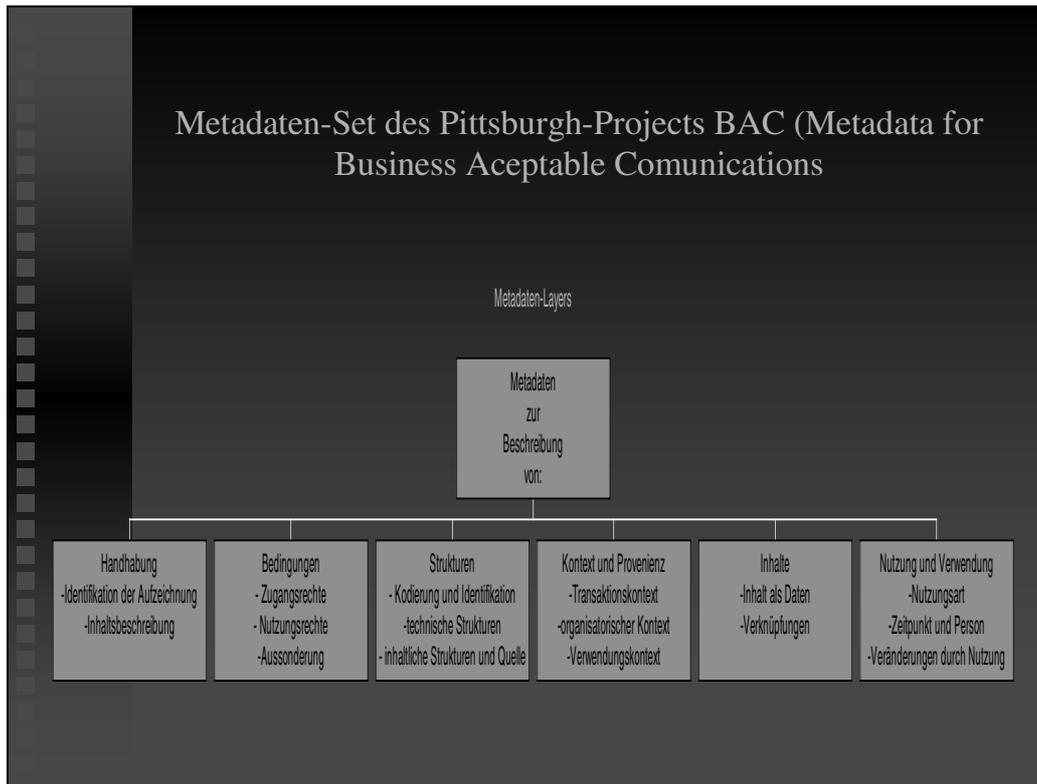
- ◆ DC.Format: Dateiart, Dateiformat, physikalisches Medium
- ◆ DC.Identifier: Z.B. URL, ISBN
- ◆ DC.Source: Quelle, Ableitungen
- ◆ DC.Language: Sprache
- ◆ DC.Relation: Z.B. Fußnoten
- ◆ DC.Coverage: Räumliche, zeitliche Bezüge
- ◆ DC.Rights: Z.B. Rechteinhaber, Copyright

....Dublin Core:

- Grundelemente werden durch Sub-Elements erweitert
- DC Qualifiers beschreiben Anwendung und inhaltliche Form der Metaelemente
- -> bibliothekarisches Arbeitsinstrument zur Verknüpfung von Formal- und Inhaltser-schliessung
- Quelle: http://purl.oclc.org/metadata/dublin_core
- In Deutschland: German Metadata Registry als Clearingstelle
<http://www.mpib-Berlin.mpg.de/dok/gmr/>

Archivische Adaption:

- David Bearman, University of Pittsburgh:
- Erweiterung des Ansatzes über den Inhaltsbezug hinaus
- Metadaten zur Beschreibung von:
 - ◆ Inhaltsbezügen
 - ◆ Kontextbezügen
 - ◆ Strukturbezügen



... Pittsburgh-Project:

- Warrants (Garantien): Externe und interne qualitative Anforderungen in Bezug auf
- Rechtliche und administrative Umgebung
- Rechenschaftsfähigkeit
- Erfassung und Eigenschaften der Aufzeichnungen
- <http://www.lis.pitt.edu/~nhprc/>

Projekt der NARA: Persistent Object Preservation

- Teilprojekt des DOCT (Distributed Object Computed Testbed)
- Bedingung: digitale Objekte, die über längere Zeit hin nicht verändert werden, werden unzugänglich. Das ist Resultat der elektronischen Obsoleszenz. Deshalb gibt es eine neuartige Spannung zwischen dem Ziel unveränderter Aufbewahrung und der Erhaltung der Nutzbarkeit.

Projekt der NARA: Persistent Object Preservation

- Einsatz zur Archivierung von...
 - ◆ 1 Mio Botschaftsberichte pro Jahr
 - ◆ 25 Mio E-Mails der Clinton-Administration
 - ◆ Große Datenbanken: Vietnam, GIS

Projekt der NARA: Persistent Object Preservation

- Automatische Verknüpfung von E-Records mit DTD's auf XML-Basis
- Aufbewahrung von E-Records in Verknüpfung mit den DTD's (Einkapselung)
- Verwendung von XSL-Stylesheets für die Darstellung
- Beschreibung der Struktur mit Metadaten

Projekt der NARA: Persistent Object Preservation

- Metadaten-Set:
 - ◆ Automatisch generiert
 - ◆ Auf der Ebene der E-Records und der Bestände
 - ◆ Ergänzung durch Nutzungshinweise
 - ◆ Provenienz und Zugang als Kernelemente
 - ◆ <http://www.sdsc.edu/NARA/Publications.html>

Australische Metadaten- Standards

- Basierend auf dem „Records-Continuum“ (Frank Upward) und Erschließungsstrategien für das australische Serienaktensystem (Chris Hurley)
- Einheitliche Metadaten-Sets für Schriftgutverwaltung und Archive
- Dreistufiger Ansatz...

Australische Metadaten-Standards

- Basis ist die australische Schriftgutverwaltungsnorm AS 4390
- 1. Ebene: AGLS (Australian Government Locator Service) Online-Zugang zu Informationen der australischen Zentralregierung
 - ◆ 2 Elemente: Agents und Services, ergänzt um 2 Thesauri

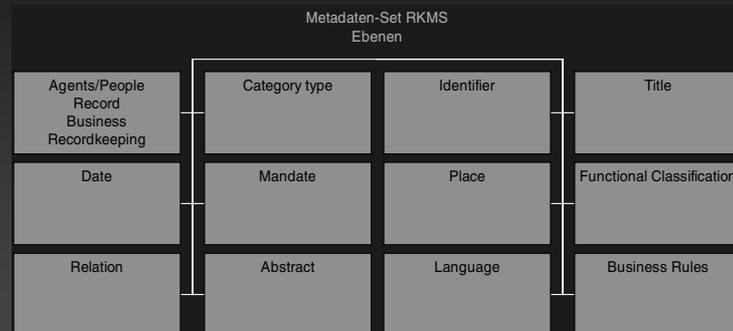
Australische Metadaten-Standards

- 2. Ebene: RKMS (Recordkeeping Metadata Schema)
- Nationaler Standard, komplementär zu AGLS, im Juli 1999 fertiggestellt
- Ziel: Verknüpfung der Geschäftsprozesse mit den bei ihrer Erledigung entstandenen Unterlagen
- Einsatz von Metadaten-Sets auf der Ebene der E-records und auf der Ebene des „Record-System“ (= Schriftgutverwaltungssystem und Archiv)

Australische Metadaten-Standards

- 3. Ebene: Standards auf der Ebene der Bundesstaaten, z.Zt. fertiggestellt:
- VERS: Für den Bundesstaat Victoria
- NRKMS: Für den Bundesstaat New South Wales
- RKMS ist Leitlinie, regionale Ansätze sind detaillierter
- Im Augenblick: Mapping-Initiative

RKMS Metadata-Element-Set:

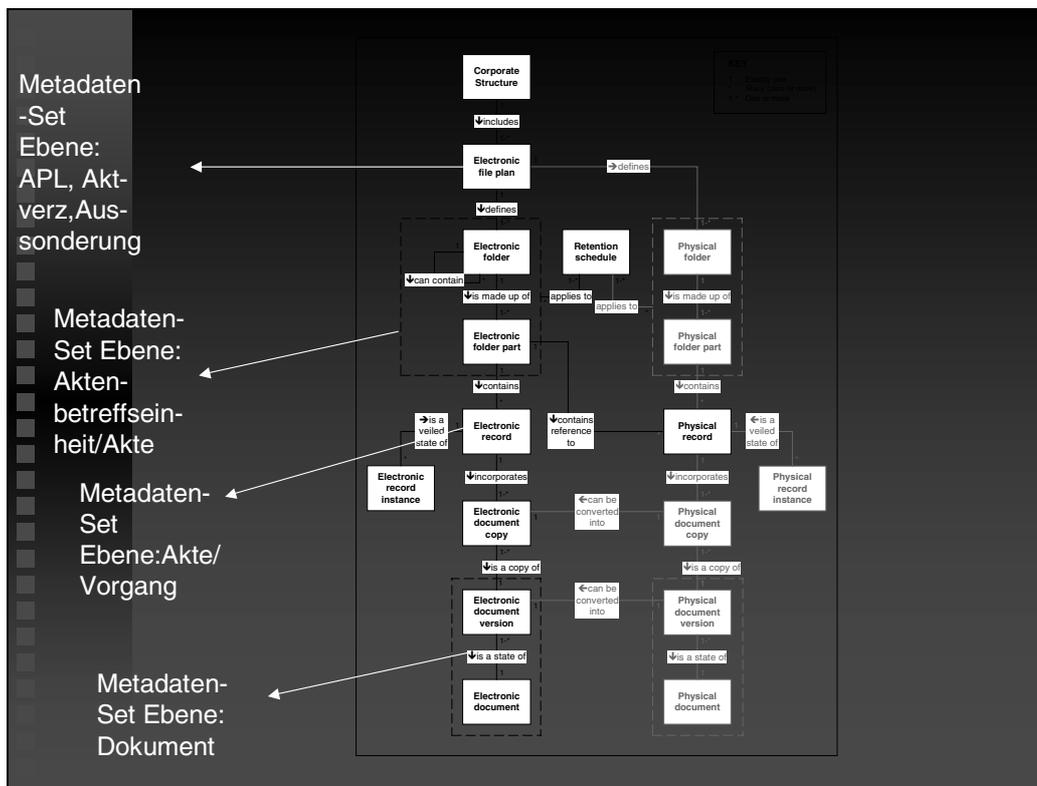


Metadaten-Elemente bei RKMS:

- Ergänzt durch Sub-Elements
- Orientierung an Dublin-Core und:
- ISAD-G/ISAAR
- <http://www.naa.gov.au/>

Aktuelle europäische Entwicklungen:

- Metadaten für den Einsatz im Bereich Schriftgutverwaltung:
 - ◆ EROS: Großbritannien
 - ◆ MOREQ: Funktionsmodell EU
 - ◆ Normentwurf Archivage Electronique
 - ◆ DOMEA
- Metadaten für den Bereich der Archive:
 - ◆ ISAD(G)
 - ◆ EVA (European Visual Archives)



Metadaten und ihre Funktion:

- Schriftgutverwaltung
 - ◆ Management von E-Records
 - ◆ Management von elektronischen Schriftgutverwaltungssystemen
 - ◆ Funktionales Instrument zur Beschreibung von Inhalten, Kontexten und Strukturen
 - ◆ Technisches Instrument zur Protokollierung von Transaktionen
- Archiv
 - ◆ Übernahme der Metadaten zum Transparentmachen der Strukturen und Kontexte im SchrGVerw-System
 - ◆ Eigenständige archivische Metadaten-Sets zu:
 - ◆ Bewertung
 - ◆ Erschließung und
 - ◆ Zugang



Bruebach@mail.uni-marburg.de

Vielen Dank !

Empfehlungen für die elektronische Archivierung. Ein Normentwurf der *Association Française de Normalisation* *

Von Frank M. Bischoff

Am 11. Januar 2001 wurde vom französischen Verband für Normung – *Association Française de Normalisation* (AFNOR) – der Entwurf zu einer nationalen Norm unter der Bezeichnung NF Z 42-013 vorgelegt. Das Normungsvorhaben trägt den Titel „Elektronische Archivierung – Empfehlungen für die Konzeption und den Betrieb von Datenverarbeitungsanlagen zur Erhaltung und unversehrten Sicherung darin gespeicherter Unterlagen“.¹ Bei dem Entwurf handelt es sich um die Erweiterung einer bereits im Juli 1999 veröffentlichten Norm.²

Dem Begleitkommentar zum Entwurf ist zu entnehmen, dass die Anregung zur Revision der bestehenden Norm von Verbänden und Unternehmen im Bereich der Verwaltung elektronischer Unterlagen ausging. Der Normenkommission gehören neben Vertretern der Industrie auch Mitarbeiter der *Bibliothèque Nationale de France* sowie des *Centre Historique des Archives Nationales* an. Ende Januar 2001 wurde der Normentwurf im Rahmen einer Prüfanfrage anderen nationalen Normungsinstitutionen – darunter das Deutsche Institut für Normung e.V. (DIN) – zugeleitet, um diesen Gelegenheit zu geben, innerhalb einer festgesetzten Frist (20. März 2001) Stellung zu nehmen. Nach Abschluss dieser Phase kann die Norm in Frankreich umgesetzt werden.

Ziel des Normungsvorhabens ist es, eine Standardisierung im Bereich technischer und organisatorischer Maßnahmen zu erreichen, die bei der Speicherung, Aufbewahrung und Wiederherstellung von elektronischen Dokumenten zu beachten sind, um dauerhaft den Erhalt und die Integrität digitaler Unterlagen sicherzustellen. Gegenüber der früheren Version sind vor allem neue Definitionen von Dateiformaten, Empfehlungen zur Wiederherstellung bzw. Darstellung von Unterlagen, zu Sicherungs- und Kontrollmaßnahmen und zur Auftragsarchivierung durch Dritte ergänzt worden.

Hintergrund

Eines der wichtigsten Motive für die Änderung der bestehenden Norm dürfte die Anpassung des französischen Beweisrechts an die Möglichkeiten der Informationstechnologie sein. Mit Gesetz vom 13. März 2000 ist die elektronische Unterschrift der handschriftlichen im Beweisrecht gleichgestellt,³ mithin auch der Stellenwert elektronischer Unterlagen generell aufgewertet worden.⁴ Damit sind in Frankreich zugleich die notwendigen gesetzlichen Weichen gestellt, um analoge Unterlagen in großem Stil zu digitalisieren und lediglich die digitale Version für eine Langzeitaufbewahrung vorzusehen, sofern entsprechende Sicher-

* Der Vortrag wurde für die Drucklegung leicht überarbeitet. Die französische Norm zur elektronischen Archivierung ist im Dezember 2001 in Kraft getreten. Die hier referierten Teile des Normentwurfs wurden in der verabschiedeten Fassung der Norm inhaltlich nicht oder nur geringfügig geändert.

¹ Archivage électronique – Recommendations relatives à la conception et à l'exploitation de systèmes informatiques en vue d'assurer la conservation et l'intégrité des documents stockés dans ces systèmes. Kurzinformation: http://normesenligne.afnor.fr/cgi-bin/normesenligne.storefront/1754428609/Ext/Fiche_Produit/PR009348.

² Kurzinformation: <http://normesenligne.afnor.fr/cgi-bin/normesenligne.storefront/1937236963/Ext/FicheProduit/NF041429>

³ Loi n° 2000-230 du 13 mars 2000 portant adaptation du droit de la preuve aux technologies de l'information et relative à la signature électronique http://www.legifrance.gouv.fr/citoyen/jorf_nor.ow?numjo=JUSX9900020L

⁴ Zur Begründung des Gesetzes durch das französische Justizministerium vgl. Exposé des motifs du Projet de Loi portant adaptation du droit de la preuve aux technologies de l'information et relative à la signature électronique, présenté en Conseil des Ministres le 1er septembre 1999, <http://www.justice.gouv.fr/publicat/signmoti.htm>

heitsstandards zur Wahrung von Integrität, Authentizität und Reproduzierbarkeit der digitalen oder digitalisierten Unterlagen erfüllt werden. Ein zentrales – und vielleicht das interessanteste – Anliegen der Norm ist der Erhalt der Rechtssicherheit im Rahmen einer Langzeitaufbewahrung digitaler Unterlagen.

Zum Hintergrund der Norm zählen auch wirtschaftliche Interessen der Industrie. Sind anerkannte Verfahren und Regeln zur Wahrung der Rechtssicherheit von Unterlagen in digitalen Umgebungen definiert, so ist der Weg frei für den Umstieg von Papier- und Mikrofilmspeichern zu elektronischen Speichern. Es geht dabei nicht allein um Kostensenkung bei der Verwaltung jeweils eigener Registraturen, sondern auch um die Ausdehnung einer noch relativ jungen Dienstleistungsbranche im Bereich des Informationsmanagements, die sich mit Auftragsarchivierung befasst und sich über Normen am Markt qualifizieren will. Nur so lässt sich das bereits erwähnte große Interesse von Verbänden und Unternehmen an einer Revision der bestehenden Norm interpretieren. Die Norm soll – so wird in der Einleitung erklärt – in erster Linie der Orientierung von Organisationen und Unternehmen dienen, die große Mengen von Unterlagen in digitalen Speichern sichern. Dabei wird explizit vorausgesetzt, dass die Aufbewahrung auch im Rahmen eines *Outsourcing* durch Dritte erfolgen kann.⁵

Angesichts dieser Zielsetzung lässt sich bereits erahnen, dass eine klare Scheidung des Lebenszyklus' digitaler Unterlagen zwischen (Alt)Registraturgut und Archivgut nicht vorgenommen wird und auch nicht vorgegeben werden soll. Der Normentwurf zielt auf nicht mehr in Bearbeitung befindliche, geschlossene elektronische Unterlagen, die nach deutschem Verständnis sowohl als Altregistratur bzw. Zwischenarchivgut als auch als Archivgut aufgefasst werden können. Es geht um Unterlagen, die nicht mehr ständig benutzt werden, die aber gleichwohl noch längere Zeit aufbewahrt und ggf. recherchiert und reproduziert werden müssen, sei es aufgrund rechtlicher Vorgaben für eine begrenzte Zeit oder im Rahmen einer Archivierung auf Dauer. Die Verwendung des Begriffs „Archivierung“ orientiert sich an dem in der Datenverarbeitung eingeführten Verständnis und weniger an dem archivfachlichen. Wenngleich angenommen werden darf, dass die Norm vor allem eine aus Gründen der Rechtssicherheit notwendige, mehr oder weniger befristete Konservierung und Nutzung von elektronischen Unterlagen regeln möchte, schließt sie ihre Anwendung auf eine dauerhafte Archivierung nicht aus. Insofern stellt die Norm de facto eine klassische Aufgabe von Archiven, nämlich die physische Aufbewahrung und Konservierung von Archivgut, zur Disposition, öffnet das Aufgabenfeld für private Dienstleister, denen zwar technische, jedoch keine archivfachlichen Qualifikationen abverlangt werden.

Inhalt

Der Entwurf ist in elf Abschnitte gegliedert. Eingangs werden neben einer Beschreibung des Anwendungsgebiets der Norm Hinweise auf weitere zu berücksichtigende nationale und internationale Normen sowie Erläuterungen zur verwendeten Terminologie und zu den Grundprinzipien der Norm gegeben. Der umfangreichste fünfte Abschnitt befasst sich mit den technischen Rahmenbedingungen einer elektronischen Archivierung. Es folgen Vorgaben zu den Sicherheits- und Verfahrensanforderungen, zu den Arbeitsabläufen und zu den durchzuführenden Prüfmaßnahmen. Der Entwurf schließt in den Abschnitten 10 und 11 mit Empfehlungen für den Bereich der Auftragsarchivierung.

