

Keitel/Naumann,
Digitale Archivierung in der Praxis

WERKHEFTE
DER STAATLICHEN ARCHIVVERWALTUNG
BADEN-WÜRTTEMBERG

Herausgegeben
vom Landesarchiv Baden-Württemberg

Serie A Heft 24

2013
Verlag W. Kohlhammer Stuttgart

Digitale Archivierung in der Praxis

16. Tagung des Arbeitskreises „Archivierung von
Unterlagen aus digitalen Systemen“ und
nestor-Workshop „Koordinierungsstellen“

Herausgegeben von
Christian Keitel und Kai Naumann

2013
Verlag W. Kohlhammer Stuttgart



Diese Publikation ist auf alterungsbeständigem, säurefreiem Papier gedruckt.

Alle Rechte vorbehalten

© 2013 by Landesarchiv Baden-Württemberg, Stuttgart

Umschlaggestaltung: agil > Visuelle Kommunikation, Pforzheim

Kommissionsverlag: W. Kohlhammer GmbH, Stuttgart

Gesamtherstellung: NINO Druck GmbH, Neustadt/Weinstraße

Printed in Germany

ISBN 978-3-17-022534-3

Inhalt

Vorwort.....	9
<i>Christian Keitel und Kai Naumann</i>	
Einleitung.....	11
16. Tagung des Arbeitskreises „Archivierung von Unterlagen aus digitalen Systemen“	
Kosten	
<i>Karlbeinz Schmitt</i>	
Kosten der digitalen Archivierung. Ein mögliches Vorgehensmodell und erste Erfahrungen.....	19
<i>Susanne Fröhlich</i>	
Kostenfragen in digitalen Archiven. Erfahrungen des Digitalen Archivs Österreich.	31
<i>Gabriele Stüber</i>	
Archivierung von digitalisierten Archivalien in einem Rechenzentrum	51
<i>Peter Sandner</i>	
10 FAQs. Argumente zu Bedarf und Notwendigkeiten der digitalen Archivierung.	57
Übernahme	
<i>Bernhard Rieder</i>	
Datenübernahme aus dem Polizei-Informationssystem (POLIS).....	71
<i>Sigrid Schieber</i>	
Entwicklung einer Aussonderungsschnittstelle für das DMS DOMEA.....	85
<i>Corinna Knobloch</i>	
Neu im Digitalen Archiv. Ein Streifzug durch Bestände und Methoden	97
Systeme und Tools	
<i>Katharina Ernst und Heike Maier</i>	
Vom Magnetband auf Content Addressed Storage. Speichermigration im Stadtarchiv Stuttgart	111

<i>Burkhardt Nolte und Karsten Huth</i>	
Einführung der elektronischen Archivierung im Sächsischen Staatsarchiv.	119
<i>Rolf Lang</i>	
Die elektronische Grundakte in G-DIMAG	129
<i>Christoph Schmidt</i>	
Zwischen User-Group und Entwicklungsgemeinschaft. Konstituierung, Aufgaben und Arbeitsperspektiven einer spartenübergreifenden Nutzergruppe	143
<i>Christian Keitel</i>	
DIMAG-Kooperationen	147
<i>Ilka Stahlberg und Jörg Homberg</i>	
Das digitale Urkundszwischenarchiv (EL.UZA) in Brandenburg. Ein Arbeitsbericht. . . .	157
Objektarten	
<i>Mike Zuchet</i>	
Pilotprojekt zur Langzeitarchivierung digitaler E-Mail-Korrespondenzen des Bundesvorstands der Vereinigten Dienstleistungsgewerkschaft ver.di	165
<i>Kai Naumann</i>	
Auf dem Weg zum Retro-GIS? Dauerhafte Erhaltung und Nutzbarmachung digitaler Geobasisdaten beim Landesarchiv Baden-Württemberg und darüber hinaus	171
<i>Claire Röthlisberger-Jourdan</i>	
Formaterkennung und Formatvalidierung. Theorie und Praxis	193
<i>Peter Bohl und Johannes Renz</i>	
AV-Unterlagen für DIMAG.	211
Grundsätze und Perspektiven	
<i>Felix Stadler</i>	
Erfahrungen bei der Übernahme von digitalen Unterlagen aller Art im Staatsarchiv St.Gallen – und kritischer Rückblick	221
<i>Steffen Schwalm</i>	
Der Nachfolger des DOMEA®-Konzepts. Das Organisationskonzept elektronische Verwaltung und seine Auswirkungen auf die elektronische Archivierung	231

Lambert Kansy und Markus Loch

Archivierung digitaler Unterlagen im Staatsarchiv Basel-Stadt. Rückblick auf den
Infrastrukturaufbau und erste Erfahrungen 253

Christian Keitel

Der nestor-Leitfaden zur Digitalen Bestandserhaltung und seine Folgen für die Archive . . 267

nestor-Workshop „Brauchen wir Koordinierungsstellen für die digitale Archivierung?“

Christian Keitel

Warum ist Kooperation bei der digitalen Archivierung unumgänglich? 281

Georg Büchler

Die KOST: Sieben Jahre Good Practice 289

Peter Worm

Erfahrungen des LWL-Archivamts bei der Koordination von Projekten zum
Langzeiterhalt elektronischer Daten 297

Ulrich Schludi

Brauchen wir Koordinierungsstellen für die digitale Archivierung?
Zusammenfassung der Diskussion 311

Autorinnen und Autoren 317

Vorwort

Mit der vorliegenden Publikation werden die Beiträge und Ergebnisse von zwei Veranstaltungen im Druck vorgelegt, die das Landesarchiv Baden-Württemberg 2012 zum *Thema Archivierung digitaler Unterlagen* durchgeführt hat.

Am 13. und 14. März 2012 fand in Ludwigsburg die 16. Tagung des *Arbeitskreises Archivierung von Unterlagen aus digitalen Systemen* statt. Damit kam im dortigen Staatsarchiv ein Gremium zusammen, das schon seit mehr als anderthalb Jahrzehnten einen archivübergreifenden und praxisbezogenen Austausch über alle Fragen pflegt, die mit der Archivierung genuin digitaler Unterlagen verbunden sind, und auf diesem Arbeitsfeld regelrechte Pionierarbeit geleistet hat. Denn das Thema hat als solches erst in der allerjüngsten Zeit im *mainstream* des Fachdiskurses die Aufmerksamkeit gewonnen, die ihm zukommt. Lange Zeit noch wurde es eher in kleinen und spezialisierten Zirkeln, zu denen der *Arbeitskreis Archivierung von Unterlagen aus digitalen Systemen* zählt, behandelt. So blieb auch die in diesen Kreisen als selbstverständlich gebrauchte spezifische Terminologie vielen Kolleginnen und Kollegen, die noch ganz in der Welt der klassischen Archive lebten, weitgehend unverständlich. Erst jetzt, nachdem eine Reihe von Archiven den Schritt in die Praxis der digitalen Archivierung gewagt hat oder unmittelbar davor steht, sieht man das Thema breiter aufgestellt und fasst die Erkenntnis Fuß, dass sich alle Archivarinnen und Archivare damit auseinandersetzen müssen. Dass sich im Jahr 2012 der 72. Südwestdeutsche Archivtag in Bad Bergzabern unter dem Motto *Das neue Handwerk. Digitales Arbeiten in kleinen und mittleren Archiven* erstmals in seiner Geschichte ausschließlich mit der digitalen Archivierung befasst hat, ist Ausdruck davon.

Bei der breiteren Aufstellung bleibt freilich noch viel zu tun. Das zeigte sich auf dem Südwestdeutschen Archivtag in Bad Bergzabern und ist nicht zuletzt an der Resonanz auf Fortbildungsangebote auf diesem Arbeitsfeld zu erkennen, wie sie unter anderem der *VdA – Verband deutscher Archivarinnen und Archivare e.V.* seit einigen Jahren auf dem Deutschen Archivtag anbietet. Sicher muss auch in der Ausbildung das Thema noch stärker verankert werden.

In welchem Maße hier ein Informations- und Diskussionsbedarf besteht, belegt auch die hohe Teilnehmerzahl an der zweiten Veranstaltung, deren Beiträge und Ergebnisse im vorliegenden Band veröffentlicht sind: Am 18. Juli 2012 fand im Hauptstaatsarchiv Stuttgart ein *nestor-Workshop* statt zu der Frage *Branchen wir Koordinierungsstellen für die digitale Archivierung?* Sie war bereits auf mehreren Tagungen zuvor angesprochen worden – so schon 2011 auf einem Kolloquium der Archivschule Marburg / Hochschule für Archivwesen, sodann 2012 in der Ludwigsburger Zusammenkunft des *Arbeitskreises Archivierung von Unterlagen aus digitalen Systemen* wie auch auf dem 72. Südwestdeutschen Archivtag. Besonders bei kommunalen Archiven war sie auf großes Interesse gestoßen. So standen Überlegungen zur Notwendigkeit und möglichen Ausgestaltungsformen von Koordinierungsstellen im Zentrum des *nestor-Workshops*, in dessen Verlauf sich immer deutlicher abzeichnete, dass in Deutschland auf dem Feld der digitalen Archivierung ein starkes Bedürfnis nach Koordination besteht, die am besten regional über bereits bestehende oder neue zu schaffende Verbände auszugestalten ist. Die vorliegende Publikation lässt diese Diskussion nachvollziehen. Es bleibt zu hoffen, dass sie an verschiedenen Stellen in konkrete Konzeptionen einmündet.

Wegen des engen inhaltlichen Bezugs zur Tagung des *Arbeitskreises Archivierung von Unterlagen aus digitalen Systemen* wie auch angesichts der sehr grundsätzlichen Fragestellung zur Organisation und Kooperation der Archive in der digitalen Welt, erschien es sinnvoll, die Beiträge und Ergebnisse des *nestor-Workshops* zusammen mit den Referaten des Arbeitskreises in einem Band zu publizieren. Die Veröffentlichung schließt sich damit den bisherigen Tagungsbänden des Arbeitskreises an. In der Reihe der *Werkhefte der staatlichen Archivverwaltung Baden-Württemberg* gesellt sie sich zugleich zu deren Band A 13, in dem die Ergebnisse der zweiten Tagung des Arbeitskreises publiziert wurden, die 1998 ebenfalls im Staatsarchiv Ludwigsburg stattfand.

Mein herzlicher Dank gilt allen, die an den beiden Veranstaltungen mitgewirkt und ihre Beiträge für den Druck zur Verfügung gestellt haben. In gleicher Weise danke ich den beiden Herausgebern, Herrn Dr. Christian Keitel und Herrn Dr. Kai Naumann, sowie Dr. Regina Keyler, die im Landesarchiv die Drucklegung besorgt hat.

Ich würde mich freuen, wenn der Band die notwendige Diskussion weiter befruchtet.

Stuttgart, den 1. Dezember 2012

Prof. Dr. Robert Kretzschmar
Präsident des Landesarchivs Baden-Württemberg

Einleitung

Von CHRISTIAN KEITEL und KAI NAUMANN

Vor gut zehn Jahren galt unter Archivarinnen und Archivaren die Beschäftigung mit der digitalen Archivierung als sehr abseitiges und geradezu fragwürdiges Unternehmen. Was bringe es, so die mehrheitlich geäußerte Meinung, sich mit einem sehr kleinen und begrenzten Spezialgebiet zu beschäftigen, das zwar sicherlich wichtig sei, auf absehbare Zeit aber nur ein Kümmerdasein fristen werde. Wer auch nur überlegte, ob er sich näher mit der digitalen Archivierung beschäftigen sollte, wurde gewarnt. Die Hinwendung zu einem derartigen Orchideenfach behindere nur den Erfolg in wichtigeren Arbeitsfeldern, und im Übrigen würden sich die Informatiker schon der Sache annehmen.

Heute ist die digitale Archivierung längst zu einem wesentlichen Thema innerhalb der deutschsprachigen Archivwelt geworden. Beigetragen hat zu dieser Entwicklung nicht zuletzt der Arbeitskreis *Archivierung von Unterlagen aus digitalen Systemen*. Bereits vor fünfzehn Jahren hat der Arbeitskreis seine ersten Tagungen abgehalten. Am 13. und 14. März 2012 konnte das Landesarchiv nun schon die 16. Tagung ausrichten. Sechzehn Tagungen in fünfzehn Jahren machen deutlich, dass der Arbeitskreis kontinuierlich sein Thema beackert hat. Kein Jahr verging ohne eine Tagung des Arbeitskreises. Seit dem letzten Jahr können die bisherigen Ergebnisse auf den Internetseiten des Staatsarchivs St. Gallen nachgelesen werden (<http://www.staatsarchiv.sg.ch/home/auds/>). Die Übersicht zeigt in den ersten Jahren eine beeindruckende Vielfalt an Themen, danach eine Konzentration auf Fragen der Schriftgutverwaltung und Systemeinführungen und seit etwa 2007 eine zunehmende Hinwendung zur praktischen Übernahme und Archivierung digitaler Unterlagen. Zeitgleich ging auch eine Ausweitung der interessierenden Objekte, über die elektronischen Akten hinaus, hin zu den Fachverfahren, Webseiten und anderen digitalen Objektarten einher. Der *Arbeitskreis Archivierung von Unterlagen aus digitalen Systemen* kann daher zu Recht als ein Ort verstanden werden, an dem sich Archivarinnen und Archivare der deutschsprachigen Archivwelt, aber auch Spezialisten aus angrenzenden Gebieten, besonders intensiv mit den Fragen beschäftigen, die sich im Themenkontext stellen. So besteht ein Zusammenhang mit dem Kooperationsprojekt *nestor*. *Nestor* hat als ein Netzwerk interessierter Archive, Bibliotheken und Museen seit 2002 zahlreiche Hürden auf dem Weg zur digitalen Archivierung beiseite geräumt. Das Landesarchiv hat daher den Workshop, den es am 18. Juli 2012 zum Thema *Brauchen wir Koordinierungsstellen zur digitalen Archivierung* veranstaltete, bewusst als *nestor-Workshop* annonciert. Es lag nahe, die Ergebnisse von beiden Veranstaltungen in einem Band zu vereinigen, zumal bereits auf der Frühjahrstagung des Arbeitskreises über das Thema diskutiert und der Workshop angekündigt worden war.

16. Tagung des Arbeitskreis Archivierung von Unterlagen aus digitalen Systemen

Die Frühjahrstagung wurde durch den Präsidenten des Landesarchivs *Robert Kretzschmar* eröffnet. Ein erster inhaltlicher Block widmete sich den Kosten. Seit langem wird einerseits beklagt, dass keine verlässlichen Kostenberechnungen zur digitalen Archivierung vorliegen. Andererseits haben verschiedene Publikationen neben einzelnen bruchstückhaften Überlegungen zu dem Thema vor allem gezeigt, wie komplex das Thema sich stellt. Es war daher besonders erfreulich, dass wir gleich mehrere Vorträge zu dieser Thematik anhören konnten. *Karlbeinz Schmitt* berichtete über die Ergebnisse des *dp4lib Projekts*, das von der Nationalbibliothek und der Staats- und Universitätsbibliothek Göttingen ausgerichtet wurde. Einmal mehr manifestierte sich auf dem Gebiet der allgemeinen und übertragbaren Kostenmodelle der Eindruck, dass Bibliotheken neue Themen häufig früher aufgreifen als die klassischen Archive. Allerdings konnte im Anschluss mit *Susanne Fröblich* eine Archivarin schon ein sehr konkretes praktisches Kostenmodell für das Österreichische Staatsarchiv vorstellen. Vergleichbare Überlegungen wurden bislang im deutschsprachigen Raum unseres Wissens nach noch nicht veröffentlicht. Besondere Probleme stellen sich bei der Frage, welche Kosten mit den vom Archiv erstellten Digitalisaten einhergehen. Dieser Frage geht *Gabriele Stüber* in ihrem Beitrag nach. Am Ende des Blocks antwortet *Peter Sandner* nicht nur auf die Kostenfrage, sondern auch auf die häufigsten anderen Fragen, die den Kolleginnen und Kollegen den Weg zur digitalen Archivierung zu erschweren oder sogar zu verbauen scheinen. Sein Aufsatz ist eine Pflichtlektüre für alle, die ihren Träger noch von ihren Vorhaben überzeugen müssen.

Lange Zeit wurde unter *digitale Archivierung* der gesamte Lebenszyklus digitaler Unterlagen verstanden, begonnen mit der Einführung von IT-Systemen über deren Betrieb und den sogenannten *Ingest* bis hin zur späteren Benutzung der archivierten Unterlagen. Erst nach der Übernahme werden digitale Unterlagen aber wirklich archiviert und für künftige Generationen gesichert. In der Praxis beginnt digitale Archivierung daher erst mit der Übernahme, und so ist es nur erfreulich, dass auch diese besonders schwierige Phase im Lebenszyklus digitaler Objekte durch einige Beiträge näher beleuchtet wurde. In diesen Block gehört der Beitrag von *Felix Stadler*, der zunächst die bisherigen Erfahrungen des Staatsarchivs St. Gallen vorstellte und dann auch einen kritischen Rückblick auf die bisherigen Erfolge wagte. *Bernhard Rieder* und *Sigrid Schieber* berichteten detailliert und konkret über die Übernahme von Informationen aus einzelnen Systemen. Rieder stellte die Übernahme aus dem Polizei-Informationssystem (POLIS), Schieber die Entwicklung einer Aussonderungsschnittstelle zum DMS DOMEA vor. Abschließend benannte *Corinna Knobloch* anhand konkreter Beispiele die spezifischen Anforderungen der digitalen Archivierung.

Nach der Übernahme müssen die digitalen Unterlagen inhaltlich und technisch aufbereitet und schließlich sicher archiviert werden. Gerade bei digitalen Archivalien sind dafür IT-Systeme und Tools notwendig. In dem gleichnamigen dritten Block beschrieben zunächst *Katharina Ernst* und *Heike Maier* die ganz konkreten Probleme, die sich im Stadtarchiv Stuttgart bei der Migration von einem Speicher auf dessen Nachfolger ergaben. *Burkhard Nolte* und *Karsten Hutb* stellten im Anschluss den Projektstand bei der Einführung der elektronischen Archivierung im Sächsischen Staatsarchiv dar, bevor *Rolf Lang* die Einführung der elektronischen Grundakte und ihre Speiche-

rung im speziell angepassten G-DIMAG des Landesarchivs Baden-Württemberg skizzierte. Dass auch bei den Archivierungssystemen die Zusammenarbeit mit anderen Archiven im Mittelpunkt stehen sollte, führten die beiden letzten Vorträge dieses Blocks aus. *Christoph Schmidt* beschrieb die Gründung und Entwicklung einer spartenübergreifende Nutzergruppe des Systems von HP/SER. *Christian Keitel* sprach im Anschluss über die möglichen Kooperationen für eine Nachnutzung des Systems DIMAG. Der Tag wurde abgeschlossen durch den Beitrag von *Ilka Stahlberg* über das digitale Urkundszwischenarchiv ELUZA in Brandenburg.

Der zweite Tag begann mit einem Block zu den verschiedenen Objektarten. Dabei standen nicht zuletzt die archivischen Anforderungen im Vordergrund, die sich je nach Objektart und IT-System erheblich unterscheiden können. Den Anfang machte *Mike Zuchet*, indem er einen Weg zur Archivierung von E-Mails vorstellte, was zu den Pionierarbeiten auf diesem Gebiet zählt. *Kai Naumann* berichtete im Anschluss über die verschiedenen Möglichkeiten, aber auch Probleme, die mit der dauerhaften Erhaltung von Geodaten bei Vermessungsbehörden und Archiven einhergehen. *Claire Röhlsberger-Jourdan* machte bereits im Titel ihres Beitrags klar, dass Formaterkennung und -validierung sich in der Theorie und in der Praxis erheblich unterscheiden können, und erläuterte sehr anschaulich die Unterschiede. Abschließend berichteten in diesem Block *Peter Bohl* und *Johannes Renz* darüber, wie digitalisierte AV-Unterlagen in DIMAG übernommen werden. Damit wurde vom Arbeitskreis ein weiteres neues Thema aufgegriffen. Zwar sind AV-Unterlagen nicht unbedingt genuin digitale Unterlagen. Sie sind aber zumeist nur durch Digitalisierung dauerhaft zu erhalten. Da AV-Unterlagen bereits heute in den klassischen Archiven sehr verbreitet sind, ist es nicht ganz unwahrscheinlich, dass sie in vielen Pilotprojekten zu den ersten digitalen Archivalien zählen werden.

Im letzten thematischen Block der Tagung wurden dann einige sehr grundsätzliche Fragen gestellt. *Lambert Kansy* und *Markus Loch* berichteten von den Erfahrungen beim Aufbau einer geeigneten Infrastruktur beim Staatsarchiv des Kantons Basel-Stadt. Sie gehen in ihrem Vortrag besonders auf die Anforderungen an eine archivische Speicherinfrastruktur ein. *Steffen Schwalm* stellte das Organisationskonzept elektronische Verwaltungsarbeit als Nachfolger des DOMEA-Konzepts vor und berichtete über dessen mögliche Auswirkungen auf die digitale Archivierung. Abschließend stellte *Christian Keitel* den *nestor-Leitfaden zur Digitalen Bestandserhaltung* vor. Erstmals werden hier die verschiedenen Ansätze zur Erhaltung der digitalen Archivalien in einem übergreifenden Modell vereinigt und praktisch Vorschläge zu dessen Umsetzung gemacht.

Zum Abschluss der Tagung diskutierten die Teilnehmer die Frage, ob auch in Deutschland eine oder mehrere Koordinierungsstellen zur digitalen Archivierung hilfreich wären. In den Wortbeiträgen wurde dies ohne Ausnahme bejaht, unklar war allerdings sowohl der Zuschnitt (eine oder mehrere?) als auch die Finanzierung der Stelle. Es bestand Einigkeit, dass die Diskussion nicht so bald in einem Konsens münden werde, aber dennoch weiter zu führen sei.

An der Tagung haben etwa 70 Kolleginnen und Kollegen teilgenommen. Ragna Boden hat über die Tagung insgesamt und auch die geführten Diskussionen im Archivar berichtet (Archivar 65 (2012), S. 186f).

nestor-Workshop

Brauchen wir Koordinierungsstellen für die digitale Archivierung? Wieder aufgegriffen wurde die Frage der Koordinierungsstellen von dem bereits erwähnten *nestor-Workshop* am 18. Juli 2012. Eröffnet wurde auch der Workshop durch den Präsidenten des Landesarchivs *Robert Kretzschmar*. *Christian Keitel* kam dann zum Ergebnis, dass Kooperation in der digitalen Archivierung zwar in zahlreichen Vorträgen angemahnt wird, aus fachlichen und finanziellen Gründen und in vielen Felder zwingend erforderlich ist, zugleich aber noch viel zu selten praktiziert wird. *Georg Büchler* stellte mit der KOST ein Beispiel für eine gelungene Kooperation in diesem Gebiet vor. Im Anschluss berichtete *Peter Worm* über Erfahrungen in Westfalen, die ebenfalls beispielhaften Charakter haben. In der hier von *Ulrich Schludi* wiedergegebenen lebhaften Diskussion wurden verschiedene Modelle erörtert. Die Teilnehmer kamen überein, dass sich die Auseinandersetzung lohne und als nächstes in den Gremien der staatlichen und kommunalen Archive Deutschlands weitergeführt werden sollte. Im Ergebnis haben sich dann die Bundeskonferenz der Kommunalarchivare beim Deutschen Städtetag und die Archivreferentenkonferenz gegen eine zentrale Koordinierungsstelle mit bundesweiter Zuständigkeit und für eine Koordinierung im Rahmen regionaler Verbände für die digitale Archivierung ausgesprochen.

Vom Call for Papers bis zur Publikation

Im Call for Papers zur Frühjahrstagung des *Arbeitskreises Archivierung von Unterlagen aus digitalen Systemen* hatten die Veranstalter nach Vortragsvorschlägen zu folgenden Themen gefragt:

- Kosten der digitalen Archivierung oder eines ihrer Teilbereiche.
- Erfahrungen bei der Übernahme von digitalen Unterlagen aller Art (z.B. Dateisammlungen / Fachverfahren / Webseiten / elektronische Akten).
- Erfahrungen aus der Ausschreibung und Einführung einer Archivierungssoftware.
- Kritische Rückblicke (z.B. auf 20 Jahre Behördenberatung einschließlich Systemeinführung, DMS-Systeme).
- Konzeptionelle und grundsätzliche Überlegungen zu einzelnen archivfachlichen Fragestellungen.

Es ist sehr erfreulich, dass alle im Call for Papers angefragten Themen wenigstens durch einen Vortrag abgedeckt werden konnten. Zwar gilt dies nicht im gleichen Maße für alle genannten Beispiele (so wäre es sehr wichtig, wenn die bisherige Geschichte der Behördenberatung eine kritische Darstellung und Würdigung erfahren könnte). Insgesamt aber können sich die Beiträge mit Blick auf die Themenvielfalt und die konkrete, teilweise bereits praxisgesättigte Herangehensweise sehen lassen. Sehr erfreulich ist die Offenheit, mit der einzelne Texte auch Fehlentscheidungen und Misserfolge thematisieren, weil solche Schilderungen den Blick für übergreifende strategische Sichtweisen schärfen. Ebenso wesentlich ist die hohe Rate von Informatikern und Dokumentaren unter den Verfassern (9 von 29), die uns verdeutlicht, dass das Thema nur

im interdisziplinären Austausch vorangebracht werden kann. Die Beiträge zeigen, dass digitale Archivierung im Jahr 2012 den Status als Orchideenfach längst abgelegt hat. Es wird zunehmend schwierig, in den Archiven Bereiche zu finden, die noch nicht von den Fragen tangiert sind, die in diesem Band behandelt werden.

Im Call for Papers wurde zwischen Vorträgen und Kurzvorträgen unterschieden. Durch letztere sollten sich auch Kolleginnen und Kollegen beteiligen können, die nur kurz über eine interessante Erfahrung oder Entwicklung informieren wollten. Für die Publikation wurde keine vergleichbare Vorgabe gemacht. Es liegt aber in der Natur der Sache, dass die Manuskripte von Kurzvorträgen eher kürzer als jene von zwanzigminütigen Vorträgen sind.

Ein solcher Band ist immer ein Ergebnis vieler Hände. Herzlich bedanken möchten wir uns zunächst bei den Referentinnen und Referenten, die die beiden Veranstaltungen erst möglich gemacht und durch ihre Beiträge bereichert haben. Inge Nesper, Elvira Grammer und Hans-Joachim Schulz vom Staatsarchiv Ludwigsburg und Andreas Sturitis vom Hauptstaatsarchiv Stuttgart danken wir sehr für die so wichtige logistische Unterstützung der Veranstaltungen. Unser besonderer Dank gilt schließlich auch Regina Keyler für die umsichtige Redaktion und Betreuung der Publikation.

16. Tagung des Arbeitskreises
„Archivierung von Unterlagen aus digitalen Systemen“

Kosten der digitalen Archivierung

Ein mögliches Vorgehensmodell und erste Erfahrungen

Von KARLHEINZ SCHMITT

Dieser Beitrag stellt ein Vorgehensmodell vor, mit dem Organisationen Kosten der digitalen Archivierung prognostizieren und berechnen können. Am Beispiel einer konkreten Langzeitarchivierungsdienstleistung, die im Rahmen des von der DFG geförderten Projektes: Digital Preservation for libraries (DP4lib) aufgebaut worden ist, wird ein generelles Vorgehensmodell verwendet, um für die im Projekt konzipierten Dienste der Langzeitarchivierung (LZA) individuelle Kostenmodelle zu erstellen. Mit Hilfe dieser Kostenmodelle werden Organisationen zum einen in die Lage versetzt, Kosten den jeweiligen LZA-Dienstleistungen verursachungsgerecht zuzuweisen und zum anderen den Aufbau, den Betrieb und die Weiterentwicklung der LZA-Dienstleistungen wirkungsvoll zu managen.

Einleitung

Die Aufgabe, digitale Publikationen, Datenbanken oder Webseiten über einen unbestimmten Zeitraum hinweg nutzbar zu halten, stellt Organisationen vor komplexe Aufgaben und Entscheidungen. Dieses liegt u.a. in der Thematik der Langzeitarchivierung selber begründet, da es immer noch eine weitgehend offene Frage ist, welche Maßnahmen oder Aktivitäten in der Gegenwart notwendig sind, um digitale Objekte langfristig nutzbar zu erhalten. Abhängig von den in den jeweiligen Organisationen vorhandenen Voraussetzungen, Sichtweisen und Zielen, sind weltweit in den vergangenen Jahren deshalb eine Reihe von sehr unterschiedlichen Langzeitarchivierungssystemen entstanden.

Was kostet aber nun die Einführung und der Betrieb des neuen organisationseigenen Service: *Digitale Langzeitarchivierung*?

Fast ebenso komplex erscheint die Frage nach der langfristigen Finanzierbarkeit dieser individuellen LZA-Systeme zu sein. IT-Dienstleistungen, und zu einem nicht unerheblichen Teil bestehen LZA-Dienstleistungen für digitale Objekte aus ihnen, gelten als sehr kritische Komponenten innerhalb einer Organisation. Die rasante Entwicklung der Technologien und deren zunehmende Komplexität bewirken meist, dass die Kosten für deren Bereitstellung und Betrieb wesentlich schneller und unberechenbarer wachsen als andere Kosten.

Von Vorteil wäre in dieser Hinsicht, wenn das Rechnungswesen einer Organisation die Kosten den jeweiligen Dienstleistungen verursachungsgerecht zuordnen kann. Meist ist jedoch ein solch detailliert aufgeschlüsseltes Rechnungswesen sehr aufwändig aufzubauen und zu pfle-

gen, so dass es oftmals erst gar nicht aufgebaut wird. In diesen Fällen übersteigen die Kosten für den Aufbau und die Pflege des Rechnungswesens, den erwarteten Nutzen. Ohne ein solches System werden Kostenschätzungen auf Grundlage bestehender Erfahrungen mit ähnlich komplexen Aufgaben abgegeben. Doch was tun, wenn diese Erfahrungen fehlen? Wie können Kostenschätzungen für Verbesserungen an Service oder neue Service dem Management gegenüber vorher und nachher gerechtfertigt werden?

Der Ruf nach einem – oder gar *dem* – Kostenmodell für digitale Langzeitarchivierung wird deshalb immer lauter. In den letzten Jahren ist die Entwicklung von Kostenmodellen und die Suche nach Gemeinsamkeiten intensiv behandelt worden¹. Allen Kostenmodellen gemeinsam ist, dass sie zu Recht auf die individuellen LZA-Systeme und die Gegebenheiten der jeweiligen Organisationen angepasst worden sind und deshalb nur sehr bedingt auf andere Organisationen übertragen werden können. Es sei an dieser Stelle auch schon angemerkt, dass dasselbe auch für das Kostenmodell aus DP4lib gilt.

In diesem Beitrag wird aus diesem Grunde auch nicht das Ergebnis – das resultierende DP4lib-Kostenmodell – in den Mittelpunkt der Ausführungen gestellt, sondern das Vorgehen, die Arbeiten und die Entscheidungen, die im Projekt notwendig waren, um das auf den speziellen LZA-Service angepasste Kostenmodell zu entwickeln. Hierbei besteht die Hoffnung, dass aus diesem individuellen Vorgehen ein allgemeines Vorgehensmodell abgeleitet werden kann, um Organisationen in die Lage zu versetzen, ein eigenes auf die individuellen Bedürfnisse angepasstes Kostenmodell der digitalen Langzeitarchivierung aufzustellen.

Der spezielle Langzeitarchivierungsdienst aus DP4lib

Das von der DFG geförderte Projekt: Digital Preservation for libraries (DP4lib) hatte die Aufgabe, ein Langzeitarchivierungssystem zur kommissarischen Archivierung digitaler Publikationen aufzubauen. Hierbei kam dem klassischen aus der Privatwirtschaft bekannten Dienstleistungsgedanken ein besonderer Stellenwert zu. Im aufgebauten Kooperationsmodell nahmen die Deutsche Nationalbibliothek (DNB) und die Niedersächsische Staats- und Universitätsbibliothek Göttingen (SUB) als Projektpartner die Rolle des Dienstleisters eines Portfolios von LZA-Diensten ein, die es im Projektverlauf aufzubauen galt. Sechs weitere externe Partner, nämlich:

- Bibliotheksservice-Zentrum Baden-Württemberg (BSZ)
- Deutsches Institut für Pädagogische Forschung, Frankfurt (DIPF)

¹ B. Hole, L. Lin, P. McCann und P. Wheatley: LIFE. *A Predictive Costing Tool for Digital Collections*. *Vortragspapier*, iPRES2010, 19.–24.9.2010, <http://www.ifs.tuwien.ac.at/dp/ipres2010/papers/hole-64.pdf>; Königliche Bibliothek und Nationalarchiv Dänemark: *Cost Model for Digital Preservation*, <http://www.costmodellfordigitalpreservation.dk/contact/cmdp-2--ingest-and-archival-storage>; Producer-Archive Interface Methodology Abstract Standard (PAIMAS), 2002, ISO 20652-2006, http://www.us-vo.org/pubs/files/CCSDS_651_W2.pdf (Alle Links wurden am 16.10.2012 überprüft).

- Sächsische Landesbibliothek – Staats- und Universitätsbibliothek Dresden (SLUB)
- Technische Informationsbibliothek Hannover (TIB)
- Thüringer Landes- und Universitätsbibliothek Jena (ThULB)
- Verbundzentrale des GBV (VZG)

nahmen im Projektverlauf die Rolle von potenziellen Dienstnehmern der LZA-Dienste ein. Nach einer intensiven Phase der Anforderungsdefinition sind zwischen Dienstleistern und Dienstnehmern die gewünschten LZA-Dienste spezifiziert worden. Tabelle 1 gibt einen Überblick über das LZA-Dienst-Portfolio. Zu sehen ist, dass im Projekt generell zwischen drei Hauptdienstleistungen unterschieden wird: Ingest, Curation und Access. Jede Hauptdienstleistung besteht aus mehreren Teildiensten die auf Wunsch individuell von jedem Dienstnehmer in Anspruch genommen werden können. Auf diese Weise ist es innerhalb des Dienstleistungsmodells möglich, auf die individuellen Voraussetzungen und Ziele der Dienstnehmer bzgl. der erforderlichen LZA-Funktionalitäten einzugehen und individuelle Ingest-, Curation- und Access-Prozesse für jeden Dienstnehmer aufzubauen und anzubieten.

Ingest	Curation	Access
Übernahme von Transferpaketen	Erhaltungsmaßnahmen	Authentifizierung
Dateiintegritätsprüfung	Integritätsprüfung und Erhaltung	Suche
Technische Metadaten	Retrieval	Bereitstellung
Deskriptive Metadaten		
Qualitätsprüfung		
Berichtswesen		
Fehlerprotokoll		
Incident-Management		

Tab. 1: LZA-Dienste innerhalb des DP4lib-Projektes.

Die Bereitstellung dieser modularen LZA-Dienstleistungen erforderte bei den Dienstleistern den Auf- und Ausbau einer ausreichend dimensionierten technischen Infrastruktur, die Programmierung der geforderten LZA-Dienste und den Aufbau eines Organisations- und Betriebsmodells, um die LZA-Dienstleistungen langfristig anbieten und organisatorisch betreuen zu können. Eine ausführliche Beschreibung der Infrastruktur und Projektergebnisse ist auf der Projekt-Homepage² zu finden und soll deshalb an dieser Stelle nicht weiter erfolgen.

² <http://dp4lib.langzeitarchivierung.de>.

Vorgehensmodell

Entscheidung über das Ziel des Kostenmodells

Um die wahren Kosten der LZA-Servicebereitstellung zu kennen und diese Kosten darüber hinaus auch in jeder Phase professionell managen zu können, stehen im Allgemeinen verschiedene Mittel zur Verfügung. Gewöhnlich können IT-Rechnungswesen- und Budgetierungs-Prozesse innerhalb einer Organisation dazu verwendet werden, um Kosten zu identifizieren, kontrollieren und den Verursachern zuzuweisen. Häufig gehen diese beiden Prozesse einem weiteren voraus, dem Prozess der Leistungsverrechnung für spezifische Dienstleistungen.

Mit Hilfe der Budgetierung ist es u.a. möglich, den Bedarf an finanziellen Mitteln für den Betrieb eines Service über einen bestimmten Zeitraum zu prognostizieren. Ebenso kann eine Organisation sicherstellen, dass es jederzeit möglich ist, die tatsächlichen Ausgaben mit den vorhergesagten zu vergleichen. Ein hierauf basierendes und ausgebautes Rechnungswesen versetzt Organisationen im Anschluss daran in die Lage, über die verwendeten Gelder jederzeit Rechenschaft abzulegen, Kosten für die Bereitstellung und Änderung eines Service zu kalkulieren und zu identifizieren. Leistungsverrechnungsprozesse ermöglichen es letztendlich, dass Kosten für einen Service durch Entgelte der Kunden gedeckt werden können.

Änderungen an vorhandenen internen Budgetierungsprozessen und dem Rechnungswesen oder gar die Einführung der Leistungsverrechnung sind Gegenstand strategischer Entscheidungen, da die Änderungen stets weitreichende Auswirkungen auf die gesamte Organisation haben. Alle drei Prozesse erfordern die Existenz eines Kostenmodells, welches beginnend mit der Budgetierung von einem relativ einfachen bis hin zu einem immer komplexeren Kostenmodell ausgebaut werden muss.

Das Ziel innerhalb des Projektes bestand darin, die aufgebauten LZA-Servicedienste den zukünftigen Dienstnehmern in Rechnung stellen zu können, so dass das aufzubauende Kostenmodell fähig sein musste auch den Leistungsverrechnungsprozess zu unterstützen.

Für die Entscheidung über das Ziel des aufzustellenden Kostenmodells unterstützen soll, muss das Kosten-Nutzen-Verhältnis zwischen dem erhofften Nutzen des Kostenmodells und den Kosten für die Einführung und Pflege des Kostenmodells jederzeit beachtet werden.

Aufschlüsselung des Kostenmodells

Es gibt verschiedene Möglichkeiten für die Aufschlüsselung eines Kostenmodells.

1. Cost-by-Customer,
2. Cost-by-Service und
3. Cost-by-Location.

Üblicherweise beruhen Kostenmodelle auf der Berechnung der Kosten für jeden einzelnen Kunden. Kunden können in diesem Zusammenhang sehr breit interpretiert werden. Kunden können interne Abteilungen sein, die Service von anderen internen Abteilungen entgegennehmen, aber

auch externe Kunden, die Dienstleistungen von einer Organisation beauftragen. Es kann aber auch sinnvoll sein die Berechnung der Kosten so zu gestalten, dass ein Kostenmodell Aussagen darüber geben kann, wie viel eine spezielle Service-Bereitstellung kostet. In diesem Fall muss die Aufschlüsselung der Kosten so gestaltet werden, dass alle Kostenelemente den spezifischen Services zugeordnet werden können. Als weitere, seltenere Variante ist die Aufschlüsselung der Kosten nach verschiedenen Standorten. Für eine Evaluation welcher Standort einer Organisation die meisten Kosten verursacht, ist dies eine Möglichkeit relativ schnell die kostentreibenden Orte zu identifizieren. Innerhalb des DP4lib-Projektes wurde ein Cost-By-Service-Kosten-Modell entwickelt, um die tatsächlichen Kosten der Bereitstellung und des Betriebs der eigenen LZA-Dienstleistungen kennen zu lernen.

Erstellen eines Kostenmodells

Wenn das Kostenmodell erstmalig erstellt wird, müssen zunächst die Kostenkategorien und die Kostenelemente des zu betrachtenden LZA-Services identifiziert werden. Hierfür ist es entscheidend, welchen der drei oben genannten Prozesse das Kostenmodell unterstützen soll. Der Detaillierungsgrad und damit auch der Aufwand müssen den jeweiligen Erfordernissen des Rechnungswesens oder gar der gewählten Leistungsverrechnung entsprechen.

Es kann für einen besseren Überblick nützlich sein, Kosten Kategorien zuzuordnen, um sicherzustellen, dass sie korrekt identifiziert und gehandhabt werden. Die Wahl der Kategorien ist vollkommen willkürlich und sollte am besten den jeweiligen Erfordernissen angepasst werden. Der Nutzen von Kostenkategorien kann in der Herstellung einer internen Konsistenz und einer Vergleichbarkeit von Kostenmodellen von verschiedenen Organisationen gesehen werden. Üblicherweise werden folgende Kostenkategorien verwendet:

- Hardware-Kosten
- Software-Kosten
- Personal-Kosten
- Kosten für Räumlichkeiten
- Kosten für externe Service und
- Transfer-Kosten

Die Kategorie der Transfer-Kosten bedarf einer näher gehenden Erklärung. Unter Transfer-Kosten werden Leistungen und auch Waren verstanden, die von einem Teil der Organisation an einen anderen Teil *verkauft* werden. Transfer-Kosten sollten immer in einem Kostenmodell bedacht werden, auch wenn ihnen keine eigene Kategorie zugewiesen wird. Die Erbringung interne Service innerhalb einer Organisation bildet oftmals einen nicht unerheblichen Teil der Kosten.

Sofern ausschließlich die Budgetierung mit dem Kostenmodell unterstützt werden soll, ist dieser Detaillierungsgrad vollkommen ausreichend. Sämtliche Kostenelemente, die unter die jeweiligen Kostenkategorien zusammengefasst werden können, können zu einer einzigen nicht näher aufgeschlüsselten jährlichen Summe zusammengefasst werden.

Wenn für die Berechnung der Kosten ein höherer Detaillierungsgrad erforderlich ist, müssen die gewählten Kosten-Kategorien weiter ausdifferenziert werden. So kann es beispielsweise für eine Organisation sinnvoll sein festzustellen, wie hoch die jährlichen Netzwerk-Kosten oder die Kosten des Backup-Systems sind. Innerhalb der Kostenkategorie Hardware ist es dann möglich, Netzwerk und Backup-System als Kostenelemente gesondert aufzuführen und die Kosten für alle Teilkomponenten unter den jeweiligen Kostenelementen zu subsummieren. Beispiele für Kostenelemente sind in Tabelle 2 dargestellt.

Kosten-Kategorie	Kostenelemente
Hardware	Festplatten, Server, Router, ...
Software	Betriebssysteme, Lizenzen, Datenbanken, ...
Personal	Gehalt, Firmenwagen, Spesen, ...
Räumlichkeiten	Büro, Serverraum, ...
Externe Service	Hausmeister-Service, Provider, ...
Transfer	Interne Geldforderungen

Tab. 2: Beispiele für Kostenelemente bzgl. der gewählten Kosten-Kategorien.

Nach Aufstellung der Kosten-Kategorien und der weiteren Detaillierung der Kategorien durch deren Unterteilen in Kostenelemente besteht prinzipiell die Aufgabe, die einzelnen Kostenelemente zu betrachten und nach direkte und indirekte Kosten zu unterteilen.

Direkte Kosten sind Kosten, die eindeutig einem einzelnen Service – oder auch Kunden oder Ort – zugeordnet werden können. Ein Server, der ausschließlich von diesem Service verwendet wird, würde zu den direkten Kosten gezählt werden. Indirekte Kosten, sind Kosten, die für mehrere oder alle Service entstehen. Das Netzwerk oder ein Server, auf dem alle Service laufen, wären gute Beispiele hierfür. Oftmals werden bei der Ermittlung aller indirekten Kosten auch solche festgestellt, die keinem Service zugeordnet werden können. In solchen Fällen muss versucht werden, diese Kosten in möglichst fairer Weise (Aufschlag) auf alle Service zu verteilen.

Abbildung 1 zeigt das bisher beschriebene Vorgehensmodell.

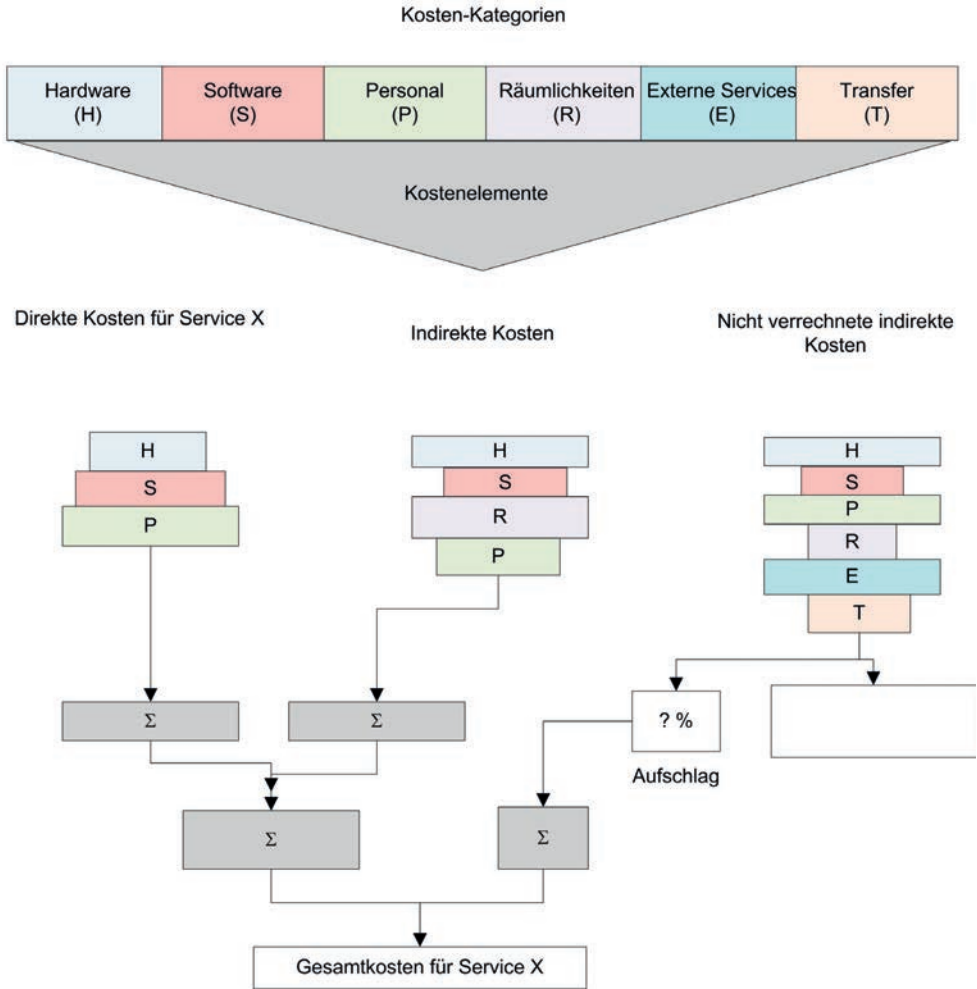


Abb. 1: Kostenmodell – nach Service aufgeschlüsselte Kosten.³

³ Angelehnt an: ITIL Best Practice for Service Delivery. Van Haren Publishing 2006. S. 90.

Das Kostenmodell aus DP4lib

Nach einer Einführung in die grundlegende Vorgehensweise einer Kostenmodellerstellung wird im Folgenden diese Vorgehensweise auf die LZA-Service des Projektes DP4lib angewendet. Das Ziel des Kostenmodells bestand darin, die Kosten für den Gesamt-Service der digitalen Langzeitarchivierung zu ermitteln und darüber hinaus das Kostenmodell derart aufzubauen, dass eine Leistungsverrechnung darauf aufgesetzt werden konnte. Auf diese Weise war der Detaillierungsgrad des Kostenmodells derart eingegrenzt, dass er zum einen mit den vorhandenen Ressourcen noch umgesetzt, aber zum anderen auch eine anfängliche Leistungsverrechnung durchgeführt werden konnte.

Wie in der Projektdarstellung bereits erläutert worden ist, wurden generell drei Haupt-Service: Ingest, Curation und Access aufgebaut. Für alle drei Prozesse galt es nun, jeweils ein Cost-By-Service Kostenmodell aufzustellen.

Kosten-Kategorien und Kostenelemente des Projektes

Der erste Schritt bestand aus der Festlegung der Kategorien und deren zugehörigen Kostenelemente. Abbildung 2 zeigt die endgültige Aufteilung aller identifizierten Kostenelemente für die neue Dienstleistung der Langzeitarchivierung.

Kostenarten	Kostenelemente	Jahres-kosten
Hardware		
	Server (Ingest/Retrieval/ Preservation/Datenbank)	1,00 €
	Festplatten-Cache (2 TB) (Dienstleister DNB)	1,00 €
	DIAS-Neukonfiguration	1,00 €
	200 LTO5 Bänder je 1TB (100TB mit 2fach Backup)	1,00 €
	200 Stellplätze (gemietet)	1,00 €
	2 Bandlaufwerke	1,00 €
	WAN-Verbindung	1,00 €
Software		
	Datenbank-Lizenzen (Open Source)	1,00 €
	OLAP-System (Audit-Trail)	1,00 €
	IBM-Lizenzen (1 Mandant)	1,00 €
	IBM-Lizenzen (Dias-System)	1,00 €
	Email-Ticket-System (JIRA)	1,00 €
	Monitoring (Nagios, Logging...) (Open Source)	1,00 €
Personal		
	Techn. Support (E9 bis E12)	1,00 €
	Manager (E13 bis E14)	1,00 €
	Assistent (E2 bis E9)	1,00 €
	Entwickler (E9 bis E12)	1,00 €
	Sachbearbeiter (E9 bis E12)	1,00 €
Räumlichkeiten		
	Büros	1,00 €
Externe Services		
	IBM-Support (Third-Level)	1,00 €
	Provider-Vertrag (GWDG)	1,00 €

Abb. 2: *Kosten-Kategorien und Kostenelemente des Projekts DP4lib.*

Kosten der einzelnen Kostenelemente: Die Wahl der Verteilungsschlüssel und Abschreibung

Im nächsten Schritt wurden dann die Kosten der jeweiligen Kostenelemente ermittelt (Spalte Jahres-Kosten in Abbildung 2). In diesem Zusammenhang wurden im Wesentlichen zwei Entscheidungen pro Kostenelement getroffen: Die Wahl der Abschreibungsmethode sofern gewünscht, und die Wahl eines Verteilungsschlüssels der Kosten.

Abschreibung ist das Maß für die Abnutzung, den Verbrauch oder eine andere Wertminderung über den Verlauf der ökonomischen Nutzungsdauer eines fixen Assets (Vermögensgegenstands). Um Abschreibungen vornehmen zu können, werden Informationen über den Abschreibungswert, den Abschreibungszeitraum sowie die zur Erfassung des Werteverzehrs geeignete Abschreibungsmethode benötigt.⁴ Generell sollte die Abschreibungsmethode verwendet werden, die einen fairen Anteil der Kosten auf den gesamten Abschreibungszeitraum ermöglicht.

Im Bereich der Hardware wurden zunächst die Anschaffungskosten als Abschreibungswert eingesetzt. Später sollte hier jedoch die Summe als Abschreibungswert verwendet werden, die am Ende der Nutzungsdauer aufzuwenden ist, um ein neues Anlagegut zu beschaffen. Dieser Wechsel sollte allerdings nur unter der Voraussetzung, dass die Bestimmung dieser Summe möglich ist, stattfinden. Die Wahl des Abschreibungszeitraums für die jeweiligen Kostenelemente beruht auf Erfahrungswerten. Zur Berechnung der Abschreibungen wurde die lineare Abschreibungsmethode eingesetzt, da sich diese speziell für Kalkulationszwecke eignet.

Beispiel:

Eine lineare Abschreibung vorausgesetzt, lautet eine übliche Formel:

$$a_t = \frac{A - R_n}{n}$$

mit:

- a_t = Abschreibung in Periode t
- A = Anschaffungskosten
- R_n = Restwert am Ende der Nutzungsdauer
- n = Nutzungsdauer

Ein Server, der in der Anschaffung bspw. 6.000 € gekostet hat, soll laut Unternehmensentscheidung auf die Art abgeschrieben werde, dass nach Ablauf von 5 Jahren ein Restwert von 0 € verbleibt. Dann betragen sich die jährlichen Kosten für den Server auf 1.200 €.

⁴ O. Plötner, B. Sieben und T. Kummer: Kosten- und Erlösrechnung. Anschaulich, kompakt, praxisnah. 2., korrigierte und aktualisierte Aufl., Berlin 2010; T. Schildbach und C. Homburg: Kosten- und Leistungsrechnung. 10., bearb. Aufl., Stuttgart 2009.

Damit eine verursachungsgerechte Aufteilung der Kosten eines LZA-Services auf die drei Prozesse möglich ist, werden weiterhin Verteilungsschlüssel benötigt. Eine Auswahl möglicher Verteilungsschlüssel ist in Abb. 3 dargestellt.⁵

Schlüssel für die Kostenverteilung bzw. -zurechnung	
Mengenschlüssel	Werteschlüssel
Zählgrößen (z. B. Zahl der abgesetzten Stücke oder Zahl der Buchungen)	Kostengrößen (z. B. Fertigungsmaterialkosten oder Herstellkosten)
Zeitgrößen (z. B. Fertigungszeit oder Maschinenstunden)	Einstandsgrößen (z. B. Wareneingangswert oder Lagerzugangswert)
Raumgrößen (z. B. Länge oder Fläche)	Absatzgrößen (z. B. Warenumsatz oder Kreditumsatz)
Gewichtsgrößen (z. B. Produktmengen oder Transportgewichte)	Bestandsgrößen (z. B. Anlagenbestandswert)
Technische Maßgrößen (z. B. kWh oder PS)	Verrechnungsgrößen (z. B. Verrechnungspreise)

Abb. 3: Verteilungsschlüssel für die Kostenverteilung

Die gewählten Verteilungsschlüssel sollten dabei einen Kompromiss zwischen benötigter Detaillierung und Abstraktion aufgrund fehlender Detailinformationen über die Systemnutzung der einzelnen Prozesse darstellen.

Ein Beispiel: Für die LZA-Dienstleistung wurde ein neuer Server mit einem Anschaffungswert von 6.000 € bereitgestellt. Alle drei Hauptdienstleistungen, Ingest, Curation und Access verwenden diesen Server. Wie sollen die jährlichen Kosten von 1.200 € auf die drei Hauptdienstleistungen verteilt werden? Als Verteilungsschlüssel wurde hierfür die jeweils von den einzelnen Dienstleistungen genutzte CPU-Zeit festgelegt.

Unter der Annahme, dass für die Ingest-Dienstleistung 40% der Gesamt-CPU-Zeit verwendet wird, für die Curation-Dienstleistung 30% und den Access-Dienst die verbleibenden weiteren 30%, müssen demnach für das Cost-By-Service Kostenmodell der Ingest-Dienstleistung 40% der jährlichen 1.200 € in Rechnung gestellt werden.

Auswertung

Das Kostenmodell zeigt, dass ca. 37% auf den Ingest-Prozess entfallen. Mit nur 3% weniger macht der Curation Prozess den zweitgrößten Teil der Kosten für den LZA-Service aus. Auf den Access entfallen 29%. Die Kostenverteilung ist in Abbildung 4 dargestellt. Diese Verteilung der Kosten auf die drei Prozesse weicht leicht von den bisher bekannten Modellen ab, bei denen

⁵ M. Schweitzer und H.-U. Küpper: Produktions- und Kostentheorie. Grundlagen – Anwendungen, 2., vollständig überarb. und wesentlich erw. Aufl., Wiesbaden 1997.

der Ingest deutlich den größten Kostenanteil ausmacht. Eine Ursache hierfür ist die teilweise weit fortgeschrittene Vorverarbeitung der Objekte durch die Dienstnehmer, wodurch eine Verschiebung der Kosten vom Ingest des Dienstleisters zu den Dienstnehmern erfolgt. Bspw. liefern einige Dienstnehmer Transferpakete, für die auf Seiten des Dienstleisters nur noch eine marginale Weiterverarbeitung notwendig ist.

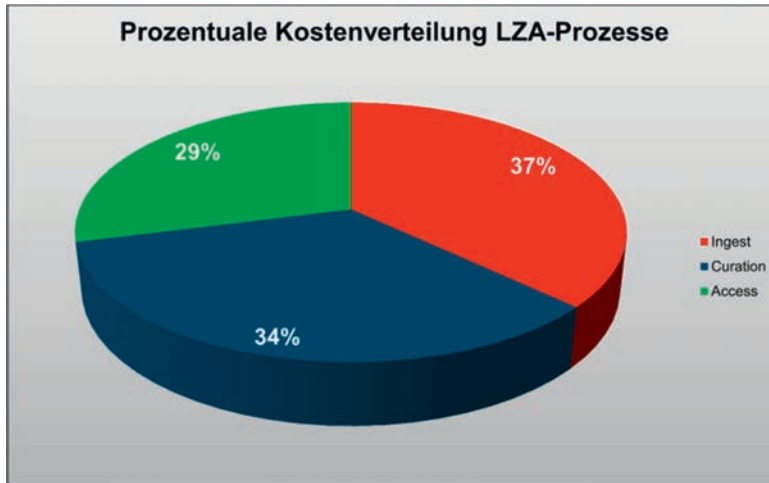


Abbildung 4: Prozentuale Kostenverteilung zwischen den LZA-Prozessen.

Schlusswort

Die Aufgabe, betriebsinterne Kostenmodelle für die Langzeitarchivierung zu erstellen, wird stets eine Herausforderung bleiben, der sich Organisationen im Dreieck zwischen eigenem LZA-System, den organisationseigenen Gegebenheiten und den Zielsetzungen für das Kostenmodell stellen müssen. Jedoch sind die Vorgehensweisen und Entscheidungen, der sich Organisationen gegenübersehen, meist vergleichbar, so dass die Hoffnung besteht, dass in Zukunft einheitliche Vorgehensmodelle für die Erstellung von Kostenmodellen die Arbeit erleichtern können.

Kostenfragen in digitalen Archiven

Erfahrungen des Digitalen Archivs Österreich

Von SUSANNE FRÖHLICH

Der vorliegende Beitrag skizziert die grundlegenden Rahmenbedingungen an ein digitales Archiv, welche in Summe die Zeitplanung sowie Kostengestaltung jedes Projektes beeinflussen. Aus rechtlichen, organisatorischen, fachlichen und technischen Vorgaben entsteht ein komplexes Anforderungsprofil, das als Framework für die Hauptfaktoren *Personal, Software, Hardware* und *Betrieb* gesehen werden kann. Als Beispiel dient die in Österreich umgesetzte Lösung des *Digitalen Archivs Österreich*. Die daraus resultierenden Erfahrungswerte werden dargelegt, etwaige Kostentreiber aufgezeigt und die theoretischen Aufzählungen durch einige praktische Rechenbeispiele veranschaulicht.

Einleitung

Jede Organisation, die sich in irgendeiner Weise mit der Erhaltung und Pflege von Schriftgut beschäftigt, steht heutzutage vor der großen Herausforderung, auch digitales Archivgut zu übernehmen und langfristig aufzubewahren. Hierbei steht nicht nur die Einlagerung (= Speicherung) der Daten im Vordergrund, sondern vor allem die Bemühung bzw. die Notwendigkeit die Inhalte der Unterlagen langfristig lesbar (= benutzbar) zu erhalten. Hinzu kommen allgemein gültige Anforderungen an die Datensicherheit, die archivinhaltliche Erschließung, allenfalls eine Veröffentlichung der Daten über das Medium Internet, sowie rechtliche und administrative Belange, welche ebenso berücksichtigt werden müssen. Dies alles firmiert in den letzten Jahren im Allgemeinen unter dem Schlagwort der *digitalen Langzeitarchivierung*.