⁵ Die Beauftragung privater Unternehmen mit der Archivierung digitaler Unterlagen wird bemerkenswerter Weise selbst bei der Direction des Archives de France (*DAF*) als mögliche Lösung für eine langfristige Sicherung von Archivgut in Erwägung gezogen. Vgl. Catherine Dherent, *Les archives électroniques. Manuel pratique*, Paris 2002, S. 61.

Der Gliederung des Entwurfs folgend werden im weiteren vor allem Empfehlungen hervorgehoben, die

den Charakter der Norm verdeutlichen,

von besonderem Interesse im Rahmen einer digitalen Archivierung im archivfachlichen Sinne sein könnten oder

diesbezüglichen Anforderungen geradezu widersprechen.

Anwendungsgebiet und Grundprinzipien

Der Normentwurf gibt Empfehlungen zu technischen und organisatorischen Maßnahmen, die im Zusammenhang mit der Speicherung, Aufbewahrung und Wiederherstellung elektronischer Unterlagen zu beachten sind. Die Maßnahmen sollen die Erhaltung (*conservation*) und die Integrität (*intégrité*) der Unterlagen sicherstellen. Daneben wird auch die Echtheit (*fidélité*) und die Unversehrtheit (*non-altération*) hervorgehoben. Für den Begriff der Echtheit oder Vertrauenswürdigkeit (*fidélité*) eines Dokuments im Vergleich zum Original- oder Ursprungsdokument wird im terminologischen Kapitel des Normentwurfs eine Definition gegeben. Demnach gilt ein Dokument als echt oder vertrauenswürdig, wenn es erlaubt, alle notwendigen Informationen zu dem Gebrauch bzw. Verwendungszweck wiederherzustellen, zu dem das Original- oder Ursprungsdokument geschaffen wurde. Die komplizierte Formulierung dieser Definition beinhaltet Spielräume bei der archivischen Behandlung digitaler Dokumente, über deren Konsequenzen aber wohl erst in der Praxis Klarheit zu gewinnen ist.

Weitere Begriffsklärungen werden von dem Normentwurf leider nicht geboten. Zum Teil drängt sich dem Leser auch der Eindruck auf, dass die Begriffe vereinzelt synonym verwendet werden. Damit bleiben zentrale Teile der Terminologie zu sehr im Vagen. Hier wäre eine Präzisierung und Differenzierung der Anforderungen wünschenswert, wie sie in der ISO-Norm „Schriftgutverwaltung“ realisiert sind. Als unverzichtbare Merkmale von Unterlagen werden dort Authentizität (*authenticity*), Zuverlässigkeit (*reliability*), Integrität (*integrity*) und Benutzbarkeit (*useability*) definiert und eingefordert.⁶ Damit sind die Anforderungen der ISO-Norm 15489-1 zugleich strenger bzw. terminologisch prägnanter gefasst, als jene des französischen Normentwurfs.

Der Normentwurf bezieht sich ausschließlich auf digitale Unterlagen, schließt unter diesen aber gewisse Typen aus, nämlich digitales Ton- und Filmmaterial, animierte Bilder, aber auch Röntgenaufnahmen. Gegenstand der Norm ist also typisches Verwaltungs- und Geschäftsschriftgut in digitaler Form. Allerdings müssen die Unterlagen nicht digital entstanden sein. Es kann sich auch um nachträglich digitalisierte Unterlagen handeln, Unterlagen also, die auf Papiervorlagen oder auf Mikroformvorlagen fußen. Die Norm nimmt an verschiedenen Stellen immer wieder Bezug auf Mikroformvorlagen, also auf mikroverfilmte oder -verfichte Unterlagen. Offenbar versucht man damit Bestrebungen zu unterstützen, bereits verfilmte Registraturen in eine digitale Umgebung zu konvertieren.

Die französische Norm geht von einer online-Speicherung elektronischer Dokumente aus und steht damit im Gegensatz zur Praxis in den meisten öffentlichen Archiven in Europa, in denen die Magazinierung digitaler Unterlagen offline erfolgt.⁷ Vorausgesetzt wird der Einsatz von Robots zur Verwaltung großer Speicher, in denen die Datenträger abgelegt sind. Damit ist ein wenn nicht sekunden- so doch minutenschneller Zugriff auf die Unterlagen gewährleistet, was bei Registraturgut eine sinnvolle Anforderung sein kann, bei Archivgut aber an der Benutzungspraxis vorbeizieht. Mit der online-Speicherung sind zugleich höhere Risiken

⁶ ISO 15489-1, Kap. 7.2 Merkmale von Schriftgut.

⁷ Z.B. Bundesarchiv in Koblenz und Centre des Archives Contemporaines (CAC) in Fontainebleau. Für die Benutzung wird jeweils eine Kopie angefertigt

verbunden, denen die Norm mit verschiedenen technischen und organisatorischen Vorgaben begegnet.

Zu den Grundprinzipien der französischen Norm gehört eine Beschränkung der zulässigen Speichermedien. Die Norm erklärt sich nur für Systeme anwendbar, die im Bereich der Langzeitaufbewahrung mit nur einmal beschreibbaren optischen Speichermedien arbeiten (WORM = Write Once Read Many). Diese optischen Speichermedien müssen bereits herstellerseitig mit einer eindeutigen Identifikationsnummer versehen sein. Jede Form von wiederbeschreibbaren Speichermedien, also magnetische oder magneto-optische Träger, wird ausgeschlossen.

Mit dieser technischen Beschränkung auf einen bestimmten Typus von Datenträgern soll Fälschungssicherheit gewährleistet werden. Einmal auf einem optischen Speichermedium aufgezeichnete Unterlagen können im nachhinein nicht mehr geändert werden, zumindest nicht auf demselben Medium. Allerdings bestehen anderweitige Möglichkeiten der Manipulation. Der Zugriff auf ein Dokument lässt sich nur über ein Referenzierungssystem realisieren. Dieses Referenzierungssysteme kann beispielsweise ein datenbankgestütztes Verzeichnis sein, das über Adressen oder Pointer auf das jeweilige Dokument bzw. die jeweilige Datei verweist. Ohne ein derartiges Referenzierungssystem, ohne Findmittel also, können die Unterlagen nicht genutzt werden. Während die gespeicherten digitalen Unterlagen unverändert bleiben sollen und deshalb auf optischen Datenträgern abgelegt werden können, unterliegen Referenzierungssysteme Veränderungen, im einfachsten Fall durch Fortschreibung. Eine Ablage auf WORMs kann folglich nicht verlangt werden. Damit sind aber faktisch Fälschungsmöglichkeiten durch Manipulation der Adressen bzw. Pointer gegeben. Im Referenzsystem zu einem Dokument kann die Adresse oder der Pointer verändert werden, so dass auf ein anderes, z.B. gefälschtes Dokument verwiesen wird. Ohne entsprechende organisatorische Vorkehrungen wäre eine derartige Fälschung nicht mehr nachweisbar und nachvollziehbar. Der französische Normentwurf versucht diesen Manipulationsmöglichkeiten durch umfassende Anforderungen an eine automatische Protokollierung für alle Vorgänge innerhalb eines elektronischen Archivsystems zu begegnen. Darüber hinaus sollen Operateure verpflichtet werden, den ordnungsgemäßen Verlauf der von ihnen durchgeführten Maßnahmen zu attestieren. Auf einzelne Vorgaben wird später noch einzugehen sein. An dieser Stelle bleibt zunächst nur festzuhalten, dass im Rahmen einer Archivierung digitaler Unterlagen, die die Echtheit, Zuverlässigkeit und Unversehrtheit der Dokumente gewährleisten will, auch besondere Vorkehrungen für die Echtheit, Zuverlässigkeit, Unversehrtheit und Benutzbarkeit der Findmittel getroffen werden müssen.

Zu den Grundprinzipien, die eine ordnungsgemäße Archivierung der elektronischen Unterlagen sicherstellen sollen, gehören auch regelmäßige Überprüfungen des Archivierungssystems und der in diesem Zusammenhang notwendigen Vorgänge und Tätigkeiten. Überhaupt werden Organisationen und Unternehmen, die eine elektronische Archivierung vornehmen wollen, verpflichtet, eine umfassende Sicherheitspolitik zu entwickeln und einzuführen. In diesem Punkt dürften auch viele Archive einen nicht zu unterschätzenden Nachholbedarf haben.

Technische Vorgaben

Dokumentation und Historien

Im Bereich der technischen Vorgaben werden zunächst umfangreiche und detaillierte Dokumentationen eingefordert. Zu dokumentieren sind

alle zum Einsatz gelangenden Soft- und Hardwarekomponenten einschließlich ihrer Funktionsbeschreibungen, darunter auch die Komponenten zur Konvertierung analoger Unterlagen,

alle Datenstrukturen,

alle Interaktionen zwischen Systemkomponenten sowie

alle Arbeitsabläufe.

Darüber hinaus sollen die physischen Bedingungen, also Temperatur und Luftfeuchtigkeit, spezifiziert werden, die der Betrieb des elektronischen Archivierungssystems voraussetzt,

andere technische Rahmenbedingungen, wie Stromversorgung, Feuermeldeanlagen usw.

und schließlich alle sonstigen technischen Sicherungsmaßnahmen, wie elektromagnetischer Schutz, Tresor- und Schließanlagen usw.

Der Normentwurf sieht weiter vor, dass alle Ereignisse innerhalb des Archivierungssystems mit genauer Datums- und Zeitangabe automatisch protokolliert werden. In diesen Historien soll festgehalten werden,

wann, von wem und ggf. an welcher Arbeitsstation ein Ereignis angestoßen wurde,

um welches Ereignis es sich handelt und

zu welchem Resultat das Ereignis geführt hat.

Solche Protokollierungen sind insbesondere auch bei allen Archivierungsvorgängen vorzunehmen. Dabei sind die behandelten Dokumente jeweils eindeutig zu identifizieren. Unter Archivierungsvorgängen sind im wesentlichen das Scannen von Papier- oder Mikroformdokumenten und das Abspeichern von digitalen bzw. digitalisierten Dokumenten auf geeigneten Speichermedien zu verstehen. In diesen Fällen sind neben der Historie von Einzelereignissen auch Journale zu führen, in denen zusammenfassend die Archivierungsereignisse jedes Tages protokolliert werden.

Historien erlauben es, festzustellen, wer auf das System zugegriffen hat, ob die Prozeduren vorschriftsmäßig durchgeführt wurden oder ob es Grund zur Annahme gibt, dass eine Aktion in böswilliger oder nicht autorisierter Absicht durchgeführt wurde. Deshalb nehmen die Protokolle innerhalb der Norm einen hohen Stellenwert im Rahmen der Sicherungsmaßnahmen ein. Die Norm gibt sogar vor, dass auch die Historien in regelmäßigen Abständen auf optischen Speichermedien gesichert und vorgehalten werden müssen. Dabei sind alle Sicherheitsvorkehrungen zu beachten, die auch bei der Speicherung der zu archivierenden Unterlagen greifen. Zur Aufbewahrungsdauer der Historien wird allerdings keine Aussage gemacht. Entspreche die Aufbewahrungsdauer der Historien jener der archivierten Unterlagen, so würde sich das Datenvolumen innerhalb des Archivierungssystems bei einem hohen Anteil dauernd aufzubewahrender Unterlagen im Laufe der Zeit exponentiell vervielfachen, und es müssten spezielle Prozeduren entwickelt werden, um die Historien nutzbar zu machen und darin gezielt nach einzelnen Ereignissen recherchieren zu können. Im Hinblick auf die langfristigen Konsequenzen solcher Vorgaben, die gerade bei einem Archivierungssystem mitbedacht werden sollten, zeigt der Normentwurf Schwächen.

Speicherung und Sicherung der zu archivierenden Unterlagen

Für die Konvertierung analoger und die Speicherung digitaler Unterlagen gibt der Normentwurf eine Reihe von Regeln vor, die hier nicht annähernd erschöpfend aufgezählt werden können. Beispielfhaft seien nur zwei Vorgaben genannt:

- Verlustbehaftete Kompressionsalgorithmen dürfen allenfalls bei Fotomaterial und dort auch nur in einem für das Auge nicht unmittelbar wahrnehmbaren Grad zum Einsatz gelangen. Für Text- und Grafikdokumente, auch Text- und Grafikdokumente in Image-Dateien, dürfen verlustbehaftete Kompressionsalgorithmen nicht eingesetzt werden.

Für digital entstandene Dokumente sind nur bestimmte Formate zugelassen, deren Charakteristikum darin bestehen muss,

- dass die Spezifikationen für ältere Versionen auch noch bei jüngeren Versionen gelten,
- dass diese Spezifikationen veröffentlicht und frei zugänglich sind und
- dass diese Formate geeignet sein müssen, mehrseitige Dokumente mit einer Kombination von Text, Bild und Grafik darzustellen.

Die Vorgaben hinsichtlich zulässiger Formate – darunter Flatfiles, SGML und Edifact, aber auch Auszeichnungsformate, sofern sie bestimmte Bedingungen erfüllen – sind nicht so streng wie etwa im DOMEA-Konzept, aber eng genug, um eine gewisse Dauerhaftigkeit zu garantieren.

Sehr detaillierte Anforderungen werden im Hinblick auf die Speicherung der Unterlagen gestellt. Grundsätzlich darf jedes optische Speichermedium nur sequentiell und ohne Leer-räume beschrieben werden. Ungenutzte Spuren am Ende eines Datenträgers müssen derart modifiziert werden, dass sie nicht mehr beschreibbar sind.

Jeder Speichervorgang muss zu einer weiteren Erfassung von Protokolldaten und deren Abspeicherung auf demselben Datenträger führen. Unterschieden werden datenträgerbezogene Protokolle, die am Anfang und am Ende des Datenträgers, und dateibezogene Protokolle, die jeweils vor der Datei, quasi als Dateikopf gespeichert werden. Nach Abschluss des Speichervorgangs müssen die gespeicherten Dokumente auf Fehlerfreiheit geprüft werden. Alle Datenträger sind eindeutig zu beschriften. Schließlich muss sichergestellt werden, dass das Indexsystem, also das DV-technische Findmittel, die neue Speichereinheit erfasst hat.

Sicherungskopien sind in regelmäßigen Intervallen auf denselben Datenträgern anzufertigen, auf denen die ‚Originale‘ gespeichert sind. Diese Kopien sind in räumlicher Distanz zu den ‚Originalen‘ verschlossen aufzubewahren. Auch für den Fall, dass ein Datenträger nicht mehr lesbar ist, gibt der Normentwurf detaillierte Verfahrensregelungen vor. Protokolldaten sollen wiederum erkennen lassen, dass ein Datenträger durch einen anderen ersetzt wurde. Ein Refreshing auf neue Datenträger ist vor Ablauf der vom Hersteller angegebenen Lebensdauer der alten Datenträger zu gewährleisten.

Indexierung, Recherche und Benutzung

Alle archivierten Unterlagen müssen in einer Weise klassifiziert, identifiziert und indexiert werden, die es erlaubt, gezielt nach einem einzelnen Dokument oder einer Serie von Dokumenten zu recherchieren. Insbesondere muss sichergestellt werden, dass die logische und physische Adressierung einer Datei nicht verloren gehen kann.

Die Wiederherstellung, also die Darstellung eines archivierten Dokuments, muss exakt dem ursprünglichen Original gleichen. Der Normentwurf unterscheidet zwei Arten der Präsentation, nämlich die Anzeige am Bildschirm und die Druckausgabe auf Papier oder Film. Eine Ausgabe auf ein anderes digitales Speichermedium ist offenbar nicht vorgesehen, was den

Ansprüchen heutiger Archivbenutzer nicht mehr gerecht werden dürfte.⁸ Konsequenz ist aber die Forderung, dass im Zuge eines Wiederherstellungsprozesses jede Bearbeitungsmöglichkeit ausgeschlossen sein muss, die zu inhaltlichen Veränderungen an dem Dokument führen könnte. In diesem Zusammenhang gibt die Norm auch vor, dass die Anwendungen zur Wiederherstellung der Dokumente von den Werkzeugen, die zur Digitalisierung benutzt wurden, unabhängig sein müssen. Es müssen also von den Scan- und Bildverarbeitungsprogrammen unabhängige Viewer benutzt werden können.

Der Normentwurf sieht darüber hinaus vor, dass Maßnahmen zur Identifikation der Benutzer und ihrer Zugriffe eingeführt werden können, schreibt diese aber nicht bindend vor. Ebenfalls optional werden Maßnahmen zur Chiffrierung und zur Signierung von digitalen Dokumenten genannt. In diesen Fällen verlangt der Normentwurf wiederum ausführliche Dokumentationen sowie Konformität mit der ISO-Norm 9594, gibt aber ansonsten keine weiteren Hinweise zum langfristigen Handling chiffrierter oder signierter Unterlagen.

Sicherheitspolitik und Sicherheitsmaßnahmen

Die wichtigste und zugleich anspruchsvollste Anforderung zur Gewährleistung der Daten- und Systemsicherheit ist die in dem Normentwurf aufgestellte Verpflichtung für Organisationen und Unternehmen, eine Sicherheitsstruktur zu schaffen, die eine geeignete Durchführung von Sicherheitsmaßnahmen gewährleistet.

Konkrete Sicherheitsmaßnahmen werden für verschiedene Bereiche benannt. So müssen Zugangsrechte zu Räumlichkeiten definiert werden, die sensibles Material enthalten oder in denen sensible Funktionen ausgeführt werden können. Die Lagerorte der optischen Datenträger, Originale wie Sicherungskopien, müssen unter Verschluss gehalten und besonderen Zugangsrechten unterworfen werden.

Selbst das Handling von unbrauchbar gewordenen Datenträgern wird in der Norm durch besondere Sicherheitsvorgaben geregelt. Sie sollen umgehend vernichtet oder aber in einer Weise überschrieben werden, dass eine Rekonstruktion der ursprünglich auf ihnen gespeicherten Informationen unmöglich ist.

Die Sicherheit des Archivierungssystems soll mit den üblichen Maßnahmen der Rechtevergabe, Zugangskontrolle und Funktionsüberwachung gewährleistet werden. In diesem Zusammenhang empfiehlt die Norm die Einführung von Chipkarten. Gegen illegale Informationsübertragung durch elektromagnetische Wellen sind insbesondere die zentralen Rechnerräume mit geeigneten Maßnahmen zu schützen.

Verfahrensregelungen

Der Normentwurf verlangt detaillierte Verfahrensregelungen in allen Bereichen, die in mittelbarem oder unmittelbarem Zusammenhang mit den Anforderungen an die Archivierung elektronischer Unterlagen stehen. Die Verfahrensregelungen sollen insbesondere die Methoden offen legen, auf denen die Arbeitsabläufe in den verschiedenen Aufgabengebieten basieren. Im Einzelnen werden Maßnahmen im Bereich der Digitalisierung von analogem Schriftgut, der Abspeicherung von elektronischen Dokumenten, deren Indexierung sowie der Qualitätskontrolle genannt.

⁸ Das CAC etwa gibt Kopien digitaler Unterlagen auf Diskette oder CD-ROM an Benutzer heraus.

Verfahrenskontrolle, Systemüberwachung und -prüfung

Unter den Verfahrenskontrollmaßnahmen werden vor allem eine Reihe von Bescheinigungen angeführt, die von den jeweils Verantwortlichen für bestimmte Verfahrensregelungen oder durchgeführte Maßnahmen auszustellen sind. Eigentümer von Unterlagen sollen die beauftragte, Operateure die durchgeführte Digitalisierung von analogen Unterlagen bestätigen. Darüber hinaus sind Bescheinigungen über den Erfolg durchgeführter Vernichtungsmaßnahmen zu erstellen, aber natürlich auch über die Übereinstimmung einer Reproduktion mit der Vorlage.

Überwachungs- und Prüffunktionen beziehen sich auf die Eignung der zum Einsatz gelangenden Verfahren und Sicherheitsvorkehrungen. Sie sollen in regelmäßigen Intervallen stattfinden, mindestens aber einmal pro Jahr. Optional können externe Unternehmen beauftragt werden. Sofern es sich um interne Prüfer handelt, sollen diese einer von der zu überprüfenden Organisationseinheit unabhängigen Stelle angehören.

Alle Sicherheitsmaßnahmen, Verfahrensregelungen und -kontrollen müssen innerhalb der Organisation oder des Unternehmens allen Mitarbeitern bekannt gemacht werden.

Auftragsarchivierung und Serviceleistungen Dritter

Über die bisher dargelegten Anforderungen hinaus enthält der Normentwurf auch nähere Bestimmungen zur Auftragsarchivierung durch Dritte bzw. über die Einbindung von gemieteten Serviceleistungen, etwa im Bereich der Digitalisierung analoger Unterlagen. Zum Hintergrund dieser Vorgaben wurden bereits einleitend einige Erläuterungen gegeben.

Die Norm unterscheidet zwei Fälle von Auftragsarchivierung, nämlich

- die vollständige Archivierung elektronischer Dokumente und
- die Archivierung digitaler Signaturen von Dokumenten.

In beiden Fällen verlangt die Norm eine präzise Festlegung der Zuständigkeiten und Leistungen von Auftraggeber und Auftragnehmer. Für den Fall, dass digitale Signaturen eine Rolle spielen, muss der Vertrag zwischen Auftraggeber und Auftragnehmer folgende Mindestvoraussetzungen erfüllen:

Die Funktionen zur Berechnung der den digitalen Signaturen zugrundeliegenden Komprimata müssen präzise benannt werden.

Es muss angegeben werden, von wem die Komprimatbildung durchgeführt wurde.

Sofern chiffrierte Dokumente archiviert werden sollen, gelten dieselben Bedingungen für die Chiffrierschlüssel. Außerdem ist festzulegen, wem die Verantwortung für die Dechiffrierung der Dokumente obliegt.

Sowohl bei der Auftragsarchivierung als auch bei der Vergabe von Serviceaufträgen kann die Notwendigkeit bestehen, einen Datentransfer über öffentliche Netze durchzuführen. In diesen Fällen sind geeignete Authentifizierungs- und Sicherheitsmaßnahmen zu ergreifen.

Resümee

Der vorliegende französische Normentwurf trägt in Teilen den Charakter einer nationalen Spezialanwendung der neuen ISO-Norm 15489 „Schriftgutverwaltung“, steht dieser aber zugleich auch fern, da Anforderungen an eine innere Ordnung und Organisation von Schriftgut praktisch keine Rolle spielen und vollkommen hinter technische Aspekte zurücktreten. Welche Anforderungen erfüllt sein müssen, damit digitales Verwaltungsschriftgut aussagekräftig ist, welche Metadaten benötigt werden, um die einzelnen Dokumente einer Schriftgutverwaltung oder eines digitalen Archivs mit Kontext zu füllen, und in welche Strukturen Dokumente eingebunden werden sollten, damit sie und die Vorgänge, innerhalb derer sie entstanden sind, verstanden werden können, hat die Autoren des französischen Normentwurfs nicht interessiert. Dem Entwurf geht es einzig und allein darum, die Echtheit, Unversehrtheit und Benutzbarkeit von digitalen Dokumenten durch technische und organisatorische Maßnahmen zu gewährleisten. Solange es nur um die befristete Sicherung von massenhaft gleichförmigen Unterlagen, wie Versicherungspolizen, Bausparverträgen oder ähnlichem mehr geht, mag der gewählte Ansatz ausreichen. Spätestens bei komplexer aufgebautem Schriftgut dürfte die Norm aber zu kurz greifen. Insofern hätte man sich wenigstens einen knappen Hinweis auf die mögliche Bedeutung von Metadaten gewünscht.

Ungeachtet dessen umfasst der Entwurf Vorgaben, die auch im Hinblick auf eine digitale Archivierung im Sinne einer dauerhaften Aufbewahrung in öffentlichen oder sonstigen Archiven kritisch geprüft werden sollten. Zu den wichtigsten Maßnahmen zählt die eindeutige, detaillierte und schriftlich gefasste Regelung der Verantwortlichkeiten, Zuständigkeiten und Verfahrensabläufe unter expliziter Benennung aller zum Einsatz gelangender Hilfsmittel sowie aller technischen und physischen Rahmenbedingungen, unter denen eine Archivierung digitaler Unterlagen stattfinden soll. Hier liefert die französische Norm viele Anregungen, die von Archiven aufgegriffen und geprüft werden sollten. Ob dabei Ereignisprotokolle in dem Umfang notwendig sind, wie sie in der Norm verlangt werden, wie lange welche Historien aufbewahrt werden müssen und wie sie auf Dauer nutzbar sind oder sein sollen, mag hier dahingestellt bleiben. Solche Fragen relativieren sich vielleicht auch aufgrund der Tatsache, dass Archive im Rahmen einer Archivierung digitaler Unterlagen aus ökonomischen Erwägungen heraus wohl noch auf absehbare Zeit ohnehin zu der Option einer offline-Magazinierung tendieren werden, die – unter der Voraussetzung, dass geeignete Sicherungsregelungen und -verfahren definiert sind – geringere Gefahren für die Authentizität und die Integrität von elektronischen Unterlagen in sich birgt.

In diesem Zusammenhang sei die Frage des archivischen Umgangs mit digitalen Signaturen gestreift. Der Autor des vorliegenden Beitrags hat vor drei Jahren die Auffassung vertreten, dass digitale Signaturen nicht archivfähig sind, und vertritt sie immer noch aus Gründen, die in dem Ludwigsburger Tagungsband des Arbeitskreises „Archivierung von Unterlagen aus digitalen Systemen“ dargelegt sind.⁹ Der französische Normentwurf dagegen setzt die Archivierung elektronischer Unterschriften selbstverständlich voraus – wie übrigens auch die gleichfalls abzulehnende Archivierung chiffrierter Unterlagen –, äußert sich aber nicht zu den Problemen der Nachsignierung und der langfristigen Nutzung digitaler Signaturen, sondern überlässt diese Frage letztlich den bilateralen Übereinkünften zwischen Auftraggeber und Auftragnehmer im Falle einer Auftragsarchivierung. Die Norm befasst sich also auch hier nicht mit der inneren Logik der – ggf. dauerhaft – zu archivierenden Unterlagen.

Verzichtet man auf eine Archivierung und Kurrenthaltung digitaler Signaturen, so stellt sich das Problem, wie die Authentizität und die Integrität der signierten elektronischen Dokumente gewährleistet werden kann. Udo Schäfer hat in dem bereits zitierten Ludwigsburger

⁹ Frank M. Bischoff, Zur Archivfähigkeit digitaler Signaturen in elektronischen Registern. In: Udo Schäfer – Nicole Bickhoff (Hrsg.), Archivierung elektronischer Unterlagen (Werkhefte der Staatlichen Archivverwaltung Baden-Württemberg, Serie A, Landesarchivdirektion 13), Stuttgart 1999, S. 183–198

Tagungsband dargelegt, dass entsprechende Verfahrensregelungen im Archiv eine digitale Signatur ersetzen können. Er hat dabei die folgenden vier Bedingungen genannt:¹⁰

1. Die Aufzeichnungen sind offline zu speichern.
2. Jede Aufzeichnung ist auf zwei verschiedenen Trägern unterschiedlichen Typs zu speichern.
3. Der Zugriff auf die Speichermedien ist auf bestimmte Personen zu beschränken und zu protokollieren.
4. Für die Benutzung sind die Aufzeichnungen auf einem dritten Träger bereitzustellen, auf den sie zu diesem Zweck kopiert worden sind.

Vor dem Hintergrund der Regelungen des französischen Normentwurfs sollte darüber nachgedacht werden, die zweite von Schäfer genannte Bedingung zu modifizieren und eine fünfte Bedingung hinzuzufügen:

2. Als Datenträger sind nur einmal beschreibbare optische Trägermedien zu verwenden. Die Datenträger sind sequentiell zu beschreiben. Jeder Datenträger ist mit geeigneten Anfangs- und Schlussprotokollen sowie mit geeigneten Kopfprotokollen für jede auf dem Träger gespeicherte Datei zu versehen. Ggf. ungenutzte Spuren auf dem Datenträger sind in geeigneter Weise unbrauchbar zu machen. Eine zweite Sicherungskopie ist in gleicher Weise auf einem Datenträger anzufertigen, der den genannten Anforderungen entspricht.
5. Es sind geeignete organisatorische und technische Maßnahmen zu definieren und zu realisieren, die sicher stellen, dass die Findmittel nicht manipuliert werden können und die eindeutige Zuordnung zwischen Datenträger bzw. Datei und Findmittel gewährleistet ist.

Das Beispiel des französischen Normentwurfs zur Archivierung elektronischer Unterlagen bietet schließlich Gelegenheit, den möglichen Stellenwert nationaler oder internationaler Normungsprojekte auf diesem Gebiet für öffentliche Archive zu diskutieren. Anders ausgedrückt: Welchen Profit können gerade die staatlichen Archive in Deutschland aus nationalen oder internationalen Normen ziehen? Sieht man von den immerhin erheblichen Anregungen und Denkanstößen ab, die eine Mitarbeit in Normungsprojekten bzw. die Auseinandersetzung mit bestehenden Normen bringt, fällt die Antwort sehr nüchtern aus. Es ist fraglich, ob nationale oder internationale Normen – soweit sie über den engeren Bereich des öffentlichen Archivwesens hinauszielen – den staatlichen Archiven befriedigende Regelungen für ihre Aufgabenerfüllung liefern können. Dafür ist deren archivgesetzlich verankerte Stellung zu privilegiert. Bereits die eindeutige Unterscheidung von Registraturgut bzw. Zwischenarchivgut einerseits und Archivgut andererseits, die im privatrechtlichen Bereich gar nicht durchsetzbar wäre, kann in Normen kaum umgesetzt werden. Damit soll keineswegs angedeutet werden, dass eine Beteiligung an Normungsprojekten für deutsche Staatsarchive nicht sinnvoll ist. Im Gegenteil: Die Bestrebungen sollten weiter reichen und über die Beteiligung bei der Normgebung hinaus zu spezifischen Empfehlungen im Sinne von *Best-Practice*-Handreichungen für (öffentliche) Archive führen. Gefragt ist hier eine archivübergreifende Kooperation, die in entsprechenden Richtlinien für staatliche Archive münden könnte.

¹⁰ Udo Schäfer, Authentizität. Vom Siegel zur digitalen Signatur. In: Archivierung elektronischer Unterlagen, (wie Anm. 9) S. 165–181, hier S. 179.

Archivierung von digitalen Dokumenten in Bibliotheken¹

Von Marianne Dörr

Einführung

Archivbibliotheken wie Die Deutsche Bibliothek – oder auf Landesebene die Bayerische Staatsbibliothek sehen sich wie die Archive zunehmend mit Aufgabe der Verwaltung und der langfristigen Vorhaltung bzw. Verfügbarmachung von digitalen Dokumenten konfrontiert. Dabei stellt sich die rechtliche Situation für die bibliothekarische Archivierungspflicht digitaler Dokumente derzeit in Deutschland (auf Bundes- und auf Länderebene) folgendermaßen dar: Publikationen auf Datenträgern unterliegen bereits der gesetzlichen Ablieferung in Analogie zu Printpublikationen. Bei den Netzpublikationen ist die Frage der Pflichtablieferung in Deutschland, wie übrigens auch in der Mehrheit der anderen europäischen Staaten, noch nicht gesetzlich geregelt. Die Deutsche Bibliothek hat ein Projekt „Elektronische Depotbibliothek“ ins Leben gerufen, in dem sie mit Verlagen wie dem Springer Wissenschaftsverlag u.a. praktische Modalitäten eines Transfers und einer Speicherung elektronischer Netzpublikationen erprobt, um damit Erfahrungen zu sammeln, die in eine Gesetzesnovellierung einfließen können.²

Die Spanne der elektronischen Publikationen, mit denen Bibliotheken umgehen müssen, ist sehr breit. Als Beispiele seien genannt:

(Verlags-)Publikationen auf Datenträgern³

- bibliographische oder Volltext-Datenbanken,
- Ratgeber, z.B. Hilfe bei der Steuererklärung,
- Demoversionen von Programmen bzw. Einführungen in die Nutzung von Word, Excel etc.,
- Spiele, Faktensammlungen ...

Netzpublikationen

- Online-Datenbanken,
- elektronische Zeitschriften,
- elektronische Dissertationen und Hochschulschriften,
- Web-Ressourcen / thematische Web-Sites,
- retrodigitalisierte Dokumente

Für die wissenschaftliche Publikation und Dokumentation stehen innerhalb der elektronischen Publikationen die elektronischen Zeitschriften im Zentrum des Interesses – und der Nutzung. Diese gelangen physisch aber schon gar nicht mehr in die Bibliotheken, die ihren Nutzern nur den Zugang über den Verlagsserver anbieten können. Partiiell werden zur Absicherung bereits Klauseln in die Lizenzabkommen aufgenommen, die der Bibliothek bei Kündigung des Lizenzvertrags Sicherheitskopien für die „abbonnierte Laufzeit“ garantieren sollen.

¹ Zum Thema Digitale Archivierung ist in den letzten Jahren eine fast nicht mehr überschaubare Menge an Publikationen erschienen, die größtenteils online zugänglich sind. Hier sei nur auf zwei grundlegende Publikationen verwiesen, die ihrerseits zahlreiche Referenzen enthalten: Preservation Management of Digital Materials. A Handbook: Neil Beagrie – Maggie Jones <http://www.jisc.ac.uk/dner/preservation/workbook/> und RLG Digital Archives Attributes Working Group: Trusted Digital Repositories: Attributes and Responsibilities <http://www.rlg.org/longterm/repositories.pdf>

² Informationen zum Umgang der Deutschen Bibliothek mit Netzpublikationen unter: http://deposit.ddb.de/netzpub/web_netzpublikationen_root.htm

³ Als Beispiel für die Menge der Publikationen seien die Zugangszahlen der Bayerischen Staatsbibliothek genannt. Der Zugang an Datenträgern ist von der Pflichtablieferung dominiert. Bei den CD-ROMs beispielsweise kamen 2000 1.600 Stück (1999: 1.466) über die Zugangsart Pflichtablieferung in die Bibliothek; nur 269 (1999: 181) wurden gekauft.

Auf der anderen Seite gibt es partiell Absichtserklärungen großer Verlage, dauerhaft auch die Archivierung zu übernehmen. Entsprechende Verlautbarungen gibt es vom Springer-Verlag, der andererseits im Projekt der Deutschen Bibliothek *Elektronische Depotbibliothek* kooperiert. Auch die American Society of Physics als Zeitschriften verlegende Fachgesellschaft hat zugesagt, die langfristige Archivierung ihrer Zeitschriften selbst zu übernehmen. Andere Verlage, wie Elsevier, verhandeln bereits mit großen Bibliotheken, die später nationale Archivserver für die Verlagspublikationen werden sollen. Die Landschaft ist auch hier uneinheitlich.

Noch erscheint die Mehrheit der elektronischen Zeitschriften parallel zu einer weiterexistierenden Printausgabe. Von den in der Elektronischen Zeitschriftenbibliothek Regensburg⁴ nachgewiesenen rd. 8.500 Zeitschriften (Stand Anfang 2001) werden „nur“ 900 ausschließlich elektronisch publiziert. Diese Zahl kann jedoch schnell ansteigen. Außerdem nehmen Unterschiede zwischen den gedruckten und elektronischen Versionen einer Zeitschrift zu, z.B. durch die Einbindung von Applets u.Ä. in die digitale Ausgabe.

Einen weiteren zunehmend bedeutenden Posten bei den Netzpublikationen machen die Online-Dissertationen aus. Im Zuge der an den Universitäten laufenden Umgestaltung der Promotionsordnungen werden immer mehr Dissertationen online publiziert. In einem kooperativen Projekt⁵ wurden Maßgaben für eine Abfassung und XML-Konvertierung von Dissertationen entwickelt, die auch den Ansprüchen der Langzeitarchivierung entgegenkämen. Bei der Mehrheit der Universitäten und Universitätsbibliotheken kann die zusätzliche Arbeit einer Konversion von Word-, LaTeX und anderen Formaten in das „neutrale“ XML jedoch nicht geleistet werden. Die Deutsche Bibliothek hat angeboten, für Online-Dissertationen als Depotbibliothek zu fungieren und entsprechende Transfer-Optionen und Metadatenchnittstellen eingerichtet. Während im Jahr 1998 der Zugang an Online-Dissertationen bei DDB noch bei 100 lag, waren es 1999 bereits 470 und im Jahre 2000 1380. Die große Mehrheit davon liegt als PDF-Files, nicht in XML-Codierung vor.

Bei den anderen Netzpublikationen, die Bibliothekare in Analogie zum Printbereich oft als „graue Publikationen“ bezeichnen, herrscht eine bunte Breite vor: Qualität, Formate, Präsentation etc. sind ausgesprochen heterogen. Außerdem entzieht sich die wachsende Masse dieser Web-Sites einer systematischen Erschließung und Sammlung.

Versuche des automatischen Einsammelns („harvesten“) von Web-Publikationen bestimmter Domainbereiche, die im Rahmen des Nedlib-Projekts⁶ europäischer Nationalbibliotheken unternommen wurden, waren nicht sehr erfolgreich: Es gibt bereits zu viele dynamisch erzeugte Websites, die sich über Harvester nicht mehr „einfangen“ lassen.

Für den explodierenden Bereich der Web-Sites und Web-Publikationen müssen neue Selektionskriterien für die Archivierung bzw. Archivwürdigkeit entwickelt werden. Veröffentlichte Ansätze hierzu aus nationalbibliothekarischer Sicht gibt es international bereits in Australien, Finnland und einigen anderen Staaten.⁷

Bibliotheken wie die Bayerische Staatsbibliothek dürfen die Frage der Sammlung und Archivierung von Web-Publikationen aber nicht nur aus der Sicht der nationalen oder regionalen Überlieferungsbildung und Archivierungspflicht sehen, sondern müssen auch den fachlichen Aspekt berücksichtigen, der aus ihrer Verantwortung für die Literatur- bzw. Informationsversorgung in bestimmten Disziplinen im Rahmen des fachlichen Verteilungsplans der Deutschen Forschungsgemeinschaft⁸ resultiert. Hier hat die BSB, wie andere SSG-Bibliotheken

⁴ <http://rzblx1.uni-regensburg.de/ezeit/fl.phtml?bibid=UBR>

⁵ <http://www.dissonline.de/>

⁶ <http://www.kb.nl/coop/nedlib/>

⁷ Vgl. für Finnland: <http://www.lib.helsinki.fi/eva/english.html> und für Australien <http://pandora.nla.gov.au/selectionguidelines.html>

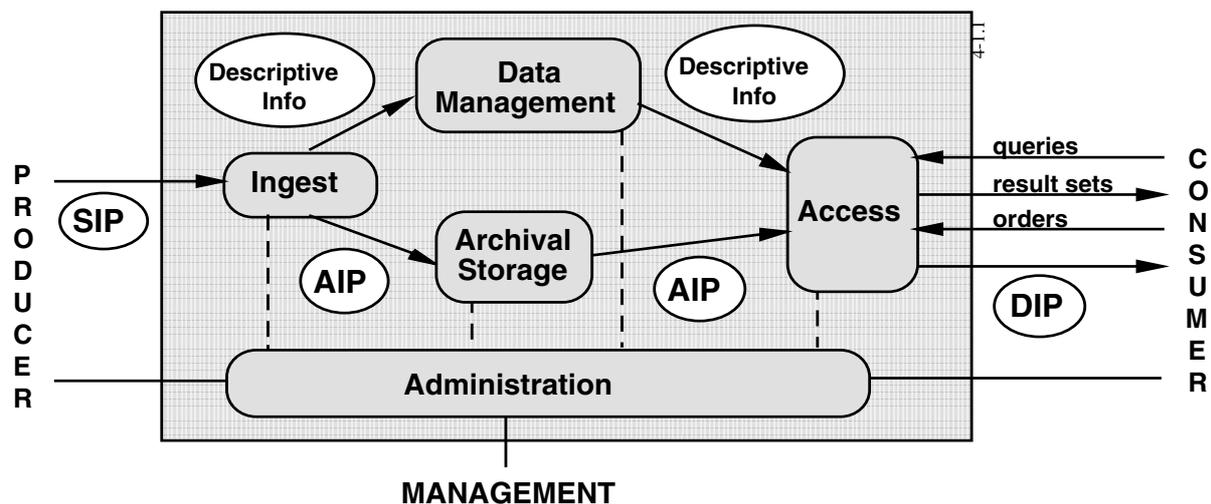
⁸ Welche Bibliothek für welches Fach zuständig ist, kann leicht über das Online-System Webis ermittelt werden, das mit Unterstützung der DFG aufgebaut wurde: <http://webis.sub.uni-hamburg.de/>

auch, in einem ihrer Sondersammelgebiete, nämlich der Geschichtswissenschaft, mit der datenbankgestützten Erschließung wissenschaftlicher digitaler Ressourcen⁹ begonnen. Es gibt aber derzeit noch kein Verfahren bzw. tragfähiges Konzept zu deren auch physischer Speicherung und langfristiger Zugangssicherung.