Doch was genau macht digitale Langzeitarchivierung aus? Welche Rahmenbedingungen bzw. Basisparameter sind dabei zu berücksichtigen? Es gibt/gab seit den 1990iger Jahren europaweit zahlreiche Projekte, die sich alle primär mit der technischen Umsetzung der digitalen Archivierung beschäftig(t)en. Doch kaum jemand hat sich mit der langfristigen Planung, insbesondere der Finanzierung von digitalen Archiven auseinandergesetzt. Wobei gerade dieses Thema ein nicht zu unterschätzender Faktor auf dem Weg zu einer professionellen und eben vor allem *langlebigen* digitalen Archivierung ist.

Jeder, der sich mit dem Archivwesen beschäftigt, weiß, dass Archiv per se *dauerhafte, ewige* Aufbewahrung und Erhaltung des übernommenen Schriftgutes bedeutet. Dies betrifft nicht nur den analogen Bereich – für den sämtliche Prozesse bereits seit Jahrhunderten erprobt und ein- gespielt sind, sondern vor allem auch die neuen Medien und Technologien. Es ist nicht damit getan, digitale Daten kurzfristig auf einem - wie auch immer geartetem Medium - zu speichern.

Dies ist im besten Fall der erste Schritt zum digitalen Archiv. Die Herausforderung beginnt erst mit der Umsetzung jener Prozesse, welche in Fachkreisen mittlerweile als *Preservation* oder *Preservation Planning* wohlbekannt sind. Ebenso ist auch für digitale Archive eine permanente Infrastruktur, personelle Betreuung und technische Adaptierung vorzusehen. Und all dies muss langfristig budgetiert bzw. finanziert werden, weshalb vor allem auf politischer Ebene noch einiges an Bewusstseinsbildung zu leisten sein wird.

Die folgenden Darstellungen sollen dem Leser einen Einblick in die im Österreichischen Staatsarchiv seit 2007 stattgefundene Umsetzung des *Digitalen Archivs Österreich* (= digLA) geben. An Hand der praktischen Erfahrungen können sowohl inhaltliche Kostenfaktoren aufgezeigt, als auch einige Rechenbeispiele angeführt werden, welche als Basis für eigene Planungen herangezogen werden können. Es muss jedoch gleich zu Anfang deutlich gemacht werden, dass jedes Projekt zur digitalen Archivierung individuell ist, immer abhängig von den eigenen freien Ressourcen, den zu leistenden Vorarbeiten, den inhaltlichen Vorgaben zur Umsetzung notwendiger Prozesse und der dafür eingeplanten Zeitspanne. Daher kann grundsätzlich keine abschließende Aussage *ein digitales Archiv kostet so und so viel* getroffen werden.

Grundlegende Planungen

Vor dem Beginn jeglicher Arbeiten an einem digitalen Archiv sollten sich die damit beauftragten Personen über die innerhalb ihrer Organisation zugrunde liegenden Anforderungen an ein solches bewusst werden. Die Erfahrungen des Österreichischen Staatsarchivs haben gezeigt, dass die internen Vorarbeiten, die zum damaligen Zeitpunkt lediglich mit dem Ziel Ausschreibungsunterlagen zu erstellen erfolgt sind, immer noch die eigentliche Basis für alle derzeit umgesetzten Prozesse des digLA bieten. Auch anfangs belächelte, weil als *unnützig* oder *unnötig* eingestufte Details, wie zum Beispiel die detaillierte Ausarbeitung von fachlichen Use Cases, werden in den aktuellen Diskussionen oder Fragestellungen bis heute regelmäßig als Grundlagen herangezogen, und es hat sich als Vorteil erwiesen, auch auf solche Aspekte bereits damals Rücksicht genommen zu haben. Dies hat bei der Umsetzung des digLA nicht nur einmal zu einer Beschleunigung mancher Entscheidungsfindung beigetragen. Daher kann aus Sicht des ÖStA als Fazit gelten: Alle Punkte, die zu Beginn eines Projektes intern genau definiert, überlegt und geplant werden, kommen der Lösung langfristig zu Gute, mag es auch anfangs noch so unbedeutend erscheinen. Denn nur, wenn man selbst weiß, was man benötigt, kann man dies auch entsprechend (nach außen) kommunizieren und vor allem budgetieren.

In diesem Sinne sollten die Fragestellungen in Bezug auf die Entwicklung eines digitalen Archivierungssystems zur Ressourcen-, System- und Zeitplanung zumindest folgende Überlegungen beinhalten:

Allgemein

- Grundsätzliche Definition der rechtlichen, technischen und organisatorischen Rahmenbedingungen (vgl. Kapitel III)

- Überlegung, welche Partner zur Umsetzung eines solchen Vorhabens benötigt werden. Können die Aufgaben alleine – sprich mit internen Kräften – bewältigt werden, oder ist Unterstützung von Extern erforderlich?
- Hat die eigene Organisation das nötige Know How für dieses Thema bzw. möchte/muss man die Umsetzung intern durchführen oder kann das Projekt (komplett) ausgelagert werden?
- Welcher technische Standard soll erreicht werden bzw. welche technischen Vorgaben (z.B. Datenbanken) gibt es bereits? Was davon muss in ein digLA eingebunden werden, was kann ersetzt werden?
- Was ist an Technologie vorhanden, was muss noch angeschafft und/oder erneuert werden.

Fachlich

- Welche Daten werden archiviert (Digitalisate, Digital Born Daten)?
- Welche Metadaten und/oder Contentdaten sind vorhanden?
- In welchen Formaten und/oder nach welchen Standards soll archiviert werden?
- Welche Datenkonvertierungen sind vor/während der Archivierung durchzuführen?
- Welche Schnittstellen werden intern (Ingest) und/oder zu andern Systemen (AIS, Internet, etc.) benötigt?
- Welche (archivfachlichen/technischen) Prozesse sollen mit dem System abgewickelt werden?

Technisch/Organisatorisch

- Welche Datenmengen fallen wie oft an?
- Wie viel Speicherplatz wird dafür benötigt?
- Werden Ressourcen wieder frei (z.B. durch Löschroutinen)?
- Welche interne und/oder externe Performance (z.B. zur alltagstauglichen Benutzung) muss das System abdecken?
- Wie oft und durch wie viele Personen wird es wann, wie benutzt werden?

Digitales Archiv oder andere Systeme?

Bei diesem Thema hat sich gezeigt, dass die Vorstellung von Langzeitarchiv nicht gleich Langzeitarchiv ist, und im Vorfeld unter Anderem eine genaue Definition der tatsächlichen Archivierungsdauer von Nöten ist. Bewegt man sich in einem Zeitraum von unter 10 Jahren Aufbewahrungsfrist ist im Normalfall kein Langzeitarchivsystem erforderlich. Hier reicht eine herkömmliche Speicherung/Sicherung auf Datenträgern bzw. Servern vollkommen aus. Auch bei einer Erhaltungsdauer von bis zu 30 Jahren ist die Überlegung, ob es eines digLAs bedarf, noch eine Frage von Zweck, Benutzung und Budget. Erst bei der Anforderung einer dauerhaften Aufbewahrung (30 Jahre bis ewig) ist die Entwicklung und der Betrieb eines umfassenden Langzeitarchivierungssystems wie es hier dargestellt wird zielführend.

Zeitplanung

Aus der Fragestellung nach der Aufbewahrungsdauer von digitalem Schriftgut sowie den zuvor genannten Faktoren kann in Folge die Zeitplanung und somit meist auch die Finanzierung für ein digitales Archivprojekt abgeleitet werden:

Kurzfristig

Planung und Umsetzung erfolgt in knappem Zeitabstand mit einmaligem Budget und klarer Zielvorgabe. Ressourcen stehen für einen bestimmten Zeitraum für einen detaillierten Aufgabenbereich zur Verfügung.

Ziel: Rasche Sicherung von digitalen Daten auf Datenträgern oder Servern, meist auch zur Verfügung Stellung im Internet; kurzfristiger Zugriff; Originale sind analog oder in Kopie vorhanden; Datenverlust unproblematisch.

Mittelfristig

Planung, Umsetzung und kurzfristiger Betrieb erfolgt in längerem Zeitabstand mit mehrmaligem Budget und schrittweisen Zielvorgaben. Ressourcen stehen über einen bestimmten Zeitraum regelmäßig für verschiedene Aufgabenbereiche zur Verfügung.

Ziel: Mehrfachsicherung von digitalen Daten auf Datenträgern oder Servern; Erschließung und regelmäßiger Zugriff; Keine Migration erforderlich; Nach Zeitablauf werden die Daten gelöscht; Originale sind nicht dauerhaft aufzubewahren oder in (analoger) Kopie vorhanden; gezielter Datenverlust.

Langfristig

Planung, Umsetzung und Dauerbetrieb eines digitalen Archivs sowie Preservation- und Migrationsvorgänge erfolgen ohne zeitliche Begrenzung mit kontinuierlichem Budget und jeweils aktualisierten Zielvorgaben. Ressourcen stehen permanent für alle Aufgabenbereiche zur Verfügung.

Ziel: Dauerhafte Archivierung von digitalen Daten; Erschließung, Benutzung und Preservation notwendig; Daten werden nie gelöscht; Daten sind das Original; Datenverlust nicht zulässig.

Nutznießung

Sobald man all diese Parameter im Detail analysiert hat, bleibt abschließend noch die Überlegung, welchen Nutzen man allenfalls aus dem digitalen Archiv ziehen kann oder möchte bzw. welche Möglichkeiten der Refinanzierung es geben könnte:

- Gibt es wirtschaftliche Anforderungen wie z.B. Vermarktung und Weiterverkauf des Produktes?
- Welche Dienstleistungen können anderen angeboten werden?

- Ist die öffentliche (= externe) Benutzung digitaler Daten kostenpflichtig? Wie können derartige Einkünfte weiterverwendet werden?
- Wo finden sich (langfristig) Synergieeffekte durch gemeinsame Folgeprojekte, günstigere Betriebsparameter etc.?

Rahmenbedingungen eines Digitalen Archivs

Im Allgemeinen setzen sich die Rahmenbedingungen eines modernen Archivs - unabhängig ob analog oder digital - aus vier großen Bereichen zusammen:

1. rechtliche Grundlagen der Organisation,
2. administrativ-organisatorische Vorgaben der Organisation,
3. fachspezifische Anforderungen und
4. technische Ausstattung bzw. Standards.

Zusätzlich sind äußere Einflüsse wie die Fluktuation personeller Ressourcen, die Änderung politischer Strategien oder die Modifikation finanzieller Bedingungen zu berücksichtigen. Für das Österreichische Staatsarchiv (= ÖStA) stellte sich die Situation in diesen vier Bereichen zu Beginn der Arbeiten am Digitalen Archiv Österreichs wie folgt dar:

1. Rechtliche Grundlagen

- Österreichische Archivgesetze¹
- Österreichische E-Government-Gesetze²
- ELAK-Richtlinien und Handbücher für den Bund³
- Büroordnung 2004 für den Bund⁴
- Verwaltungsverfahrensgesetze⁵
- Portalverbundprotokoll⁶
- Signaturgesetze (*Amtssignatur*)⁷

¹ BGBl. I/162/1999, Bundesarchivgesetz; BGBl. II/367/2002, Verordnung des Bundeskanzlers über die Kennzeichnung, Anbietung und Archivierung von Schriftgut des Bundes (Bundesarchivgutverordnung); BGBl. II/366/2002, Verordnung der Bundesregierung über nicht archivwürdiges Schriftgut des Bundes.

² BGBl. I Nr. 10/2004, Erlassung eines E-Government-Gesetzes (E-GovG); BGBl. II Nr. 289/2004, E-Government-Bereichsabgrenzungsverordnung - E-Gov-Ber.AbgrV.

³ <http://www.digitales.oesterreich.gv.at/site/5286/Default.aspx> (Alle Links in diesem Text wurden am 15.5.2012 auf Funktionsfähigkeit geprüft).

⁴ <http://www.bka.gv.at/DocView.axd?CobId=33038>.

⁵ BGBl. I Nr. 51/1991, Allgemeines Verwaltungsverfahrensgesetz (AVG).

⁶ <http://reference.e-government.gv.at/PVP2.2761.0.html>; regelt dadurch sekundär auch den Aufbau der Geschäftszahlen im ELAKimBund.

⁷ BGBl. I Nr. 190/1999, Bundesgesetz über elektronische Signaturen (Signaturgesetz - SigG).

- Zustellungsgesetze (digital/dual)⁸
 - Gleichwertigkeitsgesetze (Bürgerkarte, Barrierefreiheit)⁹
2. Administrativ-organisatorische Vorgaben
- Verwendung des OAIS-Referenzmodells
 - Konvertierung des Contents in PDF/A
 - Auslegung des Systems mandantenfähig und modular
 - Anbindung des Portalverbundprotokolls (PVP) als Rechte- und Rollenverwaltung
 - Barrierefreiheit
 - Anbindung einer AIS-Schnittstelle (= Archivinformationssystem)
 - Access-Zugang über Webshop inkl. Billingsystem
 - Langfristig: Zertifizierung als *vertrauenswürdige Archiv*
3. Fachspezifische Anforderungen
- Abbildung des gesamten Archivierungszyklus
 - Durchführung sämtlicher Prozesse auf möglichst einer (Arbeits-)Oberfläche
 - Benutzerfreundliche, einfache Handhabung (*selbsterklärende Oberflächen*)
 - Nachvollziehbarkeit jedes Arbeitsschrittes
 - Gewährleistung der Authentizität
 - Reportingfunktionen
 - Datensicherheit als oberste Priorität
 - Einhaltung bzw. Erreichung internationaler archivfachlicher Standards
4. Technische Ausstattung
- Verwendung von EDIAKT-II als Ingest- und Access-Schnittstelle¹⁰
 - Datenschutzvorgaben gemäß der (IT)-Sicherheitsrichtlinien¹¹
 - Technischer Aufbau des Archivs an zwei Standorten inkl. Backupszenarien
 - Nutzung des Zentralen Ausweichrechenzentrums des Bundes (= ZAS) in St. Johann im Pongau als Backup-Speicher

⁸ BGBl. I Nr. 200/1982, Bundesgesetz über die Zustellung behördlicher Dokumente (Zustellgesetz - ZustG).

⁹ BGBl. II Nr. 170/2010, E-Government-Gleichwertigkeitsverordnung; <http://www.buergerkarte.at/>.
<http://www.digitales.oesterreich.gv.at/site/5566/default.aspx>.

¹⁰ <http://reference.e-government.gv.at/Veroeffentlichte-Informationen.601.0.html>.

¹¹ <http://www.digitales.oesterreich.gv.at/site/5743/default.aspx>.

- Vergabe der operativen Betriebsführung an Externe (=> Bundesrechenzentrum/BRZ)
- Einhaltung sämtlicher IT-Vorgaben bzw. Nutzung der Infrastruktur des Bundes¹²
- Verwendung von SOA und Open-Source Produkten
- Deutsche Sprachführung
- Verantwortung für Gesamtbetriebsführung des digLA
- Support und Weiterentwicklung (Hotline, Releasemanagement, etc.)
- Software und Hardware auf *aktuellem Stand der Technik*¹³

Zusätzlich zu den in den Punkten 1.-4. aufgelisteten Vorgaben waren als allgemeine Faktoren zu beachten:

- Die Gesamtprojektlaufzeit beträgt 9 Jahre
- Die Umsetzungsphase bis zum Produktivstart des digLA umfasst ein Jahr
- Die operative Betriebsführung inkl. Hotline startet ab Produktivsetzung des digLA
- Ein einmaliger kompletter Hardwaretausch während der Projektlaufzeit ist vorzusehen
- Erstmalige Migration der Dokumente (Content) in PDF/A ist vorzusehen
- Erstmalige Datenübertragung aus dem AIS ist durchzuführen
- Preservation Prozesse bzw. Migrationen sind mit zu beachten/anzubieten¹⁴

Die Summe all dieser Parameter war bei der Projektplanung des digLA bzw. der Erstellung der Ausschreibungsdokumente zur Suche eines technischen Partners in Bezug auf die Budgetplanungen zu berücksichtigen und auch dementsprechend zu kalkulieren. Natürlich werden nicht für jedes digitale Archiv alle genannten Punkte zum Tragen kommen, aber die generellen Anforderungen dürften für viele Archive ähnlich sein, sodass sich daraus vier große Kostenfaktoren, die in jeden Fall ihre Gültigkeit haben, ergeben:

- Personalressourcen (intern und extern)
- Hardware (Beschaffung und Wartung)
- Software (Beschaffung und Wartung)
- Betrieb (laufend)

Diese Kostenfaktoren sind immer im jeweiligen zeitlichen Kontext der vorhin definierten Aufbewahrungsdauer bzw. Systemplanung zu betrachten: Wie lange bzw. wann stehen welche Ressourcen zur Verfügung? Wie lange benötigen welche Schritte zur Umsetzung? Für wie lange benötigt man das digitale Archiv, sprich den laufenden Betrieb?

¹² <http://www.digitales.oesterreich.gv.at/site/5242/default.aspx>.

¹³ Vgl. zu allen technischen Punkten die AVB-IT des Bundes: <http://www.bbg.gv.at/kunden/beratung/vergabekompetenz-center/gesetze-verordnungen/oesterreichische-vergabevorschriften/allgemeine-vertragsbedingungen-der-republik-oesterreich-fuer-it-leistungen-avb-it/>.

¹⁴ Vgl. zu den Ausschreibungskriterien: http://www.infodienst-ausschreibungen.at/ausschreibungen/Druckversion~id~33a9e124-6c0d-102b-91be-001a9243d5c3~pk_id~3290496.htm.

Im Weiteren ist für alle Kalkulationen zwischen Erstinvestition und laufenden Kosten zu unterscheiden. Meist wird ein Projekt mit einem für einen bestimmten Zeitraum berechneten Budget aufgesetzt, ohne dabei die langfristigen und damit stetigen Betriebskosten eines digLA mit einzukalkulieren. Dabei sind im Wesentlichen diese laufenden Zahlungen die eigentlichen Kostentreiber.

Zu den kurzfristigen bzw. einmaligen Anschaffungskosten gehören im Normalfall die Ausgaben für die notwendige Soft- und Hardware sowie für die Inbetriebsetzung des digitalen Archivs. Die regelmäßigen Aufwendungen resultieren aus den Unkosten für die Rechenzentren und/oder Server sowie dem täglichen Bedarf für Datenübernahmen, Erschließungsarbeiten und Benutzung. Langfristig ergeben sich zusätzlich Gebühren aus allfälligen Migrations- und/oder Preservation Projekten.

Kostenfaktoren im Detail

Personal

Die für ein digLA-Projekt vorzusehenden Personalressourcen werden im Allgemeinen für drei Ablaufphasen benötigt. Die Zusammensetzung der Teams resultiert aus der Überlegung ob die anfallenden Aufgaben mit internen oder externen Mitarbeitern bewältigt werden können:

Projektplanung

Projektplanung erfolgt meist in internen, fachlichen Arbeitsgruppen der Organisation und dient zur Evaluierung der Anforderungen, der Prozessdefinitionen sowie der Projektplanung (z.B. inkl. Erstellung von Ausschreibungsunterlagen wie im Falle des ÖStA)

Teilnehmerkreis: Interne (Fach)Mitarbeiter, (externe?) Berater, Techniker, Juristen

Projektumsetzung

Die Umsetzung wird im Projektteam durchgeführt und beinhaltet alle Vorgänge im Sinne eines modernen Projektmanagements wie z.B. Analysephase, Programmierungs- oder Umsetzungsphase, Testphase, Schulungsphase und Abnahmevorgang

Teilnehmerkreis: Interne und externe Mitarbeiter und Fachexperten, (externe?) Berater, Techniker, Juristen, wirtschaftliche Controllern, Projektmanager, etc.

Laufender Betrieb

Der laufende Betrieb umfasst alle Ansprechpartner für den alltäglichen Arbeitsablauf. Inkludiert sämtliche Mitarbeiter für technische (z.B. Betriebsführung, Hotline, Wartung, Support) als auch archivfachliche (z.B. Datenübernahme, Bewertung, Skartierung (Kassation), Erschließung, Benutzung, Administration, Vertrieb) Belange sowie alle Ressourcen für den Gesamtprozess des Preservation Plannings.

Teilnehmerkreis: Interne und externe Archivmitarbeiter/Fachexperten, Techniker, Juristen, Benutzer/Konsumenten, Supportmitarbeiter, Preservation Planner, etc.

Generell ist anzumerken, dass Personalressourcen nach wie vor den größten Rechnungsposten eines jeden Projektes ausmachen, wobei diese Kosten meist *fiktiv* oder *kostenneutral* aufscheinen und/oder in den sonstigen Preisposten inkludiert sind. Ebenso scheinen interne Mitarbeiter einer Organisation selten als zusätzliche Kostenfaktoren auf, weil diese die Projekte *neben* oder im Zuge ihrer täglichen Aufgaben bewältigen. So gesehen sind *lediglich* die externen Personentage oder Vollzeitäquivalente (= VZÄ) für die Budgetierung zu berücksichtigen. Dabei sind jedoch die Gehälter in entsprechender Höhe zu bemessen, da es sich fast ausschließlich um hochqualifiziertes Personal mit akademischer Bildung im Höheren oder zumindest gehobenen Fachdienst handelt, welches zu solchen Aufgaben herangezogen wird.

Im Fall des Digitalen Archiv Österreichs stellte sich die Personalverteilung folgendermaßen dar: Es waren zwei interne Organisationen - Österreichisches Staatsarchiv/Archiv der Republik (= ÖStA/AdR als Auftraggeber fachlich) und Bundeskanzleramt (= BKA für Vergabeangelegenheiten, Vertragspartner und IT-intern) - sowie drei externe Partner - Bundesrechenzentrum (= BRZ als operativer Betriebsführer für Elak und digLA), Atos IT Solutions and Services [vormals Siemens/SIS] (als Auftragnehmer und Betriebsführer des digLA) und Tessella (als Subunternehmer und Softwarelieferant) während der Umsetzungsphase zwischen 2009-2011 am Projekt beteiligt. Außerdem wurden dem Österreichischen Staatsarchiv zur Projektplanung bzw. während der Ausschreibungsphase zwei externe Berater zur Seite gestellt, welche aus dem haus-eigenen Budget finanziert wurden.

Der vorarchivische Bereich der Aktenproduzenten (Bundesministerien, Bundesdienststellen), die sonstigen IT-Partner wie Fabasoft (E-Gov-Suite, ELAKimBund) oder scope solutions ag (Archivinformationssystem) bzw. die zukünftigen Partner im Zuge der Benutzung des digLA mittels Webshop (Kreditkartenbetreiber, SMS-Gateway) und Buchhaltungsagentur (= BHAG als Verrechnungsstelle des Bundes) sind hierbei nicht berücksichtigt. Ebenso wenig ist der erhöhte Personalaufwand im ÖStA durch die Führung/Administration des digitalen Archivs mit einkalkuliert.

Hardware

In Zuge der Planung zur Hardwarebeschaffung sind primär die Anforderungen aus dem IT-Bereich in Bezug auf den aktuellen Stand der Technik sowie die notwendigen Sicherheitsrichtlinien zu berücksichtigen. Ebenso sind Ziel und Zweck der einzelnen Produkte sowie die Dauer und Häufigkeit der Benutzung der Hardware mit zu bedenken. Die Ressourcenfaktoren umfassen hierbei nicht nur die physischen Teile eines digitalen Archivs, wie z.B. Server, Cluster, Tape-Libraries oder Festplatten und Speichermedien, sondern im Wesentlichen auch die Grundkosten die allein die Stellflächen (= Rechenzentrum) der Hardware ausmachen. Wenn dann auch noch Backup-Standorte oder zusätzliche Sicherheitskomponenten hinzukommen, summieren sich die (laufenden) Kosten rasch ins Doppelte oder Dreifache. Da ein digitales Archiv aber den Anspruch an höchste Datensicherheit und geringsten Datenverlust als oberstes Gebot stellt, sind diese Investitionen wohl unvermeidlich.

Tatsache ist, dass für digitale Archive keine *Alltagsgeräte* zur Verwendung kommen, sondern es sich dabei um hoch technologisierte, mehrfach redundante Server und Speichermedien handelt, deren Kosten nicht mit jenen aus dem EDV-Fachgeschäft verglichen werden können. Das weithin *geläufige Argument Speicherplatz kostet ja beutzutage nichts mehr!* ist für den Archivbereich nicht gültig. Speichermedien für große Datenmengen sind immer noch teuer und dauernde digitale Aufbewahrung erfordert zusätzlich ein komplexes Speicherkonzept, permanente Vorkehrungen gegen die Veralterung der Hardwarekomponenten oder auch hohe Leitungsbandbreiten für eine praxistaugliche Datenübertragung in endlichen Zeitabständen. So sind z.B. große Ingestvorgänge, die noch im Laufen sind, wenn das nächste Server-Backup startet, im täglichen Betrieb nicht tragbar. Ebenso wenig wie DIP-Downloads oder Bearbeitungsvorgänge im Archiv, die den Alltagsbetrieb empfindlich behindern, da sie *ewig dauern*.

Den im Anschluss angeführten Kostenbeispielen des Digitalen Archiv Österreichs waren folgende Parameter für die Hardware-Erstbeschaffung zugrundeliegend:

- Vollausstattung für externes Rechenzentrum (nicht im ÖStA)
- Reduzierte Ausstattung für den Backup-Standort
- Vernetzung aller Standorte inkl. operativem Betriebsführer
- Schaffung bzw. Berücksichtigung aller Sicherheitseinrichtungen
- Festlegung von Backup- und Katastrophenszenarien

Zusätzlich wurde ein kompletter Hardwaretausch während der Vertragslaufzeit verpflichtend vorgesehen.

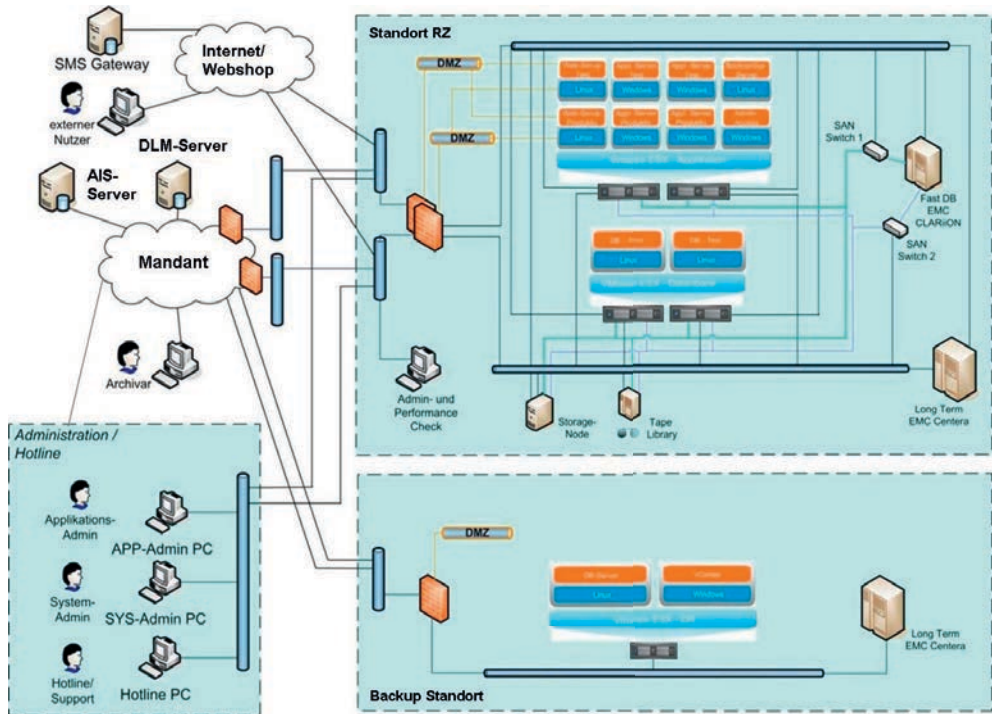


Abb. 1: Schematische Darstellung der Standorte des Digitalen Archiv Österreichs (Quelle: Atos).

Software

Die Softwarekomponenten eines digitalen Archivs, die den eigentlichen Kern des Systems darstellen und im Regelfall die aufwändigsten kundenspezifischen Anpassungen benötigen, stellen meist auch den arbeitsintensivsten – nicht unbedingt den kostenintensivsten – Teil des Archivs dar. Hier ist es besonders wichtig auf die Anforderungen der archivfachlichen Seite Rücksicht zu nehmen und der Kostenfaktor hängt sehr stark mit den zu verwirklichenden Prozessen zusammen. Ebenso muss die Diskrepanz zwischen der Verwendung von Standardprodukten im Unterschied zu Eigenentwicklungen bedacht werden. Auch die Anzahl der benötigten Schnittstellen stellt einen nicht zu unterschätzenden Aufwand dar, ebenso Multiplikatoren wie mannigfach unterschiedliche Daten- und/oder Transportformate, Datengrößen oder die zu handhabenden Arbeitsprozesse. All dies beeinflusst die durchzuführenden Anpassungen der Software und ist somit kostentragend.

Im Fall des Digitalen Archiv Österreichs wurde nach einer Standardsoftwarekomponente gesucht, in welche ohne große Anstrengungen die österreichspezifischen Arbeitsabläufe und Anforderungen integriert werden konnten. Es wurde darauf Wert gelegt, keine rein proprietäre Lösung zu finden, um z.B. im Fall eines Betreiberwechsels, möglichst unabhängig agieren zu

können, andererseits jedoch ein halbwegs durchgängiges kostenneutrales Releasemanagement sowie Wartungs- und Supportvorgänge durchführen lassen zu können.

Die Safety Deposit Box (= SDB) von Tessella erschien in diesem Zusammenhang das richtige Produkt. Alle aus dem Standardprodukt resultierenden Funktionen wurden soweit wie möglich genutzt. Gleichzeitig wurden mit den mandantenspezifischen Zusatzentwicklungen wie Ingest von EDIAKT-Paketen, Metadatenbearbeitung mittels EAD-Editor, Bewertungsvorgänge im Archiv, Löschfunktionen, erweiterten Reportfunktionen, Anbindung eines AIS, Barrierefreiheit des Browsers und nicht zuletzt der erstmaligen deutschen Übersetzung, alle Anforderungen für das Österreichische Staatsarchiv in den letzten Releases umgesetzt, welche zukünftig auch den übrigen Tessella-Kunden zu Gute kommen.

Im Weiteren wurde bei der Softwareeinbindung darauf geachtet, weitestgehend Open-Source Produkte zu verwenden, um die Gesamtlösung einerseits der Ausschreibung entsprechend, leicht anpassbar, modular und mandantenfähig zu gestalten und andererseits kostenschonend zu agieren. So sind unter anderem die Tools für Formaterkennung, Schemaprüfung oder Reporting, sowie Virens Scanner und diverse Konverter frei verfügbar. Auch bei der Anbindung des Webshops und des Billingsystems wurde eine Open-Source Lösung verwendet. Nichtsdestotrotz sind grundsätzlich allfällige Lizenzgebühren (Software, Zertifikate etc.) in die Kostenrechnung mit einzubeziehen.

Aus all den angeführten Parametern wurde in Summe eine Fülle von fachlichen, organisatorischen und administrativen Anforderungen, welche im folgenden Bild nochmals lose veranschaulicht werden, um die Komplexität eines digLA-Projektes aufzuzeigen.

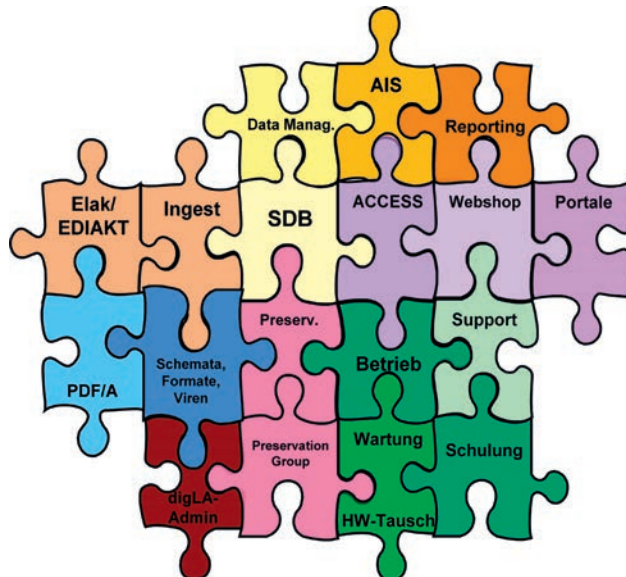


Abb. 2: digLA-Komponenten als Kostenfaktoren (Quelle: eigene Darstellung Fröhlich 2012)

Betrieb

Der vierte und letzte Kostenfaktor eines digitalen Archivs ist der Betrieb, genauer gesagt die Betriebsführung. Man darf in diesem Zusammenhang nicht nur die erstmalige Herstellung der Betriebsbereitschaft eines digLAs bzw. des physischen Betriebsstandortes sehen, sondern muss vor allem die laufende, permanente Betriebsführung mit einkalkulieren. Dies ist, wie bereits erwähnt, häufig das große Manko bei Budgetplanungen derartiger Projekte.

Zur erstmaligen Inbetriebnahme eines digitalen Archivs gelten die bereits unter dem Punkt *Hardware* dargestellten Kostenpunkte für die Bereitstellung der Serverlandschaft in einem oder mehreren Rechenzentrum/-zentren, die dazugehörenden Betriebskosten desselben für Energie, Klimabedingungen und Sicherheitseinrichtungen, die notwendigen Netzwerkanbindungen zwischen Produzenten, Konsumenten und Betreibern sowie die Erstbeschaffung von vorab kalkuliertem Speicherplatz (Festplatten, SAN, etc.). Die Betriebskosten der Standorte sind jedoch immer auch für den laufenden Betrieb zu berechnen.

Im Zuge der operativen Betriebsführung kommen als weitere Komponenten hinzu:

- die Betriebsführung an sich (Personal, Standortbetreuung, Administration)
- Personal für Hotlinebetreuung und Support
- Personal bzw. Ressourcen für laufende Wartungen (z.B. einspielen von Software-Patches)
- Betreuung von notwendigem, insbesondere mandantenübergreifendem Logging und Reporting (z.B. Ticketingsystem)
- Durchführung von sämtlichen Sicherungsvorgängen und Backups

Und nicht zuletzt zählen zu den Betriebsaufgaben auch die laufende Weiterentwicklung des Gesamtsystems (Releasemanagement) sowie die operative Durchführung von Preservation Prozessen.

Für das Digitale Archiv Österreich hat sich aus der Komplexität des Produktes, der Laufzeit des Projektes, den unterschiedlichen Vertragspartnern sowie der Vorgabe der Requirierung weiterer Mandanten, welche das System nutzen sollen, die Notwendigkeit ergeben, Gremien zur Administration des digLA einzurichten. Diese setzen sich aus den digLA-Entscheidungsgremien (Beirat bis Kundenkoordinatoren), in welchen alle Kunden vertreten sind, zusammen. Hier werden primär die kaufmännisch-wirtschaftlichen, rechtlichen und betrieblichen Themen behandelt. Im Weiteren sollen mittel- bis langfristig die in den digLA-Gremien geführten Grundsatzdiskussionen über die Weiterentwicklung des Systems in die entsprechenden Fachgremien getragen werden, um bereits vorab die Qualität des Gesamtsystems zu verbessern. Dies reicht von der direkten Einflussnahme in die Entwicklung des ELAK im Bund (Stichwort *Pflichtfelder für Metadateneingabe*) über die Verbesserung der Zusammenarbeit der Aktenproduzenten/Archive (Stichwort *Datenübergabe, Bewertung*) bis hin zur rechtlichen Notwendigkeit novellierter Archiv- und E-Government-Gesetze.

Die größten budgetentlastenden Synergieeffekte werden sich jedoch aus den zukünftigen Preservation Projekten ergeben, da sich das ÖStA nicht nur mit den bereits gewonnenen Erkenntnissen als Dienstleister für Neukunden zur Verfügung stellen wird, sondern die SDB auch

standardmäßig integrierte Preservation Prozesse vorsieht und allenfalls für alle Mandanten gemeinsame kostenschonende Migrationsprojekte – wie z.B. die Beauftragung externer Dritter – vereinbart werden können. Diesbezüglich wird die im Österreichischen Staatsarchiv verankerte *Preservation Group* als vorgelagertes digLA-Gremium national und international tätig sein.

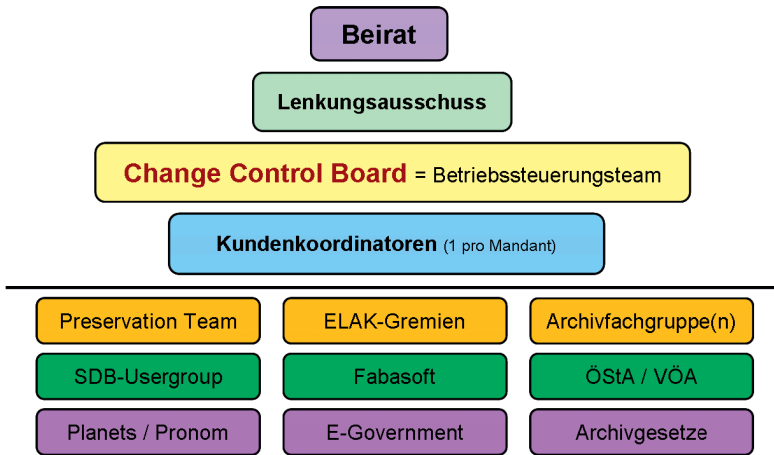


Abb. 3: digLA-Gremien und vorgelagerte Fachgruppen (Quelle: eigene Darstellung Fröhlich 2012)

Kostentreiber

Nach der Darstellung der für jedes digitale Archiv geltenden Rahmenbedingungen sowie der notwendigen Bereiche zur Entwicklung eines digitalen Archivs, sei abschließend noch auf einige kostentreibende Faktoren hingewiesen. Diese treten vor allem im vorarchivischen Bereich auf und werden häufig ebenso wenig wie die laufende Betriebsführung im Budget berücksichtigt. Aus den Erfahrungen des ÖStA kann rekapituliert werden, dass sämtliche, im Vorfeld nicht bedachten Prozesse, oder solche, die als zur Genüge umgesetzt proklamiert aber nie praxistauglich getestet waren, im Nachhinein immense Probleme bzw. Kosten aufwarfen. Die Anpassungen dieser Bereiche bzw. deren Lösungsfindung wird auf rein politisch-rechtlicher und technischer Ebene noch einige Zeit in Anspruch nehmen.

Ein großes Problemfeld stellt in Österreich die durch die Bundesarchivgutverordnungen geregelte Bewertung des Schriftgutes dar. Da de facto der Aktenproduzent für die Setzung des Bewertungsstatus verantwortlich ist, diese jedoch im seltensten Fall der Sicht des Archivars entspricht, ergeben sich hier Diskrepanzen größeren Ausmaßes. Derzeit erhält das ÖStA keinen Einblick in das Produktivsystem des ELAK, sodass über Akteninhalte oder deren Relevanz seitens des Archivs nur gemutmaßt werden kann. Eine professionelle praxistaugliche Bewertung kann vor der Datenübernahme quasi nicht stattfinden und muss daher erst nach der Einlagerung der Akten im Archiv durchgeführt werden. Dies bedeutet doppelten und vor allem kostenrele-

vanten Arbeitsaufwand bei der Erzeugung der EDIAKT-Pakete, der PDF/A-Konvertierung sowie dem anfallenden Speichervolumen bei der Übertragung der Daten in das Archiv und der dadurch bedingten anschließenden Löschung des nicht archivwürdigen Schriftgutes. Derzeit wird daran gearbeitet, eine Vorab-Bewertung sowohl gesetzlich zu verankern, als auch praktisch mit den Aktenproduzenten an Hand der Aktenpläne durchzuführen. Dieses Vorgehen bedingt jedoch den *good will* jedes Einzelnen, und gestaltet sich dementsprechend schwierig.

Zusätzlich ist die vor der Übergabe an das Archiv durch den Aktenproduzenten durchzuführende Erzeugung der PDF/A-Repräsentationen des Contents, als auch die Umwandlung des Gesamtkontaktes in ein EDIAKT-Paket kostenpflichtig. Da die Umsetzung dieser Maßnahmen nicht von den Produzenten selbst, sondern vom Betriebsführer des ELAK im Bund übernommen wird, wird jeder Arbeitsschritt den Auftraggebern in Rechnung gestellt. Dass dafür keinerlei Budgets zur Verfügung stehen, versteht sich von selbst. Daher ist derzeit die Tendenz vorherrschend, so wenig Schriftgut wie möglich an das Österreichische Staatsarchiv abzugeben. Dies entspricht natürlich weder der Intention noch dem gesetzlichen Auftrag des Archivs und es bedarf einer einheitlichen (politisch-rechtlichen) Regelung, um weiterhin substantielles Aktenmaterial zu erhalten.

Abschließend sollen auch noch einige zukünftige Kostentreiber hervorgehoben werden, auf welche im Zuge der Planung von digitalen Archiven ebenfalls selten Rücksicht genommen wird. Sie resultieren aus der langfristigen Speicherung und Pflege der digitalen Dokumente ebenso wie aus der bevorstehenden Benutzung des Systems:

- Vermehrter Personalaufwand durch zusätzliche digLA-Administration (Archivare als Techniker und/oder Administratoren)
- Entsprechende Schulungen des Personals
- Permanente Weiterentwicklung des Know-hows (*auf dem Laufenden bleiben* durch Fortbildung, Tagungen etc.)
- Notwendige Anpassungen durch Änderungen der Rahmenbedingungen (neue Prozesse, novellierte Gesetze, andere Organisation)
- Wartung, Support, Weiterentwicklung um am *aktuellen Stand der Technik* zu bleiben
- Steigender Bedarf an Speicherplatz durch wachsende Datenmengen
- Regelmäßiger Hardwaretausch, allenfalls Übersiedlung von Rechenzentren
- Aufwände durch mögliche System- und/oder Betreiberwechsel
- Aufwände für sämtliche Anforderungen des Preservation Plannings
- Notwendige Verbesserungen der Hardware-Landschaft durch steigende Benutzerzahlen (Stichwort: Performance)
- Zusatzkosten durch Webshopbetrieb (Buchhaltung, *Call Center*, etc.)

Rechenbeispiele zum Digitalen Archiv Österreich

In den angeführten Beispielen wird jeweils von einer dauernden (= langfristigen) Aufbewahrung ausgegangen. Die angegebenen Kosten sind Durchschnittswerte aus marktüblichen Gehältern, handelsüblichen Preismodellen, State of the Art-EDV-Standards sowie diversen Anbotslegungen. Sämtliche Zahlen sind gerundete Nettobeträge in Euro, exklusive einer gesetzlichen Mehrwert- oder Umsatzsteuer. Ebenso sind alle Angaben im Kontext des digLA-Projektes des Österreichischen Staatsarchivs gemäß den in den Vorkapiteln aufgezeigten Planungen zu sehen: Laufzeit des Projektes (bzw. der Vergabe) auf 9 Jahre inklusive aller angegebenen Zusatzentwicklungen und Anforderungen, die ein *einfaches* digitales Archiv sicherlich überschreiten. Die angegebenen Zahlen entsprechen nicht den realen Preisen des Digitalen Archiv Österreichs, welches pauschal mit rund 4,6 Mio. € kalkuliert wurde.¹⁵

Beispiel 1 – (fiktive) Personalkosten:

Projektlaufzeit ÖStA: April 2007 – Dezember 2018 => 140 Monate:

Beteiligte intern (3 x ÖStA, 3 x BKA, 2 x Consulter => 8 VZÄ)	≈	21.000 €
<u>Beteiligte extern (10 x SIS/Tessella, 2 x BRZ => 12 VZÄ)</u>	≈	<u>42.000 €</u>
Gesamt / Monat:	=	63.000 €
Gesamt / 140 Monate:	=	8.820.000 €

Beispiel 2 – Kosten für die Betriebserrichtung:

Projektlaufzeit ÖStA: Dezember 2009 – Dezember 2018 => 108 Monate

Rechenzentrum (physisch, Hardware)	≈	2.000 €
+ Betriebserrichtung, Software, Lizenzen	≈	62.000 €
<u>+ Support und Weiterentwicklung</u>	≈	<u>30.000 €</u>
Gesamt / Monat: ¹⁶	=	94.000 €
Gesamt / 108 Monate:	=	10.152.000 €

¹⁵ Vgl. hierzu die Presseaussendung des BKA sowie diverse Zeitungsmeldungen vom Jänner 2010: http://www.bundeskanzleramt.at/site/cob__37814/currentpage__7/6893/default.aspx.

¹⁶ Hinzu kommen noch etwaig anfallende Kosten für zusätzlichen Speicherplatz pro TB, welcher über dem derzeit einkalkulierten Jahresmaximum von 8 TB liegt. Preis pro TB siehe Beispiel 6.

Beispiel 3 – Einmalkosten/Erstinvestition:

Projektlaufzeit ÖStA: Dezember 2009 – Jänner 2012 => 25 Monate

Personal (<i>exkl. Personal intern = 42.000 x 25</i>)	≈	1.050.000 €
+ Hardware/Software (= 94.000 x 25)	≈	2.350.000 €
<hr/>		
Gesamt / 25 Monate = 2 Jahre:	=	3.400.000 €

Beispiel 4 – Laufende Kosten:

Projektlaufzeit ÖStA April 2011 – Dezember 2018 => 92 Monate

Personal (<i>10 VZÄ exkl. Personal intern</i>)	≈	35.000 €
+ Betrieb Rechenzentren	≈	18.000 €
+ Betriebsführung / Hotline / Wartung / Weiterentwicklung	≈	11.000 €
<hr/>		
Gesamt / Monat:	=	64.000 €
Gesamt / 92 Monate:	=	5.888.000 €

Beispiel 5 – digLA-Projektgesamtkosten:

Personal exkl. intern (Vorprojekte und Umsetzung)	≈	1.050.000 €
(<i>Dezember 2009 – Jänner 2012 => 25 Monate x 42.000</i>)		
+ Hardware/Software (Erstinvestition, Umsetzung, Wartung)	≈	10.152.000 €
(<i>Dezember 2009 – Dezember 2018 => 108 Monate x 94.000</i>)		
+ Betriebsführung laufend	≈	2.668.000 €
(<i>April 2011 – Dezember 2018 => 92 Monate x 29.000</i>)		
<hr/>		
Gesamt / 108 Monate = 9 Jahre	=	13.870.000 €

Beispiel 6 – Derzeitige Kosten für eine Mandanteneinrichtung im digLA:¹⁷

Diese Berechnung gilt für weitere Kunden, welche das Gesamtsystem Digitales Archiv Österreich gemäß den Konditionen für das Österreichische Staatsarchiv übernehmen. Das bedeutet Datenhaltung im gleichen Rechenzentrum, mit derselben Software- und Hardware-Ausstattung inklusive identen Schnittstellen und Basisfunktionen. Lediglich Prozesse wie Layoutgestaltung, Reportdefinitionen, Metadatenschema, Rollen-Rechteverwaltung und dergleichen werden individuell angepasst (= Mandanteneinrichtung). Allfällige Lizenzkosten für weitere Kunden entfallen, da das ÖStA/BKA eine Generallizenz für alle öffentlichen Körperschaften Österreichs ausverhandelt hat.¹⁸

Erstinvestition einmalig (= <i>Basiskosten plus Mandanteneinrichtung</i>)	≈	75.000 €
+ Laufende Kosten / Monat (<i>Basis 1 TB und 3 User</i>)	≈	8.325 €
Gesamt:	≈	83.325 €

Die Monatskosten setzen sich pro User und TB wie folgt zusammen:

Speichermedium pro übergebenem TB à 120 € (<i>2 × RZ Wien, 2 × RZ St. Johann, 1 × Bandsicherung</i>)	≈	120 €
+ Server- und Infrastrukturkosten pro TB à 1.265 € (<i>2 × RZ + Investition für österreichspezifische Softwareanpassungen</i>)	≈	1.265 €
+ Softwarewartung und Weiterentwicklung (<i>Standardsoftware SDB plus alle sonstigen Open Source Tools</i>)	≈	5.245 €
+ Betriebsführung inkl. Hotline Sockelbetrag	≈	1.500 €
+ Betriebsführung pro User 65 € (3 User Minimum)	≈	195 €
Gesamt:	≈	8.325 €

Beispiel 7 – Mandanteneinrichtung mit 3 TB und 6 Usern:

Erstinvestition einmalig (= <i>Basiskosten plus Mandanteneinrichtung</i>)	≈	75.000 €
+ Laufende Kosten/Monat (<i>Basis 1 TB und 3 User</i>)	≈	11.610 €
Gesamt:	≈	86.610 €

¹⁷ Richtwerte gemäß den zwischen BKA und Atos ausgehandelten Konditionen zur Vermarktung des digLA mit Stand März 2012. Der Mindestumfang für einen eigenen Mandaten beträgt ein Datenvolumen von 1 TB und 3 User pro Jahr. Die Useranzahl wird gestaffelt verrechnet: Der Sockelbetrag für die Betriebsführung inkl. Hotline (1st Level Support) beträgt 1.500 €. Pro User werden 65 € veranschlagt (bis 5 User). Der 2nd Level Support (= 320 € pro User) ist bis 5 User gratis. Ab dem 6. User erhöhen sich somit die Kosten für Betriebsführung und Support pro User um 385 € / Monat.

¹⁸ Dies führt zu einer beträchtlichen Kostenreduzierung für weitere Kunden, da keine eigenen Ausschreibungen durchgeführt werden müssen, sondern man mit denselben Bedingungen wie das ÖStA dem digLA-Vertrag beiträgt.

Die Monatskosten setzen sich pro User und TB wie folgt zusammen:

Speichermedium pro übergebenem TB à 120 € (2 × RZ Wien, 2 × RZ St. Johann, 1 × Bandsicherung)	≈	360 €
+ Server- und Infrastrukturkosten pro TB à 1.265 € (2 × RZ + Investition für österreichspezifische Softwareanpassungen)	≈	3.795 €
+ Softwarewartung und Weiterentwicklung (Standardsoftware SDB plus alle sonstigen Open Source Tools)	≈	5.245 €
+ Betriebsführung inkl. Hotline Sockelbetrag	≈	1.500 €
+ Betriebsführung pro User 65 € (bis 5 User)	≈	325 €
+ Betriebsführung + Support ab 6. User (320 + 65€)	≈	385 €
Gesamt:	≈	11.610 €

Archivierung von digitalisierten Archivalien in einem Rechenzentrum

Von GABRIELE STÜBER

Das Thema der Archivierung digitalisierter Archivalien berührt einen Bereich, der im Kontext einer Sicherung von Unterlagen aus einer neuen medialen Dimension zunehmende Wichtigkeit erlangt. Der kurze Beitrag stellt am Beispiel des Zentralarchivs der Evangelischen Kirche der Pfalz dar, wie diese Aufgabe angegangen und umgesetzt wird. Im Anschluss werden einige grundsätzliche Fragen formuliert, die die Archive gleich welcher Sparte aus Sicht der Verfasserin begleiten und zur Stellungnahme herausfordern.

Kurzporträt des Zentralarchivs der Evangelischen Kirche der Pfalz, Speyer

Das Zentralarchiv der Evangelischen Kirche der Pfalz ist zuständig für alle in der Landeskirche erzeugten Unterlagen inklusive die der Diakonie. Im Archiv arbeiten sieben hauptamtliche Kräfte, davon drei Fachkräfte, vier ehrenamtliche Kräfte sowie wechselnde Aushilfskräfte. Wir beschäftigen Praktikantinnen und Praktikanten aus Schule und Studium sowie im Rahmen von beruflichen Wiedereingliederungsmaßnahmen und wechselnde Projektkräfte.

Derzeit umfasst der Umfang aller aufbewahrten Unterlagen etwa 5 000 laufende Meter in drei Magazinegebäuden. In dem hier zu behandelnden Kontext sind folgende Unterlagen wichtig: 15 000 Bildquellen und 20 000 Baupläne. Aufgrund seiner regionalen Zuständigkeit und seines Bestandsprofils von etwa 800 Beständen ist das landeskirchliche Zentralarchiv mit einem Staatsarchiv vergleichbar, allerdings einem sehr kleinen. Ungeachtet dessen fallen hier alle Aufgaben an, die in jedem Archiv bewältigt werden müssen.

Erzeugung von Datenmengen durch Digitalisierung

In den vergangenen 15 Jahren wurden nach und nach Bildquellen (Fotos und Objekte der Sammlung Volksfrömmigkeit) digitalisiert, dies geschah in Eigenregie. 45 Filme aus der Ostasienmission wurden durch eine Firma digitalisiert, desgleichen die Glasdias der Mission. Alle Digitalisate wurden auf der landeskirchlichen Serverfarm gespeichert, hier kam es bald zu Engpässen, und die EDV-Abteilung wies das Archiv auf die Problematik der wachsenden Datenmengen hin.

Eine weitere Verschärfung der Lage und ein gewaltiger quantitativer Sprung wurden durch das Projekt einer Digitalisierung der Kirchenbücher erzeugt. Die Digitalisierung von ca. 6 000 Kirchenbüchern aus dem 16. bis 20. Jahrhundert mit durchschnittlich 300 Seiten war in der

Serverlandschaft nicht mehr zu realisieren. Diese Digitalisierung erfolgt seit 2011 mit einem professionellen Anbieter auf zwei Wegen.

Aus den bereits vorhandenen Sicherungsfilmern werden für Publikationszwecke und Ausdrücke im Rahmen von Beglaubigungen Dateien mit einer Auflösung von 300 dpi im JPEG-Format mit 80-prozentigem Qualitätserhalt erzeugt, für die Anwendung im Lesesaal und die Einspielung in das geplante Kirchenbuchportal der Evangelischen Kirche in Deutschland werden diese JPEG-Bilddateien mit 30-prozentigem Qualitätserhalt sowie mit einem hinterlegten Wasserzeichen gefertigt.¹

Bei mäßiger Verfilmungsqualität werden die Originalkirchenbücher nochmals mit 24 Bit Farbtiefe digitalisiert. Die Digitalisate der Sicherungsfilme sind – dem Standard der Sicherungsverfilmung entsprechend – in Schwarzweiß gehalten. Bei der Digitalisierung der Kirchenbücher werden keine DFG-Standards angewendet.

Sicherung der Digitalisate des Zentralarchivs im Rechenzentrum

Über die landeskirchliche Verwaltung wurde im März 2011 ein Hosting-Vertrag mit dem Kirchlichen Rechenzentrum in Eggenstein-Leopoldshafen bei Karlsruhe geschlossen. Dieser Vertrag hat zunächst eine Laufzeit von vier Jahren. Das Rechenzentrum stellt einen Fileserver mit 10 TB für Kirchenbücher, Fotos und Filme zur Verfügung. Der Zugang für das Archiv erfolgt über eine FTP-Verbindung mit Benutzungsrechten für derzeit vier Personen, was als ausreichend erachtet wird.

Die Daten werden mittels eines externen Datenträgers (Wechselfestplatte) importiert.

Eine Vollsicherung erfolgt im Rechenzentrum nach jedem größeren Datenzuwachs. Außerdem wird an jedem Werktag eine nächtliche Sicherung vorgenommen. Die folgende Abbildung veranschaulicht die Konfiguration.

¹ Vgl. hierzu *Gabriele Stüber*: Ein neuer Zugang zu Kirchenbüchern: www.kirchenbuchportal.de. In: *Blätter für pfälzische Kirchengeschichte und religiöse Volkskunde* 77 (2010), S. 217–223.

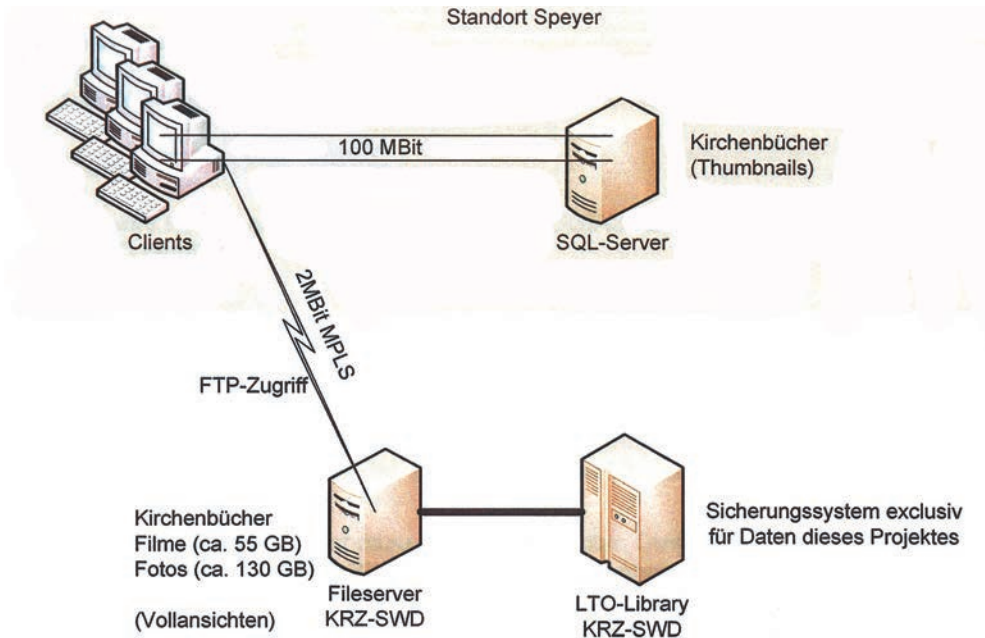


Abb. 1: Konfiguration der Datensicherung des Zentralarchivs der Evangelischen Kirche der Pfalz im Rechenzentrum Eggenstein-Leopoldshafen bei Karlsruhe, erstellt vom Rechenzentrum.

Durch die FTP-Verbindung werden die einzelnen Bilddateien hoch- bzw. heruntergeladen, wobei von einer Häufigkeit von bis zu 10 Bildern pro Tag in der Anfangsphase ausgegangen wird. Diese geringe Zahl ist erstaunlich, doch sei darauf hingewiesen, dass nur im Falle einer Publikation oder zu Beglaubigungszwecken die hochauflösenden Formate benötigt werden. Als Storage-Einheit ist exklusiv für das Archiv eine LTO-Library vorhanden.

Die JPEG-Dateien in geringer Auflösung bleiben als Arbeitsbilder auf den Servern der Landeskirche. Im Tagesgeschäft wird fast ausschließlich mit diesen gearbeitet.

Service und Support bestehen rund um die Uhr an 365 Tagen im Jahr. Die Reaktionszeit bei Störungen beträgt je nach Fehlerklasse zwei bis vier Stunden.

Die Kosten setzen sich zusammen aus

- Einmalkosten für Installation und Konfiguration von 2.135 € brutto

Dazu gehören die Bereitstellung der notwendigen Hardwareinfrastruktur, des erforderlichen Betriebssystemes sowie die Einrichtung der Sicherungsjobs für die Daten und Wartung des Fileservers.