OAIS als Modell für ein Archivierungssystem

Noch an keiner Bibliothek existiert ein Depotsystem für elektronische Publikationen, das allen vielfältigen Aspekten einer Langfristarchivierung gerecht wird. Allerdings gibt es inzwischen einen relativ breiten Konsens über ein Referenzmodell für die Konzeption und Implementierung eines solchen Systems mit seinen Funktionalitäten. Dies ist das vom Consultative Committee for Space Data Systems (CCSDS) entworfene OAIS (Open Archival Information System)-Modell¹⁰. Das OAIS-Modell will die Funktionalitäten, die für eine langfristige Archivierung von Informationen notwendig sind, in ihrem Zusammenspiel als Referenzmodell abbilden. Damit wird auch erreicht, dass eine Verständigungsbasis zwischen Institutionen, die mit dieser Aufgabe befasst sind oder befasst sein werden, gegeben ist.

Die folgende Graphik stellt nur die oberste Ebene der Funktionalitäten dar. Für alle Prozesse existieren detaillierte Feinmodelle.



(OAIS-Funktionalitätenmodell: <http://www.ccsds.org/documents/pdf/CCSDS-650.0-R-1.pdf>)

Bei der Auseinandersetzung mit dem Modell im Rahmen des bereits erwähnten Nedlib-Projekts europäischer Bibliotheken wurde moniert, dass keine Verbindungen bzw. Schnittstellen zu existierenden bibliothekarischen Arbeits- und Datenverwaltungssystemen vorgesehen sind. Außerdem wurden Erweiterungen vorgeschlagen, die die speziellen Vorgänge bzw. Erhaltungsmaßnahmen (z.B. Migration, Emulation) betreffen, die für eine Langzeitarchivierung von Daten notwendig sind. Die im Modell genannte Funktionalität „archival storage“ umfasst nur den Komplex des Backups i.S. von Datensicherung und -speicherung. Der darüber hinausgehende Bereich der Erhaltungsmaßnahmen soll unter „preservation“ gefasst werden. Die entsprechenden Modifizierungen sind in die schriftliche Fassung einer Funktionalitätenbeschreibung für ein Depotsystem elektronischer Publikationen eingegangen, die von der Koninklijke Bibliotheek in Den Haag erstellt wurde.¹¹

Bisher haben zwei Bibliotheken – nach entsprechenden Ausschreibungsverfahren – Aufträge zur Entwicklung und Implementierung eines Depot- bzw. Archivierungssystems, das auf dem

⁹ <http://www.bsb-muenchen.de/fachinfo/gesch/infoW.htm> <http://webis.sub.uni-hamburg.de/>

¹⁰ <http://www.ccsds.org/documents/pdf/CCSDS-650.0-R-2.pdf>

¹¹ <http://www.kb.nl/coop/nedlib/meetings/paris/GEN-264.pdf>

OAIS-Modell beruhen soll, erteilt. Es handelt sich um die Koninklijke Bibliotheek und um die British Library. Bei beiden ist die Firma IBM zum Zuge gekommen.

Im OAIS-Modell wird die Bedeutung von Information über die abzuspeichernden Daten schon darin deutlich, dass jeder Prozess von der Erstellung von „information packages“, also von Metadaten, begleitet wird. Elektronische Publikationen stellen besondere Anforderungen an die Verzeichnung. Nicht mehr nur bibliographische, sondern auch technische und administrative Metadaten müssen erfasst werden. Für die Verzeichnung in deutschen Bibliothekskatalogen sind die Regeln für alphabetische Katalogisierung in wissenschaftlichen Bibliotheken (RAK-WB) um Sonderbestimmungen für AV und – in einem zweiten Schritt – für elektronische Medien ergänzt worden (RAK-NBM „Non Book Materials“). Ebenso wurde das Austauschformat MAB (Maschinelles Austauschformat für Bibliotheken) um Kategorien für den Transport technischer Informationen (650–659, z.T. mit Unterfeldern) erweitert. Dies ist jedoch noch nicht ausreichend:

Für einen zielgerichteten Zugriff müssten Feldinhalte bzw. die Syntax der Eintragungen normiert werden. In der Regel sind diese Kategorien außerdem in den großen Katalogsystemen nicht indiziert, d.h. nicht dezidiert such- und zugriffsfähig. Die verzeichneten Informationen sind vorwiegend darauf ausgerichtet, einem Benutzer mitzuteilen, was er zu einer Nutzung der entsprechenden elektronischen Publikationen benötigt. Der Aspekt Verwaltung der Daten steht demgegenüber im Hintergrund.

Im internationalen Rahmen entstand inzwischen eine Reihe von Ansätzen zur Entwicklung ausgefeilterer Metadatenformate für elektronische Publikationen. Diese wurden in einem „White Paper“ von der Arbeitsgruppe der beiden großen amerikanischen Bibliotheksverbände OCLC und RLG zusammengestellt.¹² Ein gemeinsames anerkanntes und breiter angewendetes Format existiert indessen noch nicht.

Vorherrschende Ziele von Metadaten-Verzeichnung im Kontext der Bestandserhaltung bzw. Archivierung sind:

- Technische Informationen sammeln und dokumentieren, die Bestandserhaltungsentscheidungen bzw. -maßnahmen unterstützen
- Dokumentation von erfolgten Bestandserhaltungsaktivitäten
- Garantie der Authentizität der digitalen Objekte geben
- Rechte am digitalen Objekt dokumentieren
- Sicherung der Nutzbarkeit des digitalen Objekts

In Deutschland gibt es innerhalb des vom BMBF geförderten CARMEN-Projekts Arbeitspakete¹³, die sich direkt mit dem Thema Metadaten für Rechte und Archivierung beschäftigen – konkrete, i.S. von anwendbaren Ergebnissen resultierten hieraus jedoch noch nicht.

Die Mehrheit der genannten Projekte ist allgemein auf digitale Daten ausgerichtet. Es stellt sich jedoch die Frage, ob nicht noch speziellere Metadaten-Sets für bestimmte Datentypen bzw. Anwendungen definiert werden müssen. Dies gilt besonders für den ersten genannten Aspekt, die technischen Metadaten. Ansätze hierfür gibt es z.B. im Bereich der Stillbilder. Dies ist zwar nur ein kleiner Bereich der digitalen Daten; andererseits aber von zunehmender Bedeutung, da Bildagenturen, Kulturinstitutionen und Privatleute zunehmend digitale Bilder erzeugen und archivieren. Vielleicht gerade deswegen sind fast parallel Initiativen der Industrie und der genannten Kulturinstitutionen (angestoßen von der bereits genannten OCLC/RLG-Arbeitsgruppe) entstanden.

¹² http://www.oclc.org/digitalpreservation/presmeta_wp.pdf

¹³ <http://www.sub.uni-goettingen.de/carmen/>

Als Beispiel für Standardisierungsansätze sollen diese beiden kurz vorgestellt werden.

Metadaten für Still-Bilder

NISO Draft Technical Metadata for Digital Still Images¹⁴

DIG 35 Metadaten-Initiative¹⁵ der Digital Imaging Group (Adobe, Canon, Eastman Kodak, Fuji, Hewlett-Packard, IBM, Intel)
<http://www.digitalimaging.org/>

Das Ziel beider Initiativen liegt in der Schaffung einer dateiformat-unabhängigen Metadaten-Definition und eines Austauschformats für Informationen über Still-Bilder und für die Verwaltung von Bildern.

Folgende „Designprinzipien“ werden definiert:

- Aufsetzen auf existierenden Standards
- Erweiterbarkeit und Skalierbarkeit
- Konsistenz
- Internet-Fähigkeit (deshalb XML-Implementierung angestrebt)
- Fokus auf mittel-/langfristigen Perspektiven
- Erhaltung der Information – nicht der Daten

Beide Entwürfe beinhalten Kategorien zur Beschreibung der „Basic Image Parameter, der „Image Creation“ und der „History“. DIG35 umfasst weitergehend auch Metadaten zur Bildinhaltsbeschreibung (Content Description) und zu den Bildrechten (Intellectual Property Rights (IPR)).

Der NISO-Ansatz geht in den berücksichtigten Bereichen teilweise über DIG35 hinaus:

Im einzelnen werden behandelt:

- Basic Image Parameters (Format, Compression, Photometric Interpretation, Segments ...)
- File (Image Identifier, File Size, Checksum, Orientation, Display Orientation, Preferred Presentation ...)
- Image Creation (Source Type, Scanning Agency, HostComputer, Scanning System, Camera Capture Settings ...)
- Imaging Performance Assessment (Spatial Metrics – Image Width, Length ..., Color Map, TargetData – TargetType, – ID, Profiles ...)
- Change History (Image Processing-Software, Actions, Previous Image Metadata)

Inzwischen haben sich einige Institutionen zur testweisen Implementierung des NISO-Dictionaries bereit erklärt. Von den dortigen Erfahrungen wird es auch abhängen, ob das Dictionary das endgültige Standardisierungsverfahren durchlaufen wird.

Die Einsicht, dass für elektronische Publikationen ein Mehr an Metadaten notwendig ist, um alle Informationen für eine langfristige Verwaltung und Pflege zugriffsfähig zu haben, ist weit verbreitet. Ein Hauptproblem der Metadatenerhebung und Verzeichnung für elektronische Publikationen wird jedoch der Aufwand sein, der für die Institutionen damit verbunden ist. Wenn nicht Verfahren zu einer weitgehend automatisierten Metadatenverzeichnung entwickelt werden, die z.B. softwaregesteuert parallel zur Erzeugung der Daten geschieht, wird die Durchsetzung von ausführlichen Metadaten-Formaten allein aufgrund fehlender Kapazitäten in den erzeugenden und verwaltenden Institutionen schwierig bis unmöglich sein.

¹⁴ NISO (National Information Standards Organization) ist eine non-profit association, die Standardisierungsnotwendigkeiten im Bereich der Informations- und Kommunikationssysteme identifiziert und mit den jeweiligen involvierten Institutionen in die Wege leitet. Der Entwurf des technischen Dictionaries ist zugänglich unter: http://www.niso.org/standards/resources/Z39_87_trial_use.pdf

¹⁵ <http://www.bgbm.org/TDWG/acc/Documents/DIG35-v1.0-Sept00.pdf>

XML – ein Format zur dauerhaften Aufbewahrung und Nutzung einfacher, relationaler Datenbanken?

Von Björn Dehms, Andreas Engel und Ulrich Meyer

Die wachsende Verbreitung elektronischer Aufzeichnungen erfordert die Entwicklung von organisatorisch-technischen Verfahren, mit denen auch diese Unterlagen ausgesondert, dauerhaft aufbewahrt, gesichert und genutzt werden können. Während für elektronische Akten bereits Verfahren entwickelt wurden,¹ existieren zur dauerhaften Aufbewahrung, Sicherung und Nutzung von Datenbanken bislang noch keine neuen Ansätze, die über die klassische Flat-file-Archivierung hinausgehen.² Dabei scheint die Sicherung von Datenbank-Informationen zeitlich gesehen das drängendere Problem zu sein. Datenbanken sind bereits seit vielen Jahren im Einsatz, sie werden in der Regel kontinuierlich fortgeführt und aktualisiert, so dass gespeicherte Informationen schnell verloren gehen, wenn sie nicht rechtzeitig gesichert werden.

An der Forschungsstelle für Verwaltungsinformatik der Universität Koblenz wurde deshalb mit Untersuchungen begonnen, wie auch Datenbanken auf Dauer gesichert, aufbewahrt und genutzt werden können. In einem ersten Schritt wurde daher im Rahmen einer Diplomarbeit ein softwaretechnisches Verfahren zur dauerhaften elektronischen Aussonderung, Aufbewahrung, Sicherung und Nutzung *einfacher, relationaler Datenbanken unter Verwendung von XML als Archivierungs- und Nutzungsformat* entwickelt. Als Kooperationspartner für die Entwicklung der Verfahrenslösung wirkte das Bundesarchiv mit,³ aus dessen Bestand eine Datei über „Grenzzwischenfälle an den Grenzen der DDR“ bereitgestellt wurde. Das im Rahmen der Studie entwickelte Migrationsprogramm ist jedoch von der Beispielanwendung unabhängig und kann auch für die Migration beliebiger tabellenartiger Datenbanken eingesetzt werden, sofern sie nur als ASCII-Flat-file aufbereitet werden können. Im vorliegenden Beitrag wird das entwickelte Migrationsprogramm vorgestellt und eine Bewertung von XML als Archivierungs- und Nutzungsformat vorgenommen.

1. Übersicht über das entwickelte Verfahren zur Migration von einfachen, relationalen Datenbanken nach XML

Der Migrationsprozess einer einfachen, relationalen Datenbank nach XML kann aus softwaretechnischer Sicht in zwei Phasen eingeteilt werden, die Vorbereitungs- und die eigentliche Migrationsphase, wobei die Phasenübergänge durch die jeweils bearbeiteten Dateien beschrieben werden können. Die am Migrationsprozess beteiligten Dateien lassen sich bezüglich ihrer Funktion in Eingabedateien, Zwischendateien und Ausgabedateien gruppieren. Die folgende Abbildung gibt einen Überblick über den Ablauf des Migrationsprozesses.

¹ Vgl. Koordinierungs- und Beratungsstelle der Bundesregierung für Informationstechnik in der Bundesverwaltung (KBSt) (Hrsg.), Konzept zur Aussonderung elektronischer Akten, Bonn 1998 (Schriftenreihe der KBSt 40). – Engel, Andreas, Elektronische Archivierung und Nutzung (ELAN) – Eine Projektidee zur Entwicklung von Verfahren für die dauerhafte Archivierung und Nutzung von elektronischen Geschäftsunterlagen aus Behörden und Betrieben. In: Ulrich Nieß (Hrsg.), Auf der Suche nach archivischen Lösungsstrategien im digitalen Zeitalter. Beiträge zur 4. Jahrestagung des Arbeitskreises „Archivierung von Unterlagen aus digitalen Systemen“ im Stadtarchiv Mannheim, 10.–11.4.2000 (Sonderveröffentlichungen des Stadtarchivs Mannheim 26), Mannheim 2001, S. 73–84.

² Michael Wettengel, Archivierung digitaler Datenbestände aus der DDR nach der Wiedervereinigung. In: Udo Schäfer – Nicole Bickhoff (Hrsg.), Archivierung elektronischer Unterlagen, Stuttgart 1999, S. 223–239

³ Seitens des Bundesarchivs waren Frau Bettina Martin-Weber, Ulf Rathje und Dr. Michael Wettengel beteiligt.

Ausgangspunkt der Migration ist eine Datenbank, die in der Regel in einem proprietären Datenbankspeicherformat vorliegt. Damit das entwickelte Migrationsprogramm möglichst unabhängig von bestimmten Datenbankspeichersystemen eingesetzt werden kann, wird vorausgesetzt, dass die Datenbankinformationen als ASCII-Flat-file exportiert wurden, was von den gebräuchlichen Datenbankmanagement-Systemen durch eine Standardexport-Funktion unterstützt wird. Dabei muss jeder Datensatz der Datenbanktabelle eine Zeile des ASCII-Flat-Files ausfüllen. Die einzelnen Felder eines Datensatzes können formatiert oder durch Trennzeichen voneinander separiert werden. Werden keine Trennzeichen verwendet, unterstützt das Migrationsprogramm die Erfassung der Feldbeschreibungen (im wesentlichen die Länge der einzelnen Felder) per Hand.

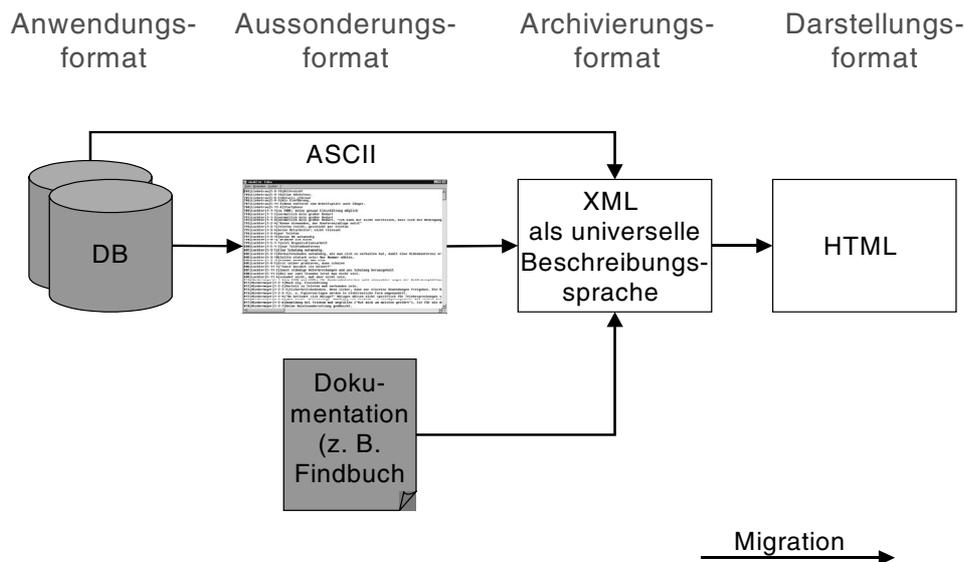


Abbildung 1: Überblick zum Ablauf des Migrationsprozesses

Häufig werden Informationen in Datenbanken in kodierter Form abgespeichert, so dass zusätzlich zur Datenbank auch ein Schlüsselverzeichnis (meist in nicht elektronischer Form) vorliegt, mit dem die Datenbankinformationen interpretiert werden können. Auch ein evtl. vorliegendes Schlüsselverzeichnis muss in das Migrationsverfahren einbezogen werden. Zu diesem Zweck muss für das entwickelte Migrationsverfahren eine Schlüssel-DTD erfasst werden. Die Schlüssel-DTD beschreibt die Klartexte der Schlüsselfelder, damit zur Präsentation der Daten auf die Klartexte der verschlüsselten Informationen zurückgegriffen werden kann. Datenbank-Schlüssel werden in der Schlüssel-DTD als Entitätsnamen, der zum Schlüssel gehörende Klartext als Inhalt der Entität abgebildet. Abb. 2 beschreibt die Dateien im Migrationsprozess.

Ergebnis des Migrationsprozesses sind drei Dateien im Archivierungsformat. Den wichtigsten Part spielt hierbei das XML-Dokument, das aus dem ASCII-Flat-File hervorgeht. Aus unverschlüsselten Feldern werden die Klartexte aus dem ASCII-Flat-File unverändert in das XML-Dokument übertragen, bei verschlüsselten Datenbankfeldern werden die Schlüssel des ASCII-Flat-Files im XML-Dokument durch Entitätsreferenzen ersetzt, die auf die in der Schlüssel-DTD befindlichen Entitäten verweisen. Der Browser ersetzt diese Entitätsreferenzen bei der Darstellung des XML-Dokuments automatisch durch den Inhalt der entsprechenden Entitäten, d. h. den Klartext.

Die dritte Ergebnisdatei des Migrationsprozesses ist die Erschließungsdatei. Sie enthält die Erschließungsdaten zur archivierten Datenbank, die vor dem Beginn des Migrationsverfahrens über Tastatur eingegeben werden müssen. XML-Dokument, Schlüssel-DTD und Erschließungsdatei sind über XML-Link-Strukturen miteinander verknüpft, so dass sie

bei Betrachtung mit einem XML-fähigen Webbrowser als ein logisches Dokument erscheinen und auch von anderen Applikationen als solches verarbeitet werden können.

Anzumerken ist noch, dass alle drei Daten ohne weitere Software von einem Standardbrowser dargestellt werden können. Bei Verwendung spezieller Algorithmen zur Umsetzung der Schlüssel in Klartext müssten diese unter Umständen an veränderte Hard-/Softwareumgebungen angepasst werden.

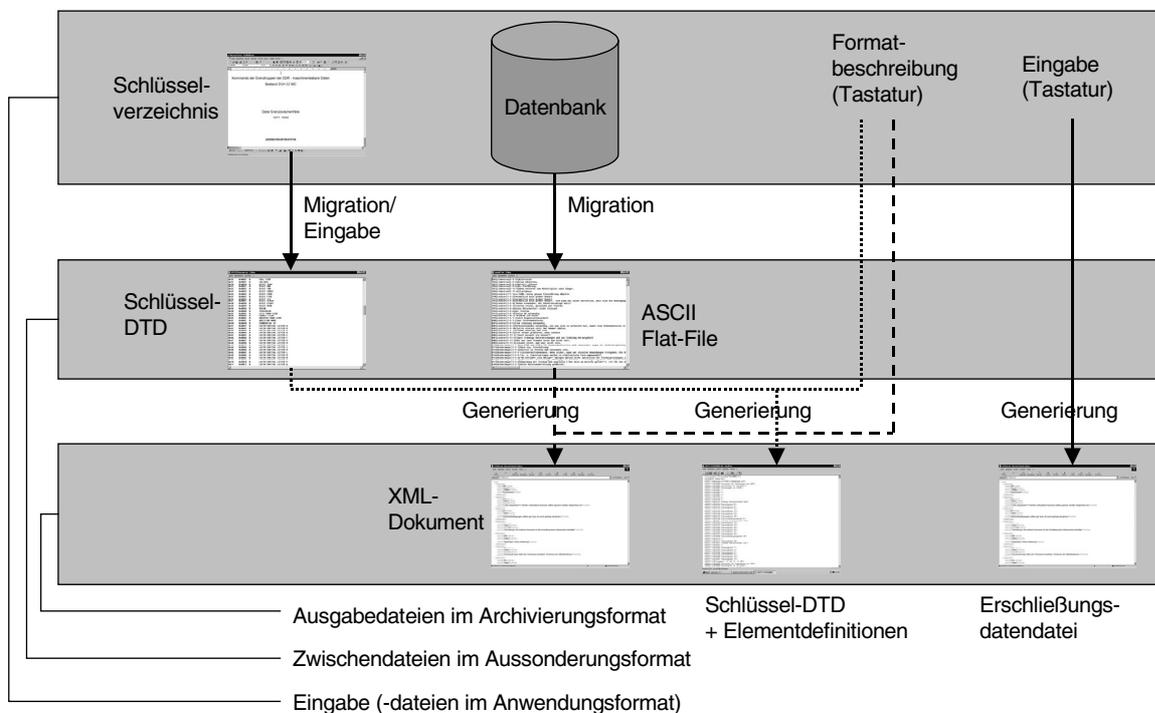


Abbildung 2: Ablauf des Migrationsprozesses mit den beteiligten Dateien

2. Grundentscheidung zur Migration von Datenbank-Inhalt in ein XML-Format

Zur Speicherung von Informationen aus einfachen, relationalen Datenbanken unter Zuhilfenahme der Extensible Markup Language bieten sich zwei Möglichkeiten an. Bei der ersten Variante werden die Feldinhalte der Datenbank als Inhalt eines Elements abgespeichert. Diese sind Kindelemente eines einen einzelnen Datensatz einrahmenden Elements, z. B.:

```
<Datensatz>
<Feld_1>
6105
</Feld_1>
<Feld_2>
000
</Feld_2>
[...]
</Datensatz>
```

Bei der zweiten Variante werden die Feldinhalte der Datenbank in Attributen abgelegt. Die Attribute gehören zu einem Element, das einen ganzen Datensatz der Datenbank aufnimmt, z.B.:

```
<Datensatz Feld_1="6105" Feld_2="000" [...] ></Datensatz>
```

Diese Lösung besitzt gegenüber der ersten Variante einen geringeren Plattenspeicherbedarf aufgrund der Verwendung einer kleineren Zahl von XML-Konstrukten zur Beschreibung der Primärinformationen. Hieraus resultiert wiederum ein geringerer Hauptspeicherbedarf bei der

Anzeige der Daten: Beim Laden in den Speicher werden XML-Dokumente intern in eine Baumstruktur gewandelt, wobei sich die zweite Variante im Hauptspeicher nicht so stark vergrößert wie Variante 1. Zudem befinden sich bei der Speicherung der Datenbankinformationen als Attributinhalt alle Primärinformationen eines Datensatzes in einer Zeile des XML-Dokuments, was auch von der äußeren Form weitestgehend dem Format des ASCII-Flat-Files entspricht.

Unverschlüsselte Feldinhalte werden bei beiden Varianten unverändert in das XML-Dokument übernommen. Zur Darstellung verschlüsselter Datenbankinhalte mit XML musste jedoch eine Lösung entworfen werden, mit der sich eine Umsetzung der Schlüssel in Klartexte auf einfache Weise vollziehen lässt. Verschiedene Alternativen mit Vor- und Nachteilen werden von Dehms⁴ beschrieben, an dieser Stelle wird nur die schließlich verwendete Version vorgestellt: Bei dieser werden die von XML zur Verfügung gestellten Entitäten verwendet, die in der entwickelten Lösung als Entitätsnamen den Feldnamen aus der Datenbank gefolgt vom Schlüssel und als Inhalt den dem Schlüssel entsprechenden Klartext besitzen, z.B.

```
<!ENTITY Feld1_1101 " Überfall mit Schußwaffen">
```

Im XML-Dokument wird anstelle des Schlüssels eine Entitätsreferenz eingetragen, die der XML-Parser durch den Inhalt der Entität, d. h. den Klartext, ersetzt. Zur Anzeige der Klartexte ist auf diese Weise kein Programmieraufwand mehr zu leisten. Sämtliche Entitäten werden in der zum XML-Dokument gehörenden Document Type Definition, der Schlüssel-DTD, definiert. Dies ist notwendig, da sich auf die in einem proprietären Format vorliegende Originaldatei des Schlüsselverzeichnis noch nicht mit den von XML zur Verfügung gestellten Funktionen zugreifen lässt. Die Generierung der Schlüssel-DTD aus dem Schlüsselverzeichnis kann jedoch nur in geringem Umfang automatisiert werden! Wie die Schlüssel-DTD erzeugt wird, bleibt letztlich dem Archivar überlassen. Nicht zu unterschätzen ist dabei der Authentizitätsverlust, der sich durch die Umwandlung des Schlüsselverzeichnis in die Schlüssel-DTD ergibt. Verschiedene Probleme sind bei Dehms⁵ ausführlich beschrieben, von denen eines hier exemplarisch angeführt wird:

Im ursprünglichen Schlüsselverzeichnis besitzen beispielsweise die Schlüssel des Feldes mit dem Feldnamen KZ06 folgenden Aufbau, wobei die erste Zeile die Überschrift der darauf folgenden Schlüssel darstellt:

Anschläge gegenüber der Zivilbevölkerung der DDR

```
1301 Überfall
1302 Beschießen
[...]
```

Würden diese in der bereits beschriebenen Form umgesetzt, so erhielten die Schlüssel die folgende Form:

```
<!ENTITY KZ6_1301 " Überfall">
<!ENTITY KZ6_1302 " Beschießen">
```

Um die Informationen aus den Überschriften des Schlüsselverzeichnis nicht zu verlieren, wurden diese in der Schlüssel-DTD zum Inhalt der Entität hinzugefügt. Ohne dieses Vorgehen sähe der Benutzer bei einer Abfrage nur die Wörter „Überfall“ und „Beschießen“, was einen erheblichen Informationsverlust gegenüber der ursprünglichen Aussage darstellen würde. Aus dem obigen Auszug der Codeliste wurden somit untenstehende Entitäten:

```
<!ENTITY KZ6_1301 " Anschläge der Zivilbevölkerung der DDR: Überfall">
<!ENTITY KZ6_1302 " Anschläge der Zivilbevölkerung der DDR: Beschießen">
```

⁴ Björn Dehms, Langzeitarchivierung einfacher, relationaler Datenbanken. Entwicklung eines Prototypen zur Migration nach XML. Diplomarbeit. Universität Koblenz, Fachbereich Informatik, Koblenz 2000.

⁵ Ebd.

Besondere Bedeutung kommt auch den fehlerhaften Datenbankeingaben (z. B. durch Zahlendreher) zu, wodurch Schlüssel (und später Entitäten) entstehen können, die nicht in der Schlüssel-DTD wieder zu finden sind und eine Fehlermeldung des XML-Parsers hervorrufen. Zu der entwickelten Fehlerbehandlung kann stark verkürzt gesagt werden, dass zu jedem fehlerhaft eingegebenen Schlüssel eine leere Entität definiert wird, z. B.

```
<!ENTITY Feld1_9999 " " >
```

Handelt es sich bei dem Schlüssel „9999“ um eine fehlerhafte Eingabe, so wird durch den XML-Parser an dieser Stelle der leere String ausgegeben, was gegenüber der Original-Datenbank einen weiteren Authentizitätsverlust darstellt. Das exakte Format zur Fehlerbehandlung ist ebenfalls bei Dehms⁶ zu finden.

Zur Beschreibung der in der entwickelten Lösung erzielten Authentizität wurde der Begriff der „Bedienungsauthentizität“ geprägt: Im besten Fall sollte nach der Migration der relationalen Datenbank in das XML-Dateiformat ein Unterschied zur Abfrage und Manipulation relationaler Datenbanken mittels SQL nicht mehr erkennbar sein. Dem Benutzer soll im besten Fall verborgen bleiben, dass es sich bei der verwendeten Datenbank um keine relationale Datenbank, sondern vielmehr um eine Datenbank im XML-Dateiformat handelt. Diese Form der Authentizität war aus Gründen, die Dehms näher erläutert⁷, nicht vollständig zu erreichen.

Die drei denkbaren Möglichkeiten zur Umsetzung der Schlüssel in Klartexte sollen nun vor allem bezüglich des Speicherbedarfs auf der Festplatte betrachtet werden: Die erste Möglichkeit wäre die Speicherung des Klartextes anstelle des Schlüssels im XML-Dokument. Hierbei würde der ursprüngliche Schlüssel nicht im Dokument abgelegt und ginge verloren. Der Platzbedarf dieser Variante wäre enorm hoch, da die u. U. sehr langen Klartexte an verschiedenen Stellen des XML-Dokumentes wiederholt abgelegt würden, wodurch sich die Menge redundanter Daten stark vergrößerte. Zweitens könnte ein Verweis auf den Klartext abgespeichert werden, was der von Dehms⁸ schließlich favorisierten Lösung entspricht. Der Klartext würde nicht direkt im XML-Dokument, sondern erst zur Anzeige vom XML-Prozessor an der Stelle des Verweises eingefügt und auf dem Ausgabegerät ausgegeben. Es müssten nicht mehr die Klartexte an verschiedenen Stellen des XML-Dokumentes redundant gespeichert werden, sondern nur noch die Verweise, die jedoch erheblich kürzer sind als die zu ihnen gehörenden Klartexte. Gegenüber Lösung 1 reduzierte sich der Speicherbedarf hierbei um fast die Hälfte (s. Berechnung in Dehms⁹). Bei der dritten Variante würden nur die Schlüssel, d.h. die Daten in ihrer ursprünglichen Form, im XML-Dokument abgelegt. Die unter Umständen sehr aufwendige Interpretation der Schlüssel fände dann während der Aufbereitung der Daten zur Anzeige statt. Gegenüber Lösung 2 würde sich der Speicherbedarf um weitere 15,5 % reduzieren.

Lösung 1 kommt somit nicht mehr in Betracht. Lösung 3 ist jedoch nur unter dem Gesichtspunkt des Speicherbedarfs auf der Festplatte die optimale Lösung und benötigt gegenüber der Lösung 2 bei der Anzeige eine aufwändigere Bearbeitung.

3. Vorstellung des Migrationsprogramms mit einer Kurzdarstellung der 7 Schritte

Das von Dehms¹⁰ entwickelte Migrationsprogramm überträgt die Informationen eines aus einer Datenbank ausgesonderten ASCII-Flat-Files in ein XML-Dokument. Da ASCII-Flat-Files in den unterschiedlichsten Formaten vorliegen können, muss der Benutzer die Struktur

⁶ Ebd.

⁷ Ebd.

⁸ Ebd.

⁹ Ebd.

¹⁰ Ebd.

der jeweiligen Datei möglichst präzise beschreiben, wozu das Programm sieben aufeinander folgende Dialoge anbietet. Als Eingabe in den Migrationsprozess erhält das Migrationsprogramm drei Dateien:

die aus der Datenbank erzeugte Aussonderungsdatei, das ASCII-Flat-File,

die Schlüssel-DTD,

die Datei mit den Namen der Felder der Datenbanktabelle, falls diese nicht über die Tastatur eingegeben werden sollen.

Zu diesen drei Dateien müssen im Startdialog Angaben gemacht werden: Das oberste Editierfeld verlangt zuerst nach einem frei wählbaren Projektnamen. In den beiden darunter liegenden Editierfeldern müssen Name und Position des ASCII-Flat-Files sowie des zu erzeugenden XML-Dokumentes im MS-DOS-Format angegeben werden.

Das darunter liegende Panel betrifft Angaben zu den Namen der im XML-Dokument zu speichernden Felder der Datenbanktabelle. Diese können sowohl aus einer bereits erstellten Datei geladen werden oder über einen Dialog von Hand eingegeben werden. Über eine Checkbox kann weiterhin festgelegt werden, ob das ASCII-Flat-File Trennzeichen zwischen den einzelnen Feldern besitzt oder ob die Felder an deren Länge erkannt werden müssen.

Im untersten Editierfeld des Dialoges „Schritt 1 von 7“ müssen Ort und Name der Schlüssel-DTD angegeben werden, damit diese vom Migrationsprogramm zum XML-Dokument hinzugefügt werden kann. Das auf den Startdialog folgende Fenster listet die Feldbezeichner der Datenbank untereinander auf. Durch Markieren einer Checkbox wird das entsprechende Feld in das erzeugte XML-Dokument übertragen, anderenfalls erscheint es nicht in der Ausgabe. Der darauf folgende Dialog besitzt den gleichen Aufbau wie Schritt 2, hier müssen jedoch die verschlüsselten Datenbankfelder markiert werden.

Screenshot of the "Schritt 1 von 7" dialog box. The dialog is titled "Schritt 1 von 7" and contains the following fields and buttons:

- Projektname:
- Name der Eingabedatei:
- Name der Ausgabedatei:
- Name der Feldbezeichnerdatei:
- Feldbezeichner von Hand eingeben
- Feldtrennzeichen vorhanden (CSV-Datei)
- Name der Schlüssel-DTD:
-

Abbildung 3: Schritt 1 von 7



Abbildung 4: Schritt 2 von 7

Der letzte Dialog (vgl. Abbildung 6) verlangt noch einmal nach Angaben über das Format des zugrundeliegenden ASCII-Flat-Files. Hier muss u. a. in einem eigenen Editierfeld das verwendete Feldtrennzeichen angegeben werden.

Eventuell sind in dem ASCII-Flat-File Begrenzerzeichen („Delimiter“) vorhanden, die die Feldinhalte einrahmen. Diese müssen im Panel „Begrenzer“ angegeben werden, damit sie bei der Ausgabe in das XML-Dokument ausgeschnitten werden können. Im Unterschied zu den Trennzeichen muss jedoch das Aussehen der Begrenzerzeichen in den Primärdaten angegeben werden, damit sie das Migrationsprogramm von den eigentlichen Begrenzerzeichen unterscheiden kann. Oft werden beispielsweise als Begrenzer Anführungszeichen verwendet, die auch als Zeichen im eigentlichen Feldinhalt auftauchen können.

Im Schritt 4 (vgl. Abbildung 5) werden Datenbankfelder beschrieben, die mehrere Schlüssel enthalten. In der Grundeinstellung stellt das einzige Fenster des Dialogs alle im Schritt 2 ausgewählten Datenbankfelder in einer Baumansicht dar, der die Namen der einzelnen Felder als Knoten besitzt. Einem Knoten können dann Informationen über die Länge der einzelnen Schlüssel innerhalb des Feldes hinzugefügt werden (s. Dehms¹¹). Mit einem weiteren Dialog, der wiederum dem Dialog aus Abbildung 4 entspricht, werden die Datenbankfelder ausgewählt, die mehrdeutige Schlüssel enthalten. Unter mehrdeutigen Schlüsseln versteht man Zahlenfolgen, die an unterschiedlichen Stellen innerhalb eines Feldes mit verschiedenen Bedeutungen verbunden sind.

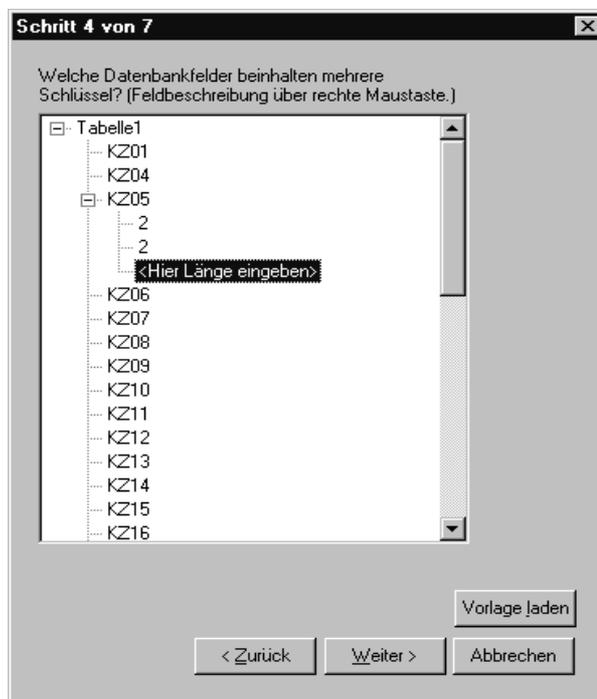


Abbildung 5: Schritt 4 von 7

¹¹ Björn Dehms (wie Anm. 4).

Die verschiedenen Zeilenumbrüche der unterschiedlichen Betriebssysteme können am linken Rand des Dialoges ausgewählt werden. Eine Projektübersicht listet zudem die am aktuellen Migrationsvorgang beteiligten Dateien auf. Der eigentliche Migrationsvorgang kann daraufhin durch Betätigen der „Migration beginnen/abbrechen“-Buttons gestartet bzw. vorzeitig beendet werden.

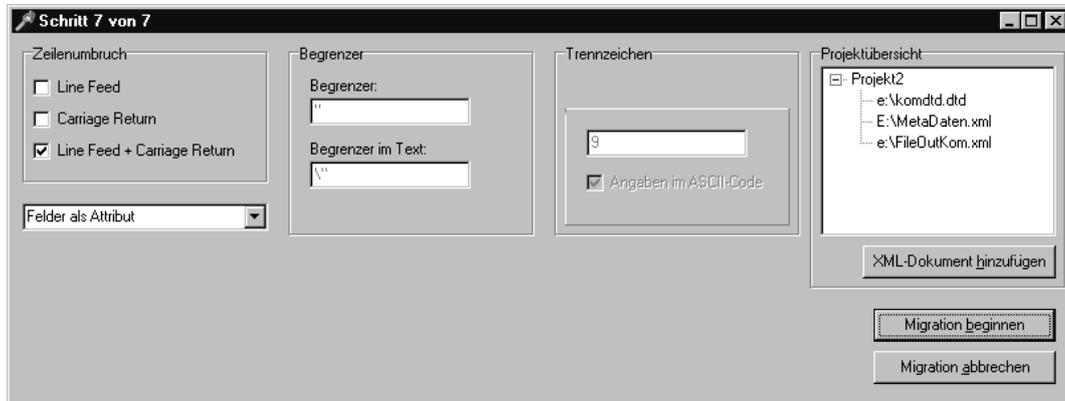


Abbildung 6: Schritt 7 von 7

4. Bereitstellung und Nutzung von Datenbank-Informationen in XML

Für die Bereitstellung und Nutzung von Datenbank-Informationen in XML sind die folgenden drei Szenarien vorstellbar. Die ersten beiden Nutzungsszenarien betreffen die Verwendung des XML-Dokuments als „Datenbank“, das letzte Szenario beschreibt die Auswertung der enthaltenen Informationen über ein Statistiksystem:

Auswahl eines oder mehrerer Datensätze ohne Einschränkung der Ausgabe

Der Benutzer könnte sich einen oder mehrere von ihm ausgewählte Datensätze komplett anzeigen lassen, d. h. es würden die Inhalte aller Felder der ausgewählten Datensätze ausgegeben. In der relationalen Algebra entspricht dies der „Selektion“ genannten Grundoperation.