- Laufende Kosten pro Monat: 878 € brutto

Aus diesen Rahmendaten ergeben sich an Gesamtkosten für das erste Betriebsjahr (März – Dezember 2011) 10.900 € brutto, für das zweite Betriebsjahr (Januar – Dezember 2012) 10.540 €. Im Anschluss an das zweite Jahr erfolgt eine Kostenprüfung und nochmalige Verhandlung mit dem Rechenzentrum.

Probleme der Datenzugänglichkeit

Via FTP ist grundsätzlich keine Voranzeige von Images möglich. Man muss also die gewünschte Seite bei einer Anforderung für eine Publikation genau bestimmen können, was aber mittels der Seitenkennung keine Schwierigkeit darstellt.

Im laufenden Betrieb sind bisher keine gravierenden Probleme aufgetreten. Der häufigere Rückgriff auf die Digitalisate in Druckqualität wird zeigen, ob es zu Engpässen kommt.

Bei den Fotos haben wir das FTP-Problem umgangen, indem wir zwar die Fotodaten im Rechenzentrum sichern, aber die Großdateien auf dem Archivserver der Landeskirche vorhalten. Da das Archiv als landeskirchliche Fotoagentur sehr häufig in Anspruch genommen wird, wäre ein anderes Verfahren mit zu hohem Zeitaufwand verbunden. Der weitere Fortschritt der Technik wird zeigen, ob diese Zugangsprobleme bei den großformatigen Dateien nur ein vorübergehendes Phänomen sind.

Ausblick und Fragen

Es wurde bereits darauf hingewiesen, dass der Digitalisierung im Zentralarchiv keine DFG-Normen zugrunde liegen. Die Umsetzung der DFG-Anforderungen erzeugt ein großes Datenvolumen, dessen Unterhaltskosten so manches Archivbudget sprengen, jedenfalls ist das im Zentralarchiv der Fall.

Ich betone diese Praxis noch einmal unter Hinweis auf die Tatsache, dass wir eines der vielen kleineren Archive repräsentieren. Aus der Sicht von Staatsarchiven mag sich die Erfüllung der DFG-Norm als unproblematisch erweisen, zumal die Drittmittelförderung der DFG an die Voraussetzung der Normeinhaltung gebunden ist. Die Kirchenarchive zahlen die Digitalisierung in der Regel aus eigenen Mitteln und befinden daher auch über die zugrunde liegende Norm.

Schreibt man diese Entwicklung fort und bezieht alle Archivsparten in die Betrachtung ein, so erhalten wir unter Umständen zwei (und mehr) Klassen von Archiven in der digitalen Welt. Das hängt vor allem mit der Verschärfung der finanziellen Situation im Kulturbereich zusammen, die sich bei den Archivträgern schon jetzt sehr unterschiedlich darstellt.

Aus der vorfindlichen Konstellation erwachsen folgende Leitfragen:

1. Welche Sicherungsanforderungen (ich vermeide bewusst den Begriff Archivierungsanforderungen) sind an Digitalisate zu stellen, die aus papierbasiertem Archivgut erzeugt wurden? Es ist sicher eine andere Messlatte anzulegen als bei digital born documents, aber welche? Sicherungsanforderungen in diesem Bereich sind zweifellos erforderlich, da die Digitalisierung Kosten erzeugt.

2. Aus der Sicht kleinerer Archive stellt sich die fast schon existentielle Frage, wie sie mit der digitalen Herausforderung umgehen und angesichts ihrer sinkenden Personal- und Sachmitteleinsatzung agieren bzw. reagieren sollen. Aber vielleicht ist das nicht nur eine Frage an die kleineren Archive... Diese digitale Herausforderung umfasst drei Bereiche:
- die Archivierung von Unterlagen aus digitalen Systemen (das Thema dieses Arbeitskreises seit 1997 und eine neue Standardaufgabe von Archiven);
 - die Erzeugung digitaler Unterlagen aus herkömmlichem Archivgut und deren Präsentation im Netz, weil der Druck der Öffentlichkeit erheblich ist und ein schneller und niedrigschwelliger Zugang zu Archivgut eine Standardforderung geworden ist. Ein Archiv, das nicht im Netz präsent ist, läuft Gefahr, mit seinen Beständen nicht mehr wahrgenommen zu werden;
 - die Sicherung der digitalisierten Archivalien.

Die Kosten für die Erzeugung, Archivierung und Sicherung von digitalen Unterlagen sind eine neue Kategorie für die Budgets aller Archive. Angesichts schrumpfender öffentlicher Haushalte und bleibender Kostenstellen im Bereich der traditionellen Kernaufgaben der Archive sind alle Kostenarten bereits mittelfristig nicht mehr zu schultern. Es könnte zu einer Verschiebung der Finanzmittel zu Lasten der traditionellen Kernaufgaben kommen.

Im Rahmen einer prinzipiell offenen Kosten-Leistungs-Rechnung drängt sich eine Frage auf, die nicht nur Rechnungshöfe stellen könnten: In welchem Rahmen können Archive ihre Leistungen zukünftig noch gebührenfrei anbieten? Müssen sie nicht Segmente ihres Services mit Gebühren belegen, um zumindest eine Teilrefinanzierung zu erzielen? Die Investitionen der Deutschen Forschungsgemeinschaft im Bemühen um ein *public brain* etwa im Kontext der Deutschen Digitalen Bibliothek lassen digitale Datenmengen zurück, die auf Dauer gesichert sein wollen – wer trägt die Folgekosten?

Angesichts dieser Herausforderungen bleibt aber auch die Hoffnung, dass mit einer leistungsfähigeren Technik unter Umständen auch die Kosten dafür sinken werden. Überdies könnte sich die Bereitschaft der Archive aller Sparten, Aufgaben gemeinsam in einem Verbund von Einrichtungen zu bewältigen, als eine zukunftsfähige Vernetzung erweisen.²

² Die Ausgabe der Zeitschrift *Archivar*, 65. Jg. Heft 1/Februar 2012 belegt diverse Verbundlösungen.

10 FAQs. Argumente zu Bedarf und Notwendigkeiten der digitalen Archivierung

Von PETER SANDNER

Es gibt keine *dummen Fragen* zur digitalen Archivierung. Zu zehn häufig gestellten Fragen (FAQs) aus der Praxis werden archivfachliche Antworten und Argumentationslinien skizziert, die die sachlichen Notwendigkeiten, den Ressourcenbedarf und die rechtlichen Rahmenbedingungen der digitalen Archivierung begründen.

Einleitung

FAQs – das sind häufig gestellte Fragen¹, auf die die Anbieter von Softwareprodukten oder Websites auf Vorrat antworten, um den Anwendern die Nutzung der Angebote zu erleichtern. Die folgenden zehn FAQs sind reale Fragen aus der fünfjährigen Planungs- und Aufbauphase des Digitalen Archivs Hessen.² Es sind Fragen, die wörtlich oder sinngemäß von politischen und administrativen Entscheidungsträgern, von Haushalts- und Personalverantwortlichen, von Archivkolleginnen und -kollegen oder von Kontaktpersonen in den betreuten Dienststellen gestellt wurden.

Mit den skizzierten Antworten wird die Absicht verfolgt, sowohl die fachlichen Notwendigkeiten als auch den materiellen Bedarf der digitalen Archivierung zu begründen. Ziel ist es, unzutreffende Annahmen zu entkräften und schlüssige Argumentationslinien aufzuzeigen.

¹ Frequently asked questions.

² Seit Beginn des Vorprojekts zum Digitalen Archiv im 3. Quartal 2007. Zum Digitalen Archiv Hessen: Peter Sandner: Digitales Archiv Hessen eröffnet. Die Archivierung originär digitaler Archivalien hat im Hauptstaatsarchiv begonnen. In: Archivnachrichten aus Hessen 11/1 (2011). S. 5; Sigrid Schieber: Das Digitale Archiv der hessischen Staatsarchive. Ein Werkstattbericht. In: Archivar 64 (2011). S. 73–78.

Administration³

Frage 1: Warum geben Sie nicht alles dem IT-Dienstleister? –

So könnte man sich doch ein zusätzliches digitales Archiv sparen ...

Die neue Aufgabe der Archivierung digitaler Aufzeichnungen erschien manchen öffentlichen Archiven zunächst als unlösbares Problem. Daher wurde häufig die Übertragung der Aufgabe an den jeweiligen IT-Dienstleister oder an die IT-Abteilung der Verwaltung überlegt. Verwaltungsinterne Richtlinien, die eine Zentralisierung der IT-Aufgaben bei der IT-Stelle vorschreiben, förderten diese Tendenz.

IT-Dienstleister sind Spezialisten für den Betrieb von IT-Systemen und für Datenspeicherung und Backups. Da auch *Datenarchivierung* zum Aufgabenspektrum von IT-Stellen zählt, schien die Beauftragung mit dem Aufbau des digitalen Archivs vordergründig sinnvoll. Letztlich geht ein solcher Plan aber von der falschen Vorstellung aus, die digitale Archivierung im Sinne der Archivgesetze sei hauptsächlich eine Frage der Datenspeicherung. Tatsächlich aber muss gelten:

Dauernde Archivingut-Aufbewahrung ist eine archivfachliche Aufgabe.

Datenarchivierung im Sinne der IT meint nämlich gerade nicht dauernde Aufbewahrung nach den Archivgesetzen, sondern deckt bestenfalls die behördlichen Aufbewahrungsfristen ab. Sie hat also eine Perspektive von 10 bis 30, nicht von 1.000 Jahren. Vor allem verfolgt sie einen anderen Zweck: *Archivierung* im Sinne der IT dient der Entlastung aktiver Systeme von seltener genutzten Daten. Das heißt: Die unwichtiger gewordenen Daten werden vor der letztlich vorgesehenen Löschung noch auf einem billigeren Speicherplatz mit einer geringeren Verfügbarkeit geparkt. Im Fokus der *IT-Archivierung* steht die Sicherung von Dateien und Verzeichnissen, nicht von digitalen Archivalien. IT-Dienstleister können damit das archivische Aufgabenspektrum nicht einmal annähernd abdecken.

Die archivfachliche Aufgabenstellung unterscheidet sich grundlegend davon. Öffentliche Archive bilden intellektuell definierte Archivalien⁴ mit dem Ziel, diese dauernd, *für die Ewigkeit* zu bewahren und verfügbar zu halten.

Das Digitale Archiv Hessen hat seit dem Beginn des Aufbauprojekts im Jahr 2009 die Erfahrung gemacht hat, dass über 90 Prozent der Aufgaben, die mit Zeit- (und damit Personalaufwand) verbunden sind, sich nicht effizient an IT-Dienstleister delegieren lassen. Insgesamt stellt sich die Verteilungen des Zeitaufwands wie im Diagramm skizziert dar.

³ Die Gliederung folgt der Einteilung des Functional Model des OAIS und dessen sechs Functional Entities: Reference Model for an Open Archival Information System (OAIS). CCSDS 650.0-M-2 (Magenta Book). Juni 2012, hier S. 4-1 f. Internet: <http://public.ccsds.org/publications/archive/650x0m2.pdf>. – Alle Links zu diesem Beitrag wurden am 15.10.2012 überprüft.

⁴ Zur Notwendigkeit der Bildung von Archivalien als Informationsobjekte im Hinblick auf deren Bestandserhaltung (*Preservation Planning*) siehe unten: Darstellung zu Frage 8.

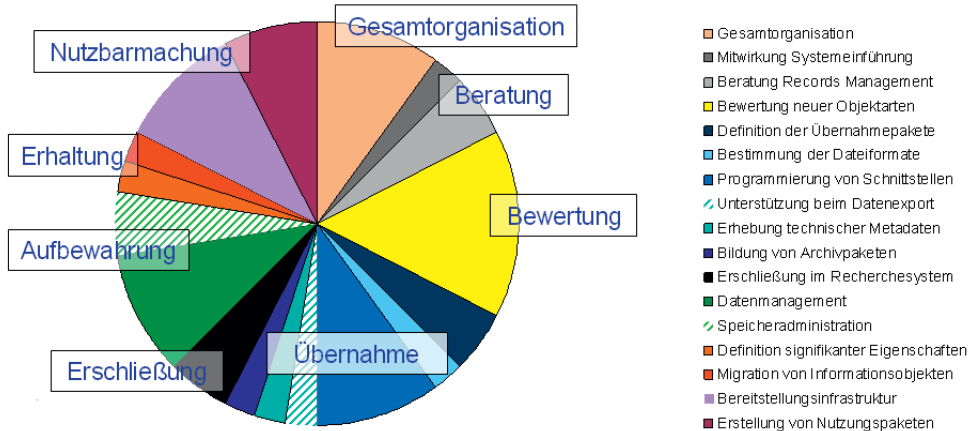


Abb. 1: Anteile am Zeit- und Personalaufwand des Digitalen Archivs Hessen (Schätzwerte für die Jahre 2010 und 2011).⁵

Zusätzlich zum Zeitaufwand sind Sachaufwendungen, vor allem für Hard- und Software und Raumnutzung (auch beim Dienstleister) sowie Gemeinkosten (Archivleitung etc.) zu berücksichtigen. An einen IT-Dienstleister können nach den bisherigen Erfahrungen effizient die Unterstützung der Behörden beim Datenexport und die Administration der Datenspeicher delegiert werden (im Diagramm schraffiert), also ein Zeitaufwand von unter 10 Prozent.

Es lassen sich sogar Argumente für eine Datenspeicherung im Archivgebäude selbst anführen, wenngleich diese voraussichtlich nur von wenigen Archiven bewerkstelligt werden kann: Der Betrieb des *Archival Storage* im direkten, auch physischen Zugriff der Archivmitarbeiter erhöht die Autarkie. Die Bergung der physischen Speichermedien in Katastrophenfällen wäre voraussichtlich nur bei Speicherung *im* Archiv selbst möglich. Beim IT-Dienstleister dagegen ist das Archiv nur ein Kunde unter vielen und genießt nicht die höchste Priorität. Generell stellt der Erfahrungsgewinn des Archivs in IT-fachlichen Fragen, der mit dem Aufbau eines digitalen Archivs einhergeht, einen unschätzbaren Wert dar. Er stärkt die archivische Kompetenz bei der Behördenberatung und bei Übernahmen von digitalen Aufzeichnungen.

Frage 2: Wozu benötigen Sie denn zusätzliches Personal? –

Die archivarischen Aufgaben bleiben doch dieselben ...

Auf den ersten Blick erscheint es nicht logisch, weshalb die Archive Zusatzpersonal für die Aufgabe der digitalen Archivierung beanspruchen müssen. Denn richtig ist ja, dass die archivarischen Aufgaben in den Grundzügen dieselben bleiben: Die Archive beraten weiterhin anbietungs-

⁵ Seit Juli 2012 wird das Digitale Archiv Hessen in der Zeit- und Mengenerfassung separat von den übrigen Archivbereichen ausgewiesen, um belastbarere Zahlen über die Kosten der digitalen Archivierung zu erzielen

pflichtige Stellen, bewerten weiterhin deren Unterlagen und übernehmen diese als Archivgut. Auch die wesentlichen Aufgaben der dauernden Aufbewahrung, der Erhaltung und der Nutzbarmachung von Archivgut bleiben bestehen. Erst auf den zweiten Blick wird deutlich, dass Zusatzaufgaben geschaffen wurden, die zwingend zu der Schlussfolgerung führen:

Neue gesetzliche Aufgaben schaffen zusätzlichen Personalbedarf.

Ein Blick auf die Aufgaben, die mit einem digitalen Archiv einhergehen, macht den Zusatzaufwand gegenüber der Archivierung konventioneller Unterlagen augenfällig:

An erster Stelle gilt es, die *Gesamtorganisation* des digitalen Archivs zu bewerkstelligen: Konzept und Richtlinien sind zu erarbeiten und aktuell zu halten, das Management des digitalen Archivs muss (zusätzlich zum Management des konventionellen Archivs) funktionieren.

Neue Aufgaben ergeben sich auch in der *Behördenberatung* in puncto E-Government (zusätzlich zu den bisherigen archivarischen Maßnahmen wie etwa Registratorenschulungen). Neu sind die Mitwirkung bei der Einführung von IT-Systemen im Hinblick auf eine spätere Aussonderung sowie die Beratung beim Records Management.

Zunächst erscheint es auch so, als ob der Aufwand für die *Bewertung* sich nicht ändere – das gilt aber nur dann, wenn eine Unterlagengruppe wie die Akten komplett von der analogen in die digitale Form überführt wird. In der Praxis aber ist die Bewertung neuer Objektarten (zusätzlich zur Aktenüberlieferung) zu berücksichtigen: Die Bewertung von Informationen aus Datenbanken von Fachverfahren macht einen erheblichen Mehraufwand aus. Zudem müssen auch Webseiten, teilweise als interaktive E-Government-Webanwendungen, oder E-Mail- und Dateisammlungen (zusätzlich zur Überlieferung im Dokumentenmanagementsystem) bewertet werden.

Besonders aufwendig ist die *Übernahme* von digitalem Archivgut, die die Archive nun zusätzlich zur Übernahme der weiterhin genutzten Papierakten erledigen müssen. Im Zuge des *Ingest* muss die Übergabeportion (das SIP⁶) definiert werden; es ist zu klären, welche Feldinhalte aus Datenbanken erhalten werden sollen. Die Dateiformate sind zu bestimmen; in der Regel müssen individuelle Exportschnittstellen programmiert werden. Meist muss das Archiv oder ein IT-Dienstleister die Behörde beim Datenexport unterstützen. Im Zuge des *Ingest*-Prozesses hat das Archiv technische Metadaten für die Archivierung zu erheben; aus den abgelieferten SIPs müssen Archivpakete (AIPs⁷) gebildet werden.

Schließlich fällt bei digitalen Archivalien die archivische *Erschließung* im Recherchesystem an (zusätzlich zur Erschließung der weiterhin übernommenen konventionellen Archivalien).

Eine Infrastruktur zur *Aufbewahrung* digitaler Archivalien muss (zusätzlich zum Magazin für Papierakten) aufgebaut und betrieben werden. Hier ist ein aufwendiges Datenmanagement erforderlich, um den Nachweis der Authentizität führen zu können. Eine komplexe Speicherstruktur muss administriert werden, um die Integrität der Daten – und damit der Archivalien – zu sichern.

⁶ Submission Information Package.

⁷ Archival Information Package.

Die digitale *Bestandserhaltung* erfordert vollständig andere Konzepte als die Bestandserhaltung für nichtdigitales Archivgut; sie muss daher komplementär betrieben werden. Notwendig ist dabei insbesondere die Definition der zu erhaltenden signifikanten Eigenschaften der Archivalien; bei drohendem Verlust der Nutzbarkeit von Archivalien muss – die Migrationsstrategie vorausgesetzt – die Überführung in neue Dateiformate betrieben werden. Bei analogen Objekten kann eine Ersatzdigitalisierung erforderlich werden.

Schließlich erzeugt die Aufgabe der *Nutzbarmachung* neue Aufwände (zusätzlich zur Bereitstellung nichtdigitaler Archivalien). Eine Infrastruktur zur Bereitstellung digitaler Objekte muss (weiter-) entwickelt werden. Individuelle Nutzungspakete (DIPs⁸) für Archivnutzer müssen erstellt werden.

Eine erste Bedarfsberechnung der hessischen Staatsarchive (Anfang 2008) ergab einen Zusatzbedarf von ca. 8 Stellen⁹ für Aufbau und Betrieb eines digitalen Archivs. Mit den Landeshaushaltsplänen für die Jahre 2009 und 2010 wurden davon fünf neue Stellen bereitgestellt: eine Stelle einer Archivarin des höheren Dienstes, zwei Stellen von Archivar/innen des gehobenen Dienstes und zwei Stellen von Informatiker/innen des höheren bzw. gehobenen Dienstes. Aufgrund der Eigenentwicklung von Software – statt einer deutlich teureren Beschaffung auf dem Markt – hat sich ein Zusatzbedarf an IT-Stellen ergeben, die zunächst befristet mit Projektfördermitteln besetzt werden konnten. Die positive Entwicklung des Digitalen Archivs Hessen hat dazu geführt, dass die Archivierung digitaler Aufzeichnungen fachgerecht begonnen werden konnte. Gleichwohl hat sich gezeigt, dass der (noch) nicht gedeckte Stellenbedarf zu deutlichen Rückständen in der Zielerreichung führt. Die Beratung der Behörden bei E-Government-Projekten müsste deutlich verstärkt werden, um spätere Aussonderungen zu erleichtern. Aufgrund von Personalmangel kann nur etwa ein Drittel der anstehenden Fachverfahrensdaten tatsächlich auch zeitnah übernommen werden.

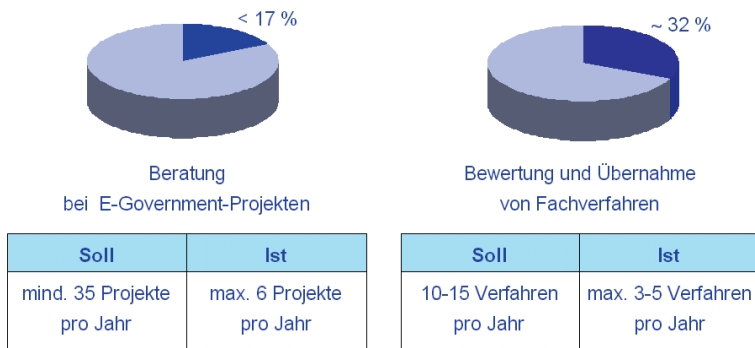


Abb. 2: Zielerreichung angesichts vorhandener Personalkapazitäten (Angaben für das Digitale Archiv Hessen 2010/2011)

⁸ Dissemination Information Package.

⁹ 2,94 Archivar/innen höherer Dienst; 2,79 Archivar/innen gehobener Dienst; 0,95 Informatiker/in gehobener Dienst (A 13).

Bei mangelnden Ressourcen für digitale Archive besteht die Gefahr, dass gesetzliche Pflichtaufgaben nicht erfüllt werden können. Unter Umständen können Überlieferungslücken entstehen, die später nicht mehr zu schließen sind, etwa weil das IT-System nicht mehr für eine Aussonderung zur Verfügung steht. Eine verzögerte Aussonderung führt zudem zur Kollision mit gesetzlichen Löschgeboten, die für bestimmte Daten gelten. Aus der Perspektive der anbieterpflichtigen öffentlichen Stellen wird das Ziel einer guten Verwaltungsführung (Good Governance) beeinträchtigt. Performanzprobleme durch überfüllte Systeme der Verwaltung können drohen; vor allem aber entstehen der Verwaltung erhöhte Kosten durch die notwendige längere Speicherung – im Extremfall sogar durch den Weiterbetrieb von Systemen, die aus Verwaltersicht nicht mehr benötigt würden. Schließlich können auch politische Ziele beeinträchtigt werden, soweit diese auf den Ausbau der E-Government-Infrastruktur ausgerichtet sind.

Frage 3: Wodurch entstehen Ihnen denn Zusatzkosten? –
Man könnte doch Ressourcen vom klassischen Archiv umwidmen ...

Die einfache Formel, ein Mehr an digitaler Archivierung könne letztlich ressourcenneutral durch ein Weniger an konventioneller Archivierung kompensiert werden, wird besonders von Gesprächspartnern geäußert, die sich noch nicht tiefer mit dem Wesen der Archivierung an sich beschäftigt haben. Tatsächlich sind ja auch bestimmte Entwicklungen nicht von der Hand zu weisen: Bei konsequenter Nutzung der E-Akte fällt langfristig weniger Papier in den Registraturen an, auch wenn sich die Vision des komplett papierlosen Büros längst als Illusion herausgestellt hat.¹⁰ Durch die erhoffte und prognostizierte Verringerung der Papierproduktion in den Verwaltungen steht zu erwarten, dass langfristig der Bedarf *am Zuwachs* neuer Regalfächen im Archiv sinkt. Grundsätzlich aber gilt das Argument:

Digitale Archivierung erfolgt zusätzlich zur bisherigen Archivierung.

Häufig bleibt bei den zitierten Einwänden unbeachtet, dass die bereits vorhandene Papierüberlieferung im Magazin der Archive sich nicht mehr verringert. Historisches Archivgut ist daher unvermindert zu sichern und nutzbar zu machen. Der archivische Auftrag der dauernden Aufbewahrung impliziert zwingend, dass Archive auf Zuwachs ausgerichtet sind: Altes wird nicht durch Neues ersetzt, sondern durch dieses ergänzt. Selbst wenn die Zuwachsraten einmal zurückgehen sollten, wird der Bedarf an Magazinraum für konventionelles Archivgut trotz Einrichtung eines digitalen Archivs stetig weiter wachsen. Bekanntlich ist längst erwiesen, dass die komplette Ersatzdigitalisierung und/oder Verfilmung und Vorhaltung der Ersatzrepräsentatio-

¹⁰ Schon die erste Version des *DOMEA*-Konzepts sprach 1999 nur noch vom *papierarmen Büro*: Konzept Papierarmes Büro (*DOMEA*-Konzept). Dokumentenmanagement und elektronische Archivierung im IT-gestützten Geschäftsgang (Schriftenreihe des KBSt 45). Bonn, November 1999.

nen deutlich teurer wäre als die Bewahrung der Originale¹¹ – einmal ganz abgesehen von dem kulturpolitischen Sündenfall, den die Vernichtung der Originale bedeuten würde.

Selbst wenn man von einer langfristigen Reduzierung der Papieraktenproduktion in den Verwaltungen ausgeht, bringt der Einsatz von Vorgangsbearbeitungs- und Dokumentenmanagementsystemen für die Archive auch einen deutlichen Mehraufwand mit sich. Hybride Akten, die sowohl konventionelle als auch digitale Aufzeichnungen umfassen, erschweren die Bewertung. Fachverfahrensdaten in Datenbanken fallen zusätzlich zur Aktenüberlieferung an – nur selten werden die Akten vollständig durch die neuen Informationssysteme ersetzt. Bereits heute lässt sich die Tendenz erkennen, dass den Archiven im Zuge der Einführung der E-Akte verstärkt Papierunterlagen angeboten werden. Aufgrund zum Teil längerer Aufbewahrungsfristen in den Behörden ist damit zu rechnen, dass ein Effekt durch zurückgehende Anbietet von Papierunterlagen erst in 20 bis 30 Jahren eintreten wird.

Ingest

Frage 4: Dürfen Sie die Daten überhaupt übernehmen? –

Das müsste erst einmal der Datenschützer genehmigen ...

Bei der Aussonderung digitaler Daten führen Behördenvertreter weitaus häufiger als bei konventionellen Aussonderungen Datenschutzbedenken ins Feld. Hintergrund ist die berechtigte Annahme, dass die Vertraulichkeit personenbezogener Daten grundsätzlich zu schützen ist. Bestimmte personenbezogene Daten unterliegen bei Behörden sogar besonderen Löschvorschriften. Die Unsicherheit gerade bei digitalen Daten mag daraus resultieren, dass die Datenschutzgesetzgebung seit den 1970er Jahren primär als legislative Antwort auf die zunehmende ADV (*Automatisierte Datenverarbeitung*, wie es damals hieß), verstanden wurde. Tatsächlich aber machen die Datenschutzgesetze längst keinen Unterschied mehr zwischen elektronisch oder herkömmlich gespeicherten Daten: So dehnte das Hessische Datenschutzgesetz bereits 1987 den Schutz, der bis dahin nur den in Dateien gespeicherten Daten gegolten hatte, auf alle personenbezogenen Daten aus – und damit auch auf Daten in Papierakten.¹² Das wesentliche Argument der Archive zum Datenschutz muss lauten:

Das Archiv als Datenschutzbehörde sichert die Vertraulichkeit.

¹¹ Gerd *Schneider*: Archivare aufgewacht! Anmerkungen eines Externen zur gegenwärtigen Situation im deutschen Archivwesen. In: *Der Archivar* 57 (2004) S. 37–44, hier S. 40. Vgl. Rechnungshof des Freistaates Sachsen: Jahresbericht 2003. Leipzig 2003. S. 101–106. Internet: <http://www.rechnungshof.sachsen.de/jb2003/jb2003.pdf>.

¹² Hessisches Datenschutzgesetz (HDSG) § 2 (2) in der ab dem 01.01.1987 gültigen Fassung (GVBl. I 1986 S. 309). Vgl. Jochen *Nungesser*: Hessisches Datenschutzgesetz. Kommentar für die Praxis. Mainz 1988. S. V.

Die Archivgesetze selbst sind als bereichsspezifische Datenschutzvorschriften entwickelt worden. Soweit personenbezogene Daten in auszusondernden Behördenunterlagen oder in Archivgut vorkommen, hat das jeweilige Archivgesetz als Spezialgesetz Vorrang vor dem allgemeinen Datenschutzgesetz.¹³ Dabei werden selbst Löschvorschriften, die der Behörde auferlegt sind, in der Regel durch Abgabe an das Archiv erfüllt. Um die bislang aufgetretenen Unsicherheiten zu vermeiden, wird bei jüngeren Gesetzgebungsverfahren bereits für eine eindeutige Klarstellung zum Verhältnis von datenschutzrechtlicher Löschvorschrift einerseits und Anbieterspflicht gegenüber den Archiven andererseits gesorgt. So hat die Bundesregierung in der Begründung zum Entwurf eines Gesetzes zur Fortentwicklung des Meldewesens (MeldFortG) festgehalten, durch den Passus zur Anbieterspflicht werde *klargestellt, dass die sich aus den jeweiligen archivrechtlichen Vorschriften ergebende Pflicht [...], die Meldedaten [...] den zuständigen öffentlichen Archiven zur Übernahme anzubieten, gegenüber dem Lösungsgebot [...] vorrangig ist.*¹⁴

Nach der Übernahme digitaler Daten ins Archiv sichern die Archivgesetze durch ihre Schutzfristen den Persönlichkeitsschutz; sie gewährleisten einen standardisierten Ausgleich zwischen den beiden hier relevanten Grundrechten: zwischen dem Recht auf informationelle Selbstbestimmung, das das Bundesverfassungsgericht 1983 aus dem Recht auf freie Entfaltung der Persönlichkeit (Art. 2 Abs. 1 GG) in Verbindung mit der Menschenwürde (Art. 1 Abs. 1 GG) abgeleitet hat,¹⁵ – und dem Recht auf Wissenschafts- und Forschungsfreiheit (nach Art. 5 Abs. 3 GG).

Frage 5: Was soll an der Datenübernahme schwierig sein? –

Man kann doch einfach alles auf einen Datenträger brennen ...

Wer sich noch nicht eingehend mit der Überführung digitaler Aufzeichnungen in ein Archiv befasst hat, äußert verständlicherweise zunächst Unverständnis über den erheblichen Aufwand, der dabei zu betreiben ist. Oft wird von der Grundannahme ausgegangen, es seien nur Dateien von einem Speicherort an einen anderen zu transferieren. Das aber kann allenfalls eine (Not-) Lösung bei einer einfachen Datenstruktur sein, denn die Datenkopie oder gar die Übergabe von Datenträgern ans Archiv ist besser als ein Datenverlust. Es bleibt aber grundsätzlich erforderlich, folgende Erkenntnis zu vermitteln:

Die Bildung digitaler Archivalien erfordert einen komplexen Prozess.

¹³ Z. B. § 3 (3) Hessisches Datenschutzgesetz (HDSG), zuletzt geändert durch Gesetz vom 20. Mai 2011 (GVBl. I S. 208): *Soweit besondere Rechtsvorschriften über den Datenschutz bei der Verarbeitung personenbezogener Daten vorhanden sind, geben sie den Vorschriften dieses Gesetzes vor.*

¹⁴ Deutscher Bundestag: Drucksache 17/7746 vom 16.11.2011. Gesetzentwurf der Bundesregierung. Entwurf eines Gesetzes zur Fortentwicklung des Meldewesens (MeldFortG). S. 38: Begründung zu § 16 (Anbieter von Daten an Archive). Internet: <http://dipbt.bundestag.de/dip21/btd/17/077/1707746.pdf>.

¹⁵ Bundesverfassungsgericht: Urteil vom 15.12.1983. Az. 1 BvR 209, 269, 362, 420, 440, 484/83.

Authentizität und Integrität der digitalen Aufzeichnungen sind schon beim Übertragungsweg nachzuweisen,¹⁶ u. a. ist ein technischer Nachweis der Identität von Übergabe- und Übernahmepaketen zu führen. Die Behördendaten dürfen erst nach geprüfter korrekter Übermittlung gelöscht werden.

Bei konventionellen Übernahmen werden die Unterlagen zunächst unverändert von der Behörde ins Archiv überführt. Dagegen müssen bei digitalen Aufzeichnungen die authentischen Archivalien aus den Behördendaten erst geformt werden. Herstellung von Archivgut – das ist ein Paradigmenwechsel, der ein Umdenken gegenüber den bisherigen Konzepten erforderlich macht. Bei der Übernahme stehen zunächst aufwendige Verhandlungen mit Behörden über die Gestaltung der Übergabepakete an. Daten müssen aus diversen Speicher- und Bearbeitungssystemen extrahiert werden; die extrahierten Informationen müssen sodann neu formiert und zu langfristig nutzbaren Archivpaketen gepackt werden. Schon bei der Übernahme müssen signifikante Eigenschaften, die künftig zu erhalten sind, definiert werden.

Als Beispiel mag die Übernahme von Daten aus der hessischen Lehrer- und Schülerdatenbank dienen.¹⁷ Der Übernahmeprozess machte in den Jahren 2010 bis 2012 bis zur Abstimmung des Übergabepakets insgesamt 19 externe und interne Termine des Digitalen Archivs Hessen erforderlich. Das lag nicht zuletzt an der Vielzahl der Beteiligten: Verantwortlich für die Daten ist der Großteil der Schulen in Hessen; federführende Fachbehörde für die Fachanwendung ist das Hessische Kultusministerium. Für Softwareprogrammierung und Verfahrensbetrieb waren der hessische IT-Dienstleister sowie eine externe Beratungsfirma einzubeziehen. Schließlich wurde die Übernahme zwischen den drei hessischen Staatsarchiven abgestimmt, die jeweils für einen Teil der betroffenen Schulen das zuständige Archiv sind.

Ein anderes Beispiel liefert die Übernahme von Orthophotos der Vermessungsverwaltung, da hier die Bildung einer großen Anzahl von Archivpaketen erforderlich ist. Pro Abgabeportion, die jeweils eine Charge an entzerrten Luftbildern der kompletten hessischen Landesfläche umfasst, fallen ca. 6.000 TIFF-Bilder (jeweils zu einer Bodenfläche von 2 x 2 km) an. Hinzu kommen ebenso viele 6.000 twf-Dateien mit Geodaten (Koordinaten). Schließlich liegen die Ortsnamen und die Konkordanz zu den Bildnummern in einer TXT-Datei vor. Ein Teil der Metadaten, die auch der Zuordnung der Dateien zueinander dienen, sind in den Dateinamen der insgesamt ca. 12.000 Dateien enthalten. Aus all diesem Quellmaterial muss das Archiv 6.000 Archivalien bilden. Dies ist nur möglich durch eine Paketierung im Massenprozess mit dem Ergebnis, dass sämtliche relevanten Informationen zu einem Orthophoto auch in einem AIP vorliegen. Erst dadurch werden diese digitalen Archivalien einzeln bestell- und nutzbar. Der Nutzer erhält damit beispielsweise ein Orthophoto der Kachel Nr. 1-0286 zur Stadt Idstein mit dem Aufnahmedatum 17.05.2003 – das Resultat eines aufwendigen *Ingest*-Prozesses.

¹⁶ Wege ins Archiv. Ein Leitfadens für die Informationsübernahme in das digitale Langzeitarchiv (nestor-materialien 10). Frankfurt a. M. 2008. Internet: http://files.d-nb.de/nestor/materialien/nestor_mat_10.pdf.

¹⁷ Sigrid Schieber: LUSD archivieren – die Lehrer- und Schülerdatenbank in Hessen. Internet: http://www.landesarchiv-bw.de/sixcms/media.php/120/52525/Workshop_Schieber_LUSD_archivieren.pdf.

Data Management

Frage 6: Wieso ist ein System für digitale Archivalien nötig? –

Das Archiv könnte doch das DMS der Verwaltung nutzen ...

Der Ansatz, das bereits für die Verwaltung angeschaffte Dokumentenmanagementsystem (DMS) zugleich für das *Data Management* des digitalen Archivguts einzusetzen, basiert auf der Überlegung, die ohnehin beschafften Softwarelizenzen effizient einzusetzen. Grundsätzlich befürworten es ja auch die Archive, wenn nicht jeder Behördenzweig zur Verwaltung seiner Unterlagen ein eigenes System einsetzt. Richtig ist auch in der Tat, dass die Funktionalitäten eines DMS und einer OAIS-konformen¹⁸ *Data-Management-Software* Schnittmengen aufweisen: Auch das DMS verwaltet Akten, Dokumente und Dateien. Die Metadaten der Objekte werden hinterlegt; die Zugriffsrechte darauf werden gepflegt. Gleichwohl gibt es ein wesentliches, auch wirtschaftlich begründetes Argument gegen die Mitnutzung des allgemeinen DMS für das *Data Management* des digitalen Archivguts:

Der Spezialbedarf macht zu viele DMS-Anpassungen erforderlich.

Die genauere Betrachtung zeigt, dass die Anforderungen an die jeweiligen Systeme letztlich erheblich voneinander abweichen. Die Struktur digitaler Archivalien ist weitaus vielfältiger als der DMS-Dreiklang Akte – Vorgang – Dokument. Auch heterogene Datenstrukturen von Datenbankinhalten, Webauftritten oder möglicherweise sogar Fileablagen müssen abbildbar sein. Wesentlich für das *Data Management* nach OAIS ist zudem, dass die Authentizität der Unterlagen über die Lebensdauer des Systems hinaus zu sichern ist – diese Anforderung ist beim DMS, das der befristeten Verwaltung von Unterlagen dient, nachrangig. Das System für das Digitale Archiv muss zudem Schnittstellen zu den archivfachlichen Prozessen ermöglichen: etwa zur Erschließung im Kontext mit nichtdigitalem Archivgut oder Bereitstellung digitaler Archivalien im virtuellen Lesesaal, also nach außerhalb des Systems.

Wollte man all das mit dem vorhandenen DMS der Verwaltung umsetzen, wäre hierfür eine aufwendige Spezialversion des DMS erforderlich. Die Mitnutzung des DMS wäre letztlich unwirtschaftlich, vor allem wenn die Folgeaufwände mitberücksichtigt werden. Schließlich müssten in jeder neuen DMS-Version die spezifischen Archivanforderungen nachgezogen werden. Ein Wechsel des DMS durch die Verwaltung wäre für das Archiv mit einem hohen Migrationsaufwand verbunden. Letztlich ist es daher weitaus günstiger, für die digitale Archivierung ein eigenes, maßgeschneidertes System für das *Data Management* einzusetzen.

¹⁸ Reference Model for an Open Archival Information System (OAIS), wie Anm. 3.

Frage 7: Ist denn dieser Aufwand bei der Datenhaltung nötig? –
 Man sollte doch nicht mehr machen als in der *analogen* Welt ...

Gerade in der Anfangszeit waren immer wieder Stimmen zu vernehmen, die die Notwendigkeit einzelner Maßnahmen beim Aufbau digitaler Archive in Frage stellten und dabei mit Analogien aus dem nichtdigitalen Archiv argumentierten. Niemand wird bestreiten, dass bestimmte wesentliche Prinzipien, die für nichtdigitales Archivgut gelten, auch auf das digitale Archivgut übertragen werden können. Gleichwohl lässt das Argument einen gravierenden Unterschied gegenüber dem nichtdigitalen Zeitalter außer Acht:

Die digitale Welt birgt viel mehr Missbrauchsmöglichkeiten.

Die Authentizität und die Integrität digitaler Archivalien müssen ebenso nachvollziehbar bleiben wie beispielsweise die der Papierunterlagen im Archiv, denn die Daten haben elementare Bedeutung für die Rechtssicherung. Anders als etwa bei Papierunterlagen mit einer Originalunterschrift des Urhebers ist die Zuordnung von Urheber zu Daten im digitalen Bereich ausschließlich durch die Metadaten möglich. Die Unverfälschtheit von Daten ist im System des digitalen Archivs unter anderem durch Zugriffsprotokolle nachzuweisen.

Auch die Vertraulichkeit, der Datenschutz, ist bei digitalen Daten einer größeren Gefährdung ausgesetzt als bei nichtdigitalen. Große Datenmengen können digital sehr viel schneller kopiert werden als in Papierform. Digitale Daten können über das Internet massenhaft verfügbar gemacht werden, während missbräuchlich verbreitete Unterlagen in herkömmlicher Form nur einem begrenzten Adressatenkreis zur Kenntnis gelangen können. Nicht zuletzt kann die Verknüpfung digitaler personenbezogener Daten, wenn sie unrechtmäßig erfolgt, einen weitaus größeren Schaden anrichten als die Auswertung konventioneller Archivalien.

Ohne Frage: Verfälschungen und Missbrauch können auch bei nichtdigitalen Archivalien vorkommen – da sie bei digitalen Archivalien einfacher und folgenschwerer sein können, sind besondere Vorkehrungen unabdingbar.

Preservation Planning

Frage 8: Wozu kümmern Sie sich noch um Datenmigration? – Es gibt doch PDF/A...

Die Erhaltungsplanung für digitales Archivgut, das *Preservation Planning* nach *OAIS*, wird beim Aufbau digitaler Archive anfangs häufig zurückgestellt. Grund dafür kann Ressourcenknappheit sein, aber auch die Annahme, die digitale Bestandserhaltung könne später noch ohne Beeinträchtigungen angeschlossen werden. Nicht selten besteht die Hoffnung, mit der Bestimmung eines Langzeitformats wie PDF/A sei den Planungsanforderungen schon Genüge getan. Tatsächlich ist die Festlegung auf offene und verbreitete Formate essentiell bei der digitalen Archivierung. Und die Reduzierung der Anzahl von Archivformaten erleichtert fraglos spätere Migrationen. Bei den Überlegungen wird aber meist folgende Erkenntnis außer Acht gelassen:

Nicht Dateien, sondern Objekte mit Eigenschaften sind zu erhalten.

Eine elektronische Akte – oder ein Vorgang daraus – stellt für das Archiv ein *Informationsobjekt*¹⁹ (eine *Intellectual Entity*²⁰) dar: Das Archiv muss aufgrund seiner Wahrnehmung²¹ entscheiden, welche signifikanten Eigenschaften der Akten für künftige Nutzergruppen erhalten werden müssen – und welche entbehrlich sind. Archivalien werden intellektuell definiert aufgrund ihrer Eigenschaften. So gehören zu den Eigenschaften der Akte nicht allein Inhalt und Form der im PDF/A-Format überlieferten Dokumente, sondern ebenso die Angaben zu den Bearbeitern oder zur Reihenfolge der Dokumente in der Akte. All diese Eigenschaften werden nicht im PDF gespeichert, sondern derzeit meist im Format XML, z. B. als *XDOMEA*-Datei.

Digitale Bestandserhaltung erfordert also eine langfristige Erhaltungsplanung,²² sie beginnt mit der Bildung von Archivalien bei der Übernahme oder im besten Fall sogar davor. Die Wahl des Archivformats bestimmt die künftigen Nutzungsmöglichkeiten: So wird mit dem Archivformat PDF/A die Weiterbearbeitbarkeit einer archivierten Akte erschwert; das ist durchaus gewollt, denn das Archiv identifiziert die Bearbeitbarkeit nicht als signifikante Eigenschaft.

Bei Migrationen (aber auch im Fall einer Emulation) müssen signifikante Eigenschaften der intellektuell definierten Archivalien berücksichtigt werden. An diesen Eigenschaften, nicht an den Dateiformaten, müssen Migrationen ansetzen. Dabei muss das Archiv immer die Anforderungen, die Nutzungsziele der künftigen Nutzergruppen berücksichtigen.

Archival Storage

Frage 9: Wieso sind viele Mittel für die Speicherung nötig? –
Speicherplatz kostet doch heutzutage nichts mehr ...

Dass die Speicherkosten nach wie vor einen erheblichen Kostenfaktor für digitale Archive darstellen, sorgt oft für Unverständnis. Sinkende Kosten für Speichermedien beim Discounter suggerieren, dass die Ausgaben für die Einheit *Archival Storage* schon bald als vernachlässigbar

¹⁹ Zum Begriff *Informationsobjekt* siehe: Information und Dokumentation – Leitfaden zur Informationsübernahme in digitale Langzeitarchive (DIN 31645:2011-11 (D)). November 2011. S. 6–10. Vgl. Wege ins Archiv, wie Anm. 16, S. 8 f., S. 26.

²⁰ In diesem Sinn verwandt in dem international erarbeiteten *PREMIS-Konzept*: PREMIS Data Dictionary for Preservation Metadata version 2.1. January 2011. S. 5–9, S. 22 f. Internet: <http://www.loc.gov/standards/premis/v2/premis-2-1.pdf>.

²¹ Grundlegend hierfür ist das *Performance Model* des Australischen Nationalarchivs: Helen Heslop, Simon Davis und Andrew Wilson: An Approach to the Preservation of Digital Records. December 2002. S. 8. Internet: http://www.naa.gov.au/Images/An-approach-Green-Paper_tcm16-47161.pdf.

²² Zu grundlegenden Konzepten siehe: Leitfaden zur digitalen Bestandserhaltung. Vorgehensmodell und Umsetzung. Version 2.0. Verfasst und herausgegeben von der nestor-Arbeitsgruppe Digitale Bestandserhaltung (nestor-materialien 15). Frankfurt a. M. 2012. Internet: http://files.d-nb.de/nestor/materialien/nestor_mat_15_2.pdf.

Größe angesehen werden könnten. In der Tat ist zu konzedieren, dass die Kosten für Speichermedien über die letzten Jahre hinweg meist gesunken sind. Allerdings zeigten die stark gestiegenen Preise für Festplatten nach der Flutkatastrophe 2011 in Thailand²³ – dem zweitgrößten Exporteur der Trägermedien –, wie fragil dieser Trend sein kann.

Wichtiger als der Hinweis auf derartige Unsicherheiten aber ist die Erkenntnis:

Bloße Speicherung gefährdet die dauernde Erhaltung von Archivgut.

Es wäre höchst fahrlässig, digitales Archivgut schlicht auf DVDs zu brennen oder auf billige externe Festplatten aus dem Supermarkt zu speichern, und diese dann im Magazin zu lagern. Um den gesetzlichen Auftrag der dauernden Aufbewahrung zu erfüllen, ist ein komplexes Speicherkonzept erforderlich. Ein solches Konzept beinhaltet u. a. die verteilte Speicherung an verschiedenen Orten als Katastrophenvorsorge. Spiegelungen und Backups verringern die Gefahr eines physischen Datenverlusts, der entweder durch Hardwareschäden oder durch Anwenderfehler drohen kann. Unumgänglich ist zudem eine laufende Hardwaremigration, also ein Ersatz der Speichermedien, um einen Datenverlust aufgrund veralteter Datenträger zu verhindern. Aus Gründen der Datensicherheit verbietet sich bei unikalem digitalen Archivgut eine Speicherung in der *Cloud*, also bei einem günstigen privaten Anbieter.

Wesentlich ist also die Erkenntnis, dass die Kosten der Datenspeicherung den Preis für die Erstbeschaffung der Speichermedien um ein Vielfaches übersteigen. Speichersysteme für große Datenmengen sind teuer, vor allem aber fallen neben den Beschaffungskosten für die Hardware die laufenden jährlichen Aufwendungen für Administration und Betrieb an.²⁴ Um einen adäquaten Zugriff auf die gespeicherten Archivalien, etwa bei der archivarisches Erschließung, zu ermöglichen, ist zudem eine hohe Leitungsbandbreite für den Transfer großer Datenmengen erforderlich.

Access

Frage 10: Wozu sollten wir überhaupt digital archivieren? –

Man findet doch alles bei Google ...

Der Hinweis, man finde doch alles bei Google, macht in all seiner Schlichtheit doch auch deutlich, welche Konkurrenz den klassischen Archiven in der digitalen Welt durch die Vielzahl verschiedener Informationslieferanten erwachsen ist. Denn in der Tat erscheint das Internet mit

²³ Produktionsstopp. Thailand-Flut macht Festplatten teurer. In: Spiegel online (27.10.2011). Internet: <http://www.spiegel.de/netzwelt/gadgets/produktionsstopp-thailand-flut-macht-festplatten-teurer-a-794341.html>.

²⁴ Für die Speicherung einer 2-stelligen Terabyte menge ist bei einem verwaltungsinternen IT-Dienstleister mit jährlichen Kosten im 5-stelligen Eurobereich zu rechnen (Hardwareabschreibung, Serverbetrieb ohne Datenadministration, ohne Redundanz, ohne Leitungskosten).

seiner globalen und verteilten Serverinfrastruktur aktuell als ein gigantischer Wissensspeicher. Was bei dem Argument allerdings nicht bedacht wird, ist die zeitliche Komponente:

Langzeitverfügbarkeit muss unabhängig vom aktuellen Bedarf sein.

Die aktuelle Verfügbarkeit von Informationen mag zu Sorglosigkeit verleiten. Es ist aber vor allem das auffindbar, was aktuell ist. Die dauerhafte Erhaltung einer Information im Internet ist dagegen nicht gewährleistet. *Historische* Internetseiten werden nur dann noch gefunden, wenn der Anbieter ein Interesse daran hat. Zudem ist die Authentizität von Informationen auf längere Sicht kaum nachvollziehbar.

Archivalien und Behördendokumente sind in der Regel nur dann im Internet zu finden, wenn Archive diese dort im Rahmen ihrer Aufgabenwahrnehmung veröffentlicht haben. Nur unabhängige öffentliche Archive, die die fachlichen Standards berücksichtigen, können darüber hinaus die Sicherung und Erhaltung digitaler Aufzeichnungen gewährleisten. Sie schaffen so die Grundlagen für die Langzeitverfügbarkeit von rechtserheblichen Unterlagen sowie von Quellen für die Forschung und sorgen damit für die Bewahrung des kulturellen Erbes.

Fazit

Es gibt keine *dummen Fragen* zur digitalen Archivierung. Die Materie ist nach wie vor für viele ungewohnt, da die bisherigen archivwissenschaftlichen Konzepte fundamental in Frage gestellt werden.²⁵ Archivarinnen und Archivare haben die Aufgabe, die Gräben zwischen der herkömmlichen und der digitalen Welt zu überbrücken, indem sie nachvollziehbare Argumentationen zu den offenen Fragen anbieten.

²⁵ Christian Keitel: Archivwissenschaft zwischen Marginalisierung und Neubeginn. In: *Archivar* 64 (2011), S. 33–37, hier S. 33.

Datenübernahme aus dem Polizei-Informationssystem (POLIS)

Von BERNHARD RIEDER

Das Polizei-Informationssystem (POLIS) ist ein digitales System zur Erzeugung und Verwaltung von Geschäften (Fallakten) der Zürcher Polizeikorps. Den zuständigen Endarchiven wurden die Fallakten gemäß § 8 des kantonalen Archivgesetzes zur Übernahme angeboten. Zusammen mit den Polizeikorps und der Firma, die POLIS entwickelt hatte, spezifizierten die zuständigen Endarchive eine Archivschnittstelle für die gesetzeskonforme Übernahme eines bewerteten Teils der POLIS-Daten.

Ursprungssystem POLIS

Das Polizei-Informationssystem (POLIS) ist ein digitales System zur Erzeugung und Verwaltung von Polizeigeschäften (Fallakten). Es wurde 1999 in Betrieb genommen und löste die papierbasierte Aktenführung ab. Die bis zur Einführung von POLIS entstandenen Akten wurden und werden den zuständigen Archiven auf Papier angeboten.

POLIS wird gemeinsam von den drei großen Polizeikorps im Kanton Zürich, der Kantonspolizei (KAPO) und den beiden Stadtpolizeien Zürich und Winterthur verwendet und betrieben. Zudem verwenden es die kleineren Kommunalpolizeien der zürcherischen Gemeinden.

Die gesetzliche Grundlage für den Betrieb und die Benutzung des Systems bildet die kantonale POLIS-Verordnung (Zürcher Gesetzessammlung LS 551.103).

POLIS wurde von der Software-Firma UNISYS entwickelt und unterstützt die Abwicklung der administrativen Aufgaben der Polizeibeamten. Die für die Archive wichtigsten Funktionalitäten sind die Aufzeichnung von eingehenden Ereignis-Meldungen in einem Journal sowie der pro Geschäftsvorfall der Polizei durch den beteiligten Beamten erstellte Rapport (Bericht) und die dazugehörigen Beilagen. Daneben werden in POLIS zu Fahndungs- und Verknüpfungszwecken verschiedene Stammdatensammlungen (z. B. Personendaten von Beteiligten) bewirtschaftet. Wie wir weiter unten bei der Bewertungsdiskussion sehen werden, besitzen diese Stammdaten keinen archivischen Wert.

Man kann POLIS als Records Management System im Sinne der Definition bezeichnen, denn die Kernserie der darin verwalteten Akten sind die nach dem Dossierprinzip gebildeten und geführten Polizeigeschäfte, mit dem jeweiligen Rapport (Bericht) als Haupt- oder Kerndokument.

Allerdings liegt dem System kein ausdifferenziertes eigentliches Ordnungssystem (Aktenplan) zugrunde. Die Hierarchie, in der die einzelnen Dossiers eingebettet sind, ist relativ flach.

Jedes an POLIS beteiligte Polizeikorps führt seine eigenen Geschäfte im System. Das heißt für die zuständigen Archive, dass sie nur diejenigen Geschäftsdossiers übernehmen, bei denen *ibid* Polizeikorps die Federführung hatte.

In der POLIS-Verordnung ist festgelegt, dass die Geschäfte und Personendaten (Stammdaten) nach Ablauf klar definierter Fristen aus datenschutzrechtlichen Gründen gelöscht werden müssen. Je nach Schwere des Delikts oder Größe des Ereignisses handelt es sich um Löschfristen von 2 bis 30 Jahren. Die Löschfrist beginnt mit dem Datum des Ereignisses und die Löschung erfolgt auf den Tag genau nach Ablauf der Frist. Da seit 1999 in POLIS Unterlagen geführt und verwaltet werden, werden theoretisch seit dem Jahr 2001 jeden Tag Geschäfte unwiederbringlich gelöscht. Es war deshalb dringend nötig, die Polizeikorps auf ihre Aktenangebotspflicht aufmerksam zu machen und eine Datenexport- und Datenübernahmелösung ins Auge zu fassen. Da die Überlieferung der Kantonspolizei im Staatsarchiv des Kantons Zürich (nachfolgend Staatsarchiv genannt) grundsätzlich sehr reichhaltig ist, hat man davon abgesehen, als Übergangslösung, bis die Exportschnittstelle in Betrieb genommen wird, sämtliche zu löschenden POLIS-Geschäfte in einem temporären Repository zwischenzulagern und zu sichern. Diese (kleine) Überlieferungslücke von 4-5 Jahrgängen kann verschmerzt werden. Dafür braucht es keine unter Umständen aufwändige Zwischen- und Übergangslösung. Außerdem konnte aufgrund dieser Entscheidung der Druck auf die Polizeikorps erhöht werden, indem wir ihnen drastisch das Bild vom täglichen, unwiederbringlichen Datenverlust und damit täglichen Verstoß gegen das Archivgesetz vor Augen führten.

Projektziel und Organisation

Ziel des Projekts war es, die Polizei-Geschäftsdossiers und die übrigen Daten in POLIS, deren Löschfrist abgelaufen ist, zu bewerten und den archivwürdigen Teil zu exportieren.

Selbstverständlich sollten die Bewertungsentscheide auch prospektiven Charakter haben und die spezifizierte Schnittstelle für die periodischen späteren Exporte ins System implementiert werden. Neben den drei Polizeikorps (KAPO, Stadtpolizeien Zürich und Winterthur) waren die jeweiligen (End-) Archive, das Staatsarchiv und die beiden Stadtarchive Zürich und Winterthur, am Projekt beteiligt.

Für das Gelingen des Projekts waren aus Sicht der drei Archive insbesondere zwei Dinge entscheidend:

1. Für alle drei Polizeikorps bzw. Archive gibt es nur eine Exportschnittstelle. Die *Verteilung* der Daten und Geschäfte erfolgt dann nach dem Federführungsprinzip.
2. Die drei Archive sprechen sich untereinander ab und treten als Einheit auf. Insbesondere gelten für alle drei Archive dieselben Bewertungsentscheide.

Es wurde eine schlanke Arbeitsgruppe gebildet, die für die Erarbeitung der Exportspezifikation verantwortlich war.

Neben einem externen Projektleiter, der bereits in mehreren Organisationsprojekten für die KAPO tätig gewesen war, war die KAPO als Vertreterin der anderen Polizeikorps und das Staatsarchiv in Person des Schreibenden als *Mandatsträger* der beteiligten Archive in der Arbeitsgruppe vertreten. Die POLIS-Herstellerfirma UNISYS stellte durch 2 Mitarbeiter die fachliche Unterstützung sicher. Martin Kaiser von der Schweizer Koordinationsstelle für die dauerhafte Archivierung elektronischer Unterlagen (KOST) konnte bei bestimmten Fachfragen hinzugezogen und konsultiert werden.

Spezifikation Export

Zeitplan

Die Datenübernahme aus POLIS wurde bereits im Jahr 2005 angedacht. Durch die Vorgänger des Schreibenden wurden insbesondere Überlegungen zur Bewertung angestellt. Die effektive Durchführung des Export-Projekts wurde dann allerdings wieder zurückgestellt. Das Staatsarchiv musste 2008 den Druck auf die Kantonspolizei erhöhen und auf die Archivgesetzgebung verweisen, damit es zu einer Wiederaufnahme des Projekts kam.

Diese eher harzige Anlaufphase zeigt, dass die zuständigen Archive einen langen Atem benötigen und über gewisse Eskalationsmöglichkeiten verfügen sollten. Bei allem Verständnis für die zahlreichen und dringlichen Tagesgeschäfte eines öffentlichen Organs muss das zuständige Archiv definieren, wann es keinen Aufschub mehr dulden will.

Der Kick-Off des wiederaufgenommenen Projekts erfolgte Ende Januar 2010. Nun wurden auch die übrigen POLIS-Anwender (Stadtpolizeien Winterthur und Zürich) und ihre Archive in die Projektorganisation eingebunden. Die mit der Übernahmespezifikation beauftragte Arbeitsgruppe tagte von Februar bis September 2010 in neun Sitzungen. Der Projektausschuss genehmigte die erarbeitete Spezifikation im November 2010.

Erst im April des Folgejahrs wurde die Spezifikation durch das oberste IT-Strategieorgan der Kantonspolizei, das nur wenige Male im Jahr tagt, behandelt. Allerdings passierte die Spezifikation diese wichtigste Hürde ohne nennenswerte Probleme.

Als letzte Instanz musste der für die Kantonspolizei zuständige Direktionsvorsteher¹ eine Verfügung zur Umsetzung des Projekts erlassen. Der Erlass dieser Verfügung verzögerte sich infolge von Neuwahlen und Abwahl des bisherigen Direktionsvorstehers bis August 2011.

Anfang November 2011 nahm das Projektteam die Umsetzung der Spezifikation an die Hand. Die Struktur der Submission Information Packages (SIP) konnte im Sommer 2012 anhand von Test-SIP überprüft werden. Die erste Echtdatenlieferung ist per Anfang 2013 zu erwarten.

Bewertung

Pro Jahr werden ungefähr 200'000 Geschäfte in POLIS eröffnet. Bei ungefähr 120'000 Geschäften besitzt die KAPO die Federführung. Gelöscht werden pro Tag ungefähr 500 Geschäfte (davon 300 der KAPO).

Die drei beteiligten Archive trafen sich bereits vor der eigentlichen Spezifikationsphase, um gemeinsam einen Bewertungsentscheid zu fällen. Man einigte sich relativ rasch auf eine archivisch vernünftige und einfache Aktenauswahl. Ein wichtiges Bewertungsziel war, den hohen Komplexitätsgrad der POLIS-Datenstruktur zu reduzieren und nur wirklich *Archivwürdiges* in die Archive zu übernehmen. Bei einem komplexen System wie POLIS zeigte sich, dass es sich bewährt, die

¹ Die kantonale Verwaltung des Kantons Zürich ist in sieben Direktionen (Ministerien) unterteilt. Der Vorsteher einer Direktion ist eines der sieben Exekutivmitglieder der Kantonsregierung (Regierungsrat). Die Kantonspolizei gehört als Amtsstelle zur Sicherheitsdirektion.

notwendige Denkarbeit bei der Bewertung zu leisten und nur diejenigen Unterlagen zu übernehmen, die das Kerngeschäft der Polizei dokumentieren. So konnte u. U. auf die Übernahme von in POLIS verwalteten Stammdaten wie Listen von beteiligten Personen mehrheitlich verzichtet werden. Auch bei der Auswahl der inhaltlichen Metadaten zu den einzelnen Geschäftsdossiers konnte stark reduziert werden. Metadaten, die für die Bearbeitung des Geschäftsvorfalles durch den federführenden Sachbearbeiter von Bedeutung waren, haben für die dauernde Überlieferung keine Relevanz mehr. Zu einem Geschäftsdossier müssen die Metadaten Titel, Ereignisdatum, Ereignisort und Abschlussdatum übernommen werden. Das Kürzel des federführenden Polizeibeamten oder die Nummer der involvierten Dienststelle können als Metadaten ohne Weiteres gelöscht werden.

Die Reduktion der Anzahl Metadaten (Felder) pro Geschäft vereinfacht die Datenstruktur erheblich.

Die Auswahl der Geschäftsdossiers (Falldossiers) erfolgt in einem zweistufigen, maschinengestützten Auswahlverfahren:

Mit einer *inhaltlichen Auswahl* werden sämtliche Geschäftsdossiers mit einer Lauffrist oder Aufbewahrungsfrist (Verjährungsfrist) grösser als 28 Jahre übernommen. Dieses Auswahlkriterium stellt sicher, dass sämtliche *schweren* Fälle wie Kapitalverbrechen und Großereignisse (Großbrände, Explosionen, Bahnunglücke, etc.) dauernd überliefert werden. Die Auswahl mittels Aufbewahrungsfrist ermöglicht zudem, dass die Auswahl in der Exportschnittstelle programmiert wird und automatisch ablaufen kann.

Mit einer *systematischen Auswahl* werden zusätzlich jeweils zwei Geschäfte pro Tag mittels Zufallszahlen² ausgewählt.

Die Auswahl der Geschäfte wird jeweils nach Ablauf der Aufbewahrungsfrist am Tag, an dem sie gelöscht werden müssten, durchgeführt.

SIP-Bildung

Die Grundlage für die Spezifikation der Exportschnittstelle bildete die Spezifikation SIP des Schweizerischen Bundesarchivs (BAR).³ Die am Projekt beteiligten Vertreter der Herstellerfirma UNISYS verlangten, dass die Ergebnisse der Spezifikationsphase in einem einzigen Dokument zusammengefasst werden. In der BAR-Spezifikation sind bewusst viele Anforderungen als *Kann-Anforderungen* definiert. Was bei einer Modell-Spezifikation mit einem relativ hohen Abstraktionsgrad sinnvoll ist, gilt nicht für ein konkretes Umsetzungsprojekt. Es hat sich sehr bewährt, dass wir auf Empfehlung des Herstellers die BAR-Spezifikation auf unsere Bedürfnisse reduziert haben. Nur die relevanten Punkte wurden aufgenommen, alles Übrige herausgestrichen. Es resultierte ein Papier, das den ausführenden Informatikern theoretisch ermöglicht, die Schnittstelle zu programmieren, ohne Rückfragen stellen zu müssen.

² Einfacher Zufallsgenerator für Pseudozufallszahlen.

³ Schweizerisches Bundesarchiv: Spezifikation Submission Information Package (SIP). Ablieferungsobjekt für digitale Unterlagen an das Schweizerische Bundesarchiv. Version 1.0. Bern 2009. Aktuell auf der Webseite des Schweizerischen Bundesarchivs ist die Version 4.0 (2012) greifbar (www.bar.admin.ch/themen). Alle Links wurden am 9.10.2012 überprüft. Zitiert: BAR-Spezifikation.

Die so entstandene gekürzte und vereinfachte SIP-Spezifikation für POLIS fand Eingang als (wichtigstes) Kapitel in die Gesamt-Exportspezifikation.

SIP-Struktur

Das Ablieferungsobjekt aus POLIS ist ein Informationspaket, das auf dem Modell des OAIS-Standards (Open Archival Information System) basiert. Dieses sieht vor, dass alle zusammengehörigen Informationen in einem sogenannten Submission Information Package (SIP) zusammengefasst werden.

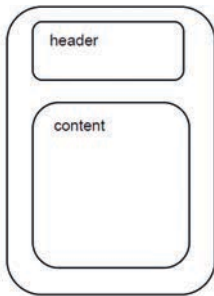


Abb. 1: Schematische Struktur eines SIP.

Basierend auf dem OAIS Modell und der BAR-Spezifikation enthält das SIP zwei separate Teile, die immer gemeinsam abgeliefert werden müssen. Der erste Teil ist der *header*. Dieser Teil enthält alle Metadaten (beschreibende Information betreffend das Paket und die Primärdaten) in der Datei *metadata.xml* im XML-Format. Ebenfalls befinden sich die Schemadateien in diesem Ordner (*header\xsd*).

Der zweite Teil ist der *content*. Dieser Teil enthält alle Primärdaten sowie in einem Ordner *dokumentationAblieferndeStelle* weiterführende beschreibende Informationen über die Ablieferung (Vgl. Abb. 2).

Ein SIP enthält die ausgewählten Geschäfte eines Löschjahres in einer sogenannten Jahressammlung.

```
header\
  metadata.xml
  xsd\
    Schemadateien
content\
  dokumentationAblieferndeStelle\
    Extrakt_fuer_Staatsarchiv_Spezifikation_V03.pdf
    Codelisten.pdf
    etc.pdf
  Jahressammlung (Primärdateien in Geschäftsdossiers\)
```

Abb. 2: Grundstruktur eines POLIS Informationspaketes.

Die beschreibenden Unterlagen im Ordner *dokumentation.Abliefernde.Stelle* dienen einem späteren Benutzer, den Inhalt und die Herkunft der eigentlichen Primärdaten des SIP besser zu verstehen und zu benutzen. Ein SIP muss aus sich selber heraus verständlich sein. Damit dies sichergestellt ist, benötigt ein Benutzer Metainformationen, die über die im *header* abgelegten *maschinenlesbaren* technischen und inhaltlichen Metadaten hinausgehen. Dazu gehören eine etwas ausführlichere Beschreibung des Ursprungssystems, Umstände des Datenexports, Hinweise zur Benutzung der Daten etc.

Metadaten

Metadaten aus dem Ursprungssystem POLIS, die über den gesamten Export- und Archivierungsprozess hinweg erhalten bleiben sollen, sind in der separaten XML-Datei namens *metadata.xml* im *header* des SIP enthalten.

In einem sogenannten Data Dictionary wurden die einzelnen Metadatenelemente nach Entitäten aufgeteilt und spezifiziert.

Beschreibung der Entitäten

Entität	Beschreibung
Paket	Die Entität <i>Paket</i> beschreibt die Verpackung der Ablieferung in einen technischen Behälter (Ordner). Das Paket trägt je nach Phase im Paket-Lifecycle einen anderen Namen (SIP/AIP/DIP). Für die Ablieferung von digitalem Archivgut handelt es sich immer um ein SIP.
Ablieferung	Die Entität <i>Ablieferung</i> beinhaltet alle Angaben zur abliefernden Stelle und zur Ablieferung ans Archiv.
Provenienz	Die Entität <i>Provenienz</i> repräsentiert den Aktenbildner.
Ordnungssystem (Aktenplan)	Die Entität <i>Ordnungssystem</i> beschreibt die innere Ordnung des Archivgutes.
Ordnungssystemposition	Die Entität <i>Ordnungssystemposition</i> beschreibt die Eigenschaften der einzelnen <i>Ordnungssystempositionen</i> im Ordnungssystem.
Dossier	Die Entität <i>Dossier</i> beschreibt die Eigenschaften des zusammengehörenden Archivgutes aus einer Geschäftsgruppe ⁴ oder die Datensätze eines Tagesjournals.

⁴ Die POLIS-Geschäftsgruppe entspricht beim Export aus POLIS einem POLIS-Geschäftsdossier. Innerhalb einer Geschäftsgruppe existieren keine weiteren (Unter-)Ordner. Gehören mehrere Einzelgeschäfte zum Selben Fall, werden sie zu einer Geschäftsgruppe zusammengefasst. Für den Export wird die Ordner-Struktur innerhalb der Geschäftsgruppe aufgehoben.

Dokument	Die Entität <i>Dokument</i> repräsentiert die kleinste Ebene, auf der Archivgut inhaltlich dargestellt wird. Diese Entität enthält die Metadaten, die zum Verständnis der eigentlichen Unterlagen resp. ihrem Inhalt notwendig sind.
Inhaltsverzeichnis	Die Entität <i>Inhaltsverzeichnis</i> repräsentiert das im metadata.xml enthaltene Inhaltsverzeichnis, das alle Ordner und Dateien umfasst.
Ordner	Die Entität <i>Ordner</i> repräsentiert die Ordner im Paket.
Datei	Die Entität <i>Datei</i> repräsentiert die Primärdaten im Dateisystem und beinhaltet auch Angaben zur Identifikation und Beschreibung der einzelnen Dateien.

Die BAR-Spezifikation sieht nur einen eingeschränkten Katalog von Metadatenelementen vor. Da im Staatsarchiv die Verzeichnung im Archivinformationssystem (AIS) auf Dossierstufe erfolgt und bei personenbezogenen Fallakten Angaben zu den beteiligten Personen und Örtlichkeiten erfasst werden, extrahiert man zusätzliche Metadaten aus dem Ursprungssystem und nimmt sie in die Metadaten-Datei metadata.xml auf.

Dazu werden diese zusätzlichen POLIS-Metadaten, die im Ursprungssystem als XML-Datei vorliegen, in einem Wrapper, bzw. einer Schema Extension, wie wir das auch von METS kennen, in die Metadaten-Datei eingefügt:

```

- <dossier id="DS-0">
  <titel>GNR10110142</titel>
  + <entstehungszeitraum>
- <bemerkung>
- <xmlData mdtype="OTHER">
  - <Geschaeftsgruppe xmlns="www.staatsarchiv.zh.ch/z/el/polis" Datum="2010-11-22" Archiv="je"
    Archivsammlung="KPZ">
      <Titel>Titel des Hauptgeschäfts</Titel>
      <Ereignisdatumende>15.01.1999 12.00</Ereignisdatumende>
      <Ereignisdatumbeginn>15.01.1999 12.00</Ereignisdatumbeginn>
      - <sachgebiete>
        <Sachgebiet>Alle Sachgebiete der Dokumente</Sachgebiet>
      </sachgebiete>
      - <Fall GNR="10110142" Korps="Kantonspolizei Zürich">
        <Titel>Titel dieses Falles</Titel>
        - <Ereignis>
          <Ereignisdatumbeginn>15.01.1999 12.00</Ereignisdatumbeginn>
          <Ereignisdatumende>15.01.1999 12.00</Ereignisdatumende>
          - <Ereignisort>
            <Ort>Zürich</Ort>
            <Stadtkreis>Altstadt</Stadtkreis>
            <Strasse-Hnr>Augustinergasse</Strasse-Hnr>
            <Ereignisortsbezeichnung>Bar zum Kalifen</Ereignisortsbezeichnung>
            <Nation>Schweiz</Nation>
          </Ereignisort>
        </Ereignis>
        <Personen />
        - <Dokumente>
          - <Dokument id="92300000104012">
            <Titel>EBD Zürich Altstadt zN. Hans-Heinrich Müller</Titel>
            <Erstelldatum>01.03.2001</Erstelldatum>
            <Korps>Kantonspolizei Zürich</Korps>
            <Organisation>Forensisch Naturwissenschaftlicher Dienst</Organisation>
            <Rapporttyp>Einbruchdiebstahl</Rapporttyp>
          </Dokument>
          + <Dokument id="92300000104216">
          </Dokumente>
        </Fall>
      </Geschaeftsgruppe>
    </xmlData>
  </bemerkung>
</aktENZEICHEN>10110142</aktENZEICHEN>
+ <eroeffnungsdatum>
+ <abschlussdatum>
<schutzfristenkategorie>Personenbezogene Dossiers</schutzfristenkategorie>
+ <dokument id="DOK-8">
+ <dokument id="DOK-9">

```

XML wrapper

Abb. 3: Zusätzliche strukturierte Metadaten in einem Wrapper in metadata.xml.⁵

⁵ Die in diesem Beispiel enthaltenen Angaben sind frei erfundene Testdaten.

Datei-Formate

Ein wichtiger Teil der Spezifikationsarbeit war die Festlegung der Zielformate. Im Ursprungssystem POLIS liegen die einzelnen Dokumentengruppen in allen möglichen Dateiformaten vor. Der Katalog reicht von XML über PDF und TIFF bis zu Microsoft-Office-Formaten.

Es wurde entschieden, als Übernahmeformate nur die beiden Formate PDF/A und XML (für tabellarische Geschäftskontrolldaten) zuzulassen.

Die eigentlichen Primärdaten werden mittels einer kommerziellen Konvertierungssoftware in PDF/A umgewandelt. Nach intensiver Diskussion wurde beschlossen, dass Dateien, die sich nicht zu PDF/A konvertieren lassen, gar nicht erst übernommen, sondern gelöscht werden. Dies werden aber nur wenige Dateien sein. Das Kerndokument eines Polizeigeschäfts ist der im Ursprungssystem POLIS im HTML-Format vorliegende Rapport. Diesen in PDF/A umzuwandeln ist kein Problem. Es wird dann Inhalt umfassender Stichproben sein, dies zu überprüfen.

Die Formatvielfalt soll anhand der nachfolgenden Tabelle dokumentiert werden:

Nr.	Format	Regeln	verifiziert mit PDF-Tools
1.	HTML	Umwandlung in PDF/A, gilt für Registerdaten und Rapporte (POLIS, FATS) sowie Journale.	ja
2.	TIFF (ZIP)	Umwandlung in PDF/A, alle Seiten eines Dokuments in einer Datei.	ja
3.	JPG	Umwandlung in PDF/A	ja
4.	PDF	Umwandlung in PDF/A, enthält auch einseitige A3 und A4-Beilagen	ja
5.	txt	Umwandlung in PDF/A	ja
6.	doc, docx	Umwandlung in PDF/A	ja
7.	xls, xlsx, ppt, pps, pptx	Umwandlung in PDF/A. Die Umwandlung erfolgt nach Vorgaben des Herstellers. Umwandlungstuning ist nicht Gegenstand des Projektes.	ja

Eine wichtige Erkenntnis war, dass bei der Konvertierung ein gewisses Maß an *Mut zur Lücke* bewiesen werden sollte. Sofern die Wahrscheinlichkeit groß ist, dass nur bei wenigen und für das Verständnis eines Geschäftsdossiers nicht zwingend relevanten Dateien eine Konvertierung unmöglich ist, sollte auf die Übernahme solcher Dokumente verzichtet werden.

Selbstredend bringt die Umwandlung in PDF/A auch Abstriche bei der Datenqualität mit sich (Fotografien als PDF/A). Die langfristige Dokumentation des polizeilichen Kerngeschäfts hängt jedoch nicht davon ab, ob eine hochauflösende, reproduzierbare Fotografie dabei ist.

Kostenteilung

Beim POLIS-Projekt handelte es sich um ein verhältnismässig kostspieliges Unterfangen. Die meisten öffentlichen, ablieferungspflichtigen Organe stehen einer Ablieferung ins zuständige Archiv nicht ablehnend gegenüber, sofern dabei weder ein personeller noch ein finanzieller Aufwand entsteht. Rechtskonforme Aktenführung, die die Archivierung eines bewerteten Teils am Ende der dritten Phase des Lebenszyklus von Verwaltungsunterlagen einschließt, ist jedoch nicht umsonst zu haben. Gerade die Kostenfrage darf eine ablieferungspflichtige Stelle nicht dazu verleiten, sich ihrer gesetzlichen Pflichten zu entziehen.

Die Kostenteilung war beim vorliegenden Projekt ein wichtiges Thema. Während der Spezifikationsphase wurde bewusst nicht über die Kosten gesprochen. Ziel war es, eine Exportschnittstelle *nach allen Regeln der Kunst* zu spezifizieren, um anschließend allenfalls in gewissen Bereichen Abstriche zu machen.

Nach der Spezifikation berechnete die Herstellerfirma UNISYS die Kosten für die Umsetzung.

Die Kosten für die Realisierung der Exportschnittstelle beliefen sich auf eine mittlere sechsstellige Summe. Davon betrug die SIP-Bildung nach unseren Vorgaben eine mittlere fünfstellige.

Selbstredend stellte sich im Projektteam die Frage, ob man diese hohen Realisierungskosten gegenüber den Entscheidungsinstanzen vertreten konnte. Die beteiligten Archive nahmen eine unmissverständliche Haltung ein. Sowohl die beiden Stadtarchive wie auch das Staatsarchiv sind der Meinung, dass es Sache der ablieferungspflichtigen Stellen ist, die Daten so aufzubereiten, damit eine Weiterverarbeitung und Archivierung möglich ist. Die Aufbereitung hat nach einem bestimmten (SIP-) Standard zu erfolgen.

Es gelang uns, die beiden wichtigsten Entscheidungsinstanzen, das IT-Strategieorgan der KAPO und den zuständigen Direktionsvorsteher⁶ zu überzeugen. Die Kostenteilungsproblematik wurde zu Gunsten der Archive entschieden. Für die öffentlichen Archive im Kanton Zürich hat dieser Entscheid richtungsweisenden Charakter. Es besteht mittlerweile innerhalb der Konferenz der Generalsekretäre⁷ der Kantonalen Verwaltung ein Konsens darüber, dass die Kosten für den Datenexport aus dem Ursprungssystem und für die Übergabe an das Archiv von den abliefernden Stellen zu tragen sind.

In der nachfolgenden Abbildung ist dargestellt, wie wir uns zukünftig Verantwortung, Aufgaben- und Kostenteilung während des gesamten Life Cycles von Unterlagen vorstellen:

⁶ Zuständiges Exekutivmitglied in der Kantonsregierung.

⁷ Stabchefs der einzelnen Direktionen (Ministerien).

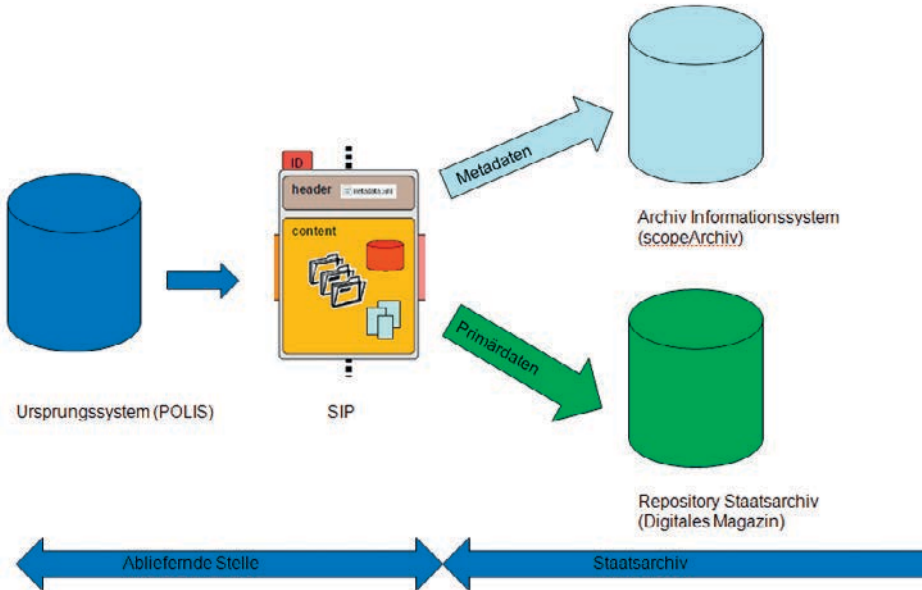


Abb. 4: Verantwortung und Kostenteilung.

Die Verantwortung für die Unterlagen geht mit der Ablieferung vollständig ans zuständige Archiv über. Die Kosten für den Export aus dem Ursprungssystem werden durch die abliefernde Stelle getragen. Dafür übernimmt das zuständige Archiv sämtliche Folgekosten für die dauernde Archivierung. Archivare wissen, dass die archivische Bearbeitung und Bewirtschaftung von digitalem Archivgut auch nach der Übernahme und Integration ins Archiv mit hohen, immer wiederkehrenden Kosten (z. B. periodische Konvertierungen) verbunden ist. So gesehen ist es fair, wenn der Aktenproduzent für die Exportkosten aufkommt.

Das öffentliche Organ erhält dafür vom Archiv kostenlos beratende Unterstützung bei der Beschaffung von Systemen zur digitalen Aktenführung und Fachwissen bei der Definition von Exportschnittstellen. Weiter stellen die Archive den öffentlichen Organen Software-Werkzeuge für den Datenexport und die Datenüberprüfung zur Verfügung.⁸

⁸ Vgl. dazu die Tools SIP-Val und csv2siard der KOST (www.kost-ceco.ch) oder den Package Handler des BAR (www.bar.admin.ch).

Ausblick

Im Verlaufe des Sommers 2012 überprüften die Archive die ersten Test-SIP's mit dem SIP-Validator der KOST. Geplant ist, dass Anfang Januar 2013 ein erstes echtes POLIS-SIP, das aus den Daten der Monate November und Dezember 2012 besteht, abgeliefert wird. Ab Anfang 2014 ist dann die periodische und vollständige Übernahme von Jahressammlungen geplant.

Da die Speicherinfrastruktur (Repository) im Staatsarchiv noch nicht fertig eingerichtet ist, werden die ersten SIP's vorderhand noch in einer simplen Netzlaufwerkinfrastruktur gespeichert.

Schlussbemerkungen

Der erfreuliche Verlauf des Projekts zeigt uns, dass die Art der Abwicklung auch für andere Projekte dieses Typs maßgebend sein sollte.

Wir sind der Meinung, dass man bei gemeinsam genutzten Systemen einheitliche Exportlösungen anstreben muss. Wir verstehen darunter Systeme, die von öffentlichen Organen der verschiedenen Staatsebenen (Gemeinden, Kanton und/oder Bund) gemeinsam betrieben und genutzt werden. Die öffentlichen Organe der verschiedenen Staatsebenen sind unterschiedlichen Archiven anbietepflichtig. Die involvierten Archive, sei es wie im vorliegenden Fall zwei große Kommunalarchive und das Staatsarchiv (Kantonsarchiv) oder auch verschiedene Staatsarchive und das Bundesarchiv bei einer durch den Bund betriebenen Applikation, haben die Chance, als Einheit aufzutreten und sich auf eine einheitliche Schnittstelle zu einigen. Die einheitliche Lösung vereinfacht den Export und senkt die Kosten. Wir sind der Überzeugung, je einheitlicher und damit einfacher die Lösung, desto eher gewinnt man die ablieferungspflichtigen Organe für den Bau einer Exportschnittstelle. Treten die Archive hingegen als heterogene Anspruchsgruppe auf, wird man wenig bewirken können.

Es empfiehlt sich bei der Bewertung, insbesondere bei der inhaltlichen und systematischen Auswahl, auf die technischen Gegebenheiten im Ursprungssystem Rücksicht zu nehmen. Der Einfachheit halber muss auf eine spezielle inhaltliche Auswahl nach bestimmten Kriterien verzichtet werden.

Wenn es sich vertreten lässt, sollte man sich bei der Konvertierung nicht dazu verleiten lassen, sicherheitshalber die nicht konvertierbaren Dateien in archivuntauglichen Dateiformaten zu übernehmen. Sofern sichergestellt ist, dass die Kerndokumente eines Geschäftsdossiers konvertiert und exportiert werden, kann man ohne Weiteres auf die (wenigen) nicht konvertierbaren Nebendokumente verzichten.

Die im vorliegenden Projekt vereinbarte Kostenteilung zwischen aktenproduzierender Stelle und dem Archiv ist nach unserem Dafürhalten zukunftsweisend. Die Archive haben dafür zu sorgen, dass diese Ansicht auch von den Entscheidungsträgern geteilt wird.

Wenn die aktenproduzierenden Stellen hier in die Pflicht genommen werden, müssen auch die Archive ihren Beitrag leisten. Die Archive stellen das fachliche Wissen bei Schnittstellenprojekten zur Verfügung und entwickeln Softwareprogramme, die die Bildung von Ablieferungspaketen unterstützen.

Weiter sollen die Archive dafür sorgen, dass sie bei Systembeschaffungen und Systemablösungen frühzeitig einbezogen werden, damit die Archivschnittstellenproblematik bereits zu Beginn thematisiert und Standards umgesetzt werden können.

Es werden allerdings noch einige Zeit aufwändige Exporte aus *alten* Systemen, die noch keine standardisierten Schnittstellen aufweisen, durchgeführt werden müssen. Die Archive müssen sich darauf einstellen, dass für solche Systeme jedes Mal eine eigene Exportschnittstelle spezifiziert werden muss.

Zwischen der KAPO und dem Staatsarchiv besteht traditionell ein ausgesprochenes Vertrauensverhältnis, das der produktiven und effizienten Zusammenarbeit im Projekt sehr zuträglich war. Der KAPO gilt der spezielle Dank des Schreibenden.

Entwicklung einer Aussonderungsschnittstelle für das DMS DOMEA

Von SIGRID SCHIEBER

Der Artikel beschreibt die Entwicklung eines Aussonderungsprozesses an einem konkreten Beispiel (DOMEA von OpenText) sowie die Nutzung des Standards xDOMEA für den Metadaten-austausch. Anhand des Abstimmungsprozesses zwischen verschiedenen DOMEA-Anwendern werden wesentliche Fragen, die sich bei einer Aussonderung aus einem Dokumentenmanagementsystem ergeben, aufgezeigt. Das Fazit weist schließlich auf einige grundlegende Aspekte bei der Einführung eines DMS hin, die eine spätere Aussonderung aus demselben erschweren oder erleichtern können.

Das Dokumentenmanagementsystem DOMEA der Firma OpenText wird in der hessischen Landesverwaltung seit 2005 eingesetzt (seit einigen Jahren unter dem Namen HeDok). Nach einer Einführungs- und Erprobungsphase wechseln seit 2010 mehr und mehr oberste Landesbehörden zu einer führenden eAkte in DOMEA, womit auch die Frage der Aussonderung immer dringlicher wird.

DOMEA ist jedoch nicht nur in Hessen, sondern auch in weiteren Landesverwaltungen, in Teilen der Bundesverwaltung sowie in kommunalen und anderen Verwaltungen im Einsatz. Diese Anwender sind in der sogenannten *Anwenderkonferenz* DOMEA (AK DOMEA) zusammengeschlossen. Da es sich bei der Aussonderung um eine übergreifende Frage handelt, bildete sich bereits 2008 als Unterarbeitsgruppe zur AK DOMEA ein *Qualitätszirkel Aussonderung* (QZ Aussonderung), der die gemeinsamen Anforderungen an ein Aussonderungsverfahren zu definieren versuchte. Der eigentliche Umsetzungsprozess mit Beteiligung der Herstellerfirma startete im Frühjahr 2010. Am QZ Aussonderung beteiligt sind Vertreter/innen von Archiven, Verwaltungen und IT der DOMEA-Anwender Hessen, Mecklenburg-Vorpommern, Nordrhein-Westfalen, Rheinland-Pfalz und Saarland sowie des Bistums München-Freising.

Ziel dieses gemeinsamen Projekts ist es, ein einheitliches Aussonderungsverfahren aus DOMEA zu definieren und so einerseits einen standardisierten Prozess zu schaffen, andererseits aber natürlich auch Kosten zu sparen. Das gemeinsame Anforderungskonzept wurde im Frühjahr 2011 fertiggestellt. Nach anfänglicher Skepsis beteiligte sich OpenText sehr engagiert und zielstrebig an diesem Projekt, so dass seitens der Firma im Juni 2011 ein erstes Umsetzungskonzept vorgelegt werden konnte. Einzig offener Punkt war zu diesem Zeitpunkt noch die Spezifikation der xDOMEA-Metadatendateien, die in Absprache und Zusammenarbeit zwischen den Anwendern und OpenText bis Mai 2012 abgeschlossen werden konnte. Nach Klärung der Kostenfrage kann das Aussonderungsverfahren im folgenden Jahr umgesetzt und in Betrieb genommen werden.

Vorentscheidungen

Zu Beginn des Prozesses stellten sich einige grundlegende Fragen, über die Einigkeit hergestellt werden musste.

Die beteiligten Archive sprachen sich schon früh für ein zweistufiges Aussonderungsverfahren aus. Die Bewertungspraxis bei analogen Akten und die Erfahrungen, die man bereits mit dem DOMEA-Einsatz gemacht hatte, ließen eine Bewertung ausschließlich anhand von Bewertungslisten als Notlösung erscheinen. Die erfassten Metadaten sind selten so aussagekräftig, dass sich die Archivwürdigkeit ganz ohne Akteneinsicht feststellen lässt.

Ebenso schnell einigte man sich darauf, dass eine Vorbewertung am Aktenplan möglich sein sollte. Diese Vorbewertung von Aktenplanpositionen dient als Hilfsmittel für die spätere Bewertung der aussonderungsreifen Akten und Vorgänge. Die Erfahrungen der Aussonderungsläufe können dabei dazu verwendet werden, diese Vorbewertung zu korrigieren und immer stärker zu präzisieren, so dass das Verfahren mit der Zeit immer effizienter wird.

Ebenfalls von Anfang an war klar, dass xDOMEA als Standard für den Export der Metadaten von Akten, Vorgängen und Dokumenten verwendet werden sollte.¹ Erst bei der Fertigstellung des Anforderungskonzepts und in Rücksprache mit der Firma OpenText wurde hingegen die Entscheidung getroffen, bei den im Aussonderungsprozess vorgesehenen Nachrichten auf die entsprechenden xDOMEA-Nachrichten zu verzichten und stattdessen einfache Excel-Tabellen zu benutzen. Da die Archivsysteme der beteiligten Archive derzeit nicht in der Lage sind, auf der Basis von xDOMEA-Nachrichten mit anderen Anwendungen zu kommunizieren, und diese Nachrichten (Anbieterliste, Bewertungsliste, Übergabeliste, Löschprotokoll) nur der Information und der Dokumentation des Aussonderungsvorfalles dienen, erschien dies ausreichend.

Während der Erstellung des Anforderungskonzepts Aussonderung wurde außerdem schnell deutlich, dass in DOMEA ein Pflichtmetadatum zur Unterscheidung zwischen Papier / Elektronisch / Hybrid auf der Ebene von Akte und Vorgang eingerichtet werden muss. Nur so kann die Anbieterung von Papierunterlagen getrennt von der Anbieterung elektronischer Unterlagen behandelt werden bzw. ist im Fall von hybriden Unterlagen sichergestellt, dass auch der Papierteil mit angeboten wird.

Weniger Einigkeit herrschte darüber, ob auf der Akten- oder auf der Vorgangsebene aussondert werden sollte, d.h. ob die beteiligten Archive als kleinste Einheiten einzelne Vorgänge oder Aktenschnitte (also alle archivwürdig bewerteten Vorgänge einer Akte in einem Aussonderungslauf) übernehmen. Um hier flexibel zu bleiben und unterschiedlichen Bedürfnissen gerecht zu werden, wurde im Prozess eine Auswahlmöglichkeit *Aussonderung aktenbezogen/vorgangsbezogen* vorgesehen.

Auch die Frage, ob die Dokumente in einem Vorgang nur als Archivformat (PDF/A) oder zusätzlich als Originaldateiformat übergeben werden sollen, konnte noch nicht abschließend beantwortet werden. Die angestrebte Lösung erlaubt es, hier je Instanz eine andere Einstellung zu wählen. Auf diese Weise könnte sich z.B. das Landesarchiv Nordrhein-Westfalen künftig

¹ Auf die genauere Spezifikation von xDOMEA wird noch gesondert eingegangen.

entschließen, Dokumente nur als Archivformat zu übernehmen, während das Digitale Archiv Hessen Originalformat und Archivformat übernimmt.

Die Wandlung der Originaldateiformate in langzeitfähige Archivformate wie PDF/A ist kein Teil des Aussonderungsverfahrens. Viele Anwender haben hierzu bereits eigene Lösungen entwickelt. Empfohlen wird, dass die Umwandlung spätestens bis zur Anbietetung und Aussonderung erfolgt ist. Sinnvoll ist jedoch eine frühzeitige Formatwandlung zu Beginn der Aufbewahrungsphase, da sichergestellt werden muss, dass die Dokumente für den Zeitraum der behördlichen Aufbewahrung lesbar bleiben.

Schließlich tauchten bei der Erstellung des Anforderungskonzepts und bei der Ausarbeitung des Umsetzungskonzepts zahlreiche Detailfragen auf, die ich hier nur kurz anreißen und im Folgenden beantworten will:

- Sind parallele Aussonderungsläufe in der gleichen Instanz nötig und müssen damit verschiedene Aussonderungsläufe technisch voneinander unterschieden werden?
- Wie sollen Versionen des gleichen Dokuments behandelt werden? Alle übernehmen? Oder nur die Endversion übernehmen?
- Wie und durch wen ist eine Fristüberwachung des Aussonderungsprozesses möglich?
- Wie sind die Lese- und Schreibberechtigungen im Aussonderungsprozess geregelt – ab wann verliert die Behörde den Zugriff auf die im Aussonderungsprozess befindlichen Objekte?
- Wie kann die Behörde vor einer Aussonderung sicherstellen, dass auch wirklich alle Akten/Vorgänge aussonderungsfähig sind (Vorabkontrolle und ggf. Korrektur der Aufbewahrungsfrist, wenn sich z.B. rechtliche Grundlagen geändert haben)?
- Wie wird der Vorgabe des Datenschutzes nach der vollständigen Löschung bestimmter Vorgänge bei gleichzeitigem Nachweis, dass gelöscht wurde, nachgekommen?
- Soll die Aussonderung im DMS selbst nachgewiesen werden, indem Rumpfdatensätze der ausgesonderten Objekte erhalten und mit dem Wert *archiviert* oder *vernichtet* versehen werden?

Das Aussonderungsverfahren – Phase 1: Vorbewertung

Die Vorbewertung wird von den zuständigen Archivaren/innen direkt im DMS der Behörde am Aktenplan vorgenommen. Eine entsprechende Berechtigung stellt sicher, dass nur auf das Metadatum *Archivierungsmerkmal* schreibend zugegriffen werden kann. Die Alternative hierzu wäre gewesen, die Vorbewertung an einem Export des Aktenplans vorzunehmen und durch die Registratur der Behörde in das DMS übertragen zu lassen, was den Arbeitsaufwand v.a. der Behörde erhöht hätte.

Das Archivierungsmerkmal besitzt fünf verschiedene Ausprägungen:

A	Archivieren	Diese Akten/Vorgänge sind voraussichtlich zu archivieren.
B	Bewerten	Diese Akten/Vorgänge müssen einzeln bewertet werden.
V	Vernichten	Diese Akten/Vorgänge sind voraussichtlich zu vernichten.
M	Massenschriftgut	Diese Akten/Vorgänge müssen bewertet werden – da es sich um Massenschriftgut handelt voraussichtlich nach einer definierten Samplebildung. Auf die Integration eines <i>Filtergenerators</i> , mit dem sich Auswahlkriterien je Aktenzeichen definieren und hinterlegen lassen, wurde aufgrund des technischen Aufwands verzichtet. Stattdessen muss wie bisher im Archiv dokumentiert werden, für welche Bereiche des Massenschriftguts welche Auswahlkriterien genutzt werden. (früher „F“ für Filtern)
L	Löschen	Diese Akten/Vorgänge können nach Ablauf der Aufbewahrungsfrist ohne Anbietung gelöscht werden. „L“ dient dazu, eine vom Archiv erteilte pauschale Löscherlaubnis (= Verzicht auf Anbietung) umzusetzen.
U	Unbelegt	Defaultwert für neue Aktenplaneinträge. Akten/Vorgänge mit dem Archivierungsmerkmal „U“ geben dem/der bewertenden Archivar/in einen Hinweis darauf, dass neue Aktenzeichen angelegt wurden, die noch nicht vorbewertet sind.

Die am Aktenplan bei der Vorbewertung gesetzten Archivierungsmerkmale vererben sich auf neu unter dem jeweiligen Aktenzeichen angelegte Akten und Vorgänge und sind während der Bearbeitungs- und der Aufbewahrungsphase schreibgeschützt. Mit der Vorbewertung wird noch keine Entscheidung über die spätere tatsächliche Bewertung getroffen, die Merkmale dienen nur zur Orientierung. Auch Akten oder Vorgänge, die das Archivierungsmerkmal „A“ oder „V“ tragen, können im Bewertungsprozess nochmals eingesehen und ggf. die Bewertung geändert werden. Einzige Ausnahme sind die auf „L“ gesetzten Aktenplanpositionen, da das Archiv mit diesem Merkmal auf eine Anbietung verzichtet. Die betreffenden Akten und Vorgänge können pauschal gelöscht werden.

Auf eine zunächst geplante rückwirkende Vererbung einer Vorbewertung auf alle bereits unter einem Aktenzeichen angelegten Akten und Vorgänge wurde aus zwei Gründen verzichtet. Zum einen hätte dies technisch nur schwer vorauszusehende Massenprozesse zur Folge gehabt. Zum anderen würden auf diese Art und Weise ältere und fachlich begründete Vorbewertungen überschrieben. Ein bestimmtes Aktenzeichen kann durchaus vor zehn Jahren mit Inhalten befüllt worden sein, die eine Vorbewertung mit „B“ oder „V“ rechtfertigten, anschließend aber so sehr an Bedeutung gewonnen haben, dass inzwischen eine Vorbewertung mit „A“ sinnvoll ist.

Das Aussonderungsverfahren – Phase 2: Bewertung

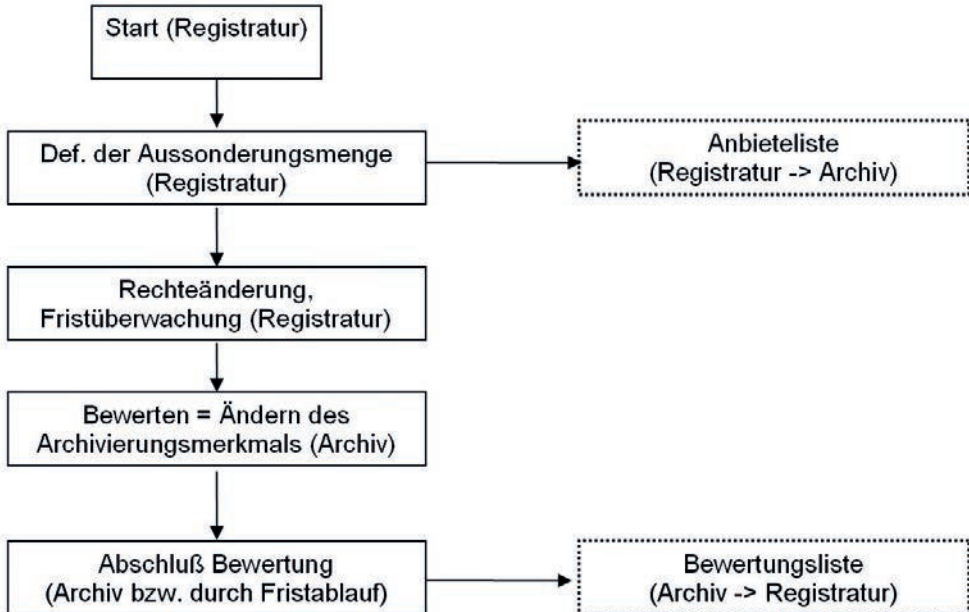


Abb. 1: Ablauf Phase 2.

Der gesamte Prozess der Bewertung und Aussonderung wird von den beiden beteiligten Parteien – Registratur und Archiv – gesteuert. Die Dokumentation erfolgt mit Hilfe von Listen, die aus den Metadaten der in der Aussonderung befindlichen Akten und Vorgänge generiert und zwischen den Beteiligten per E-Mail ausgetauscht werden. Die Behörde legt im DMS eine Akte *Aussonderung* und zu jedem Aussonderungslauf einen Vorgang an, in dem die verschiedenen Listen bis hin zum Löschprotokoll abgelegt werden.

Gestartet werden muss der Prozess der Aussonderung von der Registratur. Dem Archiv werden nicht automatisch alle Akten und Vorgänge, deren Aufbewahrungsfrist abgelaufen ist, angeboten. Dies erlaubt es der Registratur, vor dem Start eines Aussonderungslaufs zu überprüfen, ob die Aufbewahrungsfristen korrekt sind, um diese ggf. zu korrigieren.

Beim Start eines Aussonderungslaufs definiert die Registratur die Menge, die ausgesondert werden soll. Hier sind Einschränkungen möglich, z.B. kann ein Aussonderungslauf nur für eine Organisationseinheit innerhalb einer DOMEA-Instanz erfolgen. Das kann eine bestimmte Abteilung eines Ministeriums sein, oder eine nachgeordnete Dienststelle, die ihre Akten in der gleichen DOMEA-Instanz führt wie das übergeordnete Ministerium. In Rücksprache mit dem Archiv können hier noch weitere grundlegende Definitionen für den Aussonderungslauf getroffen werden, z.B. die Entscheidung über die Paketgröße (vorgangsbezogen oder aktenbezogen) und

die Entscheidung darüber, ob unterschiedliche Versionen² von Dokumenten übergeben werden sollen. Außerdem hinterlegt die Registratur die E-Mail-Adressen der beteiligten Personen für die Versendung automatisch generierter Nachrichten. Jeder Aussonderungslauf erhält eine eindeutige Nummer und alle zugehörigen Objekte werden gekennzeichnet, so dass mehrere parallele Aussonderungsläufe innerhalb einer einzelnen Instanz verwaltet werden können.

Mit dem Start des Aussonderungslaufs werden die Akten und Vorgänge, deren Aufbewahrungsfrist abgelaufen ist und auf die die von der Registratur gesetzten Kriterien (z.B. zuständige Organisationseinheit: Abteilung Z) zutreffen, aus dem Zugriff der Behörde genommen. Die Behörde verliert alle Schreib- und Leserechte,³ der/die zuständige Archivar/in erhält das Leserecht auf Metadaten und Dokumente sowie das Schreibrecht auf das Archivierungsmerkmal, um die Bewertung vornehmen zu können. Es wird eine (konfigurierbare) Excel-Tabelle der angebotenen Akten/Vorgänge erzeugt und an das Archiv geschickt. Dadurch wird das Archiv über die Anbieter und deren Umfang informiert. Sofort nach dem Start des Aussonderungslaufs kann das Archiv mit der Bewertung beginnen.

Mit dem Start des Aussonderungslaufs beginnt außerdem eine Fristüberwachung. Hier können z.B. gesetzlich vorgegebene Bewertungsfristen (in Hessen ein Jahr) hinterlegt werden.⁴

Für die eigentliche Bewertung meldet sich der/die Archivar/in am System an und erhält eine Übersicht über die verschiedenen Aussonderungsläufe, für die er/sie zuständig ist. Nach Auswahl des gewünschten Aussonderungslaufs wird eine Übersicht über die Metadaten der angebotenen Akten und Vorgänge in einer Listensicht geliefert (Aktenzeichen, Betreff, Archivierungsmerkmal). In dieser Listensicht kann nach Belieben gefiltert und sortiert werden. Bei Bedarf ist die Einsicht in den jeweiligen Vorgang oder die jeweilige Akte (Metadaten und Dokumente) möglich. Die Änderung des Archivierungsmerkmals kann einzeln bei Akte oder Vorgang oder durch Auswahl mehrerer Objekte in der Übersichtsliste in einem Massenprozess erfolgen.

Der Bewertungsprozess kann durch das Archiv nach erfolgter Bewertung beendet werden – oder er endet mit dem Ablauf der gesetzten Bewertungsfrist automatisch.

² DOMEA kann unterschiedliche Versionen von Dokumenten verwalten. Diese können von dem/der Bearbeiter/in eines Dokuments absichtlich erzeugt werden (wenn z.B. ein Entwurf in der Akte erhalten bleiben soll) bzw. sie werden bei Bearbeiterwechsel vom System automatisch erzeugt, um Änderungen, die eine weitere Person vorgenommen hat, nachvollziehbar zu halten.

³ Solange sich die digitalen Akten und Vorgänge in der Phase der Aufbewahrung befinden, hat die Behörde darauf einen Lesezugriff (wie auch bei analogem Schriftgut in der Altregistratur). Bei einzelnen Anwendern – z.B. Hessen – ist es zudem möglich, abgeschlossenen Akten und Vorgänge wieder in Bearbeitung zu nehmen. Ist der Aussonderungslauf einmal gestartet, erlischt diese Möglichkeit. Der Lesezugriff wäre aus Sicht der Archive unschädlich gewesen, wurde aber auf Wunsch der Verwaltung ebenfalls entzogen. Damit wird der in manchen Fällen gegebenen datenschutzrechtlichen Forderung nach sofortiger *Löschung* mit Ablauf der Aufbewahrungsfrist Genüge getan. Die betreffenden Informationen sind dann zwar noch nicht physisch gelöscht, befinden sich jedoch nicht mehr im Zugriff der Behörde.

⁴ Die Fristen können von Seiten der Behörde verlängert werden, nicht jedoch von Seiten des Archivs, da kein Rechtsanspruch auf eine Verlängerung besteht.

Das Aussonderungsverfahren – Phase 3: Datenexport und Löschen

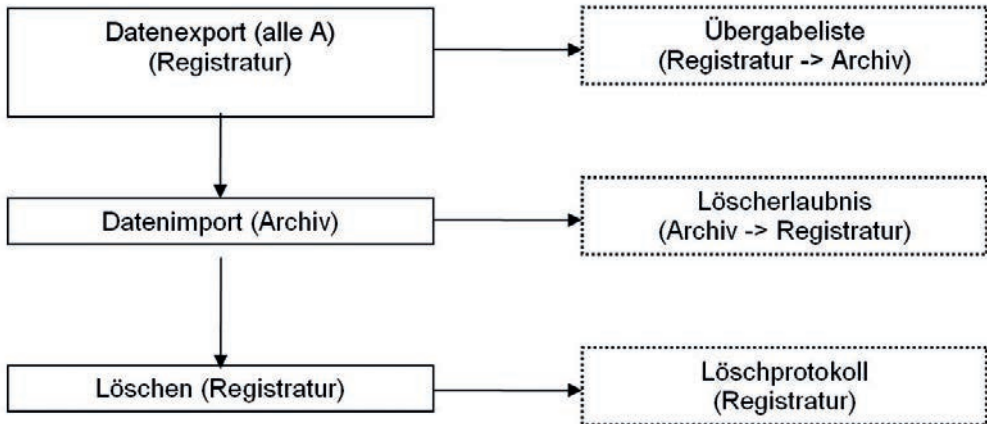


Abb. 2: Ablauf Phase 3.

Nach Abschluss der Bewertung stößt die Registratur den Datenexport an. Dabei werden die Dokumente eines archivwürdigen Vorgangs bzw. einer Akte mit den zugehörigen Metadaten (als xDOMEA-Datei) ausgegeben. Die eigentliche Datenübermittlung geschieht außerhalb des definierten Aussonderungsverfahrens und wird zwischen den Behörden und dem jeweils zuständigen Archiv vereinbart. Zu Kontrollzwecken wird außerdem automatisch eine Liste der übergebenen Akten und Vorgänge erzeugt und dem Archiv übermittelt.

Im Archiv wird das Übergabepaket auf Vollständigkeit geprüft und ins jeweilige Archivsystem aufgenommen. Im Anschluss erteilt das Archiv der Behörde per Mail eine Löschfreigabe für die in diesem Aussonderungslauf angebotenen Unterlagen.

Der eigentliche Löschvorgang wird wiederum von der Registratur angestoßen. Dabei werden alle Dokumente sowie sämtliche Metadaten aus dem System gelöscht und der Prozess in einem Löschprotokoll dokumentiert. Es bleiben also keine *Rumpfmetadaten* zurück, die das DMS über kurz oder lang überfrachten würden. Aus Gründen des Datenschutzes sind die Löschprotokolle konfigurierbar, d.h. sensible Informationen wie z.B. Betreffe (*Disziplinarverfahren Heinz Müller*) können vor der Ablage gelöscht werden. In solchen Fällen wird die Löschung ggf. nur noch durch das Geschäftszeichen nachgewiesen.

Die Nutzung von xDOMEA

Für die Aussonderung der Metadaten zu Akten, Vorgängen und Dokumenten sollte ein definierter Standard genutzt werden – xDOMEA. Verwendet wird die xDOMEA-Nachricht 503 *Aussonderung.Aussonderung*. Dass der Standard vielgestaltig genug ist, um alle nötigen Metadaten abzubilden, war schnell klar. Schwieriger war es, den Überblick über die komplexe Struktur des Standards zu gewinnen und zu definieren, welche Metadaten aus DOMEA exportiert und auf welche xDOMEA-Elemente abgebildet werden sollen. Durch einen Workshop mit Nicol Feske (Sächsisches Staatsministerium der Justiz und für Europa), der an der Entwicklung von xDOMEA beteiligt war, gewannen die Mitglieder des QZ Aussonderung einen tieferen Einblick in den Metadatenstandard und seine Verwendungsmöglichkeiten. Eine große Hilfe war auch das frei verfügbare Werkzeug *xDOMEA-Suite* (mit xDOMEA-Designer und xDOMEA-Viewer), das eine komfortable Möglichkeit zum Entwerfen von xDOMEA-Beispieldateien bietet.⁵

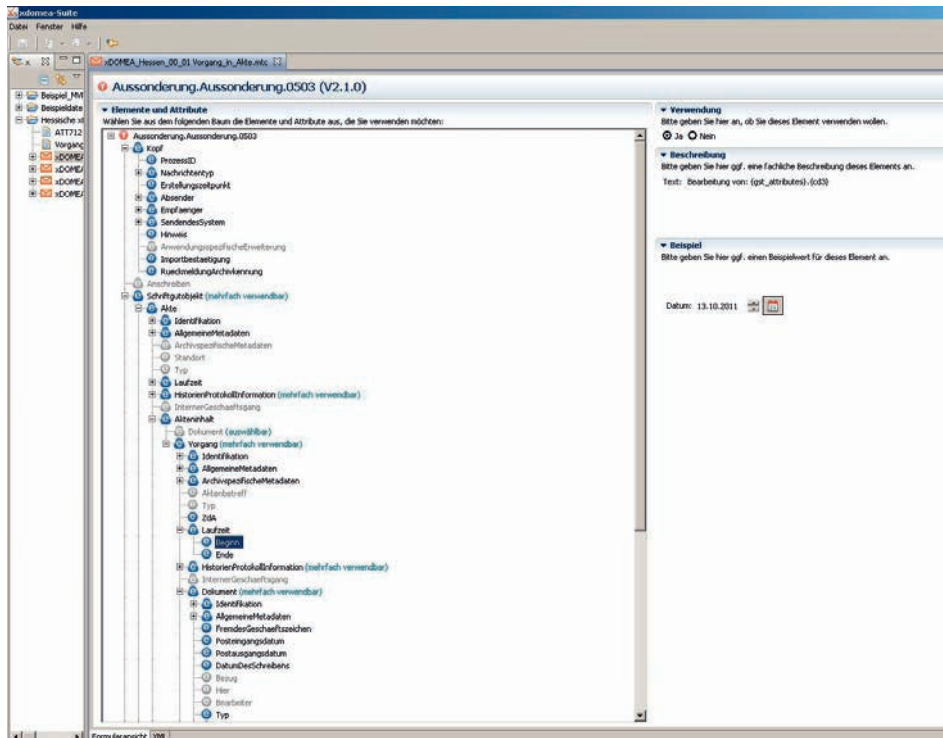


Abb. 3: xDOMEA-Suite.

⁵ Download der xDOMEA-Suite (Designer und Viewer) inkl. Handbuch unter: <https://joinup.ec.europa.eu/software/xdv/release/160> (Der Link wurde am 17.10.2012 überprüft).

Vorentscheidungen:

1. Ein wesentliches Ziel war es, einen einheitlichen Metadatenexport für alle Anwender zu realisieren. Da es DOMEA dem Anwender erlaubt, Datenbankfelder für Metadaten unterschiedlich zu nutzen,⁶ war eine konfigurierbare Lösung nötig. Definiert wurde, welche xDOMEA-Elemente in der Ausgabe verfügbar sein sollen, die Zuordnung der jeweiligen DOMEA-Datenbankfelder auf die xDOMEA-Elemente ist jedoch flexibel. Dadurch ist es möglich, in das xDOMEA-Element *Vorgangsbetreff* jeweils den Inhalt des korrekten Datenbankfelds zu exportieren, auch wenn es sich dabei in Hessen um das Feld X, in Nordrhein-Westfalen jedoch um das Feld Y handelt.
2. Alle für eine spätere Nutzung im Archiv interessanten Metadaten sollten soweit möglich in xDOMEA exportiert werden. Die denkbare Alternative wäre gewesen, nur grundlegende Metadaten in xDOMEA zu exportieren (damit diese z.B. automatisiert für die archivische Erschließung verwendet werden können) und alle weiteren Metadaten (z.B. Historien und Laufwegsinformationen) zusätzlich als PDF/A zu exportieren und zu den jeweiligen Objekten (Dokumente / Vorgänge / Akten) zu legen. Damit wäre jedoch ein Verlust an Weiterverarbeitbarkeit der Daten einhergegangen. Zudem wäre die Struktur der Primärdokumente noch komplexer geworden: Während jetzt ein logisches Element, etwa ein bestimmtes Ausgangsschreiben, aus drei unterschiedlichen Dokumentversionen (erster Entwurf, abschließender Entwurf des ersten Bearbeiters, Endversion der zweiten Bearbeiterin) bestehen kann, die jeweils als Word-Datei und als PDF/A-Datei übernommen werden, kämen dann noch die Historien und Laufwegsinformationen als PDF/A-Datei zu jeder der drei Versionen hinzu, insgesamt hätte man also neun Dateien. Auch für eine denkbare spätere Anzeige mit Hilfe eines xDOMEA-Viewers, der Dokumente, Vorgänge und Akten wieder in ihrer ursprünglichen Struktur samt Metadaten anzeigen kann, ist eine extensive xDOMEA-Nutzung sinnvoll.
3. xDOMEA-Dateien können eine unterschiedliche Bezugsgröße haben. Es ist möglich,
 - die gesamte Aussonderung (mit allen Akten und allen Vorgängen)
 - jede einzelne Akte (samt Vorgängen)
 - oder jeden einzelnen Vorgang
 in einer xDOMEA-Datei darzustellen. Hier wurde aus archivfachlichen Gründen die Lösung einer xDOMEA-Datei je Vorgang gewählt. Jedes spätere Archivpaket mit den Dokumenten eines Vorgangs beinhaltet dadurch eine xDOMEA-Datei, die nicht nur die Metadaten zu den Dokumenten und zum Vorgang enthält, sondern auch die Informationen zur Akte und zum Aussonderungslauf.⁷

⁶ Unter anderem durch sogenannte customizing Elemente, also Metadatenfelder, die vom Anwender nach den jeweiligen Bedürfnissen unterschiedlich definiert werden können.

⁷ Die xDOMEA-Datei enthält pauschale Informationen über den Aussonderungslauf, z. B. zur anbietenden Behörde. Die Inhalte einer Anbietung werden durch die im Prozess generierten Excel-Tabellen (insbesondere die Übergabeliste) nachgewiesen, die zusätzlich zur Aktenführung im Archiv im Sinne einer Dokumentation auch als digitale Archivalien archiviert werden können.

4. Aus Gründen der Wirtschaftlichkeit wurde die Metadatenstruktur vereinfacht (Abb. 4). Abgebildet wird nicht die Vielfalt, die xDOMEA erlaubt, sondern die Struktur Akte/ Vorgang/ Dokument. Das zwingt die Anwender zwar dazu, im Vorfeld der Aussonderung die Struktur der Aktenführung zu vereinheitlichen⁸ – andererseits wäre dem Wildwuchs an Ordnungsstrukturen ansonsten mit vertretbarem Aufwand nicht beizukommen gewesen.

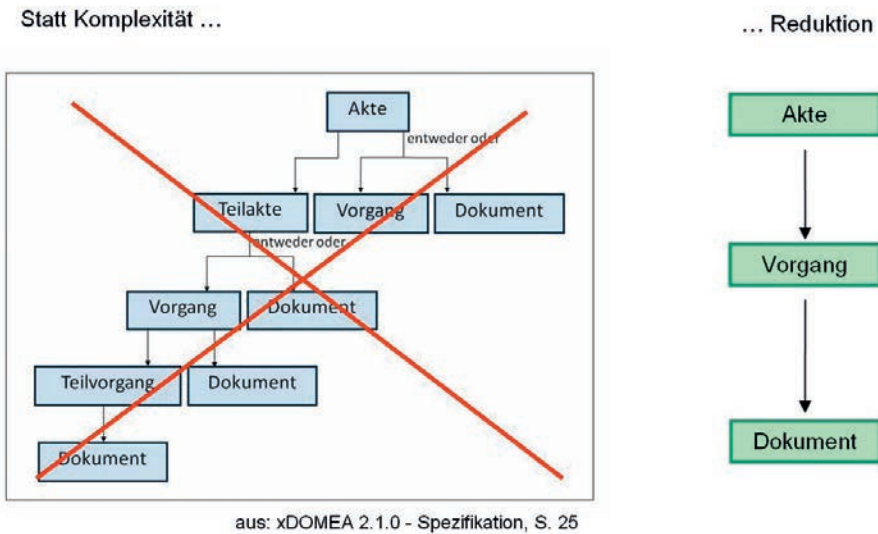


Abb. 4: Vereinfachung der Metadatenstruktur.