Auswahl eines oder mehrerer Datensätze mit Einschränkung der Ausgabe

Diese Nutzungsmöglichkeit entspricht der unter Punkt 1 genannten, jedoch würden nur die vom Benutzer ausgewählten Felder der selektierten Datensätze ausgegeben. Die Einschränkung der Ausgabe auf bestimmte Tabellenfelder wird in der relationalen Algebra mit „Projektion“ bezeichnet.

Nutzung mit einem Statistikprogramm

Da XML als Dateiformat u. a. ausgewählt wurde, um XML-Dokumente in andere Dateiformate migrieren zu können, ist eine Umwandlung des XML-Dokuments in das von einem Statistikprogramm benötigte Dateiformat möglich. Dieses könnte von einem Statistiksystem eingelesen werden und komplexe Auswertungen des Datenbestandes an einem Einzelplatzrechner ermöglichen. Diese Nutzungsmöglichkeit wäre somit nicht für die Allgemeinheit bestimmt.

Damit eine möglichst strenge Authentizität der Datenbank auch nach der Migration in ein XML-Dokument bewahrt bleibt, sollte sich die Nutzung des XML-Dokuments möglichst wenig von der Nutzung relationaler Datenbanken über ein Datenbankmanagement-System unterscheiden. Somit müssten auch bei der Verwendung von XML-Dokumenten als Daten-

basis die bei Datenbanken üblichen SQL¹²-Statements (DELETE, UPDATE, INSERT und SELECT) angeboten werden.

Da nach der eigentlichen Archivierung in aller Regel keine Änderungen an den in der archivierten Unterlage gespeicherten Informationen mehr vorgenommen werden, kann die Umsetzung der INSERT-, DELETE- und UPDATE-Statements entfallen. Es müssen lediglich SELECT-Statements angeboten werden, womit die Nutzung der Datenbank auf eine Abfrage der Informationen beschränkt ist.

Ein SELECT-Statement hat in SQL die folgende Syntax:

```
Select <FieldList> From <Tabelle> Where <Conditionlist>
```

Durch die „Conditionlist“ geschieht die Selektion der Datensätze einer Datenbanktabelle. Hier wird die Ergebnismenge (= die Menge der von einer Datenbankabfrage zurückgelieferten Datensätze) dadurch eingeschränkt, daß zu verschiedenen Feldbezeichnern Bedingungen angegeben werden, wie z. B.

```
Select <FieldList> From <Tabelle> Where Feld1>100 And Feld2=300
```

In diesem Beispiel werden somit nur solche Einträge aus der mit <Tabelle> spezifizierten Tabelle ausgewählt, bei denen „Feld1“ Werte größer 100 besitzt und der Wert von „Feld2“ genau 300 beträgt. In der Fieldlist werden die auszugebenden Felder der Datenbanktabelle angegeben. Eine mögliche Erweiterung der in obigem Beispiel angegebenen SELECT-Anweisung wäre somit

```
Select Feld3,Feld4 From <Tabelle> Where Feld1>100 Or Feld2=300
```

Hierdurch wird die Ergebnismenge auf die Felder „Feld3“ und „Feld4“ projiziert, d.h. in der Ausgabe erscheinen von jedem Datensatz nur die Werte dieser beiden Felder. Mit der Angabe „*“ für den Parameter „Fieldlist“ werden die Werte aller Felder ausgegeben. In der Fieldlist können auch Aggregationsfunktionen, z.B. Summe, Mittelwert oder Trefferanzahl der gefundenen Datenbankeinträge, verwendet werden.

Erzeugung der Abfrage (auf dem Client)

Um in SQL unerfahrenen Benutzern eine komfortable Auswertungsmöglichkeit des XML-Dokuments zu bieten, wird eine grafische Benutzeroberfläche in Form einer HTML-Seite im WWW bereitgestellt, mit der eine Suchanfrage gestellt werden kann. Das World Wide Web wurde als Medium gewählt, um die Daten für jedermann zugänglich zu machen. Eine Alternative wäre das Verschicken der Datenbank/des XML-Dokuments auf einem Datenträger, wobei allerdings die Nutzung der Unterlage durch den Empfänger nicht sichergestellt werden könnte, da diese u. U. spezielle Programme erfordert. Dennoch ist das Bestellen der Daten prinzipiell möglich.

Das im WWW bereitgestellte Formular beinhaltet einen mit der Skriptsprache „Javascript“ erzeugten Baum. Der Aufbau des Baumes ist sehr rechenintensiv und erfordert ein wenig Geduld. Der schließlich angezeigte Javascript-Baum ermöglicht eine Auswahl von Werten zu bestimmten Datenbankfeldern. Zu jedem Feld kann nur ein Wert ausgewählt werden, die selektierten Werte der verschiedenen Felder werden in der TextBox des HTML-Formulars angezeigt. Sie können in der TextBox nicht bearbeitet werden und werden im Hintergrund zu einem SELECT-Statement zusammengestellt, indem die ausgewählten Werte in der Fieldlist mit „And“ miteinander verknüpft werden. Wird aus dem JavaScript-Baum als Staats-

¹² SQL = Structured Query Language. Mit SQL können die Datensätze einer Datenbank abgefragt und manipuliert werden.

angehörigkeit beispielsweise „DDR“ und als Wochentag „Mittwoch“ gewählt, so besitzt das generierte SELECT-Statement den folgenden Aufbau:

```
SELECT * FROM Table WHERE KZ17 = 31 And KZ08 = 3
```

Blättern	
zurück	vor
Neue Abfrage	
Treffer 1 von 3630	
Dienststelle:	6102
Meldenummer:	289
Blocknummer des Falls:	1
Sichernde bzw. feststellende Einheit:	111
Anzahl der Fälle:	1
Art der Handlung:	Festnahmen wegen Versuches des Grenzdurchbruches/Fahnenflucht - durch Grenzaufklärer. Richtung BRD / Berlin (West)- DDR
Datum:	01.12.71
Wochentag:	Mittwoch
Uhrzeit:	23:30
Witterung:	Witterungsbedingungen Nachtzeit klare Sicht
Sektorennummer an der Grenze zur BRD:	
Sicherungs-/Grenzabschnitt und Meldeabschnitt von:	84
Sicherungs-/Grenzabschnitt und Meldeabschnitt bis:	84
Anzahl der handelnden Personen:	1
Durch wen wurde Handlung verübt:	Zivilpersonen , wenn nachfolgendes nicht bekannt ist
Staatsangehörigkeit:	BRD
Dienstgrad, Einheit, Alter, Geschlecht:	weiblich, Alter von/bis: unbekannt
Beruf, Beschäftigung, soziale Stellung:	Andere / Unbekannt
Lage des Wohnortes:	Kein Staatsbürger der DDR
Bezirk, Kreis, Bundesland, Staat, Dienststelle:	Schleswig - Holstein
Festnahme durch/im Zusammenwirken mit:	Die Festnahme erfolgte im Zusammenwirken mit der. PKE / MfS
Festnahme-, Handlungs- oder Fundort:	
Anzahl der festgestellten Mittel:	0
Art des Mittels:	
Methoden der Anwendung:	
Pionier- und signaltechnischer Ausbau der Grenze:	
Wirksamkeit der eingesetzten Mittel:	
Unterstützungshandlungen durch den Gegner:	
Auswirkungen der Handlung:	
Überwundene bzw. ausgenutzte Mittel:	Bahnhöfe / Deutsche Reichsbahn
Ursachen, begünstigende Umstände:	Gute Organisation des Zusammenwirkens der Grenzposten und grenz - sichernden Einheiten
Sonstige Angaben/Motive:	
Angabe Höhe und Tiefen bei Provokationen:	Flughöhe: 0 Tiefe des Eindringens: 0 Länge des Eindringens: 0
Lage des Handlungs- oder Festnahmeortes:	0000
Anzahl abgegebener Schüsse:	0
Schadenssumme:	0.00
Dauer der Handlung:	0

Als Werte der Felder erscheinen hier die Schlüssel und nicht die Klartexte des Schlüsselverzeichnis.

Nach Betätigen des „Abschicken“-Buttons wird das im Hintergrund generierte SELECT-Statement zum Server gesandt. Die Verarbeitung des SELECT-Statements ist im folgenden Kapitel beschrieben, in Abbildung 7 ist beispielhaft das Ergebnis der Verarbeitung eines SELECT-Statements zu sehen. Das grafische Front-End bietet also die Ausgabe von einzelnen Datensätzen an; bei mehreren zur Ausgabe anstehenden Datensätzen kann zwischen diesen geblättert werden. Jeder Datensatz wird komplett angezeigt, d.h. es wird bei der Anzeige keine Einschränkung auf bestimmte Datenbankfelder ermöglicht. Dies entspricht der ersten der zu Beginn dieses Kapitels beschriebenen Nutzungsmöglichkeiten.

Abbildung 7: Die Anzeige eines Datensatzes

Erzeugung der Ausgabe (auf dem Server)

Eine der für die Ausgabe der Informationen zuständigen Komponenten ist der SQL-Parser, der in Form einer ActiveX-DLL implementiert ist. Die Form einer Dynamic Link Library (DLL) wurde gewählt, um den SQL-Parser als eigenständige Komponente auch in anderen Projekten verwenden zu können, denn die Arbeit des SQL-Parsers hängt nicht vom zugrundeliegenden XML-Dokument ab. Im Anschluß an die Beschreibung des SQL-Parsers befindet sich die Beschreibung der zur Erzeugung der Ausgabe verwendeten Active Server Pages (ASP).

SQL-Parser

Der SQL-Parser dient zur Analyse eines SQL-Statements. Er kann die oben beschriebenen SELECT-Statements verarbeiten und wurde unter Zuhilfenahme der frei erhältlichen Tools „LEX“ (zur lexikalischen Analyse) und „YACC“ (als Parsergenerator) erstellt. Diese beiden Tools erhalten als Eingabe eine Grammatik, die den SQL-Befehl beschreibt, und erzeugen als Ausgabe jeweils eine Delphi-Datei, die eine Klasse mit dem Code zur lexikalischen Analyse bzw. dem Code zum Parsen des SQL-Statements enthält. Normalerweise ist es die Hauptaufgabe eines Parsers, für ein eingegebenes Wort zu überprüfen, ob dieses aus der zugrundeliegenden Grammatik erzeugt werden kann. Es wurde jedoch vielmehr von der Möglichkeit zur Abarbeitung semantischer Aktionen Gebrauch gemacht, die bei der Erkennung bestimmter Merkmale des SQL-Statements ausgeführt werden. Zur Erzeugung der eigentlichen Ausgabe werden Active Server Pages (ASP) verwendet.

Active Server Pages (ASP)

Microsoft® Active Server Pages (ASP) sind eine serverseitige Skripttechnologie zur Erzeugung von dynamischen, interaktiven Webseiten. Eine ASP-Seite ist im Prinzip eine HTML-Seite, die zusätzlich zu den HTML-Konstrukten auf dem Server auszuführenden Code (= Skripte) enthält. Diese Skripte werden ausgeführt, wenn der Benutzer mit Hilfe eines Browsers auf eine solche HTML-Seite (deren Dateierweiterung allerdings „.asp“ ist) zugreift. Der Server arbeitet die angefragte Datei von oben nach unten ab und führt alle eingebetteten Skripte aus. Die Skripte können mit JScript oder Visual Basic Script (VBScript) verfasst sein und erzeugen meist eine HTML-Seite, die an den Client geschickt und dort vom Browser angezeigt wird.

Die für die Zwecke der Examensarbeit entwickelte Active Server Page ist speziell an das im Rahmen der Diplomarbeit migrierte XML-Dokument angepaßt und kann nicht für andere migrierte Datenbanken verwendet werden. Dies wird durch den besonderen Aufbau der zugrundeliegenden Datenbank verhindert, der keine für alle „XML-Datenbanken“ gültigen Auswertungsskripte zulässt.

Die ASP erzeugt eine HTML-Seite, die einen der vom Benutzer ausgewählten Datensätze anzeigt (vgl. Abbildung 7). XSL Style Sheets konnten zur Anzeige nicht verwendet werden, da sie noch nicht in einer „official recommendation“ festgeschrieben sind und deshalb z. Z. noch von keinem Browser unterstützt werden.

Zu Beginn der Skriptverarbeitung wird dem SQL-Parser das vom Benutzer eingegebene SELECT-Statement übergeben, der dieses auf korrekte Syntax überprüft. Dann wird das XML-Dokument in den Speicher geladen.

Es werden der Reihe nach alle im XML-Dokument enthaltenen Datensätze darauf getestet, ob sie die in dem SELECT-Statement angegebenen Bedingungen erfüllen. Der erste Datensatz, der bei dieser Überprüfung ein positives Ergebnis liefert, wird ausgegeben. Wird über den Client vorwärts geblättert, so wird wieder der Reihe nach der nächste passende Datensatz im

XML-Dokument gesucht. Beim Zurückblättern wird in der gleichen Weise rückwärts nach dem nächsten passenden Datensatz gesucht.

Die Verarbeitung von XML-Dokumenten unterliegt einigen Einschränkungen, die sich aus der Performanz des XML-Parsers ergeben.

Berechnungen zur Performanz des XML-Parsers

Sämtliche Angaben zur Performanz des XML-Parsers sind Lovett¹³ entnommen. In diesem Artikel werden folgende Variablen angegeben, die die Performanz des XML-Parsers beeinflussen:

- die Art der XML-Daten
- das Verhältnis von Tags zu Text
- das Verhältnis von Attributen zu Elementen
- die Menge der Formatierungszeichen („The amount of discarded white space.“ Lovett.)

Leider ist in der Quelle nicht genauer angegeben, in welcher Weise diese Größen die Performanz des XML-Parsers beeinflussen.

Es werden an dieser Stelle nur Berechnungen bzgl. des Hauptspeicherbedarfs angeführt, da sich dieser im Verlaufe der Arbeit als der begrenzende Faktor herausstellte. Diese Berechnungen werden am Ende des Kapitels auf die „Datei Grenzzwischenfälle“ übertragen. Die zugrundeliegenden Abschätzungen sind sehr grob, geben aber einen guten Eindruck über den Ressourcenbedarf. Bei Lovett heißt es hierzu sinngemäß:

Der Speicherbedarf beim Laden eines XML-Dokumentes in den Hauptspeicher ist etwa bis zu vier mal größer als beim Ablegen des XML-Dokumentes auf der Festplatte. Der Hauptspeicherbedarf eines XML-Dokumentes kann mit der folgenden Formel grob abgeschätzt werden:

$$ws = 32(n + t) + 12t + 50u + 2w$$

Die folgende Tabelle gibt einen Überblick über die Variablen und deren Bedeutung:

Variable	Beschreibung
ws	Größe des Hauptspeicherbedarfs in Bytes
n	Summe der Elemente und Attribute
t	Anzahl der Textknoten
u	Anzahl der verschiedenen Element- und Attributnamen
w	Anzahl der Unicode-Zeichen in den Textknoten (einschließlich der Attributinhalt). Wird der ASCII-Zeichensatz verwendet, so muß die Anzahl der Zeichen mit 2 multipliziert werden.

Die Variablen besitzen bei dem nach Dehms¹⁴ erzeugten XML-Dokument in Abhängigkeit von der Anzahl der abgespeicherten Datensätze die hier aufgelisteten Werte:

$$\begin{aligned}
 n &= 37 \times \text{Anzahl der Datensätze} \\
 t &= 73 \times \text{Anzahl der Datensätze} \\
 u &= 12 \\
 w &= (2 \times 40) \times 37 \times \text{Anzahl der Datensätze}
 \end{aligned}$$

¹³ Chris Lovett, Inside MSXML Performance, Redmont, Microsoft Corporation, 21.02.2000, <http://msdn.microsoft.com/xml/articles/xml02212000.asp>.

¹⁴ Björn Dehms (wie Anm. 4).

Für die Variable w wurde die mittlere Länge des Inhalts der Textknoten sehr grob mit 40 Bytes geschätzt. Da der ASCII-Zeichensatz zugrundeliegt, muß dieser Wert wie in der Tabelle beschrieben mit 2 multipliziert werden.

Werden die angegebenen Werte in obige Formel eingesetzt, so ergibt sich nach deren Umstellung für die Anzahl der möglichen Datensätze folgende Formel:

$$\# \text{ Datensätze} = \frac{ws - 600}{10316}$$

Um die maximale Anzahl der zu verarbeitenden Datensätze zu erhalten, muss für die Variable „ ws “ die Größe des virtuellen Speichers in Byte eingesetzt werden.

Die Gleichung wird nun auf die tatsächlich zur Verfügung stehende Hardware umgesetzt: Der auf dem verwendeten Rechner zur Verfügung stehende virtuelle Speicher besaß etwa 250 MB. Von diesen sind etwa 60 MB ständig durch das Betriebssystem und im Hintergrund laufende Programme belegt, so dass noch 190 MB (= 199229440 Byte) virtueller Speicher für den XML-Parser zur Verfügung stehen. Dieser Wert wird in die obige Formel eingesetzt, so dass sich für die maximale Anzahl mit dem Rechner zu verarbeitender Datensätze folgender Wert ergibt:

$$\# \text{ Datensätze} = \frac{199229440 - 600}{10316} \approx 19300$$

Dieser Wert stimmt ungefähr mit den Erfahrungswerten überein, nach denen bei XML-Dokumenten mit etwa 20.000 Datensätzen eine Fehlermeldung über die komplette Auslastung des virtuellen Speichers erscheint. In der Praxis werden jedoch nur etwa 10.000 Datensätze auf dem Server hinterlegt, da die Antwortzeiten aufgrund des häufigen Zugriffs auf den Swap-Speicher bei größeren XML-Dokumenten nicht mehr vertretbar sind. Nach Möglichkeit sollte das gesamte XML-Dokument in dem physisch vorhandenen Hauptspeicher Platz finden.

Unter diesen Bedingungen eignet sich das XML-Dateiformat nicht als Grundlage für die Präsentation der in der Datenbank vorhandenen Informationen im Internet, denn bei insgesamt 650.000 vorhandenen Datensätzen werden auch in absehbarer Zukunft keine Rechnersysteme zur Verfügung stehen, die eine derartige Datenmenge verarbeiten können.

5. Erfahrungen mit XML als Archivierungs- und Nutzungsformat

Einfache, nur aus einer Tabelle bestehende Datenbanken lassen sich mit XML beschreiben. Bei Datenbanken ohne Schlüsselverzeichnis ist eine Migration ohne Informationsverlust problemlos durchführbar. Erheblichen Arbeitsaufwand bereitet bei verschlüsselten Datenbanken hingegen die verlustfreie Umsetzung des Schlüsselverzeichnisses in die Schlüssel-DTD. Aufgabe des Archivars ist es, diesen Informationsverlust möglichst klein zu halten.

Mit XML archivierte Datenbanken gängiger Größenordnung können auf heutigen Rechnersystemen unter Verwendung eines Standardbrowsers nicht sinnvoll präsentiert verwendet werden, da der Hauptspeicherbedarf des XML-Parsers bei der Verarbeitung von XML-Dokumenten zu hoch ist. Zur Nutzung von XML-archivierten Datenbanken werden demzufolge spezielle Programme zur Aufbereitung der Präsentation notwendig, was den Vorteil von XML als softwareunabhängigem Archivierungsformat erheblich relativiert.