Exportiert werden bei der Aussonderung im Wesentlichen folgende Informationen:

- Kopfangaben (Angaben zum Aussonderungslauf):
Prozess-ID, Erstellungszeitpunkt, Sendendes System, Angaben zum Aussonderungslauf (Auswahl, Paketgröße etc.), abgebende Stelle
- Akte:
Typ (z.B. Personalakte), Aktenzeichen, Betreff, Laufzeit, Federführung, Vertraulichkeitsstufe, Medium, Historie und Laufwegsinformationen

⁸ Dazu sind nicht in erster Linie Vorgaben an die DOMEA-Nutzer nötig, dies kann auch durch Automatismen geschehen. Beispielsweise kann beim Abschließen einer Akte automatisch ein Behelfsvorgang eingefügt werden, der die Dokumente umfasst, die direkt unter der Akte angelegt wurden, und damit die dreistufige Struktur Akte-Vorgang-Dokument herstellt.

- Vorgang:
Typ, Aktenzeichen, Betreff, Laufzeit, Federführung, Vertraulichkeitsstufe, Medium, Aufbewahrungsdauer, Aufbewahrungsende, Historie und Laufwegsinformationen, Verfügungen (xDOMEA Element *InternerGeschäftsgang*), Verweise
- Dokument:
Betreff, Kennzeichen (GZ), Vertraulichkeitsstufe, Medium, Bemerkung, fremdes Geschäftszeichen, Posteingangsdatum, Postausgangsdatum, Datum des Schreibens, Typ, Historie und Laufwegsinformationen, Verfügungen (xDOMEA Element *InternerGeschäftsgang*), Verweise
- Version (inhaltlich unterschiedliche Dokumentversionen):
Nummer, Format
- Dateiformat (unterschiedliche Renditions (Darstellungsformen) derselben Dokumentversion, z.B. Word und PDF/A):
Formatname, Version des Dateiformats, Dateiname des Primärdokuments

Fazit

Das Aussonderungsverfahren musste so beschrieben werden, dass es von den unterschiedlichen Anwendern eingesetzt werden kann. Eine besondere Herausforderung war dabei die unterschiedliche Konfiguration der gleichen Software – DOMEA sieht bei jedem Anwender anders aus. Auch die unterschiedliche Nutzung von DOMEA – bei jedem Anwender, aber auch innerhalb einer Landesverwaltung zwischen den einzelnen Ressorts, – verursachte keine geringen Schwierigkeiten. Je freier die Aktenführung, desto komplexer und unterschiedlicher sind die abgebildeten Strukturen, und desto schwieriger gestaltet sich die Aussonderung. Aus Sicht des Archivs wäre es wünschenswert, die Aktenführung in einem DMS sehr strikt und einheitlich zu regeln und wenig Spielraum für individuelle Anpassungen zu lassen. Naturgemäß widerspricht dieses Interesse jedoch dem ebenso berechtigten Interesse der Behörden nach maßgeschneiderten Möglichkeiten der Aktenführung und dem Interesse, die Akzeptanz eines DMS bei den Benutzerinnen und Benutzern nicht durch starre Reglementierung zu beeinträchtigen. Zwischen beidem muss ein Ausgleich gefunden werden.

Da die Nutzung von DOMEA von einer übergreifenden Regulierung weit entfernt ist, mussten vielfach Lösungen gefunden werden, die trotzdem einen einheitlichen Datenexport ermöglichen. Dies war nur über nachgelagerte Vereinheitlichungen möglich, entweder im Aussonderungsprozess selbst, oder durch Vorverlagerung dieser Arbeit in die davor liegende Phase. Viele der entsprechenden Maßnahmen setzen sinnvollerweise bereits beim Abschluss von Vorgängen und Akten ein, also beim Übergang in die Aufbewahrungsphase.⁹ Ob und inwieweit das im Einzelfall zu Informationsverlusten bei der Archivierung führt, wird sich erst mit der Zeit herausstellen. Sollten sich solche Informationsverluste aus archivischer Sicht als bedenklich für

⁹ Zu diesem Zeitpunkt werden zum Beispiel Postmappen aufgelöst und die enthaltenen Dokumente Vorgängen zugeordnet.

eine spätere historische Nutzung erweisen, muss nachgebessert werden. Dies kann bei einem solch komplexen Prozess an verschiedenen Stellen geschehen – entweder bei der Aussonderung selbst, oder in einer vorgelagerten Phase (beim Abschließen, oder in der Bearbeitungsphase).

Zwei Punkte können jedoch unabhängig von diesen grundsätzlichen Herausforderungen, die sich bei komplexen Systemen immer stellen werden, beachtet werden:

Organisatorische Mängel – Vernachlässigung der Aktenführung, fehlendes Wissen über die Schriftgutverwaltung bei den Sachbearbeiter/innen und fehlende Vorgaben – führen nicht nur zu Problemen in der Bearbeitungsphase von Vorgängen und Akten, sie bereiten unter Umständen große Schwierigkeiten bei der Aussonderung. Dazu gehört zum Beispiel der fehlende Nachweis von Hybridakten oder Hybridvorgängen, oder das Fehlen einer Vorgabe zum Setzen von Aufbewahrungsfristen, was willkürliche und nicht immer sinnvolle Entscheidungen der Sachbearbeiter/innen zur Folge hat. Hier können organisatorische Regelungen, die den gesamten Lebenszyklus einer Akte im Blick haben, für Verbesserung sorgen.

Die spätere Aussonderung sollte bei der Einführung eines neuen DMS und ebenso bei der Erweiterung eines DMS durch individuelle Zusatzmodule wenigstens im Blick sein. Es ist verständlich, dass die Behörden bei der Einführung eines neuen Systems erst einmal andere Sorgen haben, als die spätere Aussonderung en detail zu durchdenken. Wenn jedoch wenigstens in Form einer knappen Checkliste Anforderungen, die eine spätere Aussonderung stellt (z.B. *alle Metadaten des Kernmoduls, aber auch denkbare Erweiterungsmodule, müssen ausgesondert werden können*), überprüft werden, ließe sich vielleicht manche ungünstige Konstellation vermeiden, die ansonsten den Aussonderungsprozess komplizierter (und teurer) macht.

Neu im Digitalen Archiv. Ein Streifzug durch Bestände und Methoden

Von CORINNA KNOBLOCH

Dieser Text soll anhand einiger Beispiele eine kurze Vorstellung darüber geben, welche Arten von Unterlagen sich in DIMAG finden. Daneben werden einige besondere Herausforderungen im Vergleich zur Archivierung analoger Unterlagen und einige praktische Überlegungen skizziert.

Beschreibung der bereits in DIMAG enthaltenen Bestände

Am Ende des Jahres 2011 waren 178,5 Millionen Datensätze elektronischer Unterlagen in DIMAG, dem Digitalen Magazin des Landesarchivs Baden-Württemberg, gespeichert. Diese entsprechen einem Speichervolumen von 313 Gigabyte. Allein im Jahr 2011 wurden 100,8 Mio. Datensätze mit 187 Gigabyte nach DIMAG übernommen. Diese 178,5 Mio. Datensätze stammen aus vier verschiedenen

Abteilungen des Landesarchivs Baden-Württemberg, nämlich dem Generallandesarchiv Karlsruhe, dem Staatsarchiv Ludwigsburg, dem Staatsarchiv Sigmaringen und dem Hauptstaatsarchiv Stuttgart.

DIMAG

- **Ende 2010:**
77,7 Mio. Datensätze (126 GB)
- **Ende 2011:**
178,5 Mio. Datensätze (313 GB)

Seit jeher ist die im Landesarchiv verwahrte Überlieferung äußerst vielfältig. Die Magazine, in denen am Ende des Jahres 2011 insgesamt 144 842 laufende Meter Unterlagen verwahrt wurden, beherbergen beispielsweise Pergamenturkunden, Lagerbücher, Akten, Karten, Pläne, Glasplatten, Tonaufnahmen und Filme. Daher ist es fast selbstverständlich, dass auch DIMAG sehr unterschiedliche Unterlagen enthält. Die Vielseitigkeit resultiert einerseits aus dem großen Aufgabenspektrum, das die verschiedenen Provenienzstellen wahrnehmen. Andererseits unterscheidet sich beispielsweise auch die Software, mit deren Hilfe die Daten entstanden sind und in der sie seit ihrer Entstehung bearbeitet wurden. Um mit diesen Unterschieden zurechtzukommen, ist es wichtig, dass DIMAG für verschiedene Objektarten offen ist.

Die abgebenden Stellen weichen normalerweise nicht von denen ab, die zuvor bereits analoge Unterlagen produziert und an das Landesarchiv abgegeben haben. In der Regel handelt es sich um baden-württembergische Landesbehörden oder sonstige landeseigene Stellen, deren archivwürdige – das heißt dauerhaft aufzubewahrende – Unterlagen von einem der Staatsarchive übernommen werden.

Zahlreiche der in DIMAG enthaltenen Bestände liegen rein digital vor. Daneben sind einige hybride Bestände vorhanden, die sich aus konventionellen und digitalen Teilen zusammensetzen. Mit der Umstellung auf elektronische Aktenführung in einer wachsenden Zahl von Behörden dürfte ihr Anteil weiter steigen. Viele Unterlagen, die teilweise über Jahrzehnte in Papierform an die Archive gelangt sind, werden künftig in digitaler Form abgegeben. Diese Entwicklung ist jedoch nicht neu. Nachdem zum Beispiel Volkszählungsunterlagen aus einem Zeitraum von mehr als einem Jahrhundert in Papierform in die Archive gelangt waren, übergab das Statistische Landesamt Baden-Württemberg im Jahr 2005 elektronische Unterlagen zur Volkszählung 1961 an das Staatsarchiv Ludwigsburg. So mancher zukünftige Benutzer wird daher sowohl analoge als auch digitale Unterlagen einsehen, wenn er sich für Ergebnisse der Volkszählungen nach dem Zweiten Weltkrieg interessiert. Besonders bedeutsam ist deshalb eine enge Verzahnung der analogen Überlieferung und der digitalen Überlieferung im Erschließungssystem der jeweils zuständigen Archivabteilung.

Die elektronischen Unterlagen, die ins Digitale Archiv übernommen werden, umfassen zum Beispiel Geoinformationen, Fachverfahren, statistische Mikrodaten, digitale Fotografien, Texte, Pläne und Videos sowie Webseiten aus Intranetangeboten. Auch bei der Anzahl der Dateien und der Speichermenge gibt es erhebliche Unterschiede. Über die Bedeutung der Ablieferung sagen diese Mengenangaben allerdings nur sehr wenig aus.

BA Ergebnisse der Bundestagswahlen

Bundestagswahl 14.August 1949

Wahlberechtigte	14.08.1949	BA 0001 4908	ab 27.05.70
Wähler ohne Briefwähler	14.08.1949	BA 0003 4908	ab 27.05.70
Stimmen			
ungültige Stimmen	14.08.1949	BA 0083 4908	ab 27.05.70
Gültige Stimmen	14.08.1949	BA 0004 4900	ab 27.05.70
Verteilung der gültigen Stimmen			
CDU *Christlich Demokratische Union Deutschlands	14.08.1949	BA 0085 4908	ab 27.05.70
SPD *Sozialdemokratische Partei Deutschlands	14.08.1949	BA 0086 4908	ab 27.05.70
FDP/DVP *Freie Demokratische Partei/Demokratische *Volkspartei	14.08.1949	BA 0087 4908	ab 27.05.70
Sonstige	14.08.1949	BA 0114 4908	ab 27.05.70

Abb. 2: Merkmalskatalog zur Struktur- und Regionaldatenbank des Statistischen Landesamts Baden-Württemberg.

Aus der Überlieferung des Statistischen Landesamts hat das Staatsarchiv Ludwigsburg beispielsweise vor einigen Monaten erste Teile der Struktur- und Regionaldatenbank (SRDB) übernommen. Bislang handelt es sich um acht digitale Objekte aus den Bereichen Fläche, Be-

völkerung und Wahlen. Demnächst sollen sie jedoch weiteren Zuwachs erhalten. Der bislang vergleichsweise kleine Umfang der digitalen Objekte darf aber nicht über die in jedem einzelnen digitalen Objekt enthaltene Datenmenge und die große Bedeutung der vollständigen SRDB als zentrale statistische Datenbank des Statistischen Landesamts hinwegtäuschen. Sie beinhaltet Daten aller Statistikbereiche für alle Gemeinden, Kreise, Wahlkreise, Regionen und Regierungsbezirke in Baden-Württemberg. Die aktuellen Daten sind über die Internetseite des Statistischen Landesamts frei abrufbar.¹ Sie lassen erahnen, welche Möglichkeiten der historischen Forschung durch eine Einsichtnahme in die erhobenen Zahlen über einen längeren Zeitraum zur Verfügung stehen werden. Da zahlreiche wesentliche Bereiche des öffentlichen Lebens in Baden-Württemberg abgebildet werden, dürften die Daten von den künftigen Historikern vermutlich für sehr unterschiedliche Fragestellungen genutzt werden. Bei der Aufbereitung in der Behörde und im Archiv musste daher darauf geachtet werden, dass die Datenmengen auch außerhalb der Großrechnerwelt, in der sie entstanden sind, nutzbar sind und dass zugleich die großen Abfragemöglichkeiten weitgehend erhalten bleiben. Mit diesen Zielen wurden bereits vor der Datenübergabe umfangreiche Vorarbeiten durch das Statistische Landesamt geleistet. Beispielsweise mussten die ursprünglich jeweils mindestens 15 000 Spalten in Datensätze überführt werden, wodurch pro Ort eine Spalte entstand. Im Archiv wurden die Daten für eine Überführung nach DIMAG in Kooperation mit der Behörde in das Format CSV migriert.

Für erste wichtige Auswertungen und Diagramme wurden die in DIMAG enthaltenen Daten der Struktur- und Regionaldatenbank bereits kurz nach ihrer Übernahme genutzt. Das landeskundliche Informationssystem LEO-BW,² das zum 60. Jubiläum des Landes Baden-Württemberg am 25. April 2012 freigeschaltet wurde, verwendet die Daten unter anderem für eine Vielzahl von orts- und personenbezogenen Statistiken. Diese betreffen beispielsweise das Gemeindegebiet, die Bevölkerungsdichte, die Geschlechterverteilung, die Altersstruktur, den Ausländeranteil, die Religionszugehörigkeit und Bundestags- und Landtagswahlen.

¹ <http://www.statistik.baden-wuerttemberg.de/SRDB>. Alle Links wurden am 15.10.2012 überprüft.

² <http://www.leo-bw.de>.

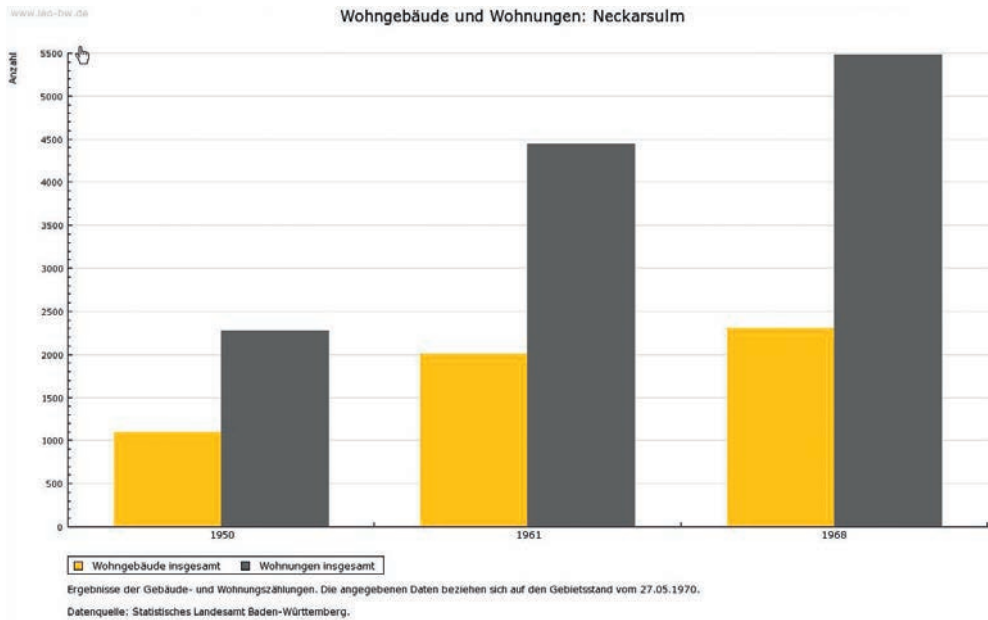


Abb. 3: Beispiel für die Nutzung der Daten der Struktur- und Regionaldatenbank in LEO-BW.

Beispiele für in DIMAG enthaltene Bestände

Aus dem Generallandesarchiv Karlsruhe, zuständig für den Regierungsbezirk Karlsruhe, befinden sich unter anderem elektronische Unterlagen zum Umweltinformationssystem der Landesanstalt für Umwelt, Messungen und Naturschutz Baden-Württemberg in DIMAG. Diese Überlieferung ist äußerst umfangreich und beinhaltet beispielsweise Unterlagen zu Altlasten, Wasserrechten, Anlagen zum Gewässerschutz, Biotopen, Abfallanlagen und zum Arbeitsschutz.

Zahlr	Objekt	Objektart (Name)	A/V	Erfassungsumfang	Erfassungsumf	Turnus Date	Geschätzte Gesamtz
162	4.1.2.3	Naturdenkmal Einzelgebilde (END)	A	Landesweit	Jeweils aktueller B keine Angaben		Ende 2005: 14 427 Objek
163	4.1.2.4	Flächenhaftes Naturdenkmal (FND)	A	Landesweit	Jeweils aktueller B keine Angaben		Ende 2005 6.174
164	4.1.2.5	Naturpark (NPK)	A	Landesweit	Jeweils aktueller B keine Angaben		Ende 2005 7
165	4.1.2.6	FFH-Gebiet	A	Landesweit	keine Angaben		363
166	4.1.2.7	Vogelschutzgebiet (SPA)	A	Landesweite Erfassung	Jeweils aktueller B keine Angaben		Exakt 73
167	4.1.2.8	Bannwald	A	Landesweit	Laufende Einarbeit jährlich		ca. 120 Einzelpolygone, E
168	4.1.2.9	Schonwald	A	Landesweit	Laufende Einarbeit keine Angaben		ca. 550 Einzelpolygone, 1
169	4.2	Landschaftsplanung					
170	4.2.1	Naturraum	B	Landesweit	Aktuelle Karte	keine Angaben	66 naturräumliche Einheit
171	4.2.2	Landschaftsökologische Einheit	B	Landesweit	Aktuelle Karte	keine Angaben	?
172	4.2.3	Landschaftszerschneidung	B	Landesweit	Aktuelle Karte	keine Angaben	?
173	4.2.4	Forstliche Standortskundliche Regionale Gliederung	B	Landesweit	ab 1995, laufende / keine Angaben		ca. 140 Einzelpolygone
174	4.3	Artenschutz					
175	4.3.1	Artenlexikon	A	Landesweit	Erfassungsstand 2 keine Angaben		geschätzt. 40.000 Arten
176	4.3.2	Artenfundort	V			keine Angaben	
177	4.4	Waldfunktionen					
178	4.4.1	Bodenschutzwald	A	Landesweit	fachlicher Stand c jährlich		ca. 5000 Einzelpolygone
179	4.4.2	Sonstiger Wasserschutzwald	A	Landesweit	Der Stand wird jäh keine Angaben		ca. 4500 Einzelpolygone
180	4.4.3	Klimaschutzwald	A	landesweit	Fachlicher Stand c keine Angaben		ca. 3000 Einzelpolygone
181	4.4.4	Lawenschutzwald	A	Landesweit	Fachlicher Stand c unregelmäßig		ca. 20 Einzelpolygone
182	4.4.5	Immissionsschutzwald	A	Landesweit	Fachlicher Stand c unregelmäßig		ca. 3200 Einzelpolygone
183	4.4.6	Sichtschutzwald	A	Landesweit	Fachlicher Stand c keine Angaben		ca. 600 Einzelpolygone
184	4.4.7	Erholungswald Stufe 1 + 2	A	Landesweit	Fachlicher Stand c unregelmäßig		ca. 4500 Einzelpolygone
185	4.4.8	Gesetzlicher Erholungswald	A	Landesweit	Fachlicher Stand c keine Angaben		ca. 85 Einzelpolygone
186	4.4.9	Gesetzlicher Schutzwald Umwelteinwirkungen	A	Landesweit	Fachlicher Stand c unregelmäßig		ca. 50 Einzelpolygone
187	4.5	Forstwirtschaft					
188	4.5.1	Waldbesitzstruktur	V	Landesweit	Fachlicher Stand 1 Aktualisierung n	ca. 110000 Einzelpolygon	
189	4.5.2	Forstliche Organisation (bis 2004)	V	Landesweit	Fachlicher Stand 1 unregelmäßig	ca. 3500 Einzelpolygone	
190	5	ARBEITSSTÄTTEN, ARBEITSSCHUTZ, IMMISSIONSSCHUTZ					
191	5.1	Tätigkeitsstatistik und Recht					
192	5.1.1	Tätigkeitsstatistik Gewerbeaufsicht	V	landesweit	anlassbezogen	monatliche	40 000 im Jahr
193	5.2	Betriebsdaten					
194	5.2.1	Arbeitsstätte	A	landesweit	Festlegung je Merk	monatliche	ca. 300.000 Arbeitsstätten
195	5.3	Arbeitsschutz					
196	5.3.1	Genehmigungen/Anzeigen nach StrlSchV	A	landesweit	anlassbezogen	monatliche	3500 Arbeitsstätten
197	5.3.2	Genehmigungen/Anzeigen nach R0V	A	landesweit	anlassbezogen	monatliche	10000 Arbeitsstätten
198	5.3.3	Besondere Arbeitsschutzmerkmale	A	landesweit	Festlegung je Merk	monatliche	

Abb. 4: Auszug aus dem WIBAS-Objektartenkatalog.

Wie der Objektartenkatalog des Informationssystems Wasser, Immissionsschutz, Boden, Abfall, Arbeitsschutz (WIBAS) zeigt, sind in diesen Bereichen sehr vielfältige Unterlagen vorhanden. Mittlerweile wurde für sie ein sehr detailliertes Bewertungsmodell ausgearbeitet, das auf eine behördliche Klassifikation aufsetzt. Daher hat es zwei Vorteile: Zum Einen ist es ein gemeinsamer Maßstab, der zentral von allen Beteiligten abgerufen werden kann. Zum Anderen kann es leichter aktuell gehalten werden als ein Einzeldokument.

Aus dem für den Regierungsbezirk Stuttgart zuständigen *Staatsarchiv Ludwigsburg* befinden sich unter anderem Daten des Statistischen Landesamts, des Landeskriminalamts, des Landesamts für Straßenwesen, des Landesdenkmalamts, des Landesamts für Geoinformation und Landentwicklung, der Akademie Schloss Solitude und der Bundesbahndirektion Stuttgart in DIMAG.



Abb. 5: Neckarbrücke bei Besigheim (StAL EL 75 VI DO 40).

Aus den Beständen des für den Regierungsbezirk Tübingen zuständigen *Staatsarchivs Sigmaringen* verwahrt das DIMAG unter anderem elektronische Unterlagen des Staatlichen Hochbauamts Tübingen, der Justizvollzugsanstalt Rottenburg, der Pädagogischen Arbeitsstelle für Erwachsenenbildung – Außenstelle Inzigkofen und der Pädagogischen Hochschule Weingarten.

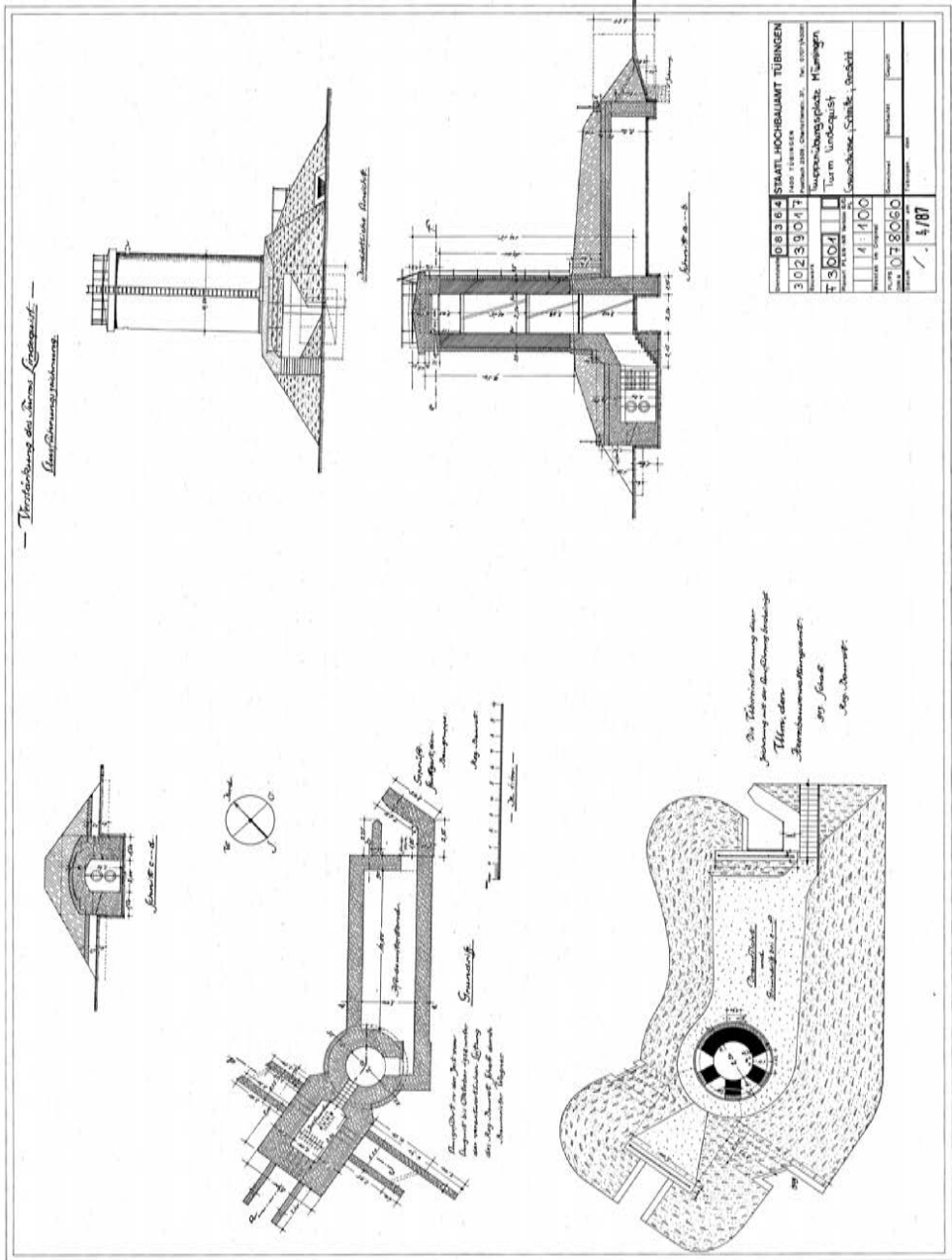


Abb. 6: Digitalisierter Plan des Turms Lindequist (StAS Wü 129/1 T 2 DO 29).

Aus dem Bereich des *Hauptstaatsarchivs Stuttgart* befinden bisher ausschließlich Daten aus der Überlieferung der baden-württembergischen Ministerien in DIMAG. Erste Daten aus dem Audiovisuellen Archiv sollen demnächst nach DIMAG überführt werden.³ Der hier präsentierte Text einer englischsprachigen Rede des ehemaligen Wissenschaftsministers Klaus von Trotha in St. Louis im Jahr 1995 wurde sowohl im damaligen Abgabeformat als Microsoft-Word-Datei in DIMAG gespeichert als auch im für die Langzeitarchivierung geeigneten Format PDF/A-1b:2005.

**Speech given by Klaus von Trotha,
Baden-Württemberg Minister of Science and Research**

**on the occasion of a visit to St. Louis
and the signing of a letter of intent with
Governor Mel Carnahan, State of Missouri**

**in appreciation of a Student Exchange Program linking
St. Louis Universities and
Berufsakademie Villingen-Schwenningen**

St. Louis/Missouri on June 5, 1995

Dear Governor Carnahan (falls zugegen),

...

Ladies an Gentlemen,

**it is a great pleasure for me to visit Missouri and to be able to
convey my personal as well as the official greetings of the State**

Abb. 7: Rede des baden-württembergischen Wissenschaftsministers Klaus von Trotha am 5. Juni 1995 in St. Louis (HStAS EA 13/102 DO 1).

³ Ausführlichere Angaben zum Audiovisuellen Archiv enthält der Beitrag von Peter Bohl und Johannes Renz in diesem Tagungsband.

Die Erhaltung des Abgabeformats hat unter anderem den Grund, allen Nutzern, die an der Authentizität der durch Migration entstandenen Dateien zweifeln, die vom Landesarchiv übernommenen Daten noch vorlegen zu können.

Besondere methodische Herausforderungen im Vergleich zur Archivierung analoger Unterlagen

Die methodischen Anforderungen, die sich Archivaren bei der Zugangsbearbeitung digitaler Unterlagen stellen, weichen zum Teil erheblich von den Anforderungen bei der Überlieferungs- bildung analoger Unterlagen ab. In anderen Teilen sind sie fast identisch.

Ein erstes Problem ist die Darstellung von Dateiinhalten. Dies gilt in besonderem Maße, wenn die *herkömmliche Software* auf einem Rechner die Dateien nicht richtig interpretieren kann. Die Ursachen des Problems können unterschiedlich sein. Einerseits sind manche Dateiformate nicht mehr gebräuchlich oder so selten, dass die Dateiinhalte nur mit spezieller Software dargestellt werden können. Andererseits stammen viele auch sehr junge Dateien aus einer anderen Soft- und Hardwareumgebung. Besondere Schwierigkeiten treten beispielsweise häufig mit Dateien auf, die mit Großrechnern erstellt wurden und in denen auch gepackte Felder auftreten können.

	A	B	C	D	E	F	G	H
1	ouññöñöññf	Y	Y	O				
2								
3								
4								
5								
6								
7								
8								
9								
10								
11								
12								
13								
14								
15								
16								
17								
18								
19								
20								
21								
22								
23								
24								

Abb. 8: Darstellung einer Großrechnerdatei mit gepackten Feldern in Microsoft Excel.

Zum Teil kann man sich durch den Einsatz von kostenlosen Tools und anderer Software behelfen. Teilweise müssen die jeweilige Behörde und das Archiv aber andere Lösungen finden, die häufig einen großen Programmieraufwand mit sich bringen. Ansonsten sieht eine Datei, die aus zahlreichen klar strukturierten Angaben bestehen müsste, unter Umständen aus wie die hier gezeigte.

Für Archivare, die sich zuvor ausschließlich mit analogen Unterlagen beschäftigt haben, erscheint ein weiteres Problem zunächst ungewöhnlich: Immer wieder kommt es vor, dass Mikrodaten vorliegen, man aber viele Angaben nicht versteht, weil man die entsprechenden Schlüssel nicht auflösen kann. Anders als in Papierakten bestehen viele Spalten- und Zeilenbezeichnungen aus kurzen Merkmalschlüsseln, für die in den Mikrodaten häufig keine Erklärung mitgeliefert wird. Man benötigt die jeweils in der Entstehungszeit eingesetzten Systematiken als Dokumentation, um viele auf den ersten Blick kryptische Bezeichnungen und Zahlenkolonnen aufzulösen. Ansonsten droht die Gefahr, dass der Archivar, der den Zugang bearbeitet, ebenso an der Interpretation der Daten scheitert wie ein Benutzer, der sich im Jahr 2250 mit ihnen beschäftigt. Dank der entsprechenden Systematik ist es dagegen fast problemlos möglich, zum Beispiel die Angabe *BB00015603* als *Wahlberechtigte* in der Struktur- und Regionaldatenbank des Statistischen Landesamts Baden-Württemberg aufzulösen.

Häufig besteht eine erste Herausforderung für die Archivare jedoch schon in der Ermittlung und rechtzeitigen Sicherung der passenden Systematiken. Immer wieder kommt es vor, dass die Unterlagen in der Behörde als veraltet und daher überflüssig angesehen und vernichtet werden. Anfragen bei vergleichbaren Behörden innerhalb und außerhalb des Bundeslandes sowie bei vor- oder nachgeordneten Stellen können unter Umständen Abhilfe schaffen. Manchmal können Behördenmitarbeiter, denen die Vorgehensweise und zahlreiche Hintergrundinformationen aus der täglichen Arbeit noch bekannt sind, einige Wissenslücken schließen. Wenn diese Personen aus dem Arbeitsleben bereits ausgeschieden sind, werden die Möglichkeiten weiter eingeschränkt. So besteht bei der digitalen Archivierung die Gefahr von *Datenruinen*, wobei die Mikrodaten zwar noch vollständig vorhanden sein können, aber nur noch zum Teil verstanden werden können. Immer wieder ist daher ein Wettlauf gegen die Zeit notwendig. Doch selbst hier unterscheidet sich die digitale Archivierung nur bedingt von der Archivierung analoger Unterlagen. So manches Amt vernichtete in der Vergangenheit ebenfalls seine Unterlagen, ohne sie vorher dem zuständigen Archiv anzubieten – nur weil der Keller voll war oder ein Umzug anstand.

Um die abgegebenen Mikrodaten übersichtlicher und für die künftige Benutzung strukturierter aufzubereiten, wurde vom Landesarchiv das Tool ToCSV entwickelt. Neben der Erstellung von CSV-Dateien ermöglicht es eine Aufteilung der fortlaufenden Zahlenkolonnen in verschiedene Spalten. Sofern die Mikrodaten mehrere Satzarten enthalten, bei denen die Angaben in unterschiedlichen Positionen stehen, können mehrere verschiedene CSV-Dateien erzeugt werden.

Die Interpretation der Daten wird zudem durch die Erstellung eines Aufbereitungsberichts im PDF/A-Format und einer Tabellenbeschreibung im XML-Format erleichtert. Das vom Landesarchiv entwickelte Tool TabBes ermöglicht die Umwandlung eines in Microsoft Excel ausgefüllten Schemas nach XML. Die entstandene Tabellenbeschreibung erläutert die in den einzelnen Spalten auftretenden Merkmalschlüssel durch eine kurze Beschreibung und durch Angaben zum Feldtyp, zur Feldlänge, zur Codierung und zu vorhandenen Werten. Die Tabellenbeschreibung kann für künftige Benutzer zu den wesentlichen Hilfsmitteln zählen, um Mikrodaten zu analysieren.

Weitere Fragen bei der digitalen Archivierung

Wie die vorgestellten Aspekte zeigen, besteht der Ingest (wie im OAIS-Standard bereits definiert) in der Regel aus weit mehr als der reinen Zugangsbearbeitung. In welchem Dateiformat elektronische Unterlagen aufbewahrt werden sollen, ist nur eine der Fragen, die beantwortet werden müssen. Beim Dateiformat ist es einerseits wichtig, dass man keine zu große Vielfalt aufkommen lässt. Ansonsten würde der schon seit einigen Jahren immer wieder skizzierte Format-Zoo drohen. Dieser wäre für die kommenden Generationen von Archivaren eine große Bürde, die Fehler fast schon provozieren würde. Das eine oder andere Format würde möglicherweise nicht rechtzeitig migriert. Daneben wäre mit einer deutlichen Kostensteigerung zu rechnen, weil automatisierte Datenmigrationen aufwändiger würden.

Zum anderen muss man sich fragen, welche Benutzungsmöglichkeiten künftig gegeben sein sollen. So bietet eine PDF/A-Datei unter Umständen weniger Optionen als eine CSV- oder XML-Datei. Die Einschränkungen reichen weit über die geringeren Möglichkeiten beim Setzen von Filtern und bei umfangreichen Abfragen hinaus. Es ist nur schwer zu prognostizieren, welche Forschungsthemen eines Tages mithilfe der Daten bearbeitet werden sollen. Auf jeden Fall sollte man aber verhindern, zu viele Möglichkeiten für die spätere Nutzung von vornherein auszuschließen und zu verbauen. Manchmal kommt man erst nach mehreren Versuchen zu dem Ergebnis, dass ein Dateiformat in einem bestimmten Fall besser geeignet ist als ein anderes. Eine allgemeingültige Regel lässt sich daher nur schwer aufstellen.

Weitere Fragen bei der Aufbereitung sind zum Beispiel:

- Welche Nutzungs- und Verknüpfungsmöglichkeiten sollen nach der Übernahme bestehen?
- Welche Angaben sind notwendig, um die elektronischen Unterlagen inhaltlich zu verstehen? Reichen die vorhandenen Metadaten aus?
- Gibt es Bezüge und Verlinkungen zu anderen Daten, die unbedingt erhalten werden müssen?
- Welche signifikanten Eigenschaften müssen bei Migrationsmaßnahmen auf jeden Fall beachtet werden?
- Welche Unterlagen sollen ins Archiv übernommen werden, wenn sie in der Behörde sowohl analog als auch digital vorliegen? Sind die analogen und die digitalen Unterlagen im jeweiligen Fall tatsächlich identisch?
- Können bzw. müssen im Vergleich zu analogen Unterlagen möglicherweise ganz neue Kriterien bei der Überlieferungsbildung herangezogen werden?
- Erfordert die digitale Archivierung eine noch engere regelmäßige Zusammenarbeit mit den Schriftgut produzierenden Behörden? Inwieweit kann das Archiv von vornherein bei der Einführung einer neuen Software in der Behörde einbezogen werden, um späteren Problemen mit dem Einrichten einer Schnittstelle vorzubeugen?
- Wie soll man bei Sonderfällen vorgehen, zum Beispiel bei geschützten PDF-Dateien?

- Sind die abgegebenen Unterlagen überhaupt vollständig angekommen? Hier empfiehlt sich die Arbeit mit Prüfsummen – ein Verfahren, das zum Beispiel durch das vom Landesarchiv entwickelte javabasierte Freeware-Programm IngestList erleichtert wird.⁴

The screenshot shows the 'Metadaten erfassen' (Collect Metadata) window of the IngestList application. It features a table with columns for file description, size, date, and MD5 hash. The table contains data for files like '006.sik', '007.doc', '008.docx', '008.eps', '010.pdf', and '011.pdf'. Below the table is an empty text input field labeled 'iliges Eingabefeld:'.

Beschreibung	Zeige:	DB Angaben	techn. Eigenschaften	Pronom / Droid	Anzahl	Jhove	Jhove s
006.sik	1920	2010-08-04 14:38:37	1807cc88e6ed4a9ebb59f637561e458				
007.doc	22628	2010-10-25 14:03:04	eb4b17e75d1a41118a7d01d0a1f7cdd1				
008.docx	10624	2010-08-04 14:51:38	713e4ddbb916dfc0786655b2052078				
008.eps	270482	2010-10-25 12:21:04	8d3dd91413090c679b6939ba847e6119			UTF-8	
010.pdf	29166	2010-10-25 12:20:46	1c1363594b2a7ea0b9a505f9014e932f				
011.pdf	29205	2010-10-25 12:21:53	e2811ec3d8a184585b864ec09e609003				

Abb. 9: Auszug aus den von IngestList erhobenen technischen Eigenschaften.

- Muss der fachliche Austausch mit Kollegen, die im gleichen Arbeitsgebiet tätig sind, weiter intensiviert werden? Der Austausch kann sehr fruchtbar für alle Seiten sein, denn häufig treten vergleichbare Probleme fast zur gleichen Zeit auch bei anderen digitalen Archiven auf. Ein Ziel muss es sein, überflüssige Mehrarbeit zu verhindern. Es kann sinnvoll sein, wesentlich mehr Besprechungen durchzuführen als es bei der Archivierung analoger Unterlagen der Fall ist, bei der die meisten Arbeitsgebiete schon seit vielen Jahren etabliert sind.

Sehr wichtig dürfte insbesondere in Erwartung immer größerer Zugangsmengen sein, dass möglichst viele Wege gefunden werden, Bearbeitungsschritte zu automatisieren und nicht zu viel manuelle Nacharbeit nötig zu machen.

Im Hinblick auf einen effektiven Einsatz der Ressourcen ist die Verwendung von bereits zur Verfügung stehenden Daten für Zwecke der Arbeitserleichterung sehr bedeutsam. Die Daten werden somit *recycelt*. Die Möglichkeiten reichen weit über die Archivierung digitaler Unterlagen hinaus und betreffen auch zahlreiche weitere archivische Arbeitsfelder, beispielsweise die Erschließung und die Überlieferungsbildung analoger Unterlagen.

Schon seit einigen Jahren können von der Behörde erstellte Ablieferungsverzeichnisse direkt in das archivische Erschließungssystem überführt werden. Für die Behörde und das Archiv besteht somit bereits im Vorfeld der endgültigen Bearbeitung des Bestands eine elektronische Recherchemöglichkeit über die vorhandenen Unterlagen. Wenn ein Archivar den Zugang bearbeitet, müssen die bereits erfassten Daten nur noch korrigiert und unter Umständen ergänzt werden.

⁴ Christian Keitel und Rolf Lang: DIMAG und IngestList. Übernahme, Archivierung und Nutzung von digitalen Unterlagen im Landesarchiv Baden-Württemberg. In: Gerald Maier und Thomas Fritzsche (Hg.): Archivische Informationssysteme in der digitalen Welt. Aktuelle Entwicklungen und Perspektiven. Stuttgart 2010. S. 53–63.

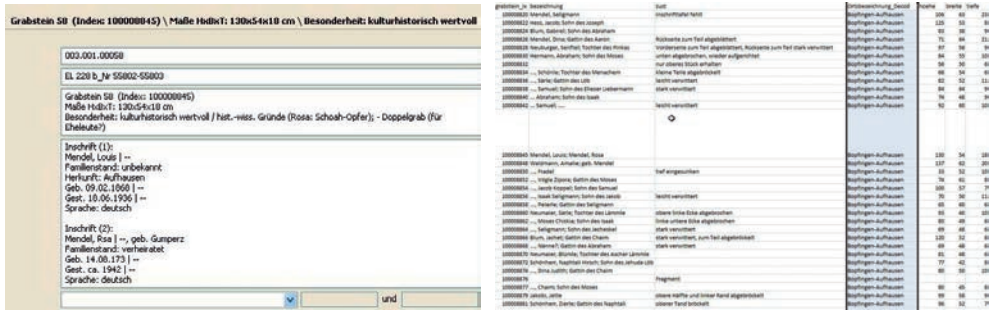


Abb. 10: Dokumentation der jüdischen Grabsteine in DIMAG und im Erschließungssystem des Landesarchivs scope.Archiv.

Bei der im Staatsarchiv Ludwigsburg verwahrten Dokumentation der jüdischen Grabsteine des Landesdenkmalamts Baden-Württemberg konnte zum Beispiel ein nach DIMAG überführter Datenbankauszug genutzt werden, um die Erschließung der Fotodokumentation voranzutreiben und dabei die bestehende Dokumentation zu erweitern. Bereits nach kurzer Zeit gab es erste Anfragen von Benutzern nach weiteren Abgleichmöglichkeiten.

Auch für die Überlieferungsbildung können die elektronischen Abgleichmöglichkeiten zu Arbeitserleichterungen und in manchen Fällen zu einer Verbesserung der Bewertungsentscheidungen führen. Seit vielen Jahren bestehen Bewertungsmodelle, die äußerst detailliert sind und teilweise sogar die Bewertungsentscheidungen in Bezug auf einzelne Aktenzeichen vorgeben. Durch Abgleiche mit den Aussonderungslisten wird das Treffen von Übernahmeentscheidungen deutlich erleichtert.⁵

Bei all den zur Verfügung stehenden Möglichkeiten darf man jedoch nicht das Thema Datenschutz aus den Augen verlieren. Die meisten Daten in DIMAG sind recht jung und unterliegen noch Sperrfristen. Diese entsprechen den Sperrfristen, die auch bei analogem Archivgut zu beachten sind. Beispielsweise können personenbezogene Unterlagen normalerweise frühestens zehn Jahre nach dem Tod des Betroffenen benutzt werden, die meisten Sachakten 30 Jahre, andere 60 Jahre nach dem Abschluss der Akte. Darin hat sich durch die Archivierung digitaler Unterlagen nichts geändert.

Da die technischen Verknüpfungsmöglichkeiten bei digitalen Unterlagen auch missbraucht werden könnten, ist aber eine erhöhte Sorgfaltspflicht notwendig und eine Grundlage für eine vertrauensvolle Zusammenarbeit mit den Schriftgut produzierenden Behörden, damit diese langfristig ebenso selbstverständlich funktioniert, wie es bei der Übernahme analoger Unterlagen schon seit langer Zeit üblich ist.

⁵ In den vergangenen Jahren entstanden einige Aufsätze, die sich mit den verbesserten Bewertungsmöglichkeiten beschäftigen. Für Personalakten sei beispielhaft genannt: Franz-Josef Ziwes: Wikipedia und Co. statt Sisyphus? Konventionelle und digitale Hilfsmittel zur qualitativen Bewertung von Personalakten. In: Archivar 63 (2010). S. 175–178.

Vom Magnetband auf Content Addressed Storage Speichermigration im Stadtarchiv Stuttgart¹

Von KATHARINA ERNST und HEIKE MAIER

Der Beitrag befasst sich mit dem Wechsel des Langzeitspeichers im Stadtarchiv Stuttgart. Er erläutert die Gründe für den Wechsel, die Anforderungen an den neuen Speicher und beschreibt den Verlauf der Speichermigration.

Der alte Speicher

Der Langzeitspeicher ist das älteste Modul der Langzeitarchivierung im Stadtarchiv Stuttgart, ihn haben wir als erstes konzipiert und umgesetzt.² Hier will ich nur kurz rekapitulieren, wie dieser Speicher aussah und warum wir ihn so gewählt hatten. Wir benötigten sehr schnell einen Langzeitspeicher, weil uns große Mengen archivwürdiger Daten angeboten worden waren. Der Speicher sollte die Daten redundant speichern, räumlich getrennt, und auf WORM-Medien. Wir konnten damals auf eine Infrastruktur aufsetzen, die in der Landeshauptstadt Stuttgart für das Backup ohnehin gerade beschafft wurde. Als wir als zweites Modul unsere Langzeitarchivierung um ein Ingest- und Datenmanagementmodul erweitert haben, haben diese Module auf dem Bandspeicher aufgesetzt.

Die Speicherung lief bis Herbst 2011 folgendermaßen ab: Der Archivierungsserver Itaflow schob das digitale Archivgut auf den Content Storage Bus (CSB). Der CSB übergab die Daten über einen Storage-Adapter an den ITA-Server. Vom ITA-Server wurden die Daten an den Tivoli Storage Manager übergeben. Der CSB kann nicht direkt mit dem TSM kommunizieren, daher war der ITA-Server dazwischengeschaltet.³ TSM brachte die Daten dann auf Magnetbänder. Die Magnetbänder waren WORM-Bänder, nur das Stadtarchiv hat auf diese Bänder geschrieben. Sie lagen in Bandbibliotheken, in denen auch die Magnetbänder vom Backup gelesen und beschrieben wurden.

¹ Der Vortragsstil wurde für die Publikation beibehalten.

² Vgl. Heike *Maier*: Der Langzeitspeicher des Stadtarchivs Stuttgart – ein strukturintegrativer Lösungsweg. In: Erfahrungen mit der Übernahme digitaler Daten. Bewertung, Übernahme, Aufbereitung, Speicherung, Datenmanagement. Hg von Katharina *Ernst* (Veröffentlichungen des Archivs der Stadt Stuttgart Bd. 99). Stuttgart 2007. S. 52–57.

³ Seit Frühjahr 2012 ist mit der Version DOXiS4 CSB V2.3. auch eine direkte Anbindung von TSM/SSAM möglich.

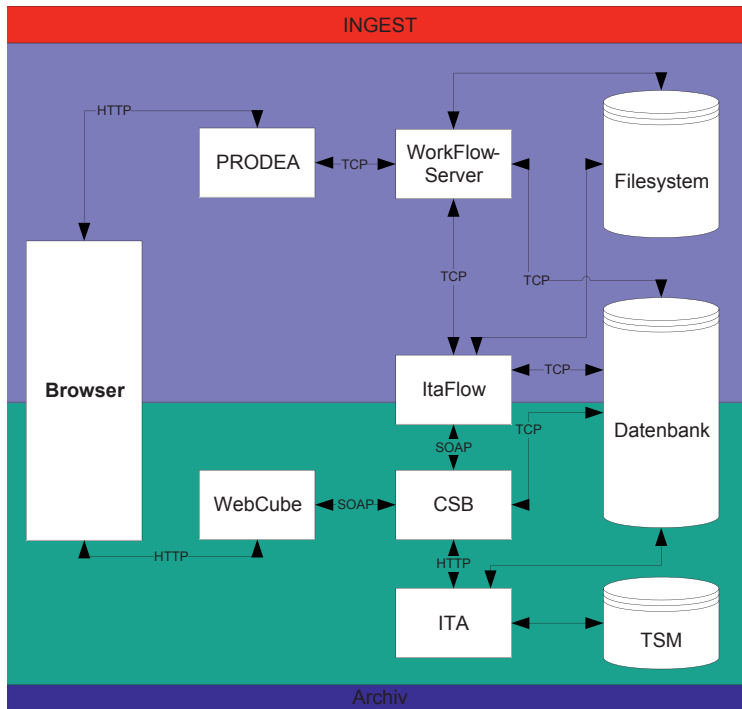


Abb. 1.: Die Architektur der Langzeitarchivierung im Stadtarchiv Stuttgart bis Herbst 2011.

Probleme des alten Speichers

Im Lauf der Zeit haben wir verschiedene Probleme mit unserem Langzeitspeicher festgestellt:

- die mangelnde Geschwindigkeit beim Zugriff
- der Zugriff war unter bestimmten Bedingungen gar nicht möglich
- die Beschränkung der AIP-Größe.

Unser Fokus lag vor sechs Jahren ganz auf der Sicherung unseres digitalen Archivguts. Dass ein reiner Bandspeicher keine besonders komfortable Nutzung zulässt, war uns von Anfang an bewusst. Ein Langzeitarchiv braucht auch nicht die Zugriffsgeschwindigkeiten, die beispielsweise ein Dokumentenmanagementsystem braucht. Aber die Aushebezeiten für digitales Archivgut müssen ebenso wie für analoges Archivgut planbar und zuverlässig sein. Wenn nun das Herunterladen von sehr großen AIPs oder einer größeren Menge von AIPs mehrere Stunden dauern kann, wird es mit der Nutzung doch schwierig. Der Zugriff war für uns gar nicht möglich, wenn auf dem betreffenden Magnetband ein Archivierungsprozess stattfand. Die Bandbibliothek wurde nicht nur für unsere Archivmagnetbänder, sondern auch bzw. insbesondere für das Backup genutzt, und wenn

sehr viele Anfragen an die IuK gingen, ein Restore vom Backup durchzuführen, dann waren auch mal alle Leseköpfe besetzt und wir konnten in dieser Zeit nicht an unsere Daten.

Die Formatmigration großer Datenmengen ist vom Magnetband aus nicht durchführbar. Wenn beispielsweise das TIFF-Format einmal obsolet zu werden droht und das Stadtarchiv Stuttgart alle TIFF-Dateien vom Speicher herunterholen muss, sie in ein Folgeformat konvertieren, neu durch den Ingest schicken und als neue Repräsentation ablegen muss, ist das mit einer reinen Magnetbandspeicherstruktur kaum darstellbar.

Ein weiterer Punkt war, dass der ITA-Server eine Beschränkung der AIP-Größe auf 2 GB mit sich brachte. Es ist zwar nicht unser Ziel, standardmäßig AIPs mit einer Größe von mehr als 2 GB zu erzeugen. Aber wir möchten die Möglichkeit dazu haben, und in bestimmten Fällen scheint es uns sinnvoll.

Anforderungen an eine neue Speicherinfrastruktur

Ein Wechsel auf einen anderen Speicher war also angezeigt. Der neue Speicher sollte folgende Anforderungen erfüllen:

Akzeptable Geschwindigkeit beim Zugriff

- Wir haben hierfür keine absolute Zeitvorgabe definiert. Das Herunterladen auch von größeren AIPs sollte sich jedoch im Minutenbereich bewegen.

Auf alle archivierten Objekte muss jederzeit zugegriffen werden können.

- Es darf nicht passieren, dass im regulären Betrieb Archivgut für uns nicht zugänglich ist, wie es früher der Fall war, wenn auf einem Magnetband ein Speicherprozess stattfand oder wenn einmal alle Leseköpfe besetzt waren.

AIPs mit einer Größe über 2 GB sollen möglich sein.

- Diese Forderung bedeutet, dass die neue Lösung ohne ITA-Server umgesetzt werden sollte.

Redundante Speicherung

- Unser digitales Archivgut muss doppelt vorhanden sein, die Kopien müssen in unterschiedlichen Gebäuden vorgehalten werden.

Integritätsschutz (WORM, CAS, ...) mindestens einer Kopie

- Um absichtliche oder unabsichtliche Veränderungen am Archivgut zu verhindern, muss eine oder alle der redundanten Kopien einen Integritätsschutz wie beispielsweise WORM oder CAS⁴ aufweisen.

⁴ WORM steht für „Write Once Read Many“, CAS für „Content Addressed Storage“. Die Nichtveränderbarkeit der Daten wird auf unterschiedliche technische Weise sichergestellt.

Von dieser Kopie aus muss Formatmigration perspektivisch möglich sein.

- Die Grundlage für künftige Migrationen muss die Kopie mit Integritätsschutz sein.

Archivierungsprotokoll

- Wenn wir neues Archivgut durch den Ingest auf den Speicher gebracht haben, benötigen wir eine Rückmeldung, dass das Archivgut dort angekommen ist.

Die neue Speicherinfrastruktur

Wie schon im Jahr 2006 traf es sich, dass unsere Wünsche – zumindest teilweise – mit dem zusammentrafen, was die Abteilung Information und Kommunikation (IuK) der Landeshauptstadt Stuttgart ohnehin plante. Aus verschiedenen Stellen der Stadtverwaltung war der Wunsch nach einem so genannten *revisionssicheren* Speicher an die IuK herangetragen worden. Die IuK der Landeshauptstadt Stuttgart hatte daher bereits die Anschaffung eines CAS ins Auge gefasst, als wir uns im Jahr 2011 mit unserem Problem an sie wandten. Sie hat dann das CAS-System *Silent Cubes* beschafft. An zwei Standorten in der Stadtverwaltung stehen die Silent Cubes, die jeweils über mehrere Speichereinheiten verfügen. Jede Speichereinheit besteht aus 12 Disks, von denen bis zu vier gleichzeitig ausfallen können, ohne dass ein Datenverlust entsteht. Sollte ein Standort oder die Verbindung zu einem Standort ausfallen, greifen wir auf den zweiten Standort zu. Die Daten werden mit einem internen Prozess täglich bitgenau überprüft. Auf den Silent Cubes ist das Stadtarchiv ein Mandant, wir verfügen über Platz für 10 TB Daten auf virtuellen Volumes. Von archivischer Seite aus waren wenige Konfigurationsentscheidungen notwendig: Silent Cubes sieht verschiedene Arten von Volumes vor: einfaches WORM, WORM mit Versionierung, WORM mit Lebensdauer und verschiedene Kombinationen dieser Parameter. Die virtuellen Volumes des Stadtarchivs sind einfache WORM-Speicher ohne Versionierung und ohne Begrenzung der Lebensdauer der Daten. In PADUA konnte nun der ITA-Server wegfallen, der CSB wurde von einem ITA-Storageadapter auf einen Filesystem-Storageadapter umgestellt.

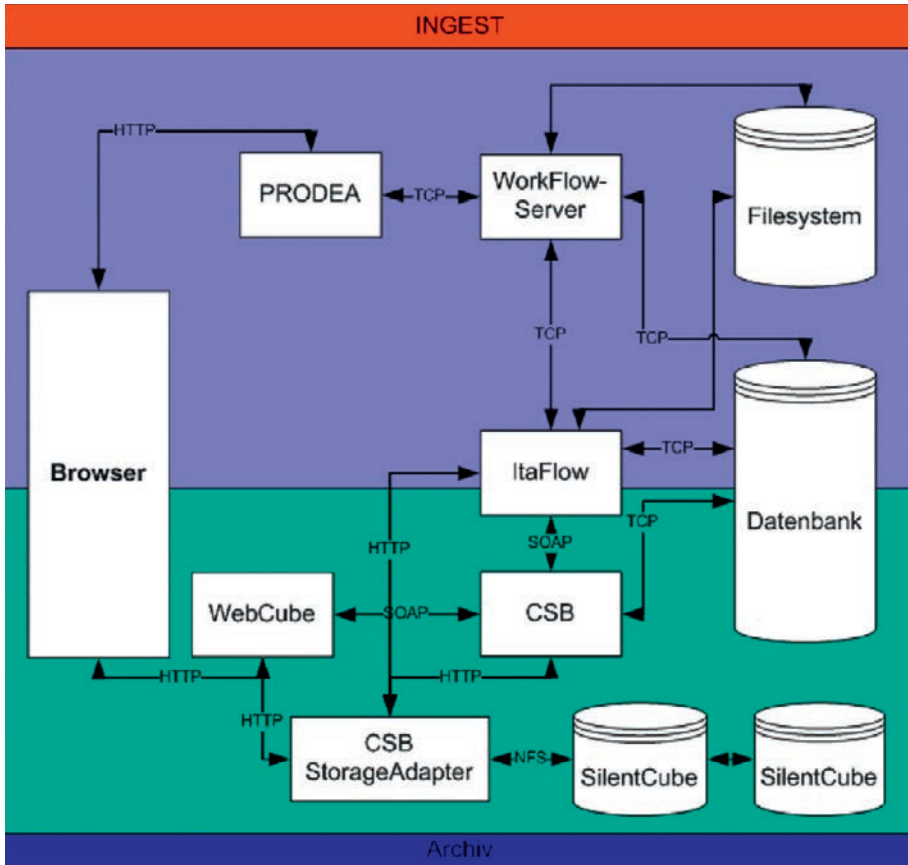


Abb. 2.: Die Architektur der Langzeitar Archivierung im Stadtarchiv Stuttgart seit Herbst 2011.

Zugriff auf Silent Cubes haben wir über die Such- und Zugriffsfunktion von PADUA, den Webclient Webcube.⁵

Verlauf der Migration

PADUA ist ein sehr flexibles System, was die Anbindung von Speichern angeht: eine simple Umkonfigurierung der Storage-Adapter bindet ein anderes Speichersystem an. Diese Umkonfigurierung musste im Produktivsystem und im Testsystem vorgenommen werden, so dass zu

⁵ Alternativ besteht die Möglichkeit, über einen CIF-Share direkt auf die Platten zu schauen. Die Daten sind dort nach Archivierungsdatum abgelegt, es ist also kein Blick auf die Daten, wie man ihn von seinem Explorer aus gewohnt ist.

archivierende Daten nicht mehr über den ITA-Server zum TSM liefern, sondern vom CSB aus direkt auf Silent Cubes. Das geschah am 14. und 15. November 2011, und schon am 15. November konnten wir neue Lieferungen mit PADUA durch den Ingest schicken und auf Silent Cubes ablegen. Die Anbindung war also problemlos erfolgt.