Zusammenfassend kann festgehalten werden, dass sich XML zwar zur Aufbewahrung und Sicherung von einfachen, relationalen Datenbanken anbietet, zur Präsentation der Informationen ohne Zusatzprogramme unter realistischen Einsatzbedingungen aber untauglich ist. Vor diesem Hintergrund ist die Verwendung des XML-Dateiformats zur Langzeitarchivierung von Datenbanken nur bedingt zu empfehlen. Vor allem bei verschlüsselten Datenbanken

sollte überlegt werden, ob sich die zeitaufwendige Umsetzung des Schlüsselverzeichnis in eine Schlüssel-DTD lohnt oder ob eine Magazinierung der Originaldatenbankdatei als ASCII-Flat-file gegenüber einer Migration nach XML nicht erheblich wirtschaftlicher ist.

Als Vorteil einer XML-Migration kann allerdings angesehen werden, dass auch mit der heute vorhandenen Technologie eine Auswahl von Datensätzen im XML-Format durchaus direkt unter Zugriff auf ein XML-Dokument präsentiert werden kann. Diese Präsentationsmöglichkeit wurde im Rahmen des Projekts zum Zweck der Aufbereitung der Datenbank als Online-Abfragemöglichkeit für interessierte Benutzer realisiert. Als Server zur Abfrage der Datenbankinformationen wurde eine Windows NT Workstation und der Personal Web Server (PWS) verwendet. Die Antwort-HTML-Seite wurde von einer Active Server Page (ASP) erzeugt. Die Online-Abfrage ist unter <http://miles.uni-koblenz.de/Query/open.html> abrufbar.

Bundesarchiv Online
▶ www.bundesarchiv.de



Suchanfrage eingeben

Die Recherche in den online verfügbaren Datensätzen ist über die Auswahl eines oder mehrerer Felder im unten stehenden Baum durch Anklicken möglich. Hierzu ist zu bemerken, dass nicht alle Felder der Datenbank recherchierfähig sind. Nach der Auswahl der Felder werden als Suchbegriffe die potentiellen Feldinhalte dargestellt. Aus den Feldinhalten kann jeweils nur ein Wert gewählt werden, "oder"-Verknüpfungen sind ausgeschlossen. Das Datum muß im angegebenen Format eingegeben werden. Nähere Informationen zu den Feldinhalten können dem [Findbuch im RTF-Format](#) oder der Beschreibung der [Suchbegriffe](#) entnommen werden. Enthält die Ergebnismenge mehrere Datensätze, so können diese durch Vorwärts- und Rückwärtsblättern angesehen werden. Soweit möglich werden von jedem Datensatz die codierten Informationen im Klartext ausgegeben.
In der Diplomarbeit wurde neben der Online-Recherche die Eignung von XML für die Langzeitarchivierung untersucht. Da für die Langzeitarchivierung die Unabhängigkeit von speziellen Anwendungsprogrammen essentiell ist, wurde für die Web-Präsentation auf die Unterstützung durch ein Datenbankmanagement-System verzichtet. Die Daten sind in einem XML-Dokument abgelegt, was bei der Recherche auch in der kleinen Auswahl an Datensätzen zu längeren Wartezeiten führen kann. Die Warnmeldung mit der Frage, ob das Java-Skript abgebrochen werden soll, kann mit "nein" beantwortet werden.

Datum(TTMMJJ): (Suche nur zwischen 1.12.1971 und 15.12.1971 möglich!)

- Dienststelle
- Wochentag
- Staatsangehörigkeit
- Durch wen wurde Handlung verübt?
- Auswirkungen der Handlung
- Überwundene bzw. ausgenutzte Mittel
- Lage des Handlungs- oder Festnahmeortes

Die von Ihnen ausgewählten Suchbegriffe (keine Bearbeitung möglich):

Abbildung 8: Die Startseite der Online-Abfrage zur Datei der Grenzzwischenfälle

Rematerialisierung digitaler Unterlagen – ein Weg zur Langzeitarchivierung?

Das Projekt HD-Rosetta

Von Karl-Ernst Lupprian

Es hieße Eulen nach Athen tragen, wenn man in diesem Kreis die Faktoren aufzählen wollte, die das Überleben von Unterlagen in digitaler Form gefährden. Charles Dollar hat sie in seinem jüngsten Buch über authentische elektronische Unterlagen¹ ausführlich diskutiert und Strategien für eine Langzeitaufbewahrung vorgeschlagen. Dabei favorisiert er mehr oder weniger die Migration, also das rechtzeitige Konvertieren der Daten in Formate, die auf jeweils neuen Hard- und Softwareplattformen verarbeitet werden können.

Zu Recht warnt Dollar vor den im voraus kaum kalkulierbaren, aber mit Sicherheit hohen Kosten für Migrationen. Schon die Aufbewahrung digitaler Unterlagen über – aus archivischer Sicht – kurze Zeiträume ist teuer genug, wie eine 1996 angestellte Schätzung des NARA Center for Electronic Records für die Archivierung digitaler Unterlagen über 10 Jahre zeigt:²

Kosten in US-\$	Kostenart
600.000	6 Arbeitsplatzrechner mit je 2 Magnetband- und -kassettenlaufwerken
928.000	Wartung
30.000	2 Prüfstationen für Magnetbandspulen
92.160	Wartung
50.000	2 Prüfstationen für Magnetbandkassetten
46.600	Wartung
10.786.260	Erste Kopieraktion des Gesamtbestands
29.876	10 Prüfläufe (einer pro Jahr)
1.030.640	Weitere Kopier- und Umformatierungsläufe
13.594.396	Gesamtkosten

Die Schätzung geht von einem Bestand von 23.000 digitalen Einheiten und einem durchschnittlichen jährlichen Zugang von 2900 Einheiten aus, ohne allerdings den Umfang einer Einheit anzugeben. Der Löwenanteil der Kosten entfällt auf Kopieraktionen (Refreshing), Migrationen sind hier noch nicht vorgesehen. Und da der Datenbestand ständig wächst, wird auch der Umfang der Kopieraktionen ständig größer, denn diese umfassen ja stets den Gesamtbestand.

Betrachtet man nun Migrationen, so zeigen sich weitere kostenträchtige Faktoren. So muss jeder Migrationslauf sorgfältig protokolliert werden, damit die Authentizität der Unterlagen

¹ Charles M. Dollar, Authentic electronic records. Strategies for long-term access, Chicago 2000. – Vgl. auch Bryan Bergeron, Dark ages II. When the digital data die, Upper Saddle River 2002.

² Charles M. Dollar (wie Anm. 1) S. 211.

nicht gefährdet wird. Hierbei entsteht eine u.U. durchaus beträchtliche Datenmenge.³ Auch sie wächst mit jeder weiteren Migration, da die Protokolle früherer Migrationen ebenfalls migriert werden müssen. Neben den eigentlichen Daten entsteht also eine sekundäre Datenlast, die jedoch genau so sorgfältig zu bewahren ist.

Hinzu kommt ein weiteres Problem, das sich erst bei Migrationen zu offenbaren pflegt. Wechselt man z.B. ein Datenbankmanagementsystem, sei es anlässlich eines Updates oder eines Umstiegs auf ein anderes Produkt, so werden Inkonsistenzen und Eingabefehler häufig erst zu diesem Zeitpunkt erkannt. Manuelle Nachbearbeitungen sind die Folge und damit erhebliche Personalkosten. Dem Autor sind zwei Beispiele derartiger Migrationen aus dem Bereich der Wirtschaft bekannt; in beiden Fällen überstiegen die tatsächlichen Kosten den Voranschlag um mehr als das Zehnfache.

Die Aufbewahrung digitalen Archivguts in digitaler Form wird also mit steigenden Kosten verbunden sein. Einsparungen, die sich aus dem im Vergleich zu Schriftgut auf Papier geringeren Platzbedarf ergeben, werden diese Kosten nicht wett machen. Man darf nicht übersehen, dass auch die Lagerung von Datenträgern Anforderungen an das Raumklima stellt. Dazu kommt die Notwendigkeit, mindestens eine Sicherungskopie des Datenbestands vorzuhalten. Bei Migrationen bedeutet dies, dass es zweier Läufe bedarf: Zuerst wird die „Masterkopie“ migriert, und erst nach der Überprüfung des migrierten Bestands auf Vollständigkeit und Korrektheit kann die Sicherungskopie migriert werden.

Auch die Benützung digitalen Archivguts ist mit einigem Aufwand verbunden. Die Vorstellung, dass ein Ausheber in das Datenmagazin geht, den oder die Datenträger herausucht, in ein Laufwerk einlegt und nach dem Auslesen der Daten wieder einstellt, ist aus organisatorischer und konservatorischer Sicht mehr als bedenklich. Ein falsch eingestellter Datenträger dürfte so unauffindbar sein wie ein verreihter Aktenband, und schon heute haucht so mancher Datenträger sein Leben aus, wenn er mit Hast und wenig Gefühl ins Laufwerk gedrückt wird. Auf der sicheren Seite ist man mit Jukeboxen oder gar Robotern, doch diese sind wartungsträchtig und von begrenzter Lebensdauer. Es ist wohl nicht zu gewagt, wenn man unter diesen Gesichtspunkten die Kosten für eine Aufbewahrung in digitaler Form höher veranschlagt als für die Lagerung und Bereitstellung klassischen Archivguts. Dann erscheint es jedoch fraglich, ob die für eine digitale Archivierungsform benötigten und kontinuierlich ansteigenden finanziellen Ressourcen von den Archivträgern überhaupt bereit gestellt werden können und ob man willens ist, dies zu tun.

Gibt es Möglichkeiten, digitale Unterlagen in einer anderen Form aufzubewahren, die dem Archiv diese – aus seiner Sicht – kurzfristigen Migrationsschübe erspart? Charles Dollar nennt zwar die Überlieferung auf Papier oder Mikrofilm als Surrogat, wenn es denn gar nicht anders zu machen sei, doch müsse in diesem Fall die Authentizität der digitalen Vorlage eigens neu geschaffen werden, und viele Funktionen könnten auf das Surrogat nicht abgebildet werden.⁴ Beispielhaft für derartige Funktionen ist der Hyperlink. Wenn er auf Film (das Surrogat Papier wird man bis auf wenige Ausnahmen nicht in Betracht ziehen) abgebildet ist, kann man ihn nicht mehr anklicken. Doch wie wird es sein, wenn man das digitale Pendant nach einigen Jahrzehnten am Bildschirm betrachtet und den Link aktiviert? Wohin wird er führen? Vermutlich ins Nichts.

Ein anderes Beispiel ist eine Datenbank. Man kann sie auf ein plattformunabhängiges flat file umsetzen, das recherchierbar bleibt, sofern man die Strukturinformationen, Kodierungen und Systemdokumentationen des Datenbanksystems dazu aufbewahrt. Theoretisch könnte man mit Hilfe dieser Informationen das flat file in ein aktuelles Datenbanksystem konvertieren,

³ Das von Charles M. Dollar (wie Anm. 1) S. 231–248 vorgestellte „Preservation metadata model“ veranschaulicht die Menge dieser Daten.

⁴ Charles M. Dollar (wie Anm. 1) S. 73 f.

das zumindest Recherchen und andere Auswertungen ermöglicht. Der Aufwand dürfte allerdings beträchtlich sein.⁵

Digitale Audio- und Videodateien auf das Surrogat Mikrofilm abzubilden ist zwar theoretisch möglich, aber wohl kaum sinnvoll. Doch zumindest im künftigen digitalen Verwaltungsschriftgut dürften solche keine nennenswerte Rolle spielen; hier dominieren wohl auch weiterhin Texte, Grafiken und Fotos. Das Bundesarchiv hat bekanntlich empfohlen – und diese Empfehlungen sind in das DOMEA-Konzept des Bundes eingegangen – Primärinformationen vor der Übergabe an das Archiv in das TIF-Format zu konvertieren und die Metadaten in Nur-Text-Dateien mit ASCII-Kodierung zu übergeben. Doch was ist eine TIFF-Datei anderes als ein digitales Foto? Der dort abgebildete Text liegt nicht mehr als Zeichenfolge, sondern als Bitmap vor. Damit enthält ein TIFF-Bild nicht mehr Informationen als eine von der Auflösung her adäquate Mikrofilmaufnahme. Was spricht also dagegen, auch die Metadaten auf Film auszubelichten, und zwar in einer Qualität, die eine fehlerfreie Rückdigitalisierung gewährleistet?⁶ Denn nur in den Metadaten kann recherchiert werden, nicht aber im TIFF-Bild, das eigentlich auch nichts anderes als ein Surrogat der ursprünglichen Word- oder Excel-Datei ist.

Die Technik für das automatisierte Ausbelichten digitaler Daten auf Mikrofilm ist vorhanden. So bietet z.B. Kodak den „Document Archive Writer“ an, und die Firma Microbox hat ein Gerät namens Polycom entwickelt, das ein konfigurierbares Selektieren und Herausziehen von Primärinformationen und Metadaten und eine strukturierte Abbildung auf Film erlaubt.

Die Benützung mikroverfilmter digitaler Unterlagen ist aufwändiger und umständlicher als der Umgang mit der digitalen Form. Doch der Mikrofilm bedarf keiner Migration. In größeren zeitlichen Abständen wird es notwendig sein, den Film zu kopieren oder – um die dabei unvermeidlichen Qualitätsverluste zu vermeiden – ihn zu redigitalisieren und davon einen neuen Film zu ziehen. Die Kosten hierfür werden nicht gering sein, doch sie sind kalkulierbar und fallen vielleicht alle 100 Jahre an. Laufende Kosten fallen natürlich für die sachgerechte Lagerung der Filme an, und so fragt man sich, ob es nicht andere Möglichkeiten gibt, ein gegenüber Umwelteinflüssen weitgehend unempfindliches und jederzeit mit einfachen Techniken lesbares Speichermedium zu erzeugen, das große Datenvolumina aufnimmt und dadurch wenig Platz beansprucht.

Schon vor etlichen Jahren gab es Pressemeldungen, wonach in Kalifornien eine Art CD aus feuer- und säurefestem Material (Iridium) entwickelt werde, in die man digitale Daten und Informationen zu ihrer Dekodierung in extremer Verkleinerung „eingravieren“ könne. Die digitalen Daten sollten in binärer Kodierung gleichsam „versteint“ werden, die Dekodierungsinformationen hingegen als analoger Klartext. Diese „CD“ sollte etwa 650 Gigabyte Daten aufnehmen können. Weiteres ließ sich zu diesem Projekt nicht in Erfahrung bringen, die Idee wurde jedoch anderweitig aufgegriffen.

Die vor einigen Jahren in den USA gegründete LongNow Foundation⁷ hat sich zum Ziel gesetzt, wichtige Schriftzeugnisse menschlicher Kultur über 10.000 Jahre hinweg zu bewahren (10kLibrary). Hinzu kam der Wunsch, diese Zeugnisse auch in sprachlicher Vielfalt zu sichern. Man wählte rund 1000 Sprachen aus und entwickelte für jede ein Informationspaket, das künftigen Linguisten die Möglichkeit geben soll, die Sprache auch nach ihrem Untergang zu rekonstruieren. Als Textbeispiel, das in jeder dieser Sprachen vorliegt, dienen die ersten drei Kapitel der Genesis.

Für einen ersten Versuch entschied man sich für eine Dokumentation der 1000 Sprachen, die jeweils 27 Druckseiten umfasst, also insgesamt 27.000 Seiten. Hinzu kam der vollständige

⁵ Vgl. den Beitrag von Björn Dehms, Andreas Engel und Ulrich Meyer in diesem Band.

⁶ Dies ist z.B. durch die Wahl eines OCR-Fonts (OCR-A oder OCR-B) vor dem Ausbelichten zu erreichen.

⁷ Vgl. <http://www.longnow.org/10klibrary/rosetta.html>

Text der Genesis in acht wichtigen Sprachen mit etwa 3000 Seiten. All dieses sollte auf einem kleinen Speichermedium Platz finden, und zwar in analoger und beständiger Form, zu lesen mit einem optischen Mikroskop mit dem Vergrößerungsfaktor tausend. Das Medium bekam den Namen HD-Rosetta: HD steht für hard disk und spielt auf die Robustheit des Mediums an, Rosetta auf die parallele Textüberlieferung.

Tatsächlich ist es der Firma Norsam Technologies⁸ gelungen, einen diesen Anforderungen genügenden Datenträger zu entwickeln. In der jetzigen Ausprägung handelt es sich um eine quadratische Nickelplatte mit 5 cm Kantenlänge und einer Dicke von knapp über 6 mm. Daten werden in analoger Form in einer Seitengröße von 21,59 mal 17,9 cm (bei normaler Vergrößerung) abgebildet. Wenn der Datenträger mit optischen Mikroskopen gelesen werden soll, fasst er 5000 bis 18.000 Seiten, steht jedoch ein Elektronenmikroskop zur Verfügung, so können etwa 196.000 Seiten auf dem Medium untergebracht werden. Der Prototyp eines Lesegeräts, das das Mikroskopbild auf einen Computermonitor überträgt (also redigitalisiert!), sollte zum Verkaufspreis von etwa 10.000 US-\$ in Serie gehen. Da die HD-Rosetta jedoch ursprünglich rund war und die Abbildungen in Form einer Spirale angeordnet waren, was sich nicht bewährt hat, wechselte man zur quadratischen Form, die auch eine einfache Navigation über ein Koordinatensystem ermöglicht. Hierfür wird zur Zeit ein spezielles Lesegerät entwickelt.

Unempfindlichkeit und Haltbarkeit des Mediums sind beachtlich. Das Material ist im Wesentlichen Nickel, das gegen elektromagnetische Felder unempfindlich ist (also auch gegen die berüchtigte „elektromagnetische Bombe“). Testberichte des Los Alamos National Laboratory zeigen, dass Temperaturen bis zu 300° C keine Veränderungen des Mediums hervorrufen; erst ab 450° C oxidierte die Oberfläche und machte die Daten teilweise unlesbar. Als man die HD 15 Wochen lang Salzwasser, Leitungswasser oder Meeresluft aussetzte, war keine die Lesbarkeit beeinflussende Korrosion festzustellen. Die Lebensdauer des Mediums wird für normale klimatische Bedingungen mit mehreren tausend Jahren angegeben.

Die HD-Rosetta wird mit einem fokussierten Ionenstrahl beschrieben, der Galliumionen auf die Nickeloberfläche schießt. Die Ionen ätzen Atome aus der Oberfläche heraus und „gravieren“ damit Vertiefungen. Als Datenquelle dienen Mikrofilme oder TIFF-Dateien. Das Schreiben einer Seite dauert weniger als eine Zehntelsekunde; in zwei Stunden sind etwa 7000 Seiten eingätzt. Den Schreibprozess kann derzeit nur die Firma Norsam durchführen; das Verfahren ist patentiert.

Leider gibt die Firma auf ihrer Webseite keine Hinweise auf Praxistests mit Anwendern und nennt auch keine Referenzen. Es bleibt also abzuwarten, ob und wann das Verfahren auf den Markt gelangen wird.

Der Traum vom Archiv, das in einen Stülpedeckelkarton passt und bedenkenlos überall hingestellt werden kann, bleibt also noch unerfüllt. Man sollte diese und andere Entwicklungen jedoch sorgfältig und unvoreingenommen beobachten.

⁸ Vgl. <http://www.norsam.com>. Leider erhielt der Autor trotz mehrerer Anfragen – schriftlich und als E-Mail – bis heute von der Firma keine Antwort.