Nun mussten noch 3698 AIPs vom Magnetband auf die neuen Platten übertragen werden.⁶ Die Übertragung des Archivguts sollte fast vollautomatisch laufen: Der CSB verfügt über einen Replikationsmechanismus, mittels dessen die AIPs direkt von LTO auf Silent Cubes übertragen werden sollten. Die Integrität der AIPs wird dabei über Hashwerte geprüft.

Eine Datenmigration von einem Speicher auf einen anderen ist in der Welt der Datenverarbeitung vielleicht nicht eine tägliche Routine, aber doch ein immer wieder vorkommendes Ereignis, und eigentlich hätte gar nichts Unerwartetes passieren dürfen.

Die Übertragung der AIPs gestaltete sich aber nicht so problemlos. Immer wieder traten Timeouts auf und unterbrachen den Übertragungsprozess. Verschiedene Maßnahmen wie ein Update der CSB-Version brachten nicht den gewünschten Erfolg. Schließlich wurde entschieden, Festplatten – normale Festplatten, kein CAS – zwischen die Magnetbänder und die Silent Cubes zu schieben, um die Daten zunächst von den Bändern auf diese Festplatten zu holen und von diesen Festplatten dann auf die Silent Cubes zu replizieren. Diese temporären Platten wurden Mitte Dezember installiert.

Von den Magnetbändern ließen sich die AIPs damit problemlos herunterholen. Die Replizierung von den temporären Platten auf die Silent Cubes stieß jedoch weiterhin auf Probleme. Die AIPs wurden in Paketen repliziert, der Schreibvorgang auf Silent Cubes wurde initiiert, aber in unregelmäßigen Abständen lief er auf Exceptions, d.h. der Zugriff auf das Dateisystem wurde verweigert und der betreffende Schreibvorgang abgebrochen. Diese Art von Fehler war auch für die Kollegen von den Firmen HP und SER eine Premiere: dass einem Prozess während des Schreibens ins Filesystem die Rechte dafür abhandenkommen, war ihnen noch nicht passiert. Normalerweise hat ein Prozess die Berechtigung, Daten zu schreiben, oder er hat sie nicht.

Nach intensiver Diagnostik wurde ein Timeout in Silent Cubes als Problem identifiziert. Wenn eine Datei auf ein CAS geschrieben wird, darf diese Datei während des Schreibprozesses noch nicht schreibgeschützt sein, sondern erst, wenn sie komplett auf das Medium geschrieben wurde. Silent Cubes löst das über eine Zeitbegrenzung: wenn eine gewisse Zeit (und diese Zeit ist konfigurierbar) nichts geschrieben wird, wird die Datei als fertig angesehen und mit einem Schreibschutz versehen. Bei der Replikation konnte es aber durchaus einmal passieren, dass über einen kurzen Zeitraum kein Schreiben auf Silent Cubes erfolgte und damit der voreingestellte Zeitraum überschritten wurde. Als der Prozess dann weiterschreiben wollte, blockte Silent Cubes das ab, weil es das als unzulässigen Manipulationsversuch an existierenden, geschützten Dateien ansah. Als das Problem einmal erkannt war, war es schnell gelöst: dieser Timeout wurde von 60 Sekunden auf 10 Minuten hochgesetzt und dann konnte die Replikation ohne weitere Probleme von statten gehen.

⁶ Das Stadtarchiv Stuttgart verfügt über wesentlich mehr Archivgut, es waren aber noch nicht alle Lieferungen durch den Ingest gelaufen. Die Lieferungen, die noch nicht durch den Ingest gelaufen sind, waren in die Speichermigration natürlich nicht einbezogen. Sie sind auf Magnetband doppelt gesichert und werden von uns nach und nach durch den Ingest auf den neuen Speicher gebracht.

Ein einziges AIP widersetzte sich allen Versuchen, es von den Zwischenfestplatten auf Silent Cubes zu replizieren. Ein Vergleich der Kopie des AIPs auf den Zwischenfestplatten und der ursprünglichen Kopie auf Magnetband zeigte, dass dieses AIP fehlerhaft auf die Zwischenfestplatten kopiert worden war. In der Ausgangskopie befanden sich vier WAV-Dateien und unsere Metadaten-XML-Datei LHSArchiv⁷, die Kopie auf Festplatte enthielt nur eine WAV-Datei und kein XML.

Das fehlerhafte AIP zeigte, und dies ist positiv zu werten, dass die automatische Qualitätskontrolle durch den Replikationsmechanismus des CSB funktionierte: Der CSB weigerte sich zu Recht, dieses AIP auf die Silent Cubes zu spielen, weil die Datei beschädigt war. Negativ ist allerdings, dass nicht festgestellt werden konnte, wodurch die Kopie beschädigt worden war. Das AIP wurde erneut vom Magnetband heruntergeholt und die fehlerfreie Kopie auf Silent Cubes gebracht.

Ergebnis der Speichermigration

Inwieweit sind unsere Anforderungen an einen neuen Speicher nun erfüllt worden?

Akzeptable Geschwindigkeit beim Zugriff

Das Herunterladen auch großer Dateien erfolgt jetzt innerhalb von Sekunden oder Minuten.

Auf alle archivierten Objekte muss jederzeit zugegriffen werden können.

Der Zugriff ist für uns jederzeit möglich.

AIPs mit einer Größe über 2 GB sollen möglich sein

Es gibt keine technische Komponente mehr, die die AIP-Größe auf 2 GB beschränkt.

Redundante Speicherung, Integritätsschutz

Das CAS-System Silent Cubes und die Art und Weise, in der es von der IuK der Landeshauptstadt Stuttgart betrieben wird, erfüllen diese Anforderungen.

Von der Kopie mit Integritätsschutz aus muss Formatmigration perspektivisch möglich sein

Beide Kopien haben die gleiche Qualität; die Geschwindigkeit und Handhabung der Silent Cubes machen eine Formatmigration technisch möglich.

Archivierungsprotokoll

Das Stadtarchiv erhält nach Beendigung eines Ingest-Prozesses eine E-Mail über die vollständige Übergabe aller AIPs an den Archivierungsprozess, der sie auf Silent Cubes bringt. Nach der

⁷ Vgl. zum Metadatenchema LHSArchiv den Beitrag von Katharina Ernst und Heike Maier: PADUA. Produktivsystem für die Archivierung digitaler Unterlagen im (Stadt)-Archiv Stuttgart. In: Entwicklung in den Bereichen Records Management/Vorarchiv – Übernahme – Langzeitarchivierung. Dreizehnte Tagung des Arbeitskreises "Archivierung von Unterlagen aus digitalen Systemen" vom 27./28. April. St. Gallen 2009. S. 89–97, hier S. 95f.

Speicherung der AIPs auf den Silent Cubes prüft ein Prozess *Check.Archive*, ob alle AIPs auf dem CSB vorhanden sind und abgerufen werden können. Diese Prüfung dient dazu, die Vollständigkeit der Lieferung im Archiv zu garantieren. Sie wird nicht mehr gesondert protokolliert. Die Anforderung *Archivierungsprotokoll* wird also durch unsere neue Infrastruktur nur teilweise erfüllt. Die Übergabe der AIPs an den Archivierungsprozess wird protokolliert, nicht aber die Speicherung auf Silent Cubes. Alle anderen von uns formulierten Anforderungen sind erfüllt worden.

Nachteil eines CAS ist, dass es nicht nur das Archivgut vor Löschungen schützt, sondern auch den Datenmüll, den wir uns während der Speichermigration eingehandelt haben, denn die abgebrochenen Schreibversuche haben eine substanzielle Menge an Daten hinterlassen: ca. 130 GB Datenschrott befindet sich auf dem Langzeitspeicher. So routinemäßig die Datenmigration vorab also auch schien, wäre im Nachhinein eine Teststellung sinnvoll gewesen, bei der man das Problem dieses Timeouts hätte identifizieren können.

Dem Stadtarchiv Stuttgart ermöglicht der neue Speicher durch den schnellen Zugriff, digitales Archivgut zeitnah bereitzustellen. Es hat die Voraussetzungen dafür geschaffen, ein Konvertierungsmodul zu implementieren und in die Konzeption der Nutzung unseres digitalen Archivguts einzusteigen.

Einführung der elektronischen Archivierung im Sächsischen Staatsarchiv

Von BURKHARD NOLTE und KARSTEN HUTH

Das Projekt *Langzeitspeicherung und elektronische Archivierung* (LeA) ist weit fortgeschritten. Nach dem derzeit gültigen Projektplan ist davon auszugehen, dass bereits Ende des Jahres 2012 elektronische Akten, die nicht mehr ständig benötigt werden, in das System zur Langzeitspeicherung eingestellt und Daten der sächsischen Verwaltung, die bleibenden Wert besitzen, im elektronischen Staatsarchiv (el_sta) archiviert werden. Zur Organisation des el_sta wurde ein eigenes Konzept erstellt, das die unterschiedlichen Tätigkeiten während des Betriebs unter drei verschiedenen Rollen aufteilt. Ziel ist eine möglichst ökonomische Aufteilung aller anfallenden Aufgaben und eine Bündelung aller relevanten Informationen an einer Stelle innerhalb des el_sta. Damit ist letztlich sichergestellt, dass auch künftige Generationen sich ihr eigenes Bild von ihrer (elektronischen) Vergangenheit machen können.

Einleitung

Die Verwaltung arbeitet zunehmend elektronisch: sie tauscht E-Mails aus, führt elektronische Akten, speichert Informationen in Datenbanken. Dadurch wird für den Bürger, der diese Daten zunehmend auch im Internet abrufen kann, vieles leichter, und die Verwaltung arbeitet schneller und effizienter. Die Daten werden gespeichert, solange sie ständig benötigt werden. Aber was passiert dann?

Die funktionelle Ausgestaltung der elektronischen Bearbeitung und Aktenführung ist nicht zu trennen vom Problem der Langzeitspeicherung und elektronischen Archivierung. Nur durch die revisionssichere Aufbewahrung elektronischer Unterlagen kann die Vollständigkeit und Rechtskonformität mit wirtschaftlich vertretbarem Aufwand gewährleistet werden.

Um den elektronischen Gedächtnisverlust zu verhindern, hat das Kabinett im Freistaat Sachsen mit Beschluss vom 30. September 2008 zur weiteren Einführung von IT-gestützter Vorgangsbearbeitung, einschließlich elektronischer Aktenführung, Langzeitspeicherung und elektronischer Archivierung in der Sächsischen Staatsverwaltung das Projekt LeA ins Leben gerufen.

Ziel des Projekts ist der Aufbau und der Betrieb eines landesweit einheitlichen Systems zur Langzeitspeicherung und elektronischen Archivierung. Die Verwaltung soll hierdurch in die Lage versetzt werden, auch elektronische Unterlagen rechtssicher aufzubewahren und nach Ablauf der Aufbewahrungsfrist an das Sächsische Staatsarchiv auszusondern. Dieses soll die archivwürdigen Unterlagen dauerhaft elektronisch vorhalten. Die hierzu erstellten Konzepte können im Internet abgerufen werden.¹

¹ Siehe <http://www.archiv.sachsen.de/7447.htm>. Alle Links wurden am 24. September 2012 überprüft.

Teilprojekt Langzeitspeicherung

Im Hinblick auf das aufzubauende System zur Langzeitspeicherung (SLZS) ist die Erarbeitung des Pflichtenhefts für das Aussonderungsmodul für Unterlagen aus dem landeseinheitlichen Dokumentenmanagement- und Vorgangsbearbeitungssystem (DMS/VBS) *eV.A.S.AX* und für die Speicherverwaltung fast abgeschlossen.

Das Aussonderungsmodul ist die zentrale Steuerungseinheit zur Durchführung einer automatisierten Anbietung von elektronischen Unterlagen und zur automatisierten Abgabe der archivwürdigen elektronischen Unterlagen an das *el_sta*. Die durchgehende Automatisierung des Prozesses gewährleistet den größtmöglichen Schutz vor

- einem versehentlichen Überschreiten der Aufbewahrungsfristen,
- einem versehentlichen Verletzen der Integrität, Authentizität und Vollständigkeit der elektronischen Unterlagen und
- dem versehentlichen Versenden von nicht archivwürdigen elektronischen Unterlagen.

Zur vollständigen Dokumentation der Aussonderung legt das Modul alle gesendeten und empfangenen Nachrichten in einem extra dafür angelegten Vorgang im DMS/VBS ab. Dadurch wird nachvollziehbar, welche elektronischen Vorgänge an das *el_sta* abgegeben wurden und welche elektronischen Vorgänge nach Ablauf der Aufbewahrungsfrist gelöscht wurden. Zudem gibt der Vorgang Auskunft darüber, wann exakt welcher Schritt des Aussonderungsprozesses eingeleitet und durchgeführt wurde.

Parallel hierzu läuft die Erstellung des Pflichtenhefts für das Konvertierungsmodul. Im Rahmen der Bearbeitung einer Datei kann es im VBS/DMS – manuell und automatisch – zur Erzeugung von Versionen dieser Datei kommen. Fraglich ist, inwieweit diese Dateiversionen vor dem Hintergrund der Sicherstellung der Nachvollziehbarkeit des Verwaltungshandelns ebenfalls in das Langzeitspeicherformat überführt werden sollen. An die Frage der Konvertierung von Datei-Versionen schließt sich die Frage nach dem Verbleib der Dateien im Ausgangsformat nach der Konvertierung an. Diese sollten noch eine zeitlang im Ermessen der jeweiligen Behörde aufbewahrt werden, um bei Bedarf weiter mit dem bisherigen Bearbeitungswerkzeug darauf zugreifen zu können. Eine erneute Bearbeitung bzw. Veränderung der Dateien ist aber nicht möglich. Innerhalb des VBS/DMS sollte konfigurierbar sein, ob dem Sächsischen Staatsarchiv die Dateien im Ausgangsformat und die Dateien in einem langzeitspeicherfähigen Format angeboten werden. Im Ergebnis einer Überprüfung der möglichen Konvertierungsvarianten unter Berücksichtigung der Prämissen der Nachvollziehbarkeit des Verwaltungshandelns, der Datensparsamkeit und der derzeitigen technischen Realisierbarkeit soll die nunmehr vorgesehene Konvertierungsplattform eine Wandlung ausschließlich unterschiedlicher Versionen einer Datei bei gleichzeitigem Verbleib aller Versionen einer Datei im Ausgangsformat im SLZS bis zum Ablauf der Aufbewahrungsfrist vornehmen (vgl. Abb. 1):

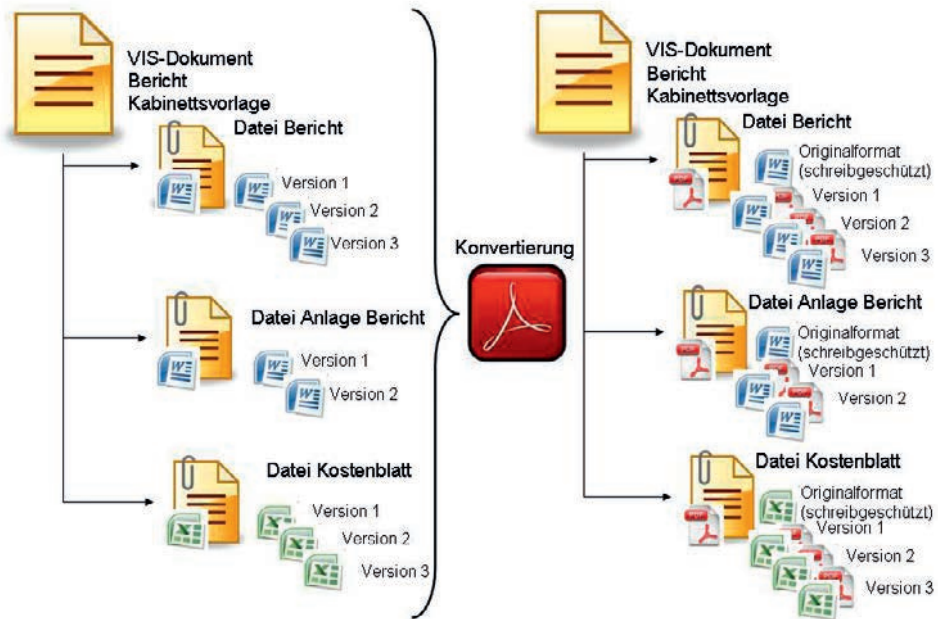


Abb. 1: Formatkonvertierung nach Dokumentenabschluss.²

Die Konvertierungsplattform soll als zentraler Dienst nicht nur der Langzeitspeicherung, sondern mittels Schnittstellen ebenfalls weiteren IT-Verfahren des Freistaates Sachsen zur Verfügung stehen. Eine manuelle Nutzung z. B. per Webseite soll ebenfalls realisiert werden. Der aktuelle Entwicklungsstand der Konvertierungsplattform des el_sta ist bei der Umsetzung einer Konvertierungsplattform für das SLZS zu berücksichtigen. Beide Konvertierungsplattformen sind möglichst einheitlich (Software, Hardware, Schnittstellen) in ein IT-System zu integrieren.

Bereits bestehende Module zur Speicherverwaltung sowie zur Aussonderung sind dementsprechend zu erweitern bzw. anzupassen. Nach Freigabe der Pflichtenhefte erfolgt dann die abschließende Fassung des Organisationskonzepts.

² Entwurf: M. Schenk / IMTB Consulting GmbH.

Teilprojekt Elektronische Archivierung

Mit Blick auf den Aufbau des el_sta wurde im vergangenen Jahr das europaweite Vergabeverfahren zur Beschaffung der hierzu notwendigen Komponenten erfolgreich abgeschlossen. Gegenstand der Ausschreibung war die technische Infrastruktur des el_sta und konkret die folgenden Module in einem Gesamtpaket:

- Übernahme
- Transformation
- Konvertierung
- Datenmanagement
- eMagazin
- Nutzung
- Bestandserhaltung
- Leitstand und
- Administration.

Der Betrieb des el_sta wird beim Staatsbetrieb Sächsische Informatik Dienste (SID) erfolgen, genutzt und fachlich geleitet wird das el_sta vom Sächsischen Staatsarchiv. Die Nutzung für andere Organe des Freistaates Sachsen soll möglich sein. Darüber hinaus ist geplant, optional die Nutzung grundsätzlich auf die zuständigen Archive der kommunalen Gebietskörperschaften im Freistaat Sachsen (470 Städte und Gemeinden sowie 10 Landkreise) zu erweitern. Deshalb müssen alle Module, die mit einem Client verbunden sind, grundsätzlich mandantenfähig sein.

Die Infrastruktur des el_sta wird so ausgerichtet sein, dass unterschiedliche Typen von elektronischem Archivgut in möglichst einheitlicher Form verarbeitet und archiviert werden können. Elektronisches Archivgut wird daher nach festgelegten Standards in einheitliche Archivinformationspakete (AIP) transformiert, die mit der ISO-Norm 14721 konform gehen. Der Transformationsprozess wird genauestens dokumentiert und stellt einen wesentlichen Aufwand bei der Übernahme dar. Dem ursprünglichen Zweck und der Authentizität des elektronischen Archivgutes wird durch festgelegte archivistische Verfahrensweisen entsprochen. Darüber hinaus findet elektronische Archivierung ausschließlich innerhalb der Infrastruktur des el_sta statt. Es wird daher keine Insellösungen für „besondere“ Arten von elektronischem Archivgut geben.

Die Zuschlagserteilung für die zum Aufbau des el_sta notwendigen Komponenten erfolgte Anfang Dezember 2011. Die technische Lösung haben die Firma T-Systems International GmbH, die Schweizer Firma scope solutions ag und die Firma H&T Greenline GmbH entwickelt. Der Startschuss für die Auftaktveranstaltung fiel am 13. Januar 2012 in Anwesenheit des Sächsischen Staatsministers der Justiz und für Europa, Herrn Dr. Martens, und des Chief Information Officer (CIO) des Freistaates Sachsen, Staatssekretär Herrn Dr. Bernhardt. Der Testserver ist installiert und die Erstellung des Pflichtenheftes ist fast abgeschlossen.

Das Projekt ist insgesamt demnach weit fortgeschritten (s. Abb. 2). Im Ergebnis werden gegen Ende des Jahres 2012 ein System zur Langzeitspeicherung und ein elektronisches Archiv für den Freistaat Sachsen in Betrieb genommen werden können.

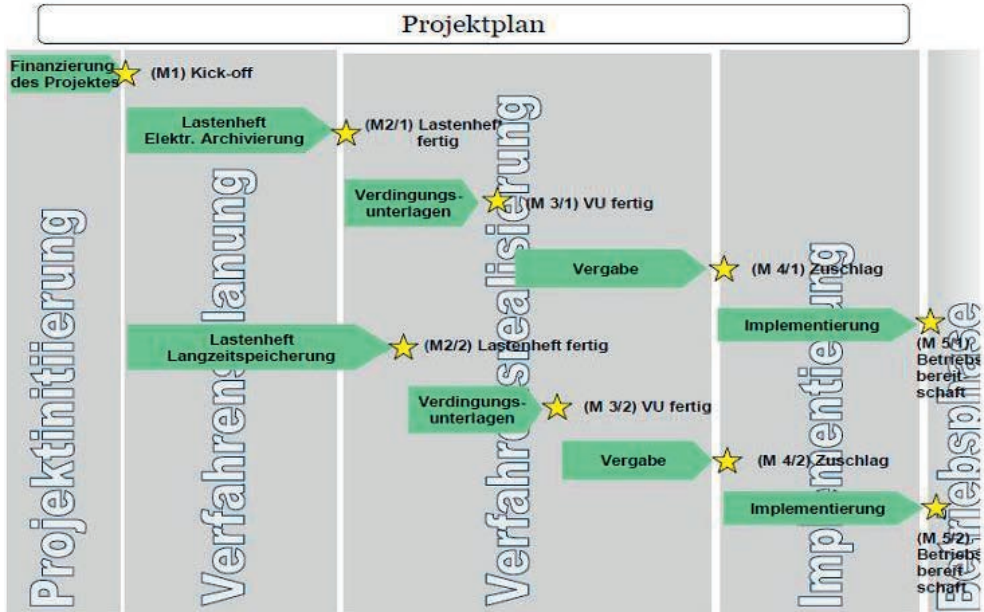


Abb. 2: Projektplan (Stand: 12. März 2012).³

Organisation und Technik

Ohne die geeignete technische Infrastruktur ist der Betrieb eines elektronischen Archivs unmöglich. Deshalb stand im vergangenen Jahrzehnt die Analyse der Anforderungen an eine geeignete Infrastruktur sowie die Beschaffung und Entwicklung einer solchen im Vordergrund. Die Definition des organisatorischen Rahmens wurde vor Allem in organisationsübergreifenden Arbeitsgruppen angegangen.⁴ Inzwischen hat sich der Markt für OAIS-konforme Archivsysteme erweitert und eine Hand voll Produkte hervorgebracht, die die Anforderungen weitestgehend erfüllen. Natürlich bedeutet die Anpassung eines Produktes an die speziellen Anforderungen eines größeren Archivs immer noch einen hohen Aufwand für ein Projekt (z.B. Implementierung eines xdomea-Aussonderungsmoduls⁵), dennoch bleibt spürbar mehr Zeit für die Planung der Organisationsstrukturen übrig.

Der Aufbau des el_sta im Freistaat Sachsen beruht auf einem Fachkonzept, in dem alle technischen Anforderungen und Prozesse niedergeschrieben sind und einem Organisations-

³ Entwurf: B. Nolte / IMTB Consulting GmbH.

⁴ Siehe nestor Arbeitsgruppen unter <http://www.langzeitarchivierung.de>.

⁵ xdomea 2.1.0 Spezifikation. XÖV Standard für den IT-gestützten Austausch und die IT-gestützte Aussonderung behördlichen Schriftgutes. 2009, S. 40 ff.

konzept,⁶ welches die unterschiedlichen Rollen und Tätigkeiten zum Betrieb des el_sta festlegt. Das Organisationskonzept bezieht sich auf die Zeit nach der Beendigung des Projekts. Die Organisationsstrukturen eines Projekts zur Einführung der elektronischen Archivierung sind zeitlich begrenzt auf ein bestimmtes Ziel hin ausgerichtet. Solche Strukturen sind grundsätzlich nicht dazu geeignet, Daueraufgaben zu gewährleisten. Auch darf man nicht dem Glauben verfallen, dass eine hervorragend funktionierende technische Plattform die Archivierung quasi von alleine regelt. Ausgehend vom OAIS (ISO 14721:2003) sind sich alle grundlegenden Standards zur elektronischen Archivierung dahingehend einig, dass ein elektronisches Archiv nur aus einem Verbund von Technik und kompetenten Fachkräften bestehen kann.⁷ Und genauso, wie ein klassisches Archiv in Organisationseinheiten mit verschiedenen Kompetenzen und Verantwortungen aufgeteilt ist, muss auch ein elektronisches Archiv geplant sein. Vor Allem, wenn sich das elektronische Archiv in eine größere Organisation eingliedern soll.

Das Organisationskonzept des el_sta geht von drei unterschiedlichen Rollen zur Aufrechterhaltung eines dauerhaften Betriebs aus:

- Die archivtechnische Leitstelle
- Die klassischen Archivtätigkeiten
- Die technische Administration

Im Folgenden sollen die Rollen näher beschrieben werden.

Die archivtechnische Leitstelle

Die Leitstelle ist verantwortlich für die fachtechnischen Belange der elektronischen Archivierung, inklusive der digitalen Bestandserhaltung. Innerhalb des OAIS liegen die Kompetenzen der Leitstelle in den Bereichen der Administration und des Preservation Plannings (s. Abb. 3).⁸ Das Personal sollte sich einerseits im Archivwesen auskennen und andererseits über technische Fähigkeiten in Bezug auf Datenstrukturen und die Planung technischer Prozesse verfügen. Zudem sollte die Leitstelle fähig sein, den Mitarbeiterinnen und Mitarbeiter des Archivs die komplexen Vorgänge innerhalb der technischen Infrastruktur und die damit verbundenen Konsequenzen für das Archivgut verständlich darzustellen.

⁶ Zur Veröffentlichung der Konzepte siehe Anm. 1.

⁷ Vgl. die Begriffsdefinition für „digitales Langzeitarchiv“ in der DIN 31644 - Information und Dokumentation - Kriterien für vertrauenswürdige digitale Langzeitarchive und der DIN 31645 - Information und Dokumentation - Leitfaden zur Informationsübernahme in digitale Langzeitarchive.

⁸ Siehe Reference Model for an Open Archival Information System (OAIS) – 2002, S. (4-10) und (4-13).

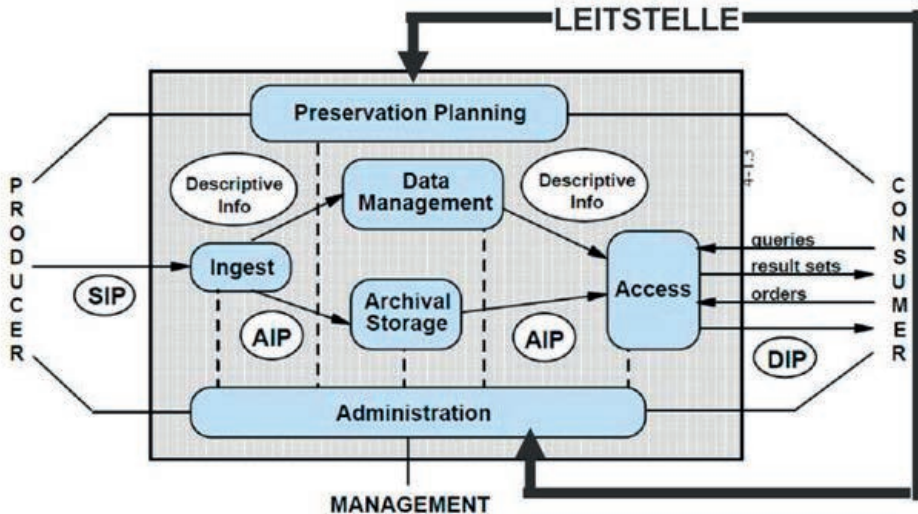


Abb. 3: Leitstelle im Verhältnis zum OAIS Functional Model.

Weitere Aufgaben der Leitstelle liegen in der Bündelung von relevanten Informationen und in der Koordinierung der Auslastung der technischen Plattform. Insbesondere die Planungen von Übernahmen durch das Ingest-Modul sind sehr aufwendig, da der Ingest konform zur DIN 31645 durchgeführt werden soll. Damit fallen für die Leitstelle im Rahmen einer ersten Übernahme von Daten aus einem Fachverfahren nach DIN Norm mindestens folgende Arbeitsschritte an:⁹

- Analyse des exportierten Datenformats (Prüfung der Archivfähigkeit des Formats),
- Eventuell Festlegung von Maßnahmen zur Migration in ein archivfähiges Datenformat,
- Festlegung und Beschaffung der notwendigen Metadaten zur späteren Identifizierung des Archivguts,
- Definition der SIP und Vereinbarung eines Weges zum Datentransfer vom Produzenten zum Archiv und
- Durchführung eines Testlaufs zum geplanten Ingest.

Die Einleitung des technischen Prozesses zum Ingest obliegt der Leitstelle. Einzelne Prozessschritte können, je nach Größe und Komplexität der zu Übernehmenden Informationen, die technischen Ressourcen mehr oder weniger stark auslasten. SIP, die mit relativ einfach strukturierten Datentabellen befüllt sind, können innerhalb weniger Minuten technisch überführt werden. SIP, die randvoll mit hochauflösenden Luftaufnahmen befüllt sind, können die techni-

⁹ Siehe DIN 31645:2011 S. 19–21.

sche Plattform hingegen für mehrere Stunden auslasten. Deshalb ist eine zentrale Planung mit zentraler Zugriffsverwaltung auf das Ingest-Modul notwendig, um eine versehentliche Überlastung der technischen Ressourcen zu vermeiden.

Das Sächsische Staatsarchiv ist in fünf Abteilungen organisiert und über fünf Standorte im Freistaat verteilt. Theoretisch kann in allen Abteilungen wertvolles Wissen über anbieterpflichtige Fachverfahren und elektronische Unterlagen vorhanden sein. Es wird Aufgabe der Leitstelle sein, das verteilte Wissen zu bündeln, um es dann strategisch einzusetzen, z. B. in Form einer Informationslandkarte für die Behörden des Freistaats, die vorhandene Verfahren in Bezug auf ihre Archivwürdigkeit darstellt.

Die klassischen Archivtätigkeiten

Die klassischen archivarisches Tätigkeiten werden auch bei der elektronischen Archivierung von den entsprechenden Fachabteilungen durchgeführt. Es ist wichtig darauf hinzuweisen, dass die Bewertung der Archivwürdigkeit allein den Mitarbeiterinnen und Mitarbeitern der Fachabteilungen obliegt. Die Leitstelle tritt nur beratend bei technischen Fragen und Fragen zur Archivfähigkeit der zu übernehmenden Informationen hinzu. Die letzte Entscheidung, ob und in welcher Form Informationen übernommen werden, wird immer von der verantwortlichen Fachabteilung getroffen.

Auch die Erschließung und die spätere Vermittlung des elektronischen Archivguts fallen in den Bereich der Fachabteilungen. Anders als beim Ingest-Modul wird der Zugriff auf das Datenmanagement und die elektronischen Bestände im elektronischen Magazin nicht von der Leitstelle koordiniert. Alle Mitarbeiterinnen und Mitarbeiter mit den entsprechenden Befugnissen können selbständig von ihrem Arbeitsplatz aus in ihren elektronischen Beständen recherchieren.

Die technische Administration

Die technische Infrastruktur mit allen notwendigen Servern und Speichermedien wird außerhalb des Sächsischen Staatsarchivs beim SID beherbergt. Alle technischen Wartungsarbeiten an der Hardware und der Software, die nicht unmittelbar das elektronische Archivgut oder die Archivierungsprozesse berühren, werden dort, natürlich in Absprache mit der Leitstelle, durchgeführt. Folgende Aufgaben sollen von der technischen Administration übernommen werden:

- Stromversorgung, Kühlung, Sicherheit
- Einspielen von Updates und Patches
- Reparatur und Austausch beschädigter Hardwarekomponenten

Informationsaustausch und organisatorische Prozesse

Das Organisationskonzept des el_sta kann nur erfolgreich umgesetzt werden, wenn die unterschiedlichen Arbeitsbereiche sich ausreichend untereinander abstimmen. Die Leitstelle spielt hier die Rolle einer Sammelstelle. Alle Fäden sollten hier zusammenlaufen. So muss die Leitstelle beispielsweise über Wartungsarbeiten an der technischen Plattform informiert sein, um nicht versehentlich eine wichtige Übernahme innerhalb der Wartungszeit durchführen zu wollen. Um solche Probleme zu vermeiden, müssen die technische Administration und die Leitstelle ihre Planungen gegenseitig abstimmen.

Von kritischer Bedeutung für die spätere Qualität des elektronischen Archivgutes ist die Kommunikation zwischen der Leitstelle und dem archivfachlichen Bereich. Dies beginnt mit dem Prozess der Bewertung. Innerhalb des Teilprojekts „Elektronische Archivierung“ wurde ein organisatorischer Ablaufplan entworfen, der das Zusammenspiel der Archivabteilung mit der Leitstelle regelt. Eine Übernahme kann nur reibungslos gelingen, wenn alle Beteiligten jederzeit über den aktuellen Stand informiert sind und jeder weiß, ob er für den nächsten Schritt etwas beisteuern muss oder nicht. Außerdem muss die archivfachliche Seite darüber in Kenntnis gesetzt werden, in welcher Art und Weise die übernommenen Informationen im elektronischen Archiv dargestellt und genutzt werden können. Die Darstellung auf der Nutzungsplattform des elektronischen Archivs kann nämlich erheblich von dem Erscheinungsbild der Darstellung innerhalb des produzierenden Fachverfahrens abweichen. Zu einer vollständigen Bewertungsentscheidung gehört deshalb auch die Zustimmung der archivfachlichen Seite zur Darstellung einer elektronischen Archive innerhalb des Archivs.

Damit alle verantwortlichen Mitarbeiterinnen und Mitarbeiter des Staatsarchivs zum Betriebsbeginn des el_sta mit ihren Rollen vertraut sind, werden im laufenden Jahr 2012 insgesamt vier Fortbildungs- und Schulungsveranstaltungen durchgeführt.

Die elektronische Grundakte in G-DIMAG

Von ROLF LANG

Das Landesarchiv Baden-Württemberg wird die elektronischen Grundakten vom Tag ihrer Entstehung in seiner Software G-DIMAG speichern und dauerhaft erhalten. Der Beitrag beschreibt die konzeptionellen Vorarbeiten und Anpassungen von DIMAG zu G-DIMAG und endet mit einem kleinen *Rundgang* durch G-DIMAG.

Einführung

Das Land Baden-Württemberg übernimmt die alleinige Verantwortung für das Grundbuchwesen, das bisher teilweise unter kommunaler Verantwortung stand. Hierzu erfolgt eine Konzentration der weit über 600 bisherigen dezentralen Grundbuchämter auf nur noch 13 zukünftige Standorte bei den Amtsgerichten Achern, Böblingen, Emmendingen, Heilbronn, Mannheim, Maulbronn, Ravensburg, Schwäbisch-Gmünd, Sigmaringen, Tauberbischofsheim, Ulm, Villingen-Schwenningen und Waiblingen.

Die Grundakten und Grundbücher werden sich aber nicht dort befinden, sondern an einer zentralen Stelle. Für die klassischen analogen Akten ist dies das Grundbuchzentralarchiv in Kornwestheim, aufgebaut in einer Kooperation mit dem Landesarchiv Baden-Württemberg. Für die elektronischen Akten ist dies die Anwendung G-DIMAG (Grundbuch Digitales Magazin). Die Amtsgerichte haben hierzu einen elektronischen Zugang über das Landesverwaltungsnetz (LVN).

Seit März 2012 ist das Grundbuchzentralarchiv produktiv und übernimmt die Einlagerung der Papierakten. Das elektronische Grundaktenarchiv wird seit dem 2. Juli 2012 sukzessiv aufgebaut. Die Einlagerung der Akten erfolgt mit der Eingliederung des jeweiligen Grundbuchamtes. Erwartet wird ein Gesamtaufkommen von 168 000 laufenden Metern, Platz für 7 Mio. Akten mit geschätzten 550 Mio. Blättern. Der Migrationsprozess wird voraussichtlich noch bis Ende 2017 andauern.

Das für das Grundbuchwesen zuständige Justizministerium beauftragte das Landesarchiv sowohl mit der Aufgabe der zentralen Lagerung für die Grundakten als auch mit der Bereitstellung und Pflege eines digitalen Magazins für die elektronischen Grundakten. Diese Unterlagen verbleiben zunächst noch im Zuständigkeitsbereich der Justiz (Registraturgut) und werden später zu Archivgut. Das Landesarchiv baut dadurch seine Rolle als zentraler Dienstleister innerhalb der baden-württembergischen Landesverwaltung zum langfristigen Erhalt von digital gespeicherten Informationen und konventionellen Dokumenten weiter aus.

Elektronischer Rechtsverkehr

Eine wesentliche Voraussetzung war das Gesetz zur Einführung des elektronischen Rechtsverkehrs (ERV) und der elektronischen Akte (eAkte) im Grundbuchverfahren sowie zur Änderung weiterer grundbuch-, register- und kostenrechtlicher Vorschriften (ERVGBG).¹ Dies trat am 1. Oktober 2009 in Kraft.

Dieses Gesetz erlaubt es, Grundakten zu digitalisieren und als elektronische Dokumente anstelle von Papierakten aufzubewahren. Damit wurden die Voraussetzungen geschaffen, dass im Grundbuchverfahren Anträge und sonstige Dokumente rechtsverbindlich auf elektronischem Wege eingereicht werden können. Die Notare sollen verpflichtet werden, ausschließlich auf diesem Wege mit den Grundbuchämtern zu korrespondieren.

Mit der Auflösung eines bisherigen *alten* Grundbuchamts werden dessen Papierakten ins Grundbuchzentralarchiv gebracht. Zugleich werden die neu eingereichten Papierdokumente digitalisiert und zusammen mit den genuin digitalen Dokumenten in der elektronischen Grundakte geführt. Damit wird ein weiteres Anwachsen der Papierakten im Grundbuchzentralarchiv begrenzt. Das weitere Anwachsen erfolgt dafür elektronisch in G-DIMAG.

Eine Digitalisierung von alten Beständen wird nur nach Bedarf erfolgen. Die Umstellung aller Papierakten wurde aus wirtschaftlichen Gesichtspunkten nicht erwogen. Die Gründe hierfür waren:

1. Bei jeder Umstellung von Altakten kommt es bei vielen Einzeldokumenten zu einer digitalen Signierung. Diese erfordert unverhältnismäßig hohen Personaleinsatz von Urkundsbeamten, welche die Übereinstimmung des elektronischen Dokumentes bestätigen.
2. Des Weiteren beurteilte das Landesarchiv Baden-Württemberg die Grundakten als historisch wertvoll und archivwürdig.
3. Hinzu kommt auch, dass viele ältere Akten selten benötigt werden. Eine Erhebung zeigt, dass 50 % der Eintragungsvorgänge ohne Hinzuziehung der Grundakten erledigt werden können. Es wird damit gerechnet, dass bis 2017 dieser Anteil sogar auf 10 % fallen wird.

DIMAG und seine Anpassung an das Grundbuch

Seit Mitte 2006 betreibt das Landesarchiv das digitale Magazin DIMAG produktiv und sammelt erste eigene Erfahrungen mit der Übernahme, Erhaltung und Nutzung von archivwürdigen, rein digital entstandenen Unterlagen. Hierbei pflegen Archivmitarbeiter mittels einer webbasierten Benutzeroberfläche die Datenbestände in den elektronischen Speicher ein. Dabei werden archivrelevante Metadaten zum Zeitpunkt der Archivierung abgefragt beziehungsweise DIMAG ermittelt automatisch noch weitere systemrelevante Metadaten. Vermieden wird damit die provisorische Ablage ohne Metadaten Beschreibung.

¹ Gesetz v. 11.8.2009 BGBl. I S. 2713 (Nr. 53), in Geltung ab 1.10.2009.

Einige wesentliche Grundfunktionen von DIMAG sind:

1. Die Trennung von logischer und physischer Beschreibung. Da in der digitalen Welt die Information nicht mehr fest mit einem Datenträger verknüpft, ja selbst die Codierung bzw. das verwendete Format äußerst variabel sind, hat DIMAG das sogenannte Repräsentationen-Modell umgesetzt. Hierbei kann ein logisches Digitales Objekt mehrere digital codierte Erscheinungsformen haben. So wird ein Zugang als erste Repräsentation im derzeit akzeptierten Archiv-Format abgelegt. Das Landesarchiv geht davon aus, dass dieses Format in wenigen Jahren oder Jahrzehnten nicht mehr unterstützt wird. Damit die Information trotzdem noch genutzt werden kann, gibt es zwei derzeitige Lösungsansätze: Emulation oder Migration. Das Landesarchiv setzt auf die Migration, sprich es wird eine weitere Repräsentation mit identischen signifikanten Eigenschaften entstehen. DIMAG unterstützt beide Ansätze.
2. Die Erstellung einer elektronischen Prüfsumme (MD5)² für alle abgespeicherten Daten zum Zeitpunkt der Ablage. Damit kann der Nachweis der Unversehrtheit der Daten sowohl bei der späteren Nutzung als auch vor der Datensicherung geführt werden. MD5 ist ein Vertreter aus einer Reihe von (kryptologischen) Hashfunktionen, die von Ronald L. Rivest am Massachusetts Institute of Technology entwickelt wurden. Es wird damit kein eindeutiger Schlüssel für jede Datei erstellt. Daher ist es theoretisch möglich, zwei speziell präparierte Dateien zu erzeugen, die sich unterscheiden, aber dennoch denselben Hashwert ergeben. Da es aber ca. $3,4 \cdot 10^{38}$ verschiedene Schlüssel³ gibt, ist es extrem unwahrscheinlich, dass ein hoffentlich selten vorkommender Bit-Fehler in einer Datei zum gleichen MD5 Wert führt. Mit anderen Worten, MD5 ist nicht für elektronische Signaturen, aber noch sehr gut zur Sicherung der Dateiintegrität geeignet.

Die Datenhaltung der Primär- und der Metadaten erfolgt im Dateisystem. Dabei ist die Ablage so organisiert, dass die entsprechenden Dateien ganz eng benachbart beieinander liegen. Der Zusammenhang erschließt sich durch einen eindeutigen Identifier aID⁴, welcher für alle vier Primär- und Metadatendateien gemeinsam Verwendung findet. Die in XML Syntax abgelegten Metadaten erhalten den Zusatz im Dateinamen ‚.xml‘, die MD5 Werte jeweils den Zusatz ‚.md5‘. Damit entsteht diese Struktur:

<aID>	die Primärdaten Datei
<aID>.xml	die Metadaten Datei
<aID>.md5	die MD5 Prüfsumme der Primärdaten Datei
<aID>.xml.md5	die MD5 Prüfsumme der Metadaten Datei

² MD5 Message-Digest Algorithm, IETF RFC 1321, <http://tools.ietf.org/html/rfc1321>.

³ Würde die Erde nur aus Staubkörnern bestehen und man würde jedes Staubkorn (mit der Masse $2 \cdot 10^{-9}$ g) zählen, reichten die Schlüssel für rund 10.000 Erden.

⁴ aID steht für Archival Identifier.

Damit nicht alle Dateien im gleichen Verzeichnis liegen, wird mittels der aID eine geeignete Unterverzeichnis-Struktur gebildet.

3. Die Datenbank-Komponente von DIMAG enthält ebenfalls die Metadaten, aber keine Primärdaten. Die Datenbank dient vorzugsweise zur Recherche, Nutzung und Steuerung der Migration. Für die Langzeitarchivierung ist die Datenbank hilfreich, aber wichtiger und vollständiger ist das Dateisystem.
4. Mehrfache Hardware Redundanz, welche vor möglichem Verlust schützen soll. Dies betrifft Dinge wie Netzteil, Stromversorgung, Festplattencontroller (z. B. RAID), Journaling-Dateisystem⁵ und das zweifache Back-up-Konzept mit unterschiedlichen Sicherungsmethoden und Standorten.

Damit waren gute Voraussetzungen für das elektronische Grundaktenarchiv gegeben, allerdings gab es Anpassungsbedarf durch veränderte Anforderungen.

1. In DIMAG entspricht das Informationsobjekt aus der klassischen Schriftgutverwaltung einer Akte oder einem Vorgang. Unterhalb des Informationsobjekts kommen dann die einzelnen Repräsentationen. Diese Zuordnung ist in der Grundbuchverwaltung nicht möglich, da bei dem Registraturgut immer wieder neue Dokumente hinzukommen. Zugleich verändert sich das einzelne Dokument vom Zeitpunkt des ersten Abspeicherns nicht mehr. In G-DIMAG bildet daher das einzelne Dokument das Informationsobjekt. Jedes Dokument hat folglich eigene Repräsentationen.
2. Im Wesentlichen sollte DIMAG nicht mehr durch eine(n) Archivar(in) befüllt werden, sondern automatisiert durch einen Workflow-Prozess, welcher vom Partner hp⁶ entwickelt wurde. Schnell hat man sich technologisch auf eine SOAP-Schnittstelle⁷ geeinigt. Dann wurden die benötigten Methoden für das Anlegen einer Akte (`storeRecord`), das Speichern eines Dokuments (`storeDocument`) und die Erstellung einer Repräsentation (`addRepresentation`) und manch weiterer Methoden festgelegt. All diese Methoden werden festgelegt in einer WSDL⁸ Beschreibung, welche die Schnittstelle in XML-Syntax spezifiziert. Ein vereinfachtes Beispiel sieht wie folgt aus: Sowohl alle Eingangsparameter als auch die Antwortparameter werden für eine Methode `getDocumentMetadata` festgeschrieben. Dabei kann nun jede Anwendung, welche sich auf einem entfernten Server befinden darf, diese Methoden als Service von G-DIMAG nutzen.

⁵ Änderungen werden vor dem eigentlichen Schreiben in einem dafür reservierten Speicherbereich, dem *Journal*, aufgezeichnet.

⁶ Hewlett-Packard, ein Technologie-Konzern.

⁷ SOAP (ursprünglich für Simple Object Access Protocol) ist ein Netzwerkprotokoll, mit dessen Hilfe Daten zwischen Systemen ausgetauscht und Remote Procedure Calls durchgeführt werden können.

⁸ Web Services Description Language.

Als Beispiel folgt hier eine SOAP-Schnittstellenbeschreibung:

```
<message name="getDocumentMetadataRequest">
  <part name="toolUser" type="xsd:string"/>
  <part name="toolPasswd" type="xsd:string"/>
  <part name="documentID" type="xsd:string"/>
</message>
<message name="getDocumentMetadataResponse">
  <part name="retStatus" type="xsd:string"/>
  <part name="retMessage" type="xsd:string"/>
  <part name="status" type="xsd:string"/>
  <part name="md5" type="xsd:string"/>
  <part name="signatur" type="xsd:string"/>
  <part name="PUID" type="xsd:string"/>
  <part name="metadataFile" type="xsd:string"/>
  <part name="references" type="tns:referenceList"/>
</message>
```

3. Wie bei allen frischen Partnerschaften war es zunächst etwas schwierig, sich einer gemeinsamen Sprache zu bedienen. So gab es insbesondere am Anfang Unterschiede im Verständnis zur Schriftgutverwaltung, welche in DIMAG nicht nur nach Akte, Vorgang, Dokument organisiert ist. DIMAG kann neben aktenförmigem Schriftgut auch anders geformte digitale Objekte archivieren, erfordert aber die Abbildung einer hierarchischen Struktur bis zum Archivale. Letzen Endes entstand aber ein gemeinsames Verständnis bezüglich der Umsetzungs-Anforderungen.
4. Im Gegensatz zu DIMAG handelt es sich bei G-DIMAG um eine Speicherlösung für aktive Dokumente, welche sich im Laufe Ihrer Zeit noch ändern. Hierbei kommt eine Funktion der Versionierung zum Einsatz, welche nicht nur das letztlich aktualisierte Dokument, sondern auch die zeitlich zuvor existierenden Versionen kennt und vorhält.
5. Eine weitere sinnvolle Ergänzung ist die *commit* bzw. *rollback* Funktion. Dieser Bedarf entstand dadurch, dass ein Aufruf der SOAP-Schnittstelle sich aus mehreren Einzelaktionen von G-DIMAG zusammensetzt. Erst wenn alle Einzelaktionen erfolgreich durchgeführt wurden, darf sich dies in G-DIMAG niederschlagen (*commit*). Scheitert eine Einzelaktion, werden alle zuvor erfolgreich durchgeführten Aktionen wieder rückabgewickelt (*rollback*).
6. Die Erfahrung hat gezeigt, dass große Datenmengen sich nicht mit der SOAP-Schnittstelle übertragen lassen. Daher wurde eine weitere Option entwickelt, große Dateien zunächst auf ein gemeinsames SFTP-Verzeichnis⁹ zu stellen und mit der SOAP Methode nur den Anstoß zur Speicherung zu geben.

⁹ SFTP (SSH File Transfer Protocol) ist eine für die Secure Shell (SSH) entworfene Alternative zum File Transfer Protocol (FTP), die eine verschlüsselte Übertragung ermöglicht.

7. Eine ebenfalls für G-DIMAG spezifische Umsetzung ist das ‚On the fly‘-Generieren von PDF/A Dokumenten für große TIFF-Dateien. Hierbei soll verhindert werden, dass eingescannte Pläne mit großer Dateigröße elektronisch über das Netzwerk zum Nutzer zurückgespielt werden. Tests haben gezeigt, dass ein sehr großer Scan im TIFF-Format nach verlustbehafteter Kompression noch brauchbar als PDF/A nutzbar ist. Die Datenmenge reduziert sich durch dieses Verfahren auf 0,3 Prozent.

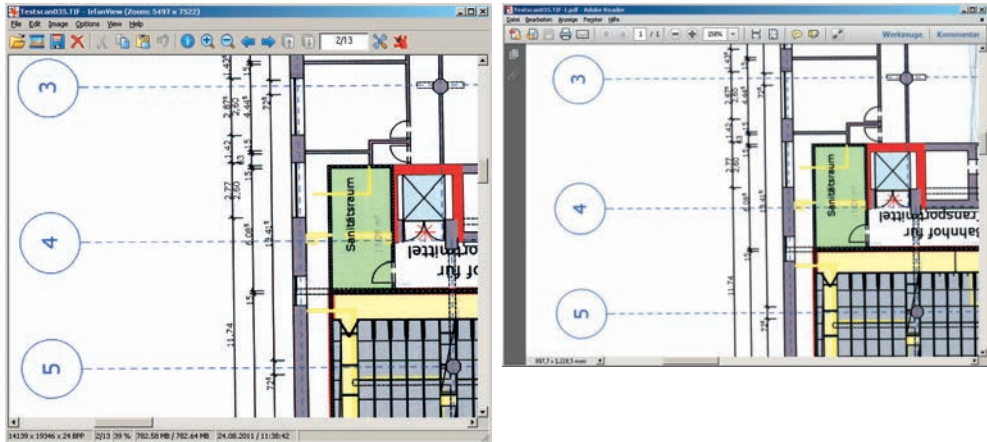


Abb. 1: Detail-Ausschnitt DIN A0 Plan in TIFF (links, unkomprimiert mit 782 MB) bzw. PDF/A (rechts, verlustbehaftet komprimiert mit 2,4 MB).

Hardware-Betrieb im Rechenzentrum

G-DIMAG wird im Rechenzentrum des LZfd¹⁰ betrieben. Dort sind derzeit 3 Installationen vorhanden. Neben der produktiven Umgebung gibt es noch eine Test-Umgebung für den Integrationstest und eine QS-Umgebung für die Endabnahme. Alle Installationen sind auf virtuellen Maschinen eingerichtet. Als Betriebssystem kommt Linux (SUSE Distribution) zum Einsatz. Die derzeitigen openSource Hauptkomponenten sind:

- Apache/2.2.17
- PHP Version 5.3.5
- MySQL 5.0.7
- Java Runtime Environment 1.6

¹⁰ Das Landeszentrum für Datenverarbeitung (LZfd) ist eine Abteilung der Oberfinanzdirektion Karlsruhe. Es erbringt IT-Dienstleistungen für Behörden und Körperschaften des öffentlichen Rechts und ist das zentrale Datenverarbeitungszentrum der Finanzverwaltung.

- IngestList 6.2
- ImageMagic 6.7.7

Der derzeitig reservierte Speicher beträgt 4,6 Terabyte auf der produktiven Umgebung.

Auch die Datensicherung wird vom LZfD übernommen. Die Kontrolle, ob und was gesichert wird, obliegt allerdings G-DIMAG. Hierzu wird ein erstes Script von G-DIMAG aufgerufen. Dies stoppt die Datenbank, exportiert deren Inhalte als SQL Dump, sichert die SFTP Dateien und prüft die Integrität des Archives mittels MD5. Erst im Anschluss wird das Backup angestoßen, damit keine beschädigten oder manipulierten Files in das Backup weitergereicht werden.

Strukturelle Abbildung in G-DIMAG

Die Tektonik beschreibt zunächst die Amtsgerichte, darunter deren Gemeinden und Grundbuchbezirke. Im ersten Schritt wird mit 5 Amtsgerichten und voraussichtlich 180.000 Akten begonnen.

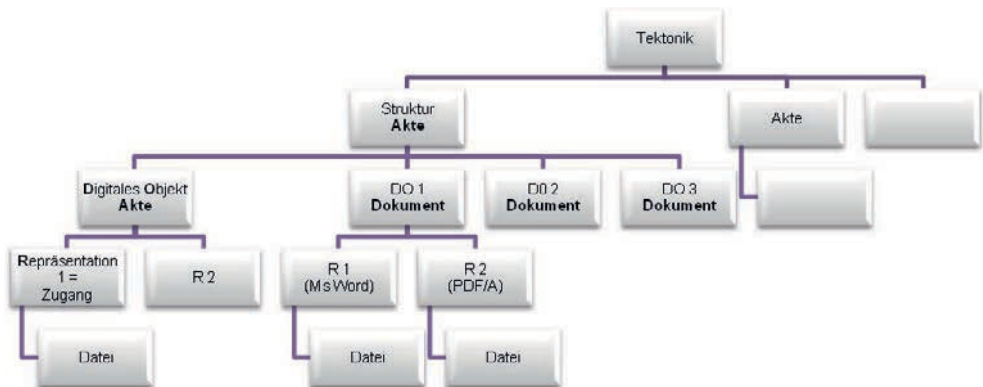


Abb. 2: Hierarchische Struktur einer Grundakte mit übergeordneter Tektonik.

DO 1, 2 und 3 gehören zu einer Akte (*Struktur Akte 1*), deren weitergehende Metadaten in einem eigenen Digitalen Objekt (*Digitales Objekt Akte*) abgelegt werden.

Eine kleine Führung durch G-DIMAG

1. Nach der Benutzer Anmeldung erscheint dieser Welcome Screen:

The screenshot shows the G-DiMag interface. On the left, there is a sidebar with the title 'G - DiMag' and a list of menu items: 'Digitalisiertes Archiv 2.3', 'Archiv', 'Anmelden', and 'Rechner'. The main content area displays a welcome message: 'Herzlich willkommen Rolf Lang!' followed by 'Welcome @ G-DiMag' and a large digital-style display showing 'Grundaktenarchiv' and the date '2012-08-16'. Below this, it says 'Elektronisches Grundaktenarchiv'.

2. Die neuen Grundbuchamt-Standorte bei den Amtsgerichten werden angezeigt, hier die ersten 5 von zukünftig 13 Amtsgerichten:

The screenshot shows the G-DiMag search results page. The sidebar on the left is identical to the previous screenshot. The main content area is titled 'Struktur Suche' and shows a breadcrumb trail: 'Tektoniki > D: Digitales Archiv > A: Archivdaten > G: Grundbuch Zentral Archiv'. There is a search input field with 'Suche aID:' and a 'GO' button. Below the search results, there is a table with columns 'Signatur / aID', 'Titel', and 'Eigenschaften'.

Signatur / aID	Titel	Eigenschaften
hp-tec210	Achern	0 7 1 1 1 X
hp-tec5	Erlenmünster	0 7 1 1 1 X
hp-tec292	Kornwestheim	0 7 1 1 1 X
hp-tec393	Tauferschiedsheim	0 7 1 1 1 X
hp-tec395	Villingen-Schwenningen	0 7 1 1 1 X

3. Die darin eingegliederten Grundbuchämter:

G - DiMag

Digitales Archiv 2.3
Archiv
Apache/2.2.17 (Unix) DAV/2
mod_ssl/2.2.17
OpenSSL/1.0.0c PHP/5.3.5
mod_spcre2/2.0.9/1.10/2.2.1
mod_perl/2.0.4 Perl/V5.10.1
Server at 10.23.140.68 Port 443
your IP=10.15.48.6
Server=10.23.140.68

Kronmüller
Rolf Lang
Reichler
DIMAG: admin
Grundbuch: admin

Struktur Suche

Tektonik: - D: Digitales Archiv
- A: Archivalien
- G: Grundbuch Zentral Archiv
- I: Villingen-Schwenningen

Suche aID:
Modus wechseln

Signatur / aID	Titel	Eigenschaften
hp-tec794	Aach	0 7 1 1 1 1 X
hp-tec800	Alten	0 7 1 1 1 1 X
hp-tec793	Albbbruck	0 7 1 1 1 1 X
hp-tec1052	Allensbach	0 7 1 1 1 1 X
hp-tec775	Bad Dürheim	0 7 1 1 1 1 X
hp-tec695	Bad Säckingen	0 7 1 1 1 1 X
hp-tec994	Bermatingen	0 7 1 1 1 1 X
hp-tec994	Bernau im Schwarzwald	0 7 1 1 1 1 X
hp-tec985	Blumberg	0 7 1 1 1 1 X
hp-tec1007	Bodman-Ludwigshafen	0 7 1 1 1 1 X
hp-tec992	Bonnndorf im Schwarzwald	0 7 1 1 1 1 X
hp-tec958	Bönigsdorf	0 7 1 1 1 1 X
hp-tec650	Braunlingen	0 7 1 1 1 1 X
hp-tec1090	Böllen	0 7 1 1 1 1 X
hp-tec729	Büdingen am Hochrhein	0 7 1 1 1 1 X
hp-tec964	Caalberg (Guttschwarzwald)	0 7 1 1 1 1 X
hp-tec958	Calandorf	0 7 1 1 1 1 X
hp-tec731	Clautingen	0 7 1 1 1 1 X
hp-tec889	Deeggenhauertal	0 7 1 1 1 1 X
hp-tec999	Dettinghofen	0 7 1 1 1 1 X
hp-tec1152	Deggern	0 7 1 1 1 1 X
hp-tec749	Donauschlingen	0 7 1 1 1 1 X
hp-tec1001	Eggingen	0 7 1 1 1 1 X
hp-tec763	Eggingen	0 7 1 1 1 1 X
hp-tec1015	Egingen	0 7 1 1 1 1 X
hp-tec1021	Frickingen	0 7 1 1 1 1 X
hp-tec1054	Fröhnd	0 7 1 1 1 1 X
hp-tec757	Furtwangen im Schwarzwald	0 7 1 1 1 1 X
hp-tec796	Gägenhofen	0 7 1 1 1 1 X
hp-tec1040	Gailingen am Hochrhein	0 7 1 1 1 1 X
hp-tec1129	Geltmadingen	0 7 1 1 1 1 X
hp-tec1029	Grafenhausen	0 7 1 1 1 1 X
hp-tec1153	Gürwühl	0 7 1 1 1 1 X
hp-tec1032	Güttenbach	0 7 1 1 1 1 X
hp-tec997	Hagnau am Bodensee	0 7 1 1 1 1 X
hp-tec762	Hesel	0 7 1 1 1 1 X
hp-tec1133	Hausen im Wiesental	0 7 1 1 1 1 X
hp-tec687	Heiligenberg	0 7 1 1 1 1 X
hp-tec1044	Hemmsried	0 7 1 1 1 1 X
hp-tec804	Hilzingen	0 7 1 1 1 1 X
hp-tec693	Hohenfels	0 7 1 1 1 1 X
hp-tec981	Hohenhingen am Hochrhein	0 7 1 1 1 1 X
hp-tec741	Hägerhärsberg	0 7 1 1 1 1 X
hp-tec802	Hilzsm	0 7 1 1 1 1 X
hp-tec689	Höchenschwand	0 7 1 1 1 1 X
hp-tec1127	Höfingen	0 7 1 1 1 1 X
hp-tec696	Ibach	0 7 1 1 1 1 X

4. Die darin enthaltenen Orte:

G - DiMag

Digitales Archiv 2.3
Archiv
Apache/2.2.17 (Unix) DAV/2
mod_ssl/2.2.17
OpenSSL/1.0.0c PHP/5.3.5
mod_spcre2/2.0.9/1.10/2.2.1
mod_perl/2.0.4 Perl/V5.10.1
Server at 10.23.140.68 Port 443
your IP=10.15.48.6
Server=10.23.140.68

Kronmüller
Rolf Lang
Reichler
DIMAG: admin
Grundbuch: admin

Struktur Suche

Tektonik: - D: Digitales Archiv
- A: Archivalien
- G: Grundbuch Zentral Archiv
- I: Villingen-Schwenningen
- I: Albbbruck

Suche aID:
Modus wechseln

Signatur / aID	Titel	Eigenschaften
hp-tec728	Albbbruck	0 7 1 1 1 1 X
hp-tec744	Birkingen	0 7 1 1 1 1 X
hp-tec746	Birndorf	0 7 1 1 1 1 X
hp-tec748	Buch	0 7 1 1 1 1 X
hp-tec745	Sachsen	0 7 1 1 1 1 X
hp-tec747	Unterlupfen	0 7 1 1 1 1 X

- 5. Und darin geführten Akten, derzeit noch wenige, später aber sehr viele.

The screenshot shows the DiMag web interface. On the left, there is a sidebar with server information for 'Digitales Archiv 2.3' and a user profile for 'Rolf Lang'. The main content area is titled 'Struktur Suche' and displays a search tree for 'D: Digitales Archiv'. The search results table shows several records:

Signatur / aID	Titel	Eigenschaften
Akte (rec137465)		0 7 1 1 1 1
Akte (rec138640)		0 7 1 1 1 1
Akte (rec138738)		0 7 1 1 1 1

- 6. Und die zugehörigen Dokumente

This screenshot shows the DiMag web interface with search results for digital objects. The search tree is expanded to show 'G: Grundbuch Zentral Archiv'. The search results table displays digital objects associated with the records:

Signatur / aID	Titel	Eigenschaften
217-7576	Digitales Objekt für Akt (rec138738)	0 7 1 1 1 1
217-7606	Digitales Objekt für Dokument (doc7335)	0 7 1 1 1 1
217-7616	Digitales Objekt für Dokument (doc7247)	0 7 1 1 1 1

- 7. Dann erfolgt der Übergang zu den Erscheinungsformen, hier die zwei Repräsentationen:

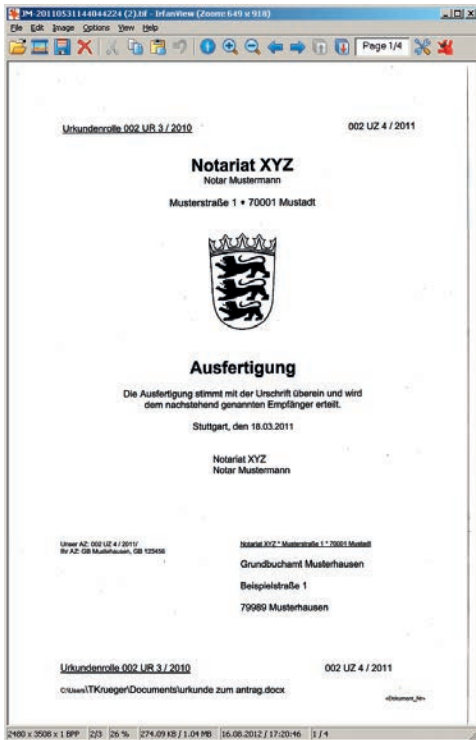
This screenshot shows the DiMag web interface with search results for two representations of a document. The search tree is expanded to show 'I: Digitales Objekt für Dokument (doc7247)'. The search results table displays two representations:

Signatur / aID	Titel	Eigenschaften
217-7617	1. Repräsentation	0 7 1 1 1 1
217-7621	2. Repräsentation	0 7 1 1 1 1

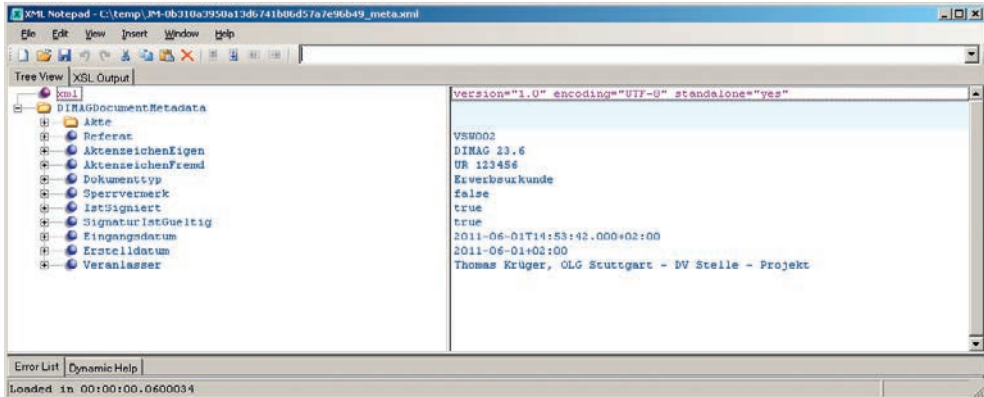
- 8. Die erste Repräsentation beinhaltet die Metadaten, welche nicht von G-DIMAG generiert wurden, sondern vom Workflow-Prozess stammen und die eigentliche Primärdatei.



- 9. Diese Primärdatei (Hauptdokument der Grundakte) als TIFF Ausgabe



10. Die Grundbuch-Metadaten, von G-DIMAG wie Primärdaten behandelt und nicht verändert



11. Die Metadaten von G-DIMAG selbst

G - DiMag

Struktur Suche

Tektonik: D: Digitales Archiv
 - A: Archivwähler
 - G: Grundbuch Zentral Archiv
 - I: Villingen-Schwenningen
 - J: Altbau
 - K: Buch
 - L: Akte (rec130738): Akte (rec130738)
 - M: Digitales Objekt für Dokument (doc7247)
 - N: 2. Repräsentation

Suche sID:

 Modus wechseln

Allgemeine Metadaten	
Titel	Primär Datei
Beschreibung / Inhalt	Primär
Typ	F
Erstellt von	Unbekannt in CA
Erstellt am	2012-08-15 17:17:33
Geändert von	DIMAG66ap
Geändert am	2012-08-15 17:17:33
Status-Code	20
XML-Version	2012-08-09
Archivallentyp	Fachverfahren
Mime-Type	application/pdf
Ursprünglicher Dateiname	20110531144044224-M.pdf
archive ID	217-7622
parent ID	217-7621
hp Dokument ID	doc7247
hp Datei Art	Primärdaten
Zugriff/Schutz	0
Inventar (dynamisch erstellt)	
Spezifische Metadaten	
Pronom Persistent ID	fmt/18
Pronom Format	Acrobat PDF 1.4 - Portable Document Format V1.4
Pronom mime	PDF
Pronom mime	application/pdf
Pronom classification	Positive (Specific Format)
Pronom Signature	DR:OID_Signaturefile_V61.xml
Letzte Datei Änderung	2012-08-15 17:17:28
Dateigröße	134226
MDS Wert der Datei	4f8ea3fc3c115784eeec158299b9e2486
Format-Version	1.4

12. Weitere Funktionen von G-DIMAG

13. Protokolle werden von allen wesentlichen Funktionen angelegt, hier das Ablegen weiterer Dokumente in die Akte

Zwischen User-Group und Entwicklungsgemeinschaft Konstituierung, Aufgaben und Arbeitsperspektiven einer spartenübergreifenden Nutzergruppe

Von CHRISTOPH SCHMIDT

Im Herbst 2011 haben sich die aktuellen Anwender der Akzessions- und Archivierungslösungen der Dienstleister HP und SER im deutschsprachigen Raum zu einem Nutzerkreis zusammengeschlossen. Ziel der bislang informell organisierten Gruppe, der das Bundesarchiv, das Stadtarchiv Stuttgart, das Landesarchiv NRW, das Historische Archiv der Stadt Köln sowie das LWL-Archivamt für Westfalen angehören, ist die Verbesserung des Informationsaustauschs, die Schaffung fachlicher und ökonomischer Synergieeffekte in der Systementwicklung sowie die Etablierung einer zentralen Kommunikationsplattform gegenüber den Dienstleistern. Perspektivisch werden die Weiterentwicklung der verwandten Einzellösungen hin zu einer möglichst homogenen Gesamtlösung und der Ausbau des Nutzerkreises zu einer Entwicklungsgemeinschaft angestrebt.

Wenn man ein Ziel des Arbeitskreises *Archivierung von Unterlagen aus digitalen Systemen* darin sieht, die Bereitschaft der deutschsprachigen Archive zu fördern, den Schritt aus der archivtheoretischen Diskussion hinaus in die archivpraktische Wirklichkeit zu wagen, so lässt sich erfreulicherweise für die letzten Jahre konstatieren, dass die langjährige Aufbauarbeit mehr und mehr Früchte zu tragen beginnt. Und auch, wenn viele Archive noch ganz am Anfang ihres praktischen Engagements beim Aufbau und Betrieb digitaler Repositorien stehen, während die Gruppe der fortgeschrittenen Einrichtungen weiterhin überschaubar ist – die Zahl der Praktiker ist deutlich gewachsen, und sie wächst weiter.

Eine wenn nicht natürliche, so doch zumindest vorhersehbare Folge dieses Wachstums ist eine langsam einsetzende fachliche wie technische Konsolidierung der vorhandenen Lösungsansätze. Zwar ist die deutschsprachige Archivlandschaft weiterhin noch weit von so etwas wie einem gemeinsamen Standard entfernt, und der Markt der konkurrierenden Systeme ist sicherlich auch noch nicht voll besetzt – und doch sind bereits heute erste Kristallisationspunkte für institutionenübergreifende Systementwicklungen erkennbar. Einer dieser Kristallisationspunkte ist derzeit der DIMAG-Entwicklungsverbund, der von den staatlichen Archiven der Länder Baden-Württemberg, Hessen und Bayern betrieben wird und dem in diesem Band ebenfalls ein Beitrag gewidmet ist. Ein zweiter Kristallisationspunkt ist der Nutzerkreis für die Akzessions- und Archivierungslösungen, die von den Firmen HP und SER als Dienstleister betreut werden. Mit diesem Nutzerkreis beschäftigt sich der vorliegende Beitrag.

Der Nutzerkreis der HP-/SER-Lösung, für die es noch keine griffige, institutionenübergreifende Bezeichnung gibt, hat sich am Rande des Deutschen Archivtags 2011 in Bremen konstituiert. Seitdem agiert er als ein formal unverbindliches Gremium, das allen aktuellen Nutzern

des Systems und seiner Varianten offensteht. Derzeit sind dies das Bundesarchiv, das Stadtarchiv Stuttgart, das Historische Archiv der Stadt Köln, das LWL-Archivamt für Westfalen sowie das Landesarchiv Nordrhein-Westfalen. Die Teilnehmerrunde, die sich vor allem aus den für das Fachgebiet *Digitale Unterlagen* zuständigen Archivarinnen und Archivaren der beteiligten Archive zusammensetzt, wird ergänzt durch Vertreterinnen und Vertreter der jeweiligen IT-Dienstleister und IT-Abteilungen der Archive, die für den technischen Betrieb der Archivierungslösungen zuständig sind. Die Firmen HP und SER als Dienstleister befürworten und unterstützen die Arbeit der Gruppe, waren aber weder an ihrer Konstituierung noch an der Festlegung von Zielen und Organisationsstrukturen unmittelbar beteiligt. Beide Dienstleister nehmen zudem nur zu konkreten Anlässen an einzelnen Sitzungen der Nutzergruppe teil.

Die bewusste Beschränkung des Teilnehmerkreises auf aktuelle Anwender der Archivierungslösung kristallisierte sich schnell als eine allgemeine formelle Anforderung des Kreises heraus. Diese Einschränkung dient allerdings nicht der fachlichen Abschottung der beteiligten Archive oder der Behinderung eines von allen beteiligten Archiven als weiterhin dringend notwendig erachteten fachlichen Meinungsaustauschs jenseits einzelner technischer Lösungen. Vielmehr dient die Beschränkung des Teilnehmerkreises der Förderung einer möglichst effektiven Arbeitsatmosphäre und eines unkomplizierten Informationsaustauschs, der, so die Überzeugung der beteiligten Archive, nur dann ungefiltert, unkompliziert und zeitnah sichergestellt werden kann, wenn die Kommunikationsstrukturen eine hinreichende Vertraulichkeit nach außen gewährleisten. Fachliche Arbeitsergebnisse des Nutzerkreises, die von überindividuellem Interesse sind, können und sollen natürlich weiterhin publiziert und weitergegeben werden, und perspektivisch wird sich der Nutzerkreis auch intensiv in bestehende oder noch einzurichtende Informationsstrukturen einbringen. Der Blick über den Tellerrand (oder vielleicht sogar: das Teilen der Suppe?) ist allerdings nicht der eigentliche Zweck des Nutzerkreises – dieser besteht vielmehr (um im gewählten Bild zu bleiben) im *gemeinsamen Kochen*. Tatsächlich versteht sich die Nutzergruppe primär als eine operative, möglichst ergebnisorientierte Arbeitsgruppe. Sie verfolgt dabei im Wesentlichen drei konkrete Ziele:

1. Den möglichst zeitnahen und unzensierten Austausch fachlicher, technischer und betriebsorganisatorischer Informationen und Erfahrungen,
2. die Schaffung fachlicher und ökonomischer Synergieeffekte bei der Weiterentwicklung des Systems und
3. die Etablierung einer gemeinsamen Kommunikationsplattform zur Vereinfachung und Verbesserung der Verhandlungsposition der Archive gegenüber den Dienstleistern.

Nach der Auftaktsitzung des Kreises in Bremen haben bislang (Stand: Mai 2012) drei Arbeitssitzungen stattgefunden, eine vierte ist für den Frühsommer 2012 vorgesehen. Die vergleichsweise enge Sitzungstaktung resultiert aus der aktuell hohen Entwicklungsdynamik der Archivierungslösung, die insbesondere die gerade systemeinführenden Archive in den Nutzerkreis mit einbringen. Ob die bislang vorgelegte hohe Schlagzahl mittelfristig beibehalten werden kann, wird sich im Zuge der weiteren Konsolidierung der Gruppenstrukturen und der Festlegung konkreter Entwicklungsziele zeigen.

Angesichts sehr unterschiedlicher Entwicklungs- und Umsetzungsstände bestand das erste

Ziel des Nutzerkreises darin, einen homogenen Informationsstand zu schaffen und geeignete organisatorische Strukturen zu etablieren, die einen möglichst reibungslosen Informationsaustausch gewährleisten können. Zu diesem Zweck wurde bei der Stadt Köln kurzfristig ein Sharepoint als *Informationszentrale* eingerichtet, der einen gemeinsamen, unkomplizierten Zugriff auf alle relevanten Spezifikationen, Schemata und sonstigen Dokumente der beteiligten Archive ermöglicht. Auf der so geschaffenen Basis eines gemeinsamen Informationsstandes konnte dann im Rahmen des zweiten Arbeitstreffens eine weitgehende Homogenisierung des Datenmodells erreicht werden, insbesondere im Hinblick auf die Umsetzung des Repräsentationenkonzepts. Obwohl weiterhin Differenzen in der Ausgestaltung der Metadatenschemata bestehen, dürfte sich diese Verständigung als eine sehr wichtige Grundlage für alle weiteren Vorhaben entwickeln, da eine gemeinsame Definition aller im Archivierungsprozess vorkommenden Objekte nicht nur die Grundlage für eine einheitliche Terminologie (und damit auch für eine glückende fachliche Kommunikation) darstellt, sondern auch die Entwicklung gemeinsamer technischer Lösungen deutlich vereinfacht.

Als möglicherweise wegweisend für die weiteren Schritte und die Entwicklung der Gruppe von einem Nutzerkreis hin zu einer Entwicklungsgemeinschaft haben die beteiligten Archive danach im Rahmen ihrer dritten Arbeitssitzung einen ersten Versuch unternommen, die Standardisierung von Systemkomponenten zu befördern. Konkret befasste sich der Nutzerkreis dabei mit der Weiterentwicklung des PreIngest Toolsets (PIT) von HP. Dieses wird derzeit in mehreren verschiedenen Programmvarianten mit unterschiedlichem Funktionsumfang von allen an der Lösung beteiligten Archiven eingesetzt. Nach intensiver Diskussion, an der auch die Dienstleister beteiligt wurden, konnte eine gemeinsame Anforderungsbasis erarbeitet werden, die auf den Systemspezifikationen des Bundesarchivs basiert und die in den nächsten Monaten technisch umgesetzt werden soll. Sollte sich die neue *Standardversion* des PIT in der Praxis bewähren, so könnte dies nicht nur die individuellen Pflege- und Entwicklungsaufwände und -kosten deutlich reduzieren, sondern auch die Vereinheitlichung, die Entwicklungsperspektiven und das Verbreitungspotential des Gesamtsystems beträchtlich befördern.

Einen zweiten inhaltlichen Schwerpunkt der dritten Arbeitssitzung bildete die Perspektivplanung für die mögliche Weiterentwicklung der Nutzergruppe hin zu einer echten Entwicklungsgemeinschaft. Hier wurde insbesondere der Austausch von Erwartungen und Einschätzungen zwischen den Archiven und den Dienstleistern gesucht und gefunden. Dabei wurde deutlich, dass trotz der bereits erzielten Ergebnisse noch ein hohes Maß an Koordination und Zusammenarbeit nötig sein wird, um die vorhandenen Lösungen sukzessive zu einem einheitlichen *Produkt* zusammenzuführen. Dabei gilt es nicht nur, bereits bestehende Lösungsansätze abzustimmen und in neuen technischen Releases zusammenzuführen, sondern auch darum, Entwicklungsaufwände möglichst effizient zwischen den verschiedenen Archiven aufzuteilen und für eine offene Nachnutzbarkeit zu Gunsten aller Lösungsteilnehmer zu sorgen. Dem entsprechend arbeitet der Nutzerkreis derzeit arbeitsteilig auf verschiedenen *Baustellen* – etwa an der Entwicklung eines Access-Moduls, an der Etablierung eines Bestandserhaltungsmoduls auf Basis signifikanter Eigenschaften oder an der Bereitstellung von Konvertierungsmechanismen im Ingestprozess. Inwieweit oder besser: bis zu welchem Grad der Gruppe eine auf diesem Wege angestrebte Standardisierung auf freiwilliger Basis gelingen wird, bleibt abzuwarten. Jede Form

von Standardisierung beschränkt naturgemäß individuelle und institutionenspezifische Handlungsspielräume und hat dort ihre Grenzen, wo unvereinbare archivspezifische Anforderungen aufeinanderstoßen. Zudem wird das Preis-Leistungs-Verhältnis zwischen Abstimmungsaufwand und standardgebundener Einschränkung auf der einen und fachlich-technischen Synergieeffekten auf der anderen Seite kontinuierlich zu überprüfen sein, wenn die bislang noch eher vage Vision einer echten Entwicklungsgemeinschaft auf freiwilliger Basis Realität werden soll. Dabei könnte sich insbesondere die Frage der Kostenübernahme für die technische Umsetzung kollektiv erarbeiteter fachlicher Anforderungen perspektivisch als eine Gretchenfrage erweisen, die nicht nur die Frage nach alternativen, formell engeren Organisationsformen nach sich zieht, sondern deren Beantwortbarkeit auch über den Gesamterfolg des gemeinsamen Vorhabens entscheiden könnte.

Allerdings, so bleibt abschließend zu bemerken, wären derartige Überlegungen zweifelsohne erst Gegenstand eines späteren Entwicklungsschrittes, den zu machen derzeit weder notwendig noch planbar erscheint. Praxisnahe, möglichst unkomplizierte operative Arbeit zu leisten war und ist das Ziel des Nutzerkreises, dem jede Form von Organisation angepasst werden sollte. Ob diese ziemlich unbürokratische Herangehensweise an ein komplexes Aufgabenspektrum sich langfristig bewähren wird, bleibt gelassen abzuwarten. Im schlechtesten Fall gewinnen alle Beteiligten Erfahrungen, wie man es nicht machen sollte. Im besten Fall kommen sie gemeinsam weiter als alleine. Das Experiment läuft, und wir sind gespannt, wo es uns noch hinführt.

DIMAG-Kooperationen

Von CHRISTIAN KEITEL

DIMAG ist die Kurzform für Digitales Magazin. Das IT-System wurde vom Landesarchiv Baden-Württemberg 2006 für die dauerhafte Erhaltung seiner digitalen Archivalien konzipiert und ist seit dem 6. Juli 2006 beim Landesarchiv im Produktivbetrieb. Seitdem wird es kontinuierlich weiterentwickelt. Das Landesarchiv Baden-Württemberg hat sich entschlossen, DIMAG unter bestimmten Bedingungen an andere Archive abzugeben. Diese Bedingungen werden als Partnerschaften verstanden, da sie über die reine Abgabe der Software weit hinausgehen. Im Folgenden soll zunächst überlegt werden, welche Rolle eine solche Software im Konzert der digitalen Archivierung spielen kann. Im zweiten Teil des Beitrags werden die verschiedenen derzeit denkbaren Formen der Partnerschaft näher erläutert.

Archivierungssysteme für Archive

Die Entwicklung von DIMAG erfolgte im Rahmen des vom Land Baden-Württemberg geförderten Projekts *Konzeption für ein digitales Landesarchiv*. Durch dieses Projekt sollten die wesentlichen Grundlagen für die digitale Archivierung beim Landesarchiv Baden-Württemberg gelegt werden.¹

Bereits zu Projektbeginn am 2. Februar 2006 war bekannt, dass in der Landesverwaltung Baden-Württembergs zahlreiche digitale Daten entweder als einzelne Dateien oder innerhalb eines Fachverfahrens archivreif waren. Ihre Bewertung und teilweise Übernahme sollten sehr schnell erfolgen, um nicht größere Überlieferungsverluste zu riskieren. Mit dieser Aussicht musste sich das Landesarchiv darauf einstellen, eine sehr große Zahl einzelner Daten- und Informationseinheiten fachlich korrekt verwalten zu können. Um die Handlungsfähigkeit des Landesarchivs zu erhalten, musste also ein IT-System eingesetzt werden,² mit dessen Hilfe derartige Mengen sinnvoll bearbeitet werden können und folglich begann das Projekt damit, nach einem solchen System Ausschau zu halten.

2006 waren in Deutschland keine speziell für Archive entwickelten Archivierungssysteme erhältlich. Als Ersatz boten sich Systeme wie DAITSS, DSpace, FEDORA oder KOPAL an, die auch als Repositories bezeichnet werden und für Bibliotheken entwickelt worden waren. Sehr schnell zeigt sich aber, dass die Archivalien der klassischen Archive sich in mehreren Punkten grundlegend von den Objekten unterscheiden, die in diesen Repositories verwahrt werden:

¹ Zum Projekt s. <http://www.landearchiv-bw.de/web/44346> (Alle Links wurden am 17. August 2012 überprüft).

² Vgl. hierzu der Beitrag des Verfassers *Warum ist Kooperation bei der digitalen Archivierung unumgänglich?* in diesem Band.

- Die archivischen Inhalte liegen nicht atomar, sondern hierarchisch geschachtelt vor, also zum Beispiel als Bestand, Klassifikationspunkt und Archivale oder als Akte, Vorgang und Dokument.
- Publikationen unterliegen einer festgefügt Struktur (Autor, Titel, Erscheinungsjahr...), die in dieser Weise bei Archivalien nicht vorausgesetzt werden kann. Dementsprechend ist bei der Übernahme von Publikationen ein wesentlich höherer Automatisierungsgrad als bei digitalen Archivalien möglich.
- Entsprechend hierzu unterscheiden sich die bislang bekannt gewordenen Ausgestaltungen der einschlägigen Informationspakete nach OAIS, also von SIP und AIP.³
- Publikationen werden kurz nach ihrem Erscheinen von Bibliotheken erworben. Dagegen liegt zwischen der Entstehung einer Verwaltungsunterlage und ihrer Anbietung an das Archiv in aller Regel ein längerer, sich über einige Jahre erstreckender Zeitraum. Dementsprechend sind weniger Informationen über die zur Übernahme anstehenden Objekte zu erwarten und auch die Einwirkungsmöglichkeiten durch die Archive geringer.
- In der Regel können digitale Publikationen sofort nach der Übernahme benutzt werden, digitale Archivalien hingegen nicht. Entsprechend unterschiedlich sollte die Rechteverwaltung des Archivierungssystems ausgestaltet werden.
- Auch die Zugriffshäufigkeit unterscheidet sich erheblich. Für digitale Archivalien muss von weitaus weniger Zugriffen ausgegangen werden. Demzufolge ist es beispielsweise aus Archivsicht kein wesentlicher Faktor, dass mehrere externe Nutzer zeitgleich auf ein und dasselbe digitale Archivale zugreifen können.⁴
- Archive beherbergen vor allem Unikate, Bibliotheken vor allem mehrfach vorhandene Publikationen. Bibliotheken konnten daher ein System wie LOCKSS (Lots of Copies Keep Stuff Safe) entwickeln,⁵ das in Archiven schon aufgrund der datenschutzrechtlichen Anforderungen wenigstens derzeit noch undenkbar ist.

Der zuerst genannte Punkt soll hier noch etwas weiter ausgeführt werden. Für die Anforderungen des Landesarchivs Baden-Württemberg war es essentiell, dass das System auch strukturierte Informationen adäquat abbilden kann. Die archivischen Anforderungen können in diesem Bereich gut anhand der beiden Standards ISAD(G) und OAIS dargestellt werden. Bekanntlich ermöglicht ISAD(G) eine Stufenverzeichnung. Die Informationen werden in einem monohierarchischen Baum abgebildet. Sobald alle Elemente einer Stufe dieselbe Eigenschaft tragen, wird diese eine Stufe höher nachgewiesen. Auf der anderen Seite verlangt OAIS bei der Definition

³ Zu OAIS vgl. Referenzmodell für ein Offenes Archiv-Informationssystem. Deutsche Übersetzung. Bearbeitet von der der nestor-Arbeitsgruppe OAIS-Übersetzung / Terminologie. Frankfurt/Main 2012.

⁴ So sah das von mehreren Bibliotheken getragene KOPAL-Projekt zunächst einen vergleichbaren Lasttest vor, <http://kopal.langzeitarchivierung.de>.

⁵ Bezeichnend die ersten Worte auf der Eingangsseite von <http://www.lockss.org/>: "The LOCKSS Program, based at Stanford University Libraries, provides libraries and publishers with award-winning, low-cost, open source digital preservation tools to preserve and provide access to persistent and authoritative digital content. Libraries keep what they buy ...".

seines Archivierungspakets (AIPs), dass dieses alles enthalten möge, was zum Aufruf und Verständnis dieses Objekts notwendig ist. Aufgrund der gestuften Erschließung kann nun nicht einfach das Informationsobjekt zusammen mit der Repräsentation und deren Dateien als ein AIP abgelegt werden, denn die relevanten Informationen könnten auch eine oder mehrere Ebenen höher erfasst sein. Auf der anderen Seite kann nicht die gesamte Tektonik ohne weiteres in dem AIP physisch abgelegt werden, da ansonsten bei jeder kleinen Änderung in der Tektonik möglicherweise tausende AIPs geöffnet, geringfügig in den Metadaten überarbeitet und wieder neu abgespeichert werden müssten. An dieser Stelle wird deutlich, dass die von Bibliotheksseite her kommende atomare Sicht auf ein AIP, wie sie ja auch in den zur Frage stehenden Repositories realisiert wurde, für Archive keine wirklich gelungene Lösung darstellt. Vielmehr musste ein Archivierungssystem entwickelt werden, das die bewährten archivischen Standards mit den neuen Anforderungen aus der Welt der digitalen Archive abgleicht. Weil ein derartiges System 2006 nicht erhältlich war, entwickelte das Landesarchiv Baden-Württemberg DIMAG.

DIMAG: Entwicklung und Eigenschaften

Bei der Entwicklung von DIMAG konnte das Landesarchiv im Wesentlichen auf drei Bereiche zurückgreifen. Schon 2002 waren in einer ersten Konzeption zu einer ähnlichen Fragestellung einige grundlegende Empfehlungen für den Einstieg in die digitale Archivierung abgegeben worden,⁶ die in der Folgezeit von der damaligen Landesarchivdirektion auch umgesetzt wurden. Die digitale Archivierung sollte zwar zunächst zentral an einer Stelle im Landesarchiv technisch begleitet werden. Zugleich sollten alle Archive auch für die digitalen Unterlagen ihres Zuständigkeitsbereichs die Aufgaben der Bewertung, Erschließung und Nutzung übernehmen. Die zentrale Archivierungsstelle wurde im Staatsarchiv Ludwigsburg untergebracht, wo auch später das DIMAG-Produktivsystem aufgebaut wurde. Der später realisierte browserbasierte Zugriff auf DIMAG stellte dann sicher, dass alle Staatsarchive auf *ihre* digitalen Archivalien zugreifen konnten und auch eine spätere Beherbergung anderer Archive möglich war. Zweitens hatte diese Konzeption die Übernahme erster archivreifer und -würdiger Unterlagen vorgeschlagen, nicht zuletzt, um Erfahrungen aufzubauen. Auf diese Erfahrungen konnte nun Anfang 2006 zurückgegriffen werden. An dritter Stelle können auch die zahlreichen Standards genannt werden, die in die DIMAG-Entwicklung einflossen:

- Ein Modell digitaler Archive gaben OAIS und der nestor-Kriterienkatalog vertrauenswürdige digitale Langzeitarchive (heute: DIN 31644) ab.
- Container für AIPs wurden von METS beschrieben.
- Die Übernahme und Aufbereitung ließ sich im nestor-Leitfaden Wege ins Archiv (heute DIN 31645) nachlesen.

⁶ Die Archivierung elektronischer Unterlagen in der baden-württembergischen Archivverwaltung. Eine Konzeption. 2002, http://www.landearchiv-bw.de/sixcms/media.php/120/47169/keitel_elektronische_konz.pdf.

- Die Erschließung wurde durch ISAD(G) und EAD skizziert.
- Erhaltende Metadaten wurden nach PREMIS und dem Metadatenmodell der neuseeländischen Nationalbibliothek NLNZ-MD ausgeführt.

DIMAG selbst soll hier nicht ausführlich vorgestellt werden.⁷ Daher nur einige Hinweise an dieser Stelle:

- DIMAG bildet die Archivalien in ihrem Kontext ab.
- Die konkreten Strukturen (Tektonik und Klassifikation) werden nicht vorgeschrieben und sind daher flexibel einsetzbar (so kann beispielsweise die Zahl der Stufen frei gewählt werden, möglich sind daher zwei-, drei- oder sechsstufige oder beliebige andere Einstellungen).
- Informationsobjekte und Repräsentationen werden systematisch voneinander unterschieden.
- Ein besonderer Wert wird auf den Erhalt der Authentizität und Benutzbarkeit gelegt. Metadaten, Protokolle und spezielle Bereiche dokumentieren den Lebenszyklus der Archivalien und machen ihn für künftige Nutzer und Archivare transparent.
- DIMAG kann alleinstehend oder zusammen mit einem Erschließungssystem eingesetzt werden.
- Die Software läuft sowohl auf Linux (Suse Enterprise Distribution) als auch auf Microsoft Windows.

⁷ Vgl. dafür Christian *Keitel* und Rolf *Lang*: DIMAG und IngestList. Übernahme, Archivierung und Nutzung von digitalen Unterlagen im Landesarchiv Baden-Württemberg. In: Archivische Informationssysteme in der digitalen Welt. Aktuelle Entwicklungen und Perspektiven. Hg. von Gerald *Maier* und Thomas *Fritz*. Stuttgart 2010, S. 53–63.

Landesarchiv Baden-Württemberg
Digitales Archiv

Apache/2.2.17 (Unix) DAV/2 mod_ssl/2.2.17 OpenSSL/1.0.0c PHP/5.3.5 mod_spwd/2.0.0/2.0.4 Perl/v5.10.1
 Server at 10.0.0.160 Port 443
 your IP=10.0.0.140
 Server=10.0.0.160

Struktur Suche
 Angemeldet: Dr. Kai Naumann; RechteU=4

Tektonik:

- Digitales Archiv
 - Archivalien
 - Staatsarchiv Ludwigsburg
 - EL-Serie: Ober- und Mittelbehörden seit um 1945
 - Landesamt für Geoinformation und Landentwicklung Baden-Württemberg: digitale Geobasisdaten
 - Liegenschaftskataster
 - Automatisierte Liegenschaftskarte (ALK)
 - 1_Stuttgart
 - 11_Stadtkreis_Stuttgart
 - 000_Stuttgart

Signatur	Titel	Eigenschaften
D A StAL EL 68_V 1 ALK 1 11SS 000 R1	R1_BGRUND	0
D A StAL EL 68_V 1 ALK 1 11SS 000 R2	R2_SHAPE	0

zurück drucken aktualisieren

Projekt "Digitales Archiv"
 Landesarchiv Baden-Württemberg
 Anregungen und Kritik sind jederzeit willkommen.

Abb. 1: Die automatisierte Liegenschaftskarte in DIMAG (Bestand StAL EL 68 V).

Nach einer gut halbjährigen Entwicklungs- und Konzeptionsphase konnte DIMAG ab dem 6. Juli 2006 im Produktivbetrieb eingesetzt werden. Zunächst galt es, mit der neuen Software Erfahrungen zu machen und die unvermeidlichen Anfangsschwierigkeiten zu minimieren. Auch mussten einige irrige Annahmen aus der ersten Konzeptionsphase überdacht werden. Besonders deutlich wurde dies in der Benennung der Dateien. In der Anfangsphase war das Landesarchiv noch davon ausgegangen, dass die Dateinamen *die Signatur enthalten und damit sprechend* sein sollten. Leider sind archivische Signaturen nicht so dauerhaft, wie dies zu wünschen wäre. Das Landesarchiv entwickelte daher DIMAG 2.0 auf der Basis eines eigenständig entwickelten Identifier Konzepts, das aller Voraussicht nach selbst nach einer künftigen Ablösung von DIMAG durch ein Nachfolgesystem noch funktionieren dürfte. DIMAG 2.0 wurde 2009 entwickelt.

Etwa zeitgleich sah sich das Landesarchiv imstande, aufgrund der gewonnenen Erfahrungen DIMAG nun auch anderen Archiven zur Weitergabe anzubieten. Zunächst wurde DIMAG anderen staatlichen Archivverwaltungen angeboten. Noch 2009 fanden erste Gespräche mit den Kollegen des Hessischen Hauptstaatsarchivs Wiesbaden statt, seit 2010 wird DIMAG dort auch eingesetzt. Etwas später kamen auch die Kollegen der Generaldirektion der Staatlichen Archive Bayerns hinzu, und am 22. Februar 2012 konnte die Entwicklungspartnerschaft zwischen den drei Archivverwaltungen auch offiziell besiegelt werden.⁸

⁸ Christian Keitel: Kooperationsvereinbarung unterzeichnet. In: Archivnachrichten 44 (2012), S. 25.

Kooperationen

Das Landesarchiv Baden-Württemberg bietet DIMAG nicht kostenlos an. Was also waren und sind die Rahmenbedingungen für eine Weitergabe von DIMAG?

Klar war von Anfang an, dass sich die Archive, die DIMAG übernehmen wollen, auch an den ursprünglichen Entwicklungskosten beteiligen sollten. Diese Position des Landesarchivs wurde nun durch den neuesten Bericht des Rechnungshofs ausdrücklich bestätigt.⁹ Außerdem benötigt ein Archiv, das in die digitale Archivierung einsteigt, sowohl für diese Aufgabe als auch für den Einsatz von DIMAG einige Hilfestellungen. Diese Hilfestellungen kann das Landesarchiv Baden-Württemberg nicht durch das Stammpersonal leisten, da ansonsten die Pflichtaufgaben des Landesarchivs nicht mehr erfüllt werden können. Es ist daher wichtig, dass die Gegenleistungen das Landesarchiv entweder durch Programmierleistungen oder durch Mittel, mit denen eine Supportstelle aufgebaut werden kann, entlasten.

Archive, die DIMAG übernehmen, sollten eine Reihe fachlicher und technischer Entscheidungen des Landesarchivs Baden-Württemberg zumindest in den Grundzügen auch teilen, da ansonsten die Zusammenarbeit schwer vorstellbar ist. Zu diesen Grundannahmen zählen unter anderem die DIMAG-Basisarchitektur in PHP, die SOAP-Schnittstelle, das Repräsentationenmodell, das Identifier-Modell, das AIP-Modell und eine gemeinsame Terminologie. Schließlich sollten die Partner die erhaltene Software nicht dazu verwenden, eigenständig ein Konkurrenzprodukt zu DIMAG aufzubauen. Mit dem Empfang von DIMAG verpflichten sich die Partner daher, dass sie DIMAG nicht weitergeben, dekompileieren oder unterlizenzieren.

Die ersten beiden Partner, denen das Landesarchiv DIMAG weitergegeben hat, waren die staatlichen Archive in Hessen und Bayern. Die hessischen Kollegen haben sich verpflichtet, als Gegenleistung eine Software zu programmieren, die im Bereich Ingest eingesetzt werden kann. Die bayerischen Kollegen sind dabei, ein vergleichbares Tool für den Access-Bereich zu entwickeln. Zusammen mit den Kolleginnen und Kollegen der beiden Archivverwaltungen sprechen wir dabei von einer Entwicklungspartnerschaft. Die Gegenleistung für die Abgabe von DIMAG besteht in der Entwicklung dieser Softwaretools. Das ursprüngliche Kernmodul und die beiden Tools zusammen bilden dann ein DIMAG-Paket, das an weitere Archive als Ganzes abgegeben werden soll.

Für das Landesarchiv Baden-Württemberg gingen mit den ersten beiden Entwicklungspartnerschaften erhebliche Änderungen einher. Vorher mussten alle fachlichen Überlegungen weitgehend im Alleingang ausgearbeitet werden. Nun profitieren wir erheblich von den Kolleginnen und Kollegen in Wiesbaden und München. Allerdings mussten für den fachlichen Austausch neue Wege gebahnt werden. Die Telefonkonferenz an jedem dritten oder vierten Freitag ist ebenso selbstverständlich geworden wie die zwei bis drei Treffen im Jahr, in denen sich die Teammitglieder der einzelnen digitalen Archive unmittelbar gegenüber sitzen, dasselbe gilt für die gemeinsamen Dokumentablagen und die gemeinsam betriebenen Systeme zu Entwicklung,

⁹ Nutzung und Finanzierung von länderübergreifenden IT-Programmen. In: Rechnungshof Baden-Württemberg, Denkschrift 2012, Karlsruhe o.J. [2012], S. 83–86.

Test und Qualitätssicherung. Klar geworden ist ebenfalls, dass die Weiterentwicklung des technischen Systems DIMAG zunächst eine Weiterentwicklung und Abstimmung der fachlichen Konzepte über Grundannahmen, Prozesse und Prozessschritte bedingt. Diese Entwicklung war und ist sehr positiv zu bewerten. Es lässt sich aber bereits absehen, dass der Kreis der Entwicklungspartner nicht unbegrenzt ausgedehnt werden kann, da ansonsten der Abstimmungsbedarf eine zügige Weiterentwicklung von DIMAG erschweren oder sogar verhindern würde. Es wird daher in Zukunft darauf ankommen, einerseits den Kreis der Entwicklungspartner begrenzt zu halten und zugleich eine Möglichkeit zu schaffen, wie die Anregungen der anderen DIMAG einsetzenden Archive dennoch aufgenommen werden können.

Nur wenige Archive dürften in der Lage sein, ein Archivierungssystem oder einzelne Tools für die digitale Archivierung zu programmieren. Auf der anderen Seite sehen sich zahlreiche Archive vor die Pflichtaufgabe gestellt, mit der digitalen Archivierung zu beginnen. Entsprechend hoch ist die Nachfrage dieser Archive beim Landesarchiv Baden-Württemberg. Mit den Entwicklungspartnern wurde daher vereinbart, weitere Partnerschaftsangebote zu entwickeln. Dabei kann zwischen Supportpartnerschaft, Magazinpartnerschaft und Dienstleistungspartnerschaft unterschieden werden.

Bei der Supportpartnerschaft bekommt das Archiv die Software von DIMAG gegen eine Gebühr. DIMAG wird als Paket, das heißt einschließlich der Module der Entwicklungspartner abgegeben. Das Archiv erhält dafür ein einfaches dauerhaftes Nutzungsrecht sowie über einen vereinbarten Zeitraum (z.B. 3 Jahre) einen definierten Support. Dieser umfasst die Installation der Software beim Nutzer, den Support im Rahmen des bestehenden Programmumfangs für drei Jahre und die Updates, die auf Veranlassung der Entwicklungspartner erfolgen. Nicht enthalten sind die Bereitstellung, Beschaffung oder Anpassung der Hardware, die Anpassung der Schnittstellen (z.B. zum Erschließungssystem), individuelle Funktionswünsche des Nutzers, individuelle Anpassungswünsche des Nutzers und die Erstellung von Backups.

Sowohl die Entwicklungspartnerschaft als auch die Supportpartnerschaft setzen voraus, dass das Archiv DIMAG selbst betreibt. Viele Archive dürften damit voraussichtlich Probleme haben. Die Entwicklungspartner haben daher die Idee der Magazinpartnerschaft entwickelt. Da DIMAG browserbasiert ist, kann es grundsätzlich von jedem Rechner der Welt aufgerufen werden. Diese grundsätzliche technische Eignung wird in der Praxis natürlich durch Schutzmaßnahmen sogleich eingeschränkt. Dennoch bleibt sie für das Angebot der Magazinpartnerschaft wichtig. Zweitens ist DIMAG mandantenfähig, d.h. es können in einer Installation von DIMAG verschiedene voneinander abgeschottete Bereiche eingerichtet werden. Ähnlich den Zimmern in einem großen Haus könnte einer größeren Zahl von Archiven jeweils ein eigenes Zimmer zugewiesen werden. Jedes Archiv hätte dann sein Zimmer, in das es seine Archivalien einstellen kann. Nur das jeweilige Archiv hat Zugang zu seinen Archivalien, die anderen Archive nicht.

Diese beiden Eigenschaften, die Aufrufbarkeit über den Browser und die Mandantenfähigkeit, ermöglichen das Angebot der Magazinpartnerschaft. In dieser legt das Partnerarchiv seine digitalen Archivalien im DIMAG eines Entwicklungspartners ab. Der Zugriff erfolgt über gesicherte Verbindungen (Virtual Private Network, VPN). Als Basiskomponenten sind zunächst die Speicherung der digitalen Archivalien, darüber hinaus aber auch eine zentrale digitale Bestandserhaltung denkbar. Hierzu ein Beispiel: Schon aus Kostengründen würde es Sinn machen, alle

Dateien vom Dateiformat A nach einem einheitlichen Verfahren in das oder die Nachfolge-Dateiformate zu migrieren. Hierfür würde aber auch die mögliche Bündelung der fachlichen Überlegungen sprechen, die mit einer solchen Migration einher gehen müssen. Vergleichbare Effekte sind auch in anderen Bereichen der digitalen Bestandserhaltung zu erwarten. Diese Basiskomponenten könnten bei Bedarf noch um weitere Angebote erweitert werden, z.B. könnte der Entwicklungspartner auch noch Übernahme und Aufbereitung, Einstellung in das DIMAG und die Ausgabe von Benutzungspaketen übernehmen.

Seit 2011 ist das Landesarchiv Baden-Württemberg außerdem noch in Gesprächen mit einem Dienstleister. Der Dienstleister könnte DIMAG in seinem Rechenzentrum installieren und kleineren Archiven dann einen Magazinplatz anbieten. Der eher technische First-Level-Support würde durch den Dienstleister erfolgen, der auch fachliche Second-Level-Support durch das Landesarchiv Baden-Württemberg. Wenn sich die Dienstleisterpartnerschaft realisieren lässt, würde das Landesarchiv keine Magazinpartnerschaft mehr anbieten.¹⁰

Einige Bemerkungen zum Abschluss

Was geschieht eigentlich im Jahr 2012, wenn ein Archiv sich ein Archivierungssystem zulegt? Noch immer gibt es kein System, das alle nach OAIS und anderen einschlägigen Standards für digitale Archive wichtigen Funktionen unterstützt. Bislang kann es auch kein solches System geben, da dafür noch nicht alle Bereiche der digitalen Archivierung konzeptionell durchdrungen sind. Dieser Umstand wirft ein Schlaglicht auf die noch ausstehenden Entwicklungen. Selbst wenn ein Archiv ein Archivierungssystem erworben oder übernommen hat, muss es damit rechnen, dass noch erhebliche Umbauten und Weiterentwicklungen nötig sein werden. Hieraus ergeben sich zwei Konsequenzen. Zunächst erscheint es gerade bei der digitalen Archivierung sinnvoll, die hierfür nötigen Kenntnisse im Archiv selbst aufzubauen und nicht an Andere zu delegieren (z.B. durch Beauftragung von Dienstleistungsfirmen). Sicherlich ist es auf den ersten Blick einfacher, eine Leistung bei einem Dienstleister einzukaufen. Da der Weg zur digitalen Archivierung aber ein langer ist und aus einer ganzen Reihe aufeinander aufbauender Lernprozesse besteht, ist dann zu befürchten, dass der einmal eingekauften Dienstleistung eine ganze Reihe weiterer Dienstleistungen folgen, die ebenfalls eingekauft werden müssen, da dann für eine eigenständige Bewältigung des Problems schlicht die Fachkenntnisse im Haus fehlen.

Problematisch sind auch die in aller Regel sehr teuren Nachprogrammierungen, also Programmierungen, die nach der Abnahme des ursprünglichen Produkts notwendig werden. Für die drei Archive der DIMAG-Entwicklungspartnerschaft erscheint es daher gerade in der derzeitigen Phase der digitalen Archivierung besonders sinnvoll, dass die unvermeidlichen Anpassungen und Verbesserungen durch die Informatiker der drei Archivverwaltungen selbst vorgenommen werden können.

¹⁰ Weitere Angaben zur Dienstleisterpartnerschaft in Christian *Keitel*: Dienstleisterpartnerschaft mit DIMAG, In: Das neue Handwerk. Vorträge des 72. südwestdeutschen Archivtags in Bad Bergzabern. Hg von Kai *Naumann* und Peter *Müller*, erscheint 2013.

Es gibt aber noch weitere Gründe, die aus Sicht der Archive, die ein Archivierungssystem suchen, für DIMAG sprechen könnten. Da ist zunächst die öffentlich-rechtliche Trägerschaft. Bereits jetzt haben drei große Archivverwaltungen ein erhebliches Eigeninteresse an der Weiterentwicklung von DIMAG.¹¹ Damit bestehen sehr gute Aussichten, dass die Software lange Zeit noch genutzt und entwickelt wird. Schließlich besteht digitale Archivierung nur zu einem Teil aus technischen Prozessen und Instrumenten. Mindestens ebenso wichtig sind die fachlichen Abstimmungen und Entwicklungen. Hier bietet eine DIMAG-Partnerschaft gerade kleineren Archiven eine gute Möglichkeit, an dem Fachwissen der staatlichen Archivverwaltungen teilzuhaben.

¹¹ Gestützt wird diese Ansicht auch durch den Umstand, dass in Baden-Württemberg künftig alle Grundakten digital geführt und vom Tag ihrer Entstehung in DIMAG abgelegt werden.

Das digitale Urkundenzwischenarchiv (EL.UZA) in Brandenburg Ein Arbeitsbericht

Von ILKA STAHLBERG und JÖRG HOMBERG

Die Autoren berichten über die Entwicklung und Umsetzung des Projektes zum Aufbau und Betrieb eines Zwischenarchivs für die elektronisch verkündeten Gesetze und Verordnungen des Landes Brandenburg. Sie gehen dabei auf die Rechtsgrundlagen und die verwendeten Standards ein.

EL.UZA – Abgrenzung und Einordnung

Einleitend möchten wir den aktuellen Stand der Entwicklung der in den vergangenen Jahren¹ vorgestellten E-Government-Projekte im Land Brandenburg darstellen. Das brandenburgische Dokumentenmanagement- und Vorgangsbearbeitungssystem (DMS/VBS) EL.DOK (basierend auf dem Produkt Prodea von SER) wird mittlerweile in sieben von zehn Ministerien in unterschiedlichem Umfang, sowie beim Zentralen IT-Dienstleister (ZITBB) und im Brandenburgischen Landeshauptarchiv (BLHA) produktiv genutzt. Somit sind die Teilprojekte Zentraler Betrieb und Zentrale Schulung, sowie die Einführungsprojekte im Innenministerium, Wissenschaftsministerium, Staatskanzlei, Wirtschaftsministerium, Arbeits- und Sozialministerium, Bildungsministerium abgeschlossen worden. Das Teilprojekt Landesreferenzmodell (DMS-Basis) arbeitet weiter und bereitet die technische Abnahme vor. Vordringlich sind dabei die Implementierung der XDOMEA-Schnittstelle 2.1.0 für den mandantenübergreifenden Austausch von Dokumenten, Vorgängen, Zeichnungsmappen und Akten und die Aussonderung in das Brandenburgische Landeshauptarchiv. Das auf dem Projekt EL.DOK aufsetzende Kabinetttinformationssystem ELKIS hat den Produktivbetrieb ebenfalls aufgenommen. Der Landtag unternimmt mit seinen Bemühungen zur Einführung eines DMS/VBS einen neuen Anlauf. Das Projekt einer zentralen digitalen Altregistratur E.LAS, für das das Fachkonzept in Schwerin 2011 vorgestellt wurde,² ist aus finanziellen Gründen vorläufig zurückgestellt worden.

¹ Zuletzt Ilka *Stahlberg*: Die Konzipierung eines digitalen Zwischenarchivs in Brandenburg und der Sachstand bei der DMS-Einführung in der Brandenburgischen Landesverwaltung. Ein Arbeitsbericht. In: Auf dem Weg zum digitalen Archiv. Stand und Perspektiven von Projekten zur Archivierung digitaler Unterlagen. 15. Tagung des Arbeitskreises *Archivierung von Unterlagen aus digitalen Systemen* vom 2./3. März 2011 in Schwerin. Hg. von M. *Manke*, S. 47f.

² *Stahlberg*, wie Anm. 1, S. 48 ff.

Die Anwendungen EL.Norm³ (ausschließlich elektronische Ausfertigung von Gesetzen und Verordnungen) und EL.Archiv⁴ – das digitale Archiv des BLHA – arbeiten im Produktivbetrieb. Die Rechtsgrundlage für die in der Anwendung EL.Norm seit 2010 erfolgende elektronische Ausfertigung und Verkündung von Gesetzen und Verordnungen im Land Brandenburg bildet das Gesetz über die elektronische Ausfertigung und Verkündung von Gesetzen und Rechtsverordnungen des Landes Brandenburg (Brandenburgisches Ausfertigungs- und Verkündigungsgesetz- BbgAusVerkG) vom 18. Dezember 2009. Nach dessen § 1 Abs. 4 sind *ausgefertigte Gesetze und Rechtsverordnungen ... mitsamt den zugehörigen Signaturen⁵ dauerhaft und unveränderlich zu archivieren und gemäß § 5 Bereithaltung des Gesetz- und Verordnungsblattes für das Land Brandenburg sind nach Abs. 1 die zum Abruf bereitgestellte Ausgaben des Gesetz- und Verordnungsblattes für das Land Brandenburg ... unter der in § 2 Absatz 1 genannten Adresse⁶ dauerhaft unverändert zugänglich zu halten.*⁷

In Umsetzung seines Archivierungsauftrages hat das BLHA deshalb mit dem Ministerium der Justiz des Landes Brandenburg (MdJ) am 18. Mai 2011 eine Verwaltungsvereinbarung über die Einrichtung und den Betrieb eines elektronischen Urkundszwischenarchivs (EL.UZA) in dem elektronischen Archivsystem EL.ARCHIV des BLHA geschlossen. Das BLHA richtet danach im Auftrag des Justizministeriums das elektronische Urkundszwischenarchiv ein und betreibt es auf der Grundlage des § 5 Abs. 5 in Verbindung mit § 14 und § 4 Abs. 8 BbgArchivG in eigener Zuständigkeit.⁸ Mit dieser Lösung wurde auch den Verzögerungen beim Aufbau eines digitalen brandenburgischen Zwischenarchivs (E.LAS) Rechnung getragen.

Insofern ist EL.UZA formal gesehen eine modulare Erweiterung des im Land Brandenburg eingesetzten Fachverfahrens zur elektronischen Langzeitarchivierung EL.ARCHIV um eine spezifische Zwischenarchiv-Komponente. Es inkludiert eine Schnittstelle zwischen dem Verfahren EL.NORM zur Verkündung und Veröffentlichung von Normen und Rechtsvorschriften und dem Verfahren EL.ARCHIV zur Langzeitarchivierung von archivwürdigem Schriftgut. EL.UZA ist ein Sonderfall eines Zwischenarchivs mit einer Langzeitarchivierungskomponente, da die Gesetze und Normen einerseits potentiell unbegrenzte Aufbewahrungsfristen haben können, die zum Zeitpunkt der Zwischenarchivierung nicht bekannt sind, und sie andererseits bereits mit der Zwischenarchivierung als archivwürdig bewertet sind. Im Unterschied zu einer klassischen

³ Simone *Stumpe*: IMTB Consulting: EL.NORM – die Elektronische Normverkündung im Land Brandenburg. In: Entwicklung in den Bereichen Records Management/Vorarchiv – Übernahme – Langzeitarchivierung. 13. Tagung des Arbeitskreises *Archivierung von Unterlagen aus digitalen Systemen* Staatsarchiv, St. Gallen 27./28. April 2009.

⁴ Zuletzt *Jörg Homberg*: Umsetzung des OAIS-Konzepts im Brandenburgischen Landeshauptarchiv. Ein Arbeitsbericht. In: Auf dem Weg zum digitalen Archiv, wie Anm. 1, S. 43ff.

⁵ Gemeint sind elektronische Signaturen im Sinne des Authentizitätsbeweises, nicht archivische Bestellnummern.

⁶ <http://www.landesrecht.brandenburg.de> (alle Links wurden am 28.9.2012 überprüft).

⁷ http://www.bravors.brandenburg.de/sixcms/detail.php?gsid=land_bb_bravors_01.c.48696.de.

⁸ Gesetz über die Sicherung und Nutzung von öffentlichem Archivgut im Land Brandenburg (Brandenburgisches Archivgesetz- BbgArchivG) vom 07. April 1994, Gesetz- und Verordnungsblatt des Landes Brandenburg Teil I/1994 S. 94 ff., http://www.bravors.brandenburg.de/sixcms/detail.php?gsid=land_bb_bravors_01.c.15130.de.

Zwischenarchivlösung wird kein Bewertungsworkflow benötigt, aber es müssen Mechanismen der nachträglichen Übermittlung von Aufbewahrungsfristen eingebaut werden. Die Software HT.DIVAS wurde von der Firma H&T Greenline GmbH entwickelt.⁹

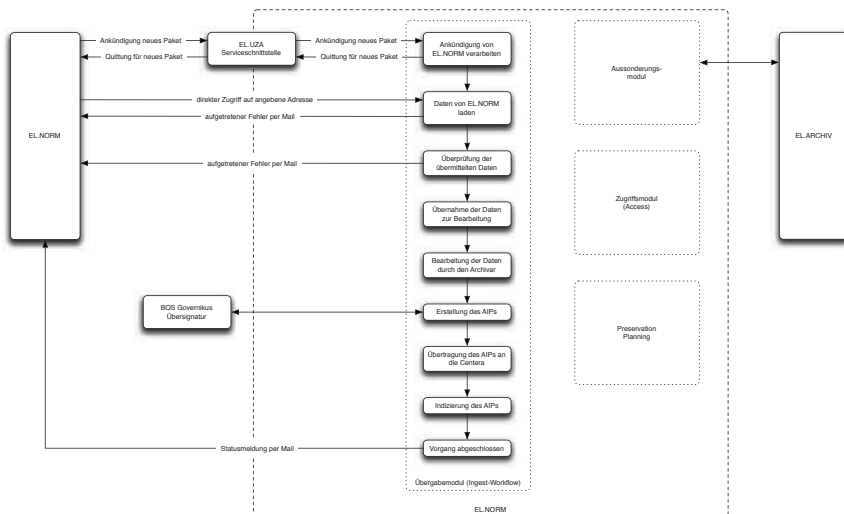
EL.UZA – Wechselwirkung

Die aus dem Verfahren EL.Norm nach EL.UZA übergebenen Normen werden jeweils in zwei Versionen übergeben. Die Originalnorm ist das digital signierte Gesetz beziehungsweise die Verordnung. Die Ausgabenorm entspricht der amtlichen Drucksache, dem Gesetz- und Verordnungsblatt Teil I und II. Nach Implementierung der XDOMEA-Schnittstelle im Verfahren EL.Norm soll auch die Verkündungsakte¹⁰ im Landes-DMS EL.DOK geführt und über die dortige Schnittstelle ausgesondert werden.

Die Nutzung des Systems als Zwischenarchiv für weitere Mandanten und Anwendungen ist perspektivisch ebenfalls möglich.

EL.UZA – Funktionsweise

Im Folgenden sei ergänzend zur beigefügten Abbildung dargestellt, wie die OAIS-Prozesse Ingest, Archival Storage und Access bei EL.UZA funktionieren.