Programm

5. März 2001, Vormittag. Moderation: Dr. Karl-Ernst Lupprian, Generaldirektion der Staatlichen Archive Bayerns		
09:30	Begrüßung	<i>Prof. Dr. Hermann Rumschöttel, Generaldirektor der Staatlichen Archive Bayerns</i>
09:45–10:15	Das Fachkonzept zur Aussonderung elektronischer Akten in FAVORIT-Office-Flow	<i>Dr. Michael Wettengel, Bundesarchiv</i>
10:15–10:45	Einführung und Einsatz von VIS im Geschäftsbereich des Bayerischen Staatsministeriums für Landesentwicklung und Umweltfragen	Ministerialrat <i>Klaus-Peter Volkmann</i> , Bayerisches Staatsministerium für Landesentwicklung und Umweltfragen
11:30–12:15	Der Elektronische Rechtsverkehr und die elektronische Akte – Bericht zum Stand der Erhebungen und der Vorgaben des Arbeitskreis „Querschnittsfragen Technik“	Regierungsdirektor <i>Jürgen Ehrmann</i> , Justizministerium Baden-Württemberg
12:15–12:45	Das Projekt ELGAR zur elektronischen Unterstützung von Geschäfts- und Aktenregistrierungsprozessen im Kanton Basel-Stadt	<i>Thomas Schärli</i> , Staatsarchiv des Kantons Basel-Stadt
5. März 2001, Nachmittag. Moderation: Dr. Carsten Müller-Boysen, Landesarchiv Schleswig-Holstein		
14:45–15:15	Bewertung elektronischer Statistiken	<i>Dr. Christian Keitel</i> , Landesarchivdirektion Baden-Württemberg
15:15–15:45	Empfehlungen für die elektronische Archivierung. Ein Normentwurf der Association Française de Normalisation	<i>Dr. Frank M. Bischoff</i> , Nordrhein-Westfälisches Staatsarchiv Münster
16:30–17:00	Zum DOMEA-Konzept und seiner Umsetzung im Freistaat Sachsen	<i>Dr. Thekla Kluttig</i> , Sächsisches Staatsministerium des Innern
17:00–17:30	Rematerialisierung digitaler Unterlagen – ein Weg zur Langzeitarchivierung? Das Projekt HD-Rosetta	<i>Dr. Karl-Ernst Lupprian</i> , Generaldirektion der Staatlichen Archive Bayerns
Dienstag, 6. März 2001. Moderation: Dr. Michael Wettengel, Bundesarchiv		
08:30–09:00	XML – ein Format zur dauerhaften Aufbewahrung und Nutzung einfacher, relationaler Datenbanken	<i>Dipl.-Inf. Ulrich Meyer</i> , Universität Koblenz
09:00–09:30	Archivierung von digitalen Publikationen in Bibliotheken	<i>Dr. Marianne Dörr</i> , Münchener Digitalisierungszentrum
10:15–10:45	Erschließung und Nutzbarmachung digitaler Überlieferung im Bundesarchiv	<i>Bettina Martin-Weber</i> , Bundesarchiv
10:45–11:15	Aktuelle internationale Ansätze für den Einsatz von Metadaten-Sets in den Bereichen Schriftgutverwaltung und Archiv	<i>Dr. Nils Brübach</i> , Archivschule Marburg

Teilnehmerverzeichnis

Name	Institution
Dr. Christoph Bachmann M.A.	Staatsarchiv München
Dr. Frank M. Bischoff	Nordrhein-Westfälisches Staatsarchiv Münster
Alois Blabsreiter	Bayerischer Verwaltungsgerichtshof – IuK-Stelle
Dr. Mechthild Black-Veldtrup	Nordrhein-Westfälisches Hauptstaatsarchiv
Dr. Nils Brübach	Archivschule Marburg
Dr. Edgar Büttner	Bundesarchiv Koblenz
Dr. Marianne Dörr	Münchener Digitalisierungszentrum, Bayerische Staatsbibliothek
Jürgen Ehrmann	Justizministerium Baden-Württemberg – IuK-Leitstelle
Myriam Erwin	Schweizerisches Bundesarchiv
Gerhard Fritsch	Bayerisches Staatsministerium des Innern
Horst Gehringer	Stadtarchiv München
Bernhard Grau M.A.	Bayerisches Hauptstaatsarchiv
Günter Handl	Bayerischer Verwaltungsgerichtshof, IuK-Stelle
Ilka Hebig	Brandenburgisches Landeshauptarchiv
Dr. Andreas Hedwig	Hessisches Hauptstaatsarchiv Wiesbaden
Dr. Barbara Hoen	Staatsarchiv Ludwigsburg
Dipl.-Ing. Dietrich Hofmaier	MFM Dipl.-Ing. Dietrich Hofmaier GmbH & Co., München
Dr. Herbert Hutterer	Österreichisches Staatsarchiv, Abt. Finanz- und Hofkammerarchiv
Werner Jürgensen M. iur. utr.	Landeskirchliches Archiv Nürnberg
Dr. Christian Keitel	Landesarchivdirektion Baden-Württemberg
Dr. Thekla Kluttig	Sächsisches Staatsministerium des Innern
Dr. Margit Ksoll-Marcon M.A.	Generaldirektion der Staatlichen Archive Bayerns
Dr. Karl-Ernst Lupprian	Generaldirektion der Staatlichen Archive Bayerns
Bettina Martin-Weber	Bundesarchiv Koblenz
Robert Meier	Oberlandesgericht München - Gemeinsame IT-Stelle der bayerischen Justiz
Ulrich Meyer	Institut für Wirtschafts- und Verwaltungsinformatik der Universität Koblenz
Dr. Carsten Müller-Boysen	Landesarchiv Schleswig-Holstein
Dr. Ulrich Nieß	Stadtarchiv Mannheim

Name	Institution
Ulf Rathje	Bundesarchiv Koblenz
Dr. Lothar Saupe	Bayerisches Hauptstaatsarchiv
Thomas Schärli	Staatsarchiv Basel-Stadt
Michael Schmidbauer	Regierung von Schwaben, Augsburg
Rudolf Steidl	Bayerischer Verwaltungsgerichtshof
Werner Stubenvoll	infowerk ag, Nürnberg
Angela Ullmann	Sächsisches Hauptstaatsarchiv
Klaus-Peter Volkmann	Bayerisches Staatsministerium für Landesentwicklung und Umweltfragen
Dr. Michael Wettengel	Bundesarchiv Koblenz
Dr. Andrea Wettmann	Sächsisches Hauptstaatsarchiv
Dr. Robert Zink	Stadtarchiv Bamberg

Anhang

Inhaltsübersicht über die bisher erschienenen Tagungsbände des Arbeitskreises „Archivierung von Unterlagen aus digitalen Systemen“

Archivierung von Unterlagen aus digitalen Systemen. Beiträge zur Tagung im Staatsarchiv Münster 3.–4. März 1997, hrsg. von Frank M. Bischoff (Veröffentlichungen der staatlichen Archive des Landes Nordrhein-Westfalen, Reihe E: Beiträge zur Archivpraxis, Heft 4), Münster 1997

WILFRIED REININGHAUS, Zur Einführung (S. 7–8)

MICHAEL WETTENGEL, Digitale Signaturen und Pilotprojekte zur IT-gestützten Vorgangsbearbeitung in der Bundesverwaltung (S. 9–20)

GUDRUN FIEDLER, Archivierung von Unterlagen aus digitalen Systemen am Beispiel des Landes Niedersachsen (S. 21–29)

UDO SCHÄFER, Büroautomation in der Landesverwaltung Baden-Württemberg. Strategisches und operatives archivarisches Handeln am Beispiel der Justiz (S. 31–48)

FRANK M. BISCHOFF, Elektronische Registratursysteme. Überlegungen zur Archivierung digitaler Unterlagen der nordrhein-westfälischen Verwaltungsgerichte (S. 49–67)

VOLKER BUCHHOLZ, Kataster – ein Auslaufmodell in Archiven? (S. 69–74)

RALF STREMMEL, Elektronische „Archivierung“ in den Industrie- und Handelskammern – Konzepte und Auswirkungen auf die Wirtschaftsarchive (S. 75–84)

MICHAEL HÄUSLER, Das Informationsmanagement-System KIDAT als Beispiel des Einsatzes elektronischer Systeme in Verwaltungen der evangelischen Kirche (S. 85–94)

ROLF-DIETRICH MÜLLER, Bericht über die Tätigkeit des Unterausschusses EDV der Bundeskonferenz der Kommunalarchive (BKK) (S. 95–101)

ULRICH NIESS, Sachstandsbericht über die Mannheimer Arbeitsgruppe Schriftgutverwaltung und Büroautomation (S. 103–115)

ANJA GUSSEK-REVERMANN, Die Einführung der „Digitalen Archivierung“ bei der Stadtverwaltung Münster – Mitwirkung des Stadtarchivs (S. 117–133)

HANS-HEINRICH EBELING, Nutzung, Verwaltung und Sicherung archivisch generierter elektronischer Unterlagen am Stadtarchiv Duderstadt – ein Rückblick 1989–1997 (S. 135–148)

HANS-REINHARD FRICKE, Digitale Erschließung von Archivbeständen am Beispiel der mittelalterlichen und frühneuzeitlichen Amtsbücher des Stadtarchivs Duderstadt (S. 149–163)

BRUNO MORS, Elektronisches Archivkonzept in der Westdeutschen Landesbausparkasse (S. 165–174)

CARSTEN MÜLLER-BOYSEN, „Elektronisk arkivering“. Die dänischen Richtlinien zur Anbietung und Übernahme maschinenlesbarer Daten aus Büroautomationssystemen (S. 175–190)

FRANK M. BISCHOFF – SUSANNE BROCKFELD, Zusammenfassung der Diskussion (S. 191–202)

Verzeichnis der Autoren (S. 203)

Archivierung elektronischer Unterlagen, 2. Arbeitstagung, hrsg. von Udo Schäfer und Nicole Bickhoff (Werkhefte der staatlichen Archivverwaltung Baden-Württemberg, Serie A Landesarchivdirektion, Heft 13), Stuttgart 1999

WILFRIED SCHÖNTAG, Vorwort (S. 7–8)

UDO SCHÄFER, Einleitung (S. 9–12)

Der Einsatz der Informationstechnik und seine Folgen für die Archive:

CARSTEN MÜLLER-BOYSEN, Das Archiv als „Informationsrecycling“. Gedanken zur Neudefinition archivischer Arbeitsfelder (S. 15–24)

RUTH KAPPEL, Neue Medien – Archiv im Wandel? Ein GEHE-Praxisbericht (S. 25–32)

Archivische Überlieferungsbildung aus elektronischen Unterlagen. Konzeption und Strategien:

THOMAS LUX (unter Mitwirkung von Thekla Kluttig und Angela Ullmann), Konzeptionen zur archivischen Betreuung von DV-Anwendungen in der Verwaltung des Freistaates Sachsen. Projekte und Perspektiven (S. 35–44)

ILKA HEBIG, Die Vorbereitung auf die Archivierung digitaler Unterlagen im Land Brandenburg (inhaltliche und organisatorische Fragen). Ein Erfahrungsbericht (S. 45–49)

KARLJOSEF KRETER, Modelle für den Umgang mit digitalem Schriftgut in Kommunalarchiven (S. 51–64)

Die IT-gestützte Vorgangsbearbeitung in der öffentlichen Verwaltung:

ANDREAS ENGEL – ANDREA KERN, Die elektronische Akte als Ziel der IT-gestützten Vorgangsbearbeitung. Erfahrungen aus dem Projekt DOMEA (Dokumentenmanagement und elektronische Archivierung im IT-gestützten Geschäftsgang) (S. 67–78)

NIKLAUS BÜTIKOFER, Anforderungen an die Aktenführung in konventionellen und digitalen Informations- und Kommunikationssystemen (S. 79–84)

ANDREAS HEDWIG, Via Internet zum papierlosen Büro? Ein Pilotprojekt zur sicheren Informationsübermittlung in der hessischen Landesverwaltung (S. 85–91)

Das elektronische Grundbuch:

ANDREAS BURKHARDT, FOLIA®. Die Basis des elektronischen Grundbuchs in Baden-Württemberg (S. 95–100)

FRANK M. BISCHOFF, Elektronisches Grundbuch in Nordrhein-Westfalen. Möglichkeiten der Überlieferungssicherung aus archivischer Perspektive (S. 101–110)

Geographische Informationssysteme in der öffentlichen Verwaltung:

UDO SCHÄFER, Geographische Informationssysteme in der Landesverwaltung Baden-Württemberg aus archivischer Perspektive (S. 113–129)

NICOLE BICKHOFF – CLEMENS REHM, Das Automatisierte Liegenschaftskataster in Baden-Württemberg (S. 131–143)

FRANZ-JOSEF ZIWES, Überlegungen zur Bewertung von digitalen Unterlagen aus dem Umweltinformationssystem Baden-Württemberg (S. 145–151)

GUDRUN FIEDLER, Archivierung digitaler Katasterunterlagen. Die Fortführung eines Archivierungsmodells der niedersächsischen Staatsarchive (S. 153–162)

Digitale Aufzeichnungen und Authentizität:

UDO SCHÄFER, Authentizität. Vom Siegel zur digitalen Signatur (S. 165–181)

FRANK M. BISCHOFF, Zur Archivfähigkeit digitaler Signaturen in elektronischen Registern (S. 183–198)

THOMAS SCHÄRLI, Authentische Überlieferungsbildung in elektronischer Systemumgebung. Eine Standortbestimmung aus schweizerischer Perspektive (S. 199–210)

Bilanzen und Perspektiven:

IVAR FONNES, Methoden zur Langzeiterhaltung elektronischer Informationen (S. 213–222)

MICHAEL WETTENGEL, Archivierung digitaler Datenbestände aus der DDR nach der Wiedervereinigung (S. 223–239)

NICOLE BICKHOFF – FRANZ MÖGLE-HOFACKER, Archivierung von Unterlagen aus digitalen Systemen. Bilanz und Perspektive (S. 241–245)

Digitale Herausforderungen für Archive. 3. Tagung des Arbeitskreises „Archivierung von Unterlagen aus digitalen Systemen“ am 22. und 23. März 1999 im Bundesarchiv in Koblenz, hrsg. von Michael Wettengel (Materialien aus dem Bundesarchiv, Heft 7), Koblenz 1999

FRIEDRICH P. KAHLENBERG, Vorwort (S. 5–6)

MICHAEL WETTENGEL, Einleitende Bemerkungen (S. 7–8)

HANS HOFMAN, Herausforderungen an die digitale Erhaltung (S. 9–13)

FRANK M. BISCHOFF, Emulation – das Archivierungskonzept der Zukunft? (S. 15–23)

RAPHAEL OSTERMANN, Potentielle Dateiformate zur Langzeitarchivierung von Dokumenten unter Berücksichtigung von Primär- und Metainformationen (S. 25–35)

KARL-ERNST LUPPRIAN, Die Archivierung digitaler Unterlagen in einer heterogenen IT-Landschaft. Ein Beispiel für archivische Informationspolitik (S. 37–39)

HORST ROMEYK, Archivische Anforderungen an die IT-gestützte Vorgangsbearbeitung und an elektronisch geführte Register bzw. Elektronische Schriftgutverwaltung, Aktenbildung und Vorgangsbearbeitung (S. 41–44)

ANDREAS HEDWIG, Tendenzen der IT-gestützten Vorgangsbearbeitung in der Hessischen Landesverwaltung (S. 45–50)

THOMAS SCHÄRLI, Das archivische „Lebensphasen“-Modell im Kontext elektronischer Informationssysteme. Ein Werkstattbericht des Staatsarchivs Basel-Stadt als Beitrag zur Methodendiskussion im Bereich der Archivierung digitaler Unterlagen (S. 51–58)

MATHIAS SCHILLER, Telekooperation in der Verwaltung – das Einzelprojekt POLIVEST als Bestandteil von POLIKOM. Erfahrungen aus dem Einsatz in Sachsen-Anhalt (S. 59–60)

UDO SCHÄFER, Das Recht auf Zugang zu Informationen des öffentlichen Sektors in seinem Verhältnis zur Archivierung analoger und digitaler Aufzeichnungen (S. 61–78)

ANDREAS ENGEL – ANDREA KERN, DOMEA. Pilotprojekt der Bundesverwaltung zum Dokumentenmanagement und zur elektronischen Archivierung im IT-gestützten Geschäftsgang (S. 79–88)

MICHAEL WETTENGEL, Aussonderung und Archivierung elektronischer Akten (S. 89–96)

STEFAN GEMPERLI, Die Arbeitsgruppe „Archivierung elektronischer Akten“ des Vereins Schweizerischer Archivarinnen und Archivare (VSA) (S. 97–100)

ULRICH NIESS, Archive als Dienstleister IT-gestützter Vorgangsbearbeitung in der Verwaltung: Praktische Beispiele aus dem Stadtarchiv Mannheim (S. 101–106)

UDO SCHÄFER, Anforderungen an die Entwicklung des Aussonderungsmoduls für die IT-gestützte Vorgangsbearbeitung (S. 107–115)

MICHAEL WETTENGEL, IT-gestützte Vorgangsbearbeitung und archivische Anforderungen (S. 117–124)

Anhang: Generaldirektion der Staatlichen Archive Bayerns (Hrsg.), Digitale Unterlagen. Entstehung – Verwaltung – Archivierung. Empfehlungen für die Behörden des Freistaats Bayern, München 1998 (Entwurfssfassung) (S. 127–136)

Kurzbiografien der Autoren (S. 137–139)

Auf der Suche nach archivischen Lösungsstrategien im digitalen Zeitalter. Beiträge zur 4. Jahrestagung des Arbeitskreises „Archivierung von Unterlagen aus digitalen Systemen“ im Stadtarchiv Mannheim 10.–11.4.2000, hrsg. von Ulrich Nieß (Sonderveröffentlichungen des Stadtarchivs Mannheim Nr. 26), Mannheim 2001 [dem Band liegt eine CD-ROM bei]

GERHARD WIDDER, Zum Geleit (S. 7)

ULRICH NIESS, Vorwort (S. 9–10)

DIETMAR H. APPEL – ULRICH NIESS, Das Mannheimer Moderatorenmodell: Von der Datenmodellierung zur Online-Recherche (S. 11–24)

THEKLA KLUTTIG – ANGELA ULLMANN, Digitale Registerdaten als Anbietungsliste – Neue Möglichkeiten bei der Bewertung von Verfahrensakten der Staatsanwaltschaften (S. 25–30)

BETTINA WISCHHÖFER, Projekt Digitalisierung Fotosammlung „Kirchen der Evangelischen Kirche von Kurhessen-Waldeck“ – eine „low-budget-Lösung“ (S. 31–34)

MICHAEL WETTENGEL, Verwaltungsmodernisierung und IT-gestützte Vorgangsbearbeitung in der Bundesverwaltung (S. 35–46)

ANDREAS HEDWIG, Zum Stand des elektronischen Grundbuchs in Hessen (S. 47–53)

HANS-WERNER LANGBRANDTNER, Derzeitiger Stellenwert digitaler Überlieferungsbildung in rheinischen Kommunen (S. 55–61)

CARSTEN MÜLLER-BOYSEN, Flat file oder virtuelle Behörde? – Was erwartet der Benutzer? (S. 63–68)

JOHANNES KEUSEKOTTEN – GABRIELE SAUERMANN, Erfolgreiche Verwaltung mit Favorit7-OfficeFlow7 (S. 69–72)

ANDREAS ENGEL, Elektronische Archivierung und Nutzung (ELAN). Eine Projektidee zur Entwicklung von Verfahren für die dauerhafte Archivierung und Nutzung von elektronischen Geschäftsunterlagen aus Behörden und Betrieben (S. 73–84)

CHRISTOPH POPP, Dienstleistung via Dokumentenmanagement am Beispiel der Mieterakten der GBG – Mannheimer Wohnungsbau mbH: konzeptionelle Überlegungen zur Frage der Langzeitarchivierung (S. 85–90)

KARL-ERNST LUPPRIAN, Ein neues „Jahrtausendproblem?“ Forschungsprojekt zur Pflege und Archivierung digitaler Unterlagen der Verwaltung (S. 91–94)

MICHAEL WETTENGEL, Australische Metadaten-Standards: Ein Thema für Archive? (S. 95–103)

NILS BRÜBACH, Internationale Normen für die Verwaltung von Unterlagen aus digitalen Systemen: Die ISO 15489 „Archives and Records Management“ und das MoReq-Projekt der Europäischen Union (S. 105–112)