⁹ <http://www.htgreenline.de>.

¹⁰ Die Verkündungsakte enthält – heute noch in Papierform – das Verwaltungshandeln um die Verkündung: Anschreiben, Bestätigungen, Nachfragen etc. Es handelt sich um Redaktionsunterlagen, die nicht zu verwechseln sind mit der Akte zur Erarbeitung der Norm, die im Fachressort entsteht.

Ingest

Der Ingest unterscheidet sich je nachdem, ob es um eine Originalnorm oder um eine Ausgabenorm geht.

Originalnorm (ON): Die signierten und validen Pakete (SIPs) zur Zwischenarchivierung werden dem BLHA aus dem Fachverfahren EL.NORM im XDOMEA-Containerformat mit inkludierten PDF/A-Dokumenten (wahlweise via Webservice oder E-mail) angeboten, daraufhin wird aus EL.NORM ein Download-Link generiert. EL.UZA bestätigt anschließend den Empfang durch synchrone Übermittlung einer Upload-ID, lädt zeitgesteuert die Pakete herunter und validiert sowohl die Containerintegrität als auch die Formattauglichkeit.

Ausgabenorm (AN): Diese werden ohne Pflege der Signatur über die Schnittstelle von EL.UZA direkt dem digitalen Archiv (EL.ARCHIV) zugeführt.

Die Zusammenhänge zwischen AN und ON bleiben über Metadaten (Archiv-ID, Archivsignatur, Upload-ID) ersichtlich.

Archival Storage

Nach erfolgreicher Archivierung wird EL.NORM asynchron eine Archiv-ID (beziehungsweise Archivsignatur der Originalnorm) übermittelt. Die SIPs werden mit Metadaten angereichert; aus SIPs werden AIPs. Die ArchivID dient der späteren Kommunikation (zum Beispiel Anstoß der Aussonderung nach Ablauf der Aufbewahrungsfrist – das heißt nach Ablauf der Gültigkeit des Gesetzes).

Access

Der Zugriff auf die *Originalnorm* erfolgt akkreditiert über Webclients. Die Software verhindert einen parallelen Zugriff auf das angefragte Objekt bzw. auf Instanzen des Objekts, um sicherzustellen, dass nur jeweils ein Nutzer zeitgleich das Original liest. Der Zugriff auf eine Ausgabenorm erfolgt in EL.UZA ebenfalls webbasiert, unterstützt über flankierende Erschließungsdaten. Primärer Nutzungskanal für Verwaltung und Öffentlichkeit ist aber das Brandenburgische Vorschriftensystem BRAVORS.¹¹

Die DIPs unterscheiden sich also bei Originalnorm und *Ausgabenorm* hinsichtlich des Erschließungsgrades über Metadaten und hinsichtlich der elektronischen Signatur. Als Viewer-Plugin wird analog zum Verfahren El.Norm der jeweils neueste funktionierende Adobe Reader genutzt. Eine mit der Ausgangsdarstellung auf Dauer identische Darstellung war eine wichtige Anforderung der Auftraggeber.

¹¹ <http://www.landesrecht.brandenburg.de>.

EL.UZA und OAIS

Die Software ist als Appliance¹² angelegt und damit sowohl unabhängig von Plattformen als auch vom Versionsstand etwaiger Runtime-Umgebungen (JAVA, .NET etc.). EL.UZA bedient sich gängiger langzeitarchivtauglicher Formate (z.B. PDF/A) als auch Transportmechanismen (derzeit XDOMEA 2.1). EL.UZA bringt unabhängig vom korrespondierenden Fachverfahren der anbietenden Instanz eigene Validierungskomponenten für Formate (PDF/A) und Container (XDOMEA) mit.¹³ Die Komponente kann sowohl automatisch über Interfaces als auch manuell angesprochen werden.

Die elektronischen Signaturen der Originalnormen (ON) werden gemäß Einschätzung von Sicherheit und Verlässlichkeit nach BSI¹⁴-Vorgaben erneuert (Übersignierung). Die Übersignatur erfolgt mit Hilfe von Governikus¹⁵ welches für EL.UZA entsprechend angepasst wurde. Nach Überführung der ON (z. B. nach Ablauf der Gültigkeit eines Gesetzes) aus EL.UZA in das Endarchiv EL.Archiv werden die elektronischen Signaturen nicht mehr gepflegt.

Die Speicherung erfolgt auf mehrfach redundanten CENTERA-Systemen der Firma EMC². Über Leasing-Verträge mit dem Hersteller ist gewährleistet, dass jeweils das neueste Hardware-Release verwendet wird. Eine Migration auf andere Systeme ist jederzeit möglich. Die Wahl des Storage-Systems CENTERA, also einer Content-Adress-Technologie, bringt einige maßgebliche Vorteile mit:

- sehr niedriger administrativer Aufwand (Blackbox)
- Planungssicherheit aufgrund von Skalierbarkeit
- völlige Unabhängigkeit von Plattformen; lediglich die Datenelemente (Binary Large Objects – BLOBs) und die in Form eindeutiger Inhaltshashwerte (C-Clips) generierter Adressinformationen werden benötigt. Etwaige Systemstrukturelemente wie Filesysteme, Verzeichnissysteme oder ähnliches sind nicht notwendig.
- Nachhaltigkeit: Technologietransfersicherheit aufgrund langfristiger budgetierter Leasing-respektive Wartungsverträge
- offengelegte API EL.UZA – CENTERA
- Zukunftssicherheit durch XAM (Extensible Access Method)-Zugriffsmethoden; d.h.: Austauschbarkeit von Informationen zwischen XAM-kompatiblen Archivsystemen ist gegeben.

¹² Appliances sind in diesem Zusammenhang in sich geschlossene Softwarepakete, die in jede handelsübliche Plattform als virtuelle Maschinen (VM) eingehängt werden können.

¹³ Validierungskomponente für DOMEA: Eigenentwicklung von H&T Greenline, für PDF/A: Produktivlizenz callas pdfaPilot SDK für Windows und Callas pdfaPilot2-GOV.

¹⁴ Bundesamt für Sicherheit in der Informationstechnik.

¹⁵ bremen online services GmbH & Co. KG (bos KG), <http://www.governicus.com>.

EL.UZA und XDOMEA¹⁶

Für die Übergabe/Übernahme der Normen aus EL.Norm nach EL.UZA wurde die Nutzung der XDOMEA-Schnittstelle 2.1.0. festgelegt und dafür u.a. folgende Festlegungen für die Datenübergabe getroffen:

- Es werden ausschließlich Dateien in den Formaten PDF/A und XML archiviert. Daher werden die Rohfassung (*.doc) und Druckversion der Ausgabenorm (PDF/X-3) nicht ans Archiv übergeben.
- Jede Norm wird in zwei Zip-Dateien ans Archiv übergeben: jeweils ein Zip-Archiv für die Originalnorm (ON) und ein Zip-Archiv für die Ausgabenorm (AN).
- Zusätzlich zu den Dateien der Norm werden technische Metadaten / Revisionsdaten in einer XML-Datei übergeben. Verabredet wurde als Name der Datei: Ausgabenorm.xml bzw. Originalnorm.xml. Diese Datei darf nicht in der XDOMEA-Datei referenziert sein. Außerdem werden die EL.Norm-Metadaten als metadata.xml in die Zip-Archive integriert und standardkonform in der XDOMEA-Datei referenziert.

Für die XDOMEA-Datei wurden die folgenden Strukturen festgelegt und Namenskonventionen getroffen:

- Der Dateiname der XDOMEA-Datei setzt sich aus <uuid>_<Aussonderungsart>.xml (Aussonderungsart: „Aussonderung.Aussonderung0503“) zusammen.
- Innerhalb der XDOMEA-Datei werden Normen in Akten gruppiert. Für die Akten werden folgende Bezeichner verwendet:
 - I_Gesetze_Originalnorm_<Verkündungsjahr>
 - I_Gesetze_Ausgabenorm_<Verkündungsjahr>
 - II_Verordnungen_Originalnorm_<Verkündungsjahr>
 - II_Verordnungen_Ausgabenorm_<Verkündungsjahr>
- Hierbei legt das Datum der Verkündung das Verkündungsjahr fest.
- Innerhalb der jeweiligen Akte werden die Normen als Vorgänge, mit dem Titel der Norm als Bezeichner, angelegt. Der Vorgang referenziert die jeweiligen Dokumente zur Originalnorm bzw. Ausgabenorm, sowie die Metadaten zur Norm.

Beispiel für die Dateien eines Vorganges der Ausgabenorm:

Hauptdokument + optionale Prüfdateien:

<Name>__an.pdf (Hauptdokument)
sowie ggf. <Name>__an_report.pdf, <Name>__an_pdf1a.pdf

¹⁶ <http://www.xoev.de/sixcms/detail.php?gsid=bremen83.c.2439.de>.

Anlagen + optionale Prüfdateien:

<Name>__an__anlg__*.pdf

Metadaten XML:

metadaten.xml (ELNORM-Metadaten)

XDOMEA XML:

<uuid>_Aussonderung.Aussonderung.0503.xml

Revisionsdaten XML:

Ausgabenorm.xml

(technische Metadaten, unter anderem Angabe der Konvertierungssoftware, der Wert für die Aufbewahrungsdauer im XML wird leer gelassen).

EL.UZA wurde im Juli 2012 in den Produktivbetrieb überführt. Derzeit erfolgen die Übernahmen der Alt-Daten aus EL.NORM (Jahrgänge 2009, 2010, 2011). Zusätzlich wird im Sinne der Standardisierung der Generierung langzeitarchivtauglicher Formate im Land Brandenburg eine Konsolidierung der PDF/A-Generatoren im Ministerium der Justiz und des Landtags durchgeführt.

Pilotprojekt zur Langzeitarchivierung digitaler E-Mail-Korrespondenzen des Bundesvorstandes der Vereinten Dienstleistungsgewerkschaft ver.di

Von MIKE ZUCHET

Der Einsatz von E-Mails für die interne und externe Kommunikation ist heute so selbstverständlich wie die früher übliche Umlaufmappe oder der klassische Brief in Papierform. Glaubt man verschiedenen Studien, die sich mit der Relevanz der E-Mail als Kommunikationsmittel in Unternehmen und Organisationen auseinandergesetzt haben, liegen zwischen 35 und 75 % aller relevanten Informationen nur noch in dieser digitalen Form vor und finden keinen papierernen Ausdruck mehr.¹ Aus diesem Grund und weil der Gesetzgeber fordert, dass steuerlich relevante E-Mails sechs resp. zehn Jahre revisionssicher vorgehalten werden müssen, liegt das Thema der Langzeitarchivierung digitaler E-Mail-Korrespondenzen mehr als auf der Hand.² Überdies werden dezidierte E-Mail-Server durch die Auslagerung in entsprechende Langzeitspeicher entlastet und der in der Regel hohe Aufwand bei der Wiederherstellung gelöschter E-Mails wird vermieden.

Das Archiv der sozialen Demokratie (AdsD) der Friedrich-Ebert-Stiftung, das sich seit längerer Zeit mit der Langzeitarchivierung digitaler Überlieferungen befasst, nahm dies zum Anlass, sich intensiv mit diesem Thema auseinander zu setzen. Es handelte sich im wahrsten Sinne des Wortes um Pionierarbeit und das AdsD darf dabei für sich in Anspruch nehmen, als erstes

¹ Vgl. dazu <http://www.silicon.de/39193609/e-mails-sind-das-wichtigste-in-deutschen-unternehmen/> (alle Links in diesem Text wurden am 31.5.2012 geprüft).

² Vgl. dazu Abgabenordnung (AO) § 146 *Ordnungsvorschriften für die Buchführung und für Aufzeichnungen*: http://www.gesetze-im-internet.de/ao_1977/__146.html und § 147 *Ordnungsvorschriften für die Aufbewahrung von Unterlagen*: http://www.gesetze-im-internet.de/ao_1977/__147.html; Grundsätze ordnungsgemäßer DV-gestützter Buchführungssysteme (GoBS), Abschnitte 5 Datensicherheit und 6 Dokumentation und Prüfbarkeit: http://www.bundesfinanzministerium.de/nr_314/DE/BMF__Startseite/Service/Downloads/Abt_IV/BMF__Schreiben/015,templateId=raw,property=publicationFile.pdf; Handelsgesetzbuch (HGB) § 257 *Aufbewahrung von Unterlagen/Aufbewahrungsfristen*: <http://www.gesetze-im-internet.de/bundesrecht/hgb/gesamt.pdf>; Grundsätze zum Datenzugriff und zur Prüfbarkeit digitaler Unterlagen (GDPdU), Abschnitte II *Prüfbarkeit digitaler Unterlagen* und III *Archivierung digitaler Unterlagen*: http://www.bundesfinanzministerium.de/nr_95356/DE/Wirtschaft_und_Verwaltung/Steuern/Veroeffentlichungen__zu__Steuerarten/Abgabenordnung/Datenzugriff_GDPdU/002,templateId=raw,property=publicationFile.pdf. Siehe dazu auch die Merksätze zur revisionssicheren elektronischen Archivierung des Verbands Organisations- und Informationssysteme e.V. (VOI): <http://www.voi.de>.

Archiv in Deutschland über seine Aktivität in diesem Bereich zu berichten.³ Eine server- oder clientgesteuerte Archivierung⁴ wurde von vorneherein ausgeschlossen. Stattdessen wurde im AdsD großen Wert darauf gelegt, die Archivierung außerhalb der Ursprungsumgebung durchzuführen sowie E-Mails und die dazugehörigen Metadaten für eine *systemunabhängige* Umgebung aufzubereiten und vorzuhalten. Neben der einseitigen Produktbindung sprechen noch weitere Gründe gegen die (Langzeit-)Archivierung mittels MS Outlook, das heißt als PST-Datei. Neben der Größenbeschränkung der zentral verwaltenden PST-Datei – näheres siehe weiter unten – sind große Dateien überproportional stark von einem Ausfallrisiko betroffen und die Datensicherung, die größtenbedingt längere Zeit in Anspruch nimmt, blockiert das Produkktivsystem. Hinzu kommt, dass die Recherche nach einzelnen E-Mails eine langwierige Angelegenheit werden kann, falls mehrere PST-Dateien durchsucht werden müssen, und dass der Speicherort auf der lokalen C-Partition des Anwenderrechners liegt, die in der Regel nicht redundant gesichert wird.

Zusammen mit dem ver.di-Bundesvorstand wurde ein Pilotprojekt gestartet, das die Machbarkeit der Übernahme, Aufbereitung und des Zugriffs digitaler E-Mail-Korrespondenzen nach dem Open Archival Information System (OAIS, ISO 14721:2003⁵) im AdsD überprüfen sollte. Nach intensiven Gesprächen, die in erster Linie als ‘vertrauensbildende Maßnahmen’ dienten, und der Ausarbeitung einer speziellen Datenschutzerklärung seitens ver.di konnte das AdsD im Oktober 2010 die gesamte vorliegende E-Mail-Korrespondenz der Leitung der Abteilung Grundsatz und im November darauf des Büros des Bundesvorsitzenden übernehmen.⁶ Bei dem verwendeten E-Mail-Programm handelte es sich um Microsoft Outlook 2007. Im ersten Fall wurden ca. 14 000 E-Mails übernommen, im zweiten Fall ca. 9 000 E-Mails. MS Outlook

³ Der Aufsatz von Beda Kupper ist dem Verfasser erst nach Fertigstellung dieses Texts bekannt geworden: Beda Kupper: E-Mail-Archivierung. In: *Actualité archivistique Suisse. Archivwissenschaft Schweiz aktuell*. Hg. von G. Coutaz, N. Meystre-Schaeren, B. Roth-Lochner und A. Steigmeier. Baden (CH) 2008. S. 88–117. In vielerlei Hinsicht stimmen die Auffassungen des Verfassers mit denen Koppers überein, allerdings vertritt letzterer hinsichtlich des Punktes *Erscheinungsbild* (vgl. S. 101) einen diametralen Standpunkt, der der anerkannten Forderung nach Erhalt der strukturellen Authentizität nicht gerecht wird. Der konkrete äußere Anlass für die Pilotierung ist die Einführung eines Dokumentenmanagementsystems (DMS) beim Bundesvorstand der Vereinten Dienstleistungsgewerkschaft ver.di. Dieser trat an das AdsD mit der Frage heran, ob Möglichkeiten existieren, die bislang entstandenen E-Mail-Korrespondenzen außerhalb des zukünftigen DMS zu archivieren. Wie oben erwähnt, ergibt sich dabei auch der positive *Nebeneffekt*, dass die E-Mail-Server bei ver.di entlastet würden.

⁴ Beim serverseitigen Ansatz, dem Journaling, werden sämtliche eingehende wie ausgehende E-Mails sofort in das systemeigene Archivsystem transferiert, weshalb der Speicherplatz entsprechend hoch bemessen sein muss. Vorteil: Vollständigkeit, Nachteil: Speicherbedarf, digitaler *Beifang* (Spam). Beim clientseitigen Ansatz entscheidet der Anwender selbst, welche E-Mails im systemeigenen Archivsystem archiviert werden. Vorteil: Auswahl/Flexibilität, Nachteil: möglicher Datenverlust.

⁵ Siehe dazu http://nestor.sub.uni-goettingen.de/handbuch/artikel/nestor_handbuch_artikel_366.pdf.

⁶ Unmittelbar verknüpft mit der Übernahme der E-Mail-Korrespondenz ist die Frage der Bewertung. Die Bewertung wurde auf der Ebene des jeweiligen E-Mail-Ordners durchgeführt, nicht auf der Ebene der einzelnen E-Mail. Sämtliche als *privat* deklarierten E-Mail-Ordner wurden nicht übernommen. An dieser Stelle wurde wieder besonders deutlich, wie wichtig bereits die Arbeit im vorarchivischen Umfeld der Langzeitarchivierung digitaler Überlieferungen ist.

2007 – wie auch die Nachfolgeversion – legt sowohl die eigentliche Nachricht (E-Mail⁷) als auch die Anhänge in einer proprietären Personal Storage-Datei (PST) ab, die je nach Mail-Volumen mehrere Gigabyte (GB) groß werden kann.⁸ Es sei an dieser Stelle angemerkt, dass das erarbeitete Konzept nicht nur dazu in der Lage ist, E-Mail-Korrespondenz aus der besagten Programmumgebung zu übernehmen und aufzubereiten, sondern aus nahezu allen gängigen E-Mail-Programmen.⁹

Nach der Überführung der PST-Dateien in das AdsD wurden zuerst Kopien der Ursprungsdateien auf mehreren voneinander unabhängigen Datenträgern erstellt. Um die Datenintegrität bei den Kopierschritten nachhalten zu können, wurde ein Prüfsummenverfahren (SHA-1; *secure hash algorithm*) eingesetzt. Da sämtliche Transferschritte fehlerfrei verliefen, wurden die PST-Dateien in ihrer Entstehungsumgebung (gemeint sind sowohl Programm als auch Betriebssystem) geöffnet, mit anschließendem inkrementellem Abgleich.¹⁰ Danach wurde jede zu archivierende E-Mail mithilfe eines eigens im AdsD entwickelten Makros aus MS Outlook 2007 exportiert und als MSG-Datei abgelegt.¹¹ In Anbetracht der Funktionen und Importmöglichkeiten des Datenbanksystems Faust 6.0 Professional, das derzeit im AdsD eingesetzt wird, wurde das Makro speziell dahin gehend programmiert, dass der Name jeder MSG-Datei aus einer fortlaufenden Nummer, dem Datum und der Uhrzeit des Eingangs resp. des Ausgangs in das E-Mail-Programm generiert wurde.¹² Als Nächstes wurde mithilfe des Tools Solid PDF Tools

⁷ Hauptbestandteile einer E-Mail sind der Header, Body (Nachrichtentext) und Anhang.

⁸ In der PST-Datei werden sämtliche E-Mail-Vorgänge (Eingang, Ausgang, Gesendet, Entwürfe), Termine, Aufgaben, Notizen, Journaleinträge, Kontakte (ausgenommen: Adressen des Persönlichen Adressbuches (PAB)) sowie alle Ordner und deren strukturelle Gliederungen gespeichert. In MS Outlook 2003 und 2007 liegt die Standardgrößenbeschränkung bei PST-Dateien bei 20 GB. Vorkonfiguriert ist MS Outlook 2010 für eine Größe der PST-Datei von 50 GB, die allerdings durch Änderungen in der Windows-Registry erhöht oder reduziert werden kann. Siehe dazu <http://support.microsoft.com/kb/982577/de>.

⁹ Derzeit wird die Übernahme digitaler E-Mail-Korrespondenzen eines Gewerkschaftssekretärs der ehemaligen Gewerkschaft Öffentliche Dienste, Transport und Verkehr (ÖTV) resp. der Vereinten Dienstleistungsgewerkschaft ver.di vorbereitet, der mit einem proprietären E-Mail-Programm eines Internetproviders arbeitet. Nach dem Transfer der E-Mail-Korrespondenz des Hinterlegers an den Mail-Server der FES werden diese an MS Outlook 2010 übergeben und wie im Text beschrieben aufbereitet. Da es bei diesem Ablauf zu entsprechenden Einträgen in den jeweiligen E-Mail-Headern kommt, werden diese entsprechend gekennzeichnet, um den letzten Transferschritt ins AdsD von den vorherigen differenzieren zu können.

¹⁰ Die bisherige Praxis hat gezeigt, dass E-Mails bereitgestellt werden, die bereits bei einer vorherigen Übernahme übernommen wurden. Da auch diesem Wege der Grad der Redundanzen sukzessive zunahme, verfolgt das AdsD den Ansatz des inkrementellen Archivierens. Dabei werden die zu übernehmenden E-Mails mit den bereits übernommenen E-Mails automatisch abgeglichen und nur der Zuwachs kommt für die Langzeitarchivierung infrage.

¹¹ Derzeit wird daran gearbeitet, den Zwischenschritt der MSG-Datei zu überspringen und die E-Mails direkt als PDF resp. PDF/A-Datei zu generieren.

¹² Der Aufbau des Dateinamens entspricht folgendem Schema: nn mm.dd.yyyy hh.mm.ss, z.B. 01.12.2004 07.35.55.

V6 jede MSG-Datei in eine PDF/A-Datei (PDF/A-1b)¹³ konvertiert, um Aufbau und Struktur der E-Mails *einzufrisieren* und somit die informationelle und strukturelle Authentizität der digitalen Quelle zu gewährleisten.¹⁴ In einigen wenigen Fällen war die PDF/A-Konvertierung nicht möglich, da der Inhalt der Ausgangsdatei nicht konform war mit den Spezifikationen der ISO-Norm. In diesem Falle fand lediglich eine Konvertierung nach PDF (1.4) statt. Diese Ausnahmen wurden vom o.g. Konvertierungsprogramm protokolliert und fanden Eingang in die dazugehörigen Metadaten.

Jede PDF/A-Datei wurde mit einer Prüfsumme versehen, um wie im Falle der PST-Datei Änderungen festzustellen und die Datenintegrität langfristig zu gewährleisten. Da es nicht nur um die E-Mail ging, sondern auch um den jeweiligen Anhang – oftmals diente die E-Mail nur als *Transportmittel* für eine zu übermittelnde Datei –, wurden die Anhänge aus der PST-Datei extrahiert und in das Langzeitarchivierungsformat PDF/A konvertiert, soweit dies technisch möglich und sinnvoll war. War die Level B-Konformität nicht zu erreichen, wurde zumindest versucht, die Dateianhänge in das PDF-Format (1.4) zu konvertieren. Schlug auch dieser Schritt aus inhaltlich-technischen Gründen fehl, wurden die Dateianhänge im Ausgangsformat vorgehalten. Redundanzen innerhalb der identische Dateianhänge wurden mit Hilfe von Prüfsummen erkannt, Dubletten werden vor der Formatkonvertierung gelöscht.¹⁵ Da jede Langzeitarchivierung digitaler Überlieferungen mit der quantitativen und qualitativen Existenz von Metadaten steht oder fällt, war es nun notwendig, auf die Metadaten der einzelnen E-Mails zuzugreifen, sie aus der proprietären PST-Hülle zu extrahieren und sie in ein programmunabhängiges Dateiformat zu überführen.¹⁶ Überdies werden auch die Metadaten des E-Mail-Headers (Internetkopfzeilen) übernommen, da sie Aufschluss über Absender und Weg einer E-Mail geben und bei der Über-

¹³ PDF/A ist ein Subset des verbreiteten Portable Document Format (PDF) und liegt seit 2005 als ISO-Norm vor (ISO 19005-1:2005). Die ISO-Norm beschreibt zwei Konformitätsebenen: A. PDF/A-1b – Level B (Basic) conformance, Ziel: eindeutige visuelle Reproduzierbarkeit. B. PDF/A-1a – Level A – (Accessible) conformance, Ziel: eindeutige visuelle Reproduzierbarkeit sowie Textabbildung nach Unicode und inhaltliche Strukturierung des Dokuments. Im Verlauf des Pilotprojekts hat sich Level B als ausreichende und zuverlässige Variante erwiesen, die auch nach der Pilotphase verwendet wird.

¹⁴ Falls der Text aus den E-Mails lediglich extrahiert und in das Faust-Datenbanksystem überführt worden wäre – auch über den Zwischenschritt einer entsprechenden XML-Metadatendatei –, wären sämtliche Textformatierungen und -strukturierungen verloren gegangen. Die jeweilige Nachricht hätte sich als unübersichtlicher Fließtext dargestellt, was einem wesentlichen Verlust der sog. *significant properties* entspricht.

¹⁵ Handelt es sich dabei zweifellos um einen wichtigen Aspekt der Bewertung, ist dieser Vorgang auch unter dem Gesichtspunkt der Speicherökonomie interessant. Sind mehrere E-Mails mit Dateien verknüpft, die ausweislich ihrer Prüfsumme identisch sind, so werden die Dateien in der Datenhaltung auf eine reduziert. Die Verknüpfung der gelöschten Dateien wird umgelenkt, sodass der Nutzer der archivierten E-Mails keinen Unterschied bemerkt, Administratoren können den Vorgang aber nachvollziehen. Bei Dateien, die unterschiedliche Namen tragen, aber den identischen Inhalt aufweisen, hat dies jedoch Auswirkungen auf die Metadaten, d. h. auf die Referenz. In diesem Falle trägt die Referenz als letzte Pfadangabe nicht den ursprünglichen Dateinamen, sondern den Namen der jeweiligen übrig gebliebenen Datei. Auch in diesem Falle ist ein entsprechender Vermerk in der Datenbank zu finden, der den Namen der gelöschten Datei beinhaltet.

¹⁶ Folgende Metadaten wurden extrahiert: Dringlichkeit, Betreff, Absender, Empfänger, Kopie an, Erhalten am, Erstellt am, Nachrichtentext, Name(n) der Anlage(n).

prüfung der Authentizität gegebenenfalls zusätzlich herangezogen werden können.¹⁷ Die Wahl für das Dateiformat fiel auf das XML-Format, nicht zuletzt deshalb, weil Faust 6.0 professional über entsprechend konfigurierbare Importfilter verfügt und es sich bereits bei früheren Datenimporten als zuverlässiges Austauschformat bewährt hatte.

Sowohl die überlieferten E-Mails der erwähnten Abteilung Grundsatz als auch die des Büros des ver.di-Bundesvorsitzenden lagen größtenteils strukturiert in entsprechenden E-Mail-Ordern vor, die eine Mischung aus Provenienz- und Pertinenzmerkmalen aufwiesen. Die jeweiligen (Ablage-)Strukturen wurden übernommen und mit Hilfe von Thesauri in Faust-Datenbanken nachgebildet, was der Ordnung und dem schnellen Zugriff diene.

Nach Abschluss der aufgeführten Arbeitsschritte lagen sämtliche Komponenten vor, um sie in das Archivdatenbanksystem Faust 6.0 professional zu überführen. Als Erstes wurden die PDF/A-Dateien, die als authentische E-Mail-Abbilder fungierten, sukzessive importiert, wobei nicht die PDF/A-Datei selbst Bestandteil der Datenbank wurde, sondern nur die Referenz zum jeweiligen Speicherort. Als Nächstes wurden die dabei generierten Datensätze um die dazugehörigen Metadaten ergänzt. Damit dieser Schritt reibungslos funktionierte und es zu keinen falschen Verknüpfungen von E-Mail-Abbildern und Metadaten kam, war es wichtig, dass das oben erwähnte Makro an besagter Stelle eingesetzt wurde. Die vorliegenden Dateianhänge wurden von einer separaten Faust-Datenbank erfasst. Beide Datenbanken – a) Referenz zum E-Mail-Abbild samt Metadaten und b) technische Metadaten und Referenz zu den Dateianhängen – wurden über Assoziativ-Referenzen miteinander verknüpft.¹⁸ Sie ermöglichen es einerseits, dass der jeweilige Dateianhang vom Faust-Datensatz aus, der das E-Mail-Abbild repräsentiert, angesteuert und anschließend angesehen werden kann. Andererseits ist es ebenso möglich, von einem recherchierten Anhang zur entsprechenden E-Mail zu gelangen. Das beschriebene Prinzip – ein zentraler Datensatz, von dem entsprechende Verweise ausgehen – ist auf eine Vielzahl von Datenbanksystemen übertragbar und deshalb nicht auf das im AdsD verwendete beschränkt – eine konsequente und zukunftsweisende Umsetzung des OAIS-Moduls *Preservation Planning*. Da sowohl der textuelle Inhalt der PDF/A-Dateien als auch der der Dateianhänge im Datenbanksystem invertiert wurden, kann direkt in den E-Mails und v.a. in den Dateianhängen recherchiert werden. Damit existieren nun Recherchemöglichkeiten, die deutlich über die Möglichkeiten im Ursprungsprogramm hinausgehen.

¹⁷ ver.di nutzt den Microsoft Exchange Server für die E-Mail-Kommunikation. Sofern E-Mails intern auf einem Exchange Server zugestellt werden, werden keine Metadaten in den Header geschrieben und die Internetkopffzeile bleibt entsprechend leer. Aus diesem Grund können aus internen E-Mails der ver.di keine entsprechenden Header-Metadaten extrahiert werden. Sobald E-Mails einen externen Server passieren, werden entsprechende Metadaten in den Header geschrieben, anhand derer die jeweilige Route rekonstruiert werden kann.

¹⁸ Die Praxis hat gezeigt, dass in zahlreichen Fällen die identischen Dateien als Anhänge verschickt wurden. Wie im Text erwähnt, wurden diese Redundanzen gelöscht. Hinzu kam der Umstand, dass unterschiedliche Dateien identische Dateinamen trugen und die Assoziativ-Referenzen nicht mehr eindeutig waren und es falschen Referenzierungen gekommen ist. Um dieser Fehlerquelle zukünftig zu beheben, wurden die bisherigen Metadaten in den Faust-Datenbanken um das Metadatum *Priifsumme Anhang* ergänzt, über das nun die Assoziativ-Referenzierung läuft.

Erfreulicherweise kann festgestellt werden, dass das vom AdsD entwickelte Verfahren auf eine breite Akzeptanz innerhalb der ver.di-Bundesverwaltung gestoßen ist. Dies wird allein dadurch deutlich, dass sich ver.di aus eigenem Antrieb an das AdsD wendet, um digitale Korrespondenzen für die Langzeitarchivierung abzugeben – für private und politische Archive kein alltäglicher Vorgang.

Die Pilotierung wurde erfolgreich abgeschlossen und die Langzeitarchivierung digitaler Korrespondenzen läuft mittlerweile im Produktivbetrieb. Die letzte Übernahme fand Anfang April 2012 statt, die nächste steht Ende Mai des Jahres bevor.

Auf dem Weg zum *Retro-GIS*?

Dauerhafte Erhaltung und Nutzbarmachung digitaler Geobasisdaten beim Landesarchiv Baden-Württemberg und darüber hinaus¹

Von KAI NAUMANN

Dieser Aufsatz versucht das Themengebiet Archivierung von Geobasisdaten im Allgemeinen darzustellen und beispielhaft zu veranschaulichen, wie in Baden-Württemberg das Landesamt für Geoinformation und Landentwicklung und das Landesarchiv die dauerhafte Sicherung von Geobasisdaten ausführen. Er betrachtet sowohl die befristete Aufbewahrung von Geodaten für die Geschäfte der Vermessung als auch die Bewertung, dauerhafte Verwahrung und Nutzung von Geodaten mit historischem Wert im Geschäftsbereich des Landesarchivs und stellt Verbindungen zwischen beiden Arbeitsbereichen dar.

*Die Archivierung in jeder Form bereitet erhebliche Schwierigkeiten. [...] Kosten und Aufwand sind enorm, die Realisierungsmöglichkeiten bei schwindenden Ressourcen fraglich.*² So wie die hier zitierten Kollegen schätzten 1998 viele Archivare die Aussichten ein, was Archivierung digitaler Geobasisdaten angeht. Wo stehen wir heute, vierzehn Jahre später?

¹ Dieser Aufsatz ist die ausgearbeitete Fassung zweier Vorträge. Der eine Vortrag unter dem Titel *Dauerhafte Erhaltung von Geodaten. Ein Rundflug* wurde am 14. März 2012 beim Arbeitskreis *Archivierung von Unterlagen aus digitalen Systemen* in Ludwigsburg, der andere mit dem Titel *Auf dem Weg zum „Retro-GIS“ – Dauerhafte Erhaltung und Nutzbarmachung digitaler Geobasisdaten beim Landesarchiv Baden-Württemberg* am 9. Mai 2012 bei der Fachtagung des Deutschen Vereins für Vermessungswesen (DVW) Landesverein Baden-Württemberg in Sinsheim gehalten. Der Aufsatz wird (mit kleineren Anpassungen an die Zielgruppen) parallel in den Mitteilungen des DVW Landesverein BW und im Tagungsband des Arbeitskreises veröffentlicht. Für wertvolle Hinweise bei der Fertigstellung und auf dem Weg dorthin danke ich allen Beteiligten.

² Nicole *Bickhoff* und Clemens *Rehm*: Das automatisierte Liegenschaftskataster in Baden-Württemberg. In: *Archivierung elektronischer Unterlagen*. Hg. von Udo *Schäfer* und Nicole *Bickhoff* (Werkhefte der staatlichen Archivverwaltung Baden-Württemberg A 13). Stuttgart 1999. S. 131–143, hier S. 132. Scans der Originalausgabe unter <http://www.landearchiv-bw.de/web/52490> und <http://www.staatsarchiv.sg.ch/home/auds.html>. Alle Hyperlinks wurden am 24.9.2012 überprüft.

Kurzvorstellung des Landesamts für Geoinformation und Landentwicklung

Das Landesamt für Geoinformation und Landentwicklung (LGL)³ ist die Obere Vermessungs-, Flurbereinigungs- und Siedlungsbehörde für die Landkreise in Baden-Württemberg. Es übt als Obere Landesbehörde die Fachaufsicht über die unteren Vermessungsbehörden und über die öffentlich bestellten Vermessungsingenieure sowie die unteren Flurbereinigungsbehörden aus. Es nimmt u.a. folgende Aufgaben wahr:

- Aufsicht über Vermessung und Flurbereinigung, Genehmigungen und zwischeninstanzliche Entscheidungen bei Streitfällen in Verwaltungsverfahren
- Landesvermessung: Grundlagenvermessung (Lage, Höhe, Schwere) und Topographische Landesaufnahme, Fernerkundung
- Erzeugung von landesweiten Geobasisdaten und zentrales Geodatenmanagement
- Herstellung und Herausgabe der amtlichen topographischen Karten sowie von Wander-, Rad- und Freizeitkarten
- Ausbildung (Vermessungstechniker/in, Geomatiker/in).
- Geodatenzentrum für die Fachbereiche im Ressort des Ministeriums für Ländlichen Raum und Verbraucherschutz (MLR).

Das LGL hat ca. 600 Mitarbeiter, die vor allem in Stuttgart, Kornwestheim und Karlsruhe, aber auch bei den Unteren Vermessungsbehörden im Lande arbeiten. Es ist ein Landesbetrieb, der u.a. mehrere Millionen Euro Gebühreneinnahmen durch den Vertrieb von Geobasisdaten erzielt. Sein Schwerpunkt im Geschäftsbereich Vermessung ist derzeit die Umstellung aller Fachverfahren auf die bundeseinheitliche AAA-Architektur, die Kartographische Produkte, das Kataster und das Festpunktnetz in einem einheitlichen Datenmodell zusammenfasst.

Digitale Archivierung von Geodaten

Handlungsbedarf und Motivation

Digitale Information bestimmt unser Leben, doch die neuen Möglichkeiten hinterlassen auch Unsicherheit. Insbesondere die Haltbarkeit der Information unterliegt Zweifeln. Dennoch lassen sich viele Firmen und auch die öffentliche Verwaltung gern auf digitale Aufbewahrensverfahren ein, denn die Technik erspart nicht nur erhebliche Lagerkosten für Papier, sondern sie erleichtert auch das Suchen und Finden älterer Information.

Zur Lösung all dieser Probleme hat sich in den letzten Jahren ein Markt von Firmen etabliert, die sogenannte Enterprise Content Management Systeme anbieten. Ziel dieser Systeme sind

³ Internetpräsenz des Landesamts für Geoinformation und Landentwicklung, <http://www.lgl-bw.de>.

stets Prozesse mit großen Fallzahlen, häufigem Rückgriff und der gesetzlichen Anforderung, die Aufzeichnungen für fünf bis zehn Jahre aufzubewahren. So gehört schon in vielen Geschäftsbereichen die Papierform von Bestellformularen, Rechnungsserien, Buchungsbelegen und anderen umfangreiche Unterlagensorten der Vergangenheit an. Für Inhalte dieser Art ist bereits ein Dateiformat namens PDF/A definiert, von dem man sich eine besondere Haltbarkeit verspricht. Auch mit Datenbanken, wenngleich viel weniger intensiv, beschäftigt sich die Softwareindustrie. Hier gibt es Lösungen, die aus überquellenden Datenbankservern planmäßig ältere Datensätze aussortieren und für den schnellen Rückgriff in haltbaren Formaten speichern.

Weniger hingegen hat die Industrie bislang, was Aufbewahrung angeht, die Archivwelt und die Geoinformatik bedacht. Für beide stehen Herausforderungen an, die anderweitig kaum relevant sind:

- Für den Auftrag der Archive geht es nicht um eine Erhaltung auf Jahrzehnte, sondern auf Jahrhunderte. Um digitale Objekte lesbar zu erhalten, sind die vorhandenen Systeme noch unzureichend auf die Techniken der Formatmigration und der Emulation vorbereitet. Die grundsätzlichen Konzepte sind in Gestalt der ISO 14721 (Open Archival Information System) und anderer Empfehlungen und Normen⁴ formuliert, aber die Entwicklung marktgängiger Produkte setzt gerade erst ein.⁵
- Für den Bereich der Geoinformatik geht es zusätzlich noch darum, sowohl den Inhalt der primären Geodaten als auch deren Georeferenzierung in künftige Systeme zu überführen. Weiter erschwert wird der Auftrag hier durch die enge Koexistenz von vektor- und rasterbasierten Inhalten, die in anderen Fachgemeinden völlig unabhängig voneinander betrachtet werden können.

Bei Geodaten müssen die digitalen Daten mitsamt ihren wesentlichen Funktionen erhalten werden. Was nützt ein digitales Luftbild auf Mikrofilm ohne maschinenlesbare Georeferenzierung? Was nützt eine Datenbank in Form von 10.000 PDF-Berichtsdokumenten, die man nicht mehr tabellarisch auswerten kann? Beides sind Beispiele dafür, dass digitale Archivierung auf längere Zeit das Konzept der signifikanten Eigenschaften⁶ berücksichtigen muss. Solche Eigenschaften müssen vorab festgelegt und im Erhaltungsprozess berücksichtigt werden.

In Zusammenhang der Aufbewahrung digitaler Daten stellen sich weitere, immer wieder ähnliche Fragen:

⁴ PREMIS Data Dictionary for Preservation Metadata (V. 2.1), <http://www.loc.gov/standards/premis/>; Wege ins Archiv - Ein Leitfaden für die Informationsübernahme in das digitale Langzeitarchiv, <http://nbn-resolving.de/urn:nbn:de:0008-2008103009> (2008, deutsch und englisch verfügbar); Leitfaden zur digitalen Bestandserhaltung (2011), <http://nbn-resolving.de/urn:nbn:de:0008-2011101804>.

⁵ Christoph H. Ferle. Marktstudie Digitale Langzeitarchivierung. Im Spannungsfeld zwischen Digital Preservation und Enterprise Information Archiving. Stuttgart 2012. Zu beziehen unter <http://www.lza2012.iao.fraunhofer.de>.

⁶ Zuerst erwähnt wurde der Begriff 1999 in einem Abschlussbericht des Cedars Projekts, der von der erstellenden Institution nicht mehr vorgehalten, aber freundlicherweise hier gespiegelt wird: <http://www.imaginar.org/dppd/DPPD/146%20pp%20Digital%20Preservation%20Strategies.pdf>.

- Was passiert, wenn das Format, in dem die Daten dargestellt sind, von künftigen Computern nicht mehr unterstützt wird?
- Wie können Inhalte gleichzeitig unverschlüsselt und datenschutzrechtlich sicher aufbewahrt werden?
- Wie können Inhalte verschiedener komplexer Anwendungen (wie z.B. eines Katasters) in ein einheitliches Speichersystem übertragen werden?
- Wie lassen sich bei einem Wechsel des produktiven Speichersystems Inhalte reibungslos auf das Folgesystem übertragen? Ist dieser Weg gut geeignet, dann kann auch eine Überführung in eine langfristige Aufbewahrung hier ansetzen.

Lösungsansätze

Als Reaktion haben inzwischen nicht nur einige staatliche und kommunale Archive, aber auch Bibliotheken und Hochschulen, angepasste Systeme geschaffen, die vor diesen Herausforderungen bestehen können und Aufbewahrung über sehr lange Zeit gewährleisten. Auch in der Geodaten-Community sind in Deutschland, Europa und weltweit einzelne Wissenschaftler angetreten, um im Dialog mit Archivaren und Informatikern an diesen Fragen zu arbeiten.

In Deutschland haben sich die Arbeitsgruppe Elektronische Systeme in Justiz und Verwaltung (AG ESys) der Archivreferentenkonferenz⁷ und der Arbeitskreis Geotopographie der Arbeitsgemeinschaft der Vermessungsverwaltungen⁸ mit dem Thema beschäftigt. In der benachbarten Schweiz besteht seit drei Jahren ein Projekt zum Aufbau eines Regelbetriebs für die Geodatenarchivierung in enger Kooperation zwischen Bundesarchiv und Bundesamt für Landestopographie swisstopo.⁹ Auf europäischer Ebene haben wir das European Spatial Data Research Network (EuroSDR) mit der Arbeitsgruppe *Geographic Data Archiving*, die ihr Ziel in der Verbreitung von Grundsätzen und Praxisbeispielen sieht.¹⁰ In den USA hat sich ein Verbund von Bundesstaaten (North Carolina, Kentucky, Montana und Utah) und der Library of Congress der Entwicklung von Lösungen verschrieben.¹¹ Beim Open Geospatial Consortium (OGC) betätigt sich die weltweit ausgerichtete Arbeitsgruppe *Data Preservation*.¹²

Digitale Bestandserhaltung kostet Geld und kann darum kein Selbstzweck sein. Warum nimmt man sich dieses Problems an? Für die staatlichen Archive ist Bestandserhaltung ein gesetzlicher Auftrag. Ohne digitale Archivierung werden sie in 100 Jahren nur noch lückenhaft Auskunft über die Jetztzeit geben können. Die Vermessungsverwaltung ist in einer ähnlichen

⁷ Beate Dorfey u.a.: Handreichung zur Archivierung elektronisch vorliegender Geodaten, 2009, http://www.bundesarchiv.de/imperia/md/content/bundesarchiv_de/fachinformation/ark/handreichung_geodaten_20090928.pdf.

⁸ Arbeitskreis Geotopographie auf der Internetpräsenz der AdV: <http://www.adv-online.de/icc/ext-deu/broker.jsp?uMen=da370061-7527-a8fe-ebc4-f19f08a07b51>.

⁹ Arbeitsbereich Archivierung von Geodaten beim Bundesarchiv Schweiz: <http://www.bar.admin.ch/themen/00876/00939/index.html?lang=de>.

¹⁰ <http://www.eurosdnet.net/archiving/>.

¹¹ GeoMAPP-Verbund: <http://www.geomapp.net>.

¹² <http://www.opengeospatial.org/projects/groups/preservdwg>.

Lage, wenngleich ihr Auftrag in der Regel 10 bis spätestens 30 Jahre nach Entstehung der Geodaten endet. Doch schon für diese Zeiträume sind Verfahren erforderlich, die über die markt-gängige Technik hinausgehen.

Hinzu kommt, dass Archivierung die Mitarbeitermotivation steigern kann. Werden Großsysteme, die nach mehreren Jahrzehnten abgelöst werden, in Phasen des Umbruchs an die Archive abgegeben, so sind manche altgediente Mitarbeiter hochmotiviert. Vermutlich fällt Personen, die ein System ein Arbeitsleben lang begleitet haben, der Abschied leichter, wenn eine Archivierung stattfindet.

Archivierung als Geschäftsmodell

Sind die Beteiligten also auf dem Weg zum *Retro-GIS*? Man sollte sich ein solches Geoinformationssystem (GIS) nicht als ein Softwaresystem im eigentlichen Sinne vorstellen, sondern als einen technisch-organisatorischen Rahmen zur möglichst vielfältigen Nutzung älterer Stände von Geobasisdaten. Retro-GIS ist keine Lösung, sondern ein Geschäftsmodell. Dies bestätigt sich nicht nur in der hier geschilderten Zusammenarbeit in Baden-Württemberg, sondern auch weltweit. Während zu Anfang des Jahrtausends große Euphorie über die Möglichkeiten komplexer und umfassender Infrastrukturen herrschte,¹³ halten sich beide Fachgemeinden heute in dieser Hinsicht eher zurück. Während die Rolle der konkreten Technik zurücktritt, gelangen auch für sonstige digitale Inhalte die Bedeutung der kompetenten Institutionen und der fördernden oder dämpfenden finanziellen, rechtlichen und politischen Faktoren in den Vordergrund.¹⁴

¹³ Als Beispiele aus den USA hier die Programme Electronic Records Archive, <http://www.archives.gov/era/about/status-accomplishments.html> oder National Geospatial Digital Archive, <http://www.ngda.org/home.html>.

¹⁴ Blue Ribbon Task Force on Sustainable Digital Preservation and Access <http://brtf.sdsc.edu/>.

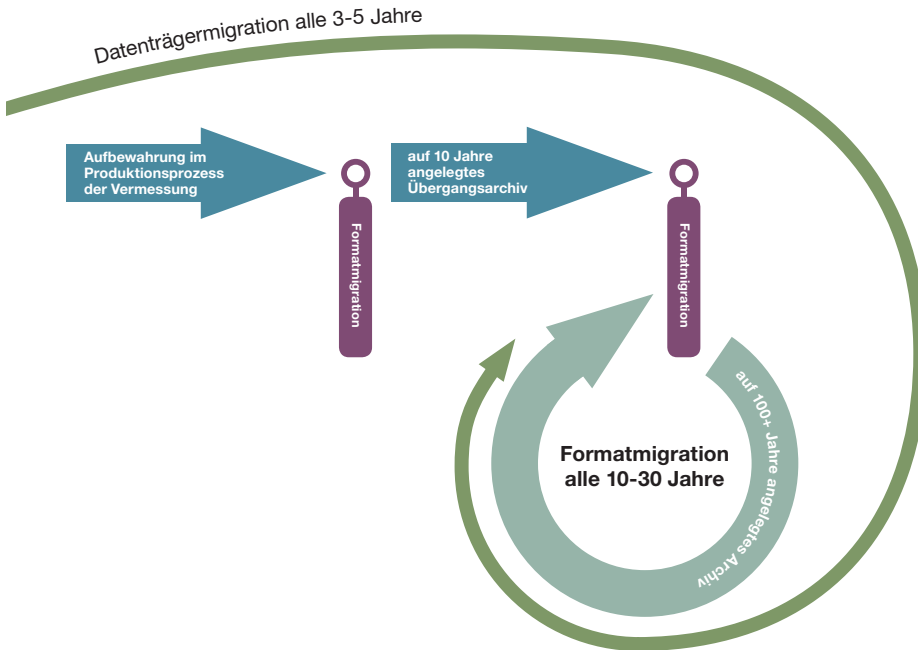


Abb. 1: Das 1-10-100 Diagramm der EuroSDR Arbeitsgruppe Geographic Data Archiving. Quelle: <http://www.eurosd.net/archiving>.

Die EuroSDR-Arbeitsgruppe Geographic Data Archiving hat diese Rolle der Geschäftsmodelle inzwischen mit der 1-10-100 Formel auf den Punkt gebracht (vgl. Abb. 1).¹⁵ Es gibt für die Historisierung von Geodaten im Grunde drei Systemkreise, die ineinandergreifen müssen. Die Formel 1-10-100 ist symbolisch zu verstehen und kann durch andere Zeiträume ersetzt werden. Die Systemkreise sind wie folgt zu umschreiben:

1. Die Produktivsysteme, die im Zeitraum eines Jahres Daten vorhalten,
2. ein Speichersystem, das wir Übergangsarchiv genannt haben und in dem die Daten für eine Dekade noch für den stetigen Rücklauf in das Produktivsystem bereitstehen, in dem das Datenformat aber bereits durch eine erste Formatmigration auf mittelfristige Lagerung optimiert ist. Ein Teil dieser Daten wird in diesem Übergangsarchiv für die langfristige Lagerung aufbereitet und schließlich

¹⁵ Carsten Rönndorf u.a.: GI+100: Langzeiterhaltung digitaler geographischer Informationen – 16 grundlegende, von staatlichen Vermessungsbehörden und Archiven vereinbarte Prinzipien, Stand 29.3.2012, S. 3/4, http://www.eurosd.net/archiving/Paper_EuroSDR%20Archiving%20Principles_DE.pdf.

3. an das Langzeitarchiv zur dauerhaften Verwahrung über 100 Jahre und mehr weitergegeben. Spätestens hier finden immer wieder kostspielige Formatmigrationen und Validierungen statt und hier ist auch der Funktionsbereich anzusiedeln, den in Deutschland die Landesarchive ausfüllen.

Man sieht in diesem Diagramm zwei Arten der Migration von Daten, zum einen die Migration von Datenträger zu Datenträger, die alle drei bis fünf Jahre auftritt, zum anderen die Migration von Format zu Format, die in selteneren Abständen, hier 10 bis 30 Jahre, wiedergegeben ist.

Was die Kosten digitaler Archivierung angeht, argumentieren Optimisten gern mit dem Moore'schen Gesetz, nach dem jede Art von Hardware alle zwei Jahre ihre Kapazität um den Faktor zwei steigert. Realisten müssen zu diesen Speicherkosten aber alle Ausgaben hinzurechnen, die eine vertrauenswürdige Aufbewahrung erfordert, nämlich Dienste für Übernahme, Backup, Recherche, Zugangsverwaltung und vor allem Bestandserhaltung, die ihrerseits stets überwacht und validiert werden müssen. Hierbei zeigt sich, dass eine sichere langfristige Verwahrung digitaler Objekte große Investitionen in kompetentes Personal und Betriebsmittel erfordert.¹⁶

Eine funktionale Aufgabentrennung zwischen Archiv und Vermessungsbehörde ist nicht nur wirtschaftlich geboten, sondern auch in den Archivgesetzen Deutschlands verankert.¹⁷ Auch aus diesem Grund wurde bei der Umsetzung der INSPIRE-Richtlinie in deutsches Recht unter Beteiligung der archivischen Gremien dafür gesorgt, dass die Pflicht zur öffentlichen Zugänglichmachung nur für Geodaten gilt, die *noch in Verwendung stehen*.¹⁸ Unter *Verwendung* sollte man in diesem Kontext eine volkswirtschaftlich relevante primäre Nutzung im Gegensatz zu einer Sekundärnutzung für kulturelle Gedächtniszwecke verstehen.

Wenn die Vermessung ältere Geodaten an einen anderen Dienstleister abtritt, bedeutet dies nicht nur eine Entlastung ihrer Speicherkapazitäten, sondern vor allem Synergieeffekte bei der Bestandserhaltung, da diese vom Landesarchiv für alle Verwaltungszweige gleichmäßig wesentlich kostengünstiger geleistet wird, als wenn hier mehrere Ressorts parallel arbeiten. Für den Zeitraum, in dem archivwürdige Daten im Rahmen der sogenannten Auftragsverwahrung¹⁹ vorab einem Landesarchiv überlassen werden, bietet das Landesarchiv überdies eine Art Backup-Funktion. Diese Backup-Funktion ist aber den Notwendigkeiten der langfristigen Erhaltung untergeordnet. Eine andere Lösung wäre ein digitales Zwischenarchiv, eine Lösung, die in Deutschland derzeit aber von den meisten staatlichen Archiven nicht aktiv angestrebt wird.

¹⁶ Vgl. in diesem Band Peter Sandner: 10 FAQs. Argumente zu Bedarf und Notwendigkeiten der digitalen Archivierung, S. 57–70.

¹⁷ <http://archivschule.de/service/archivgesetze/>.

¹⁸ § 4 Abs. 1 Geodatenzugangsgesetz vom 10. Februar 2009 (BGBl. I S. 278).

¹⁹ Auftragsverwaltungsverordnung, GBl. BW 1992, S. 685, <http://www.landearchiv-bw.de/web/46652>.

Bisherige Umsetzung in Baden-Württemberg

Das erste Übernahmeprojekt des Landesarchivs, das so etwas wie ein geographisches Informationssystem war, war die Straßendatenbank der Landesstelle für Straßentechnik beim Regierungspräsidium Tübingen. Das Datenmodell dieser Datenbank war von 1968 bis 2000 auf vielen Generationen von Großrechnern gelaufen (vgl. Abb. 2). Die Datenbank beschrieb Straßenabschnitte zwischen Netzknoten, zu denen wir Gauß-Krüger-Koordinaten mitgeliefert bekamen, und konnte 2007 auf dem Stand von 2000 übernommen werden.²⁰

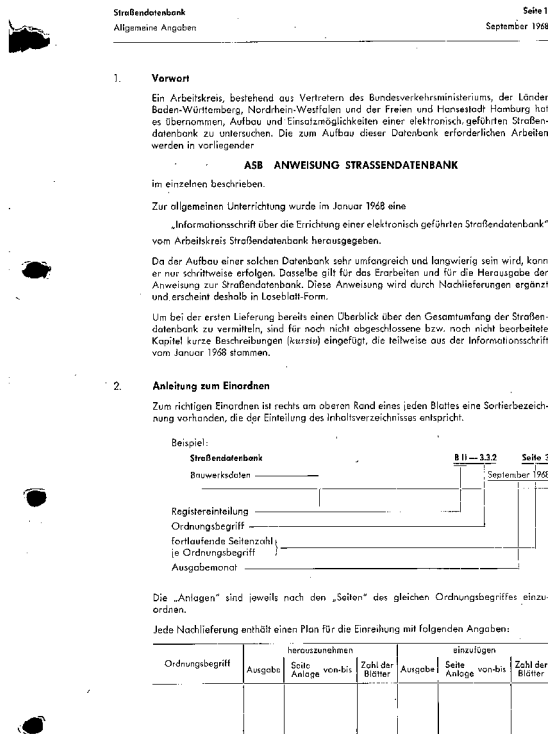


Abb. 2: Anweisung Straßendatenbank von 1968, Quelle: Staatsarchiv Ludwigsburg (StAL) EL 75 V DO 1.

²⁰ Findmittel zum Bestand EL 75 V, <http://www.landearchiv-bw.de/plink/?f=2-1023440>.

Weiterhin bestanden seit Ende der neunziger Jahre bereits Kontakte zur Umwelt- und zur Vermessungsverwaltung.²¹ Es lag daher 2007 nahe, mit der Landesanstalt für Umwelt, Messungen und Naturschutz (LUBW) eine Arbeitsgruppe zu gründen, die inzwischen ein Konzept zur Übernahme und Erhaltung von geographisch codierten Umwelfachdaten aus der Referenzdatenbank der Behörde erarbeitet hat.²² Inzwischen wird auch mit dem Geoportall Raumordnung,²³ das vor allem dem Verkehrsressort zugeordnet ist, ein entsprechendes Übernahmeverfahren erstellt.

Auf das Thema der Archivierung von Geobasisdaten wurde das Landesarchiv im Herbst 2009 durch den Präsidenten des LGL Hansjörg Schönherr angesprochen. Im März 2010 wurde dann eine *Arbeitsgruppe Archivierung von Geobasisdaten des Landesamts für Geoinformation und Landentwicklung Baden-Württemberg* (AG AGL) aus Mitarbeitern des LGL und des Landesarchivs eingerichtet, deren Ziele unter anderem waren:

- Archivwürdige Bestände an digitalen Geobasisdaten bestimmen und deren Übernahmeturnus festlegen.
- An pilothaften Datenbeständen die Übernahme und Aufbereitung für die Lagerung und spätere Nutzung erproben.
- Einen Dauerbetrieb für die Archivierung zentraler Geobasisdaten einrichten.

Die AG hat sich inzwischen insgesamt sechsmal getroffen. Beteiligt waren alle Abteilungen des LGL und aus dem Landesarchiv die Querschnittsreferate für Digitale Unterlagen, einige Kollegen mit inhaltlicher Zuständigkeit für die Vermessung sowie das Referat Informations- und Kommunikationstechnologie.

²¹ Nicole *Bickhoff* und Clemens *Rehm*: Das automatisierte Liegenschaftskataster in Baden-Württemberg, in: *Schäfer/Bickhoff*, wie Anm. 2, S. 131–143; Franz-Josef *Ziwe*: Überlegungen zur Bewertung von digitalen Unterlagen aus dem Umweltinformationssystem Baden-Württemberg. In: *Schäfer/Bickhoff*, wie Anm. 2, S. 145–151.

²² Vgl. Kai *Naumann*: Online-Findmittel zum Bestand Generallandesarchiv 518-1: Landesanstalt für Umwelt, Messungen und Naturschutz: Umweltinformationssystem (UIS), <http://www.landesarchiv-bw.de/plink/?f=4-740417>, sowie Christian *Keitel* u.a.: Langzeitarchivierung von Umweltinformationen. In: Umweltinformationssystem Baden-Württemberg. F+E Vorhaben KEWA. Kooperative Entwicklung wirtschaftlicher Anwendungen für Umwelt, Verkehr und benachbarte Bereiche in neuen Verwaltungsstrukturen. Phase V 2009/10, Karlsruhe 2010, <http://digbib.ubka.uni-karlsruhe.de/volltexte/1000019255>; Christian *Keitel*: Archiving of Environmental Data. A Cooperation Project. Vortrag zur 8. Europäischen Konferenz über digitale Archivierung, 29.4.2010, www.bar.admin.ch/aktuell/00568/00702/00861/01570/.

²³ <http://www.geoportall-raumordnung-bw.de>.

Übernahme und Aufbereitung

Während bei der Übernahme von Papierunterlagen lediglich eine Neuordnung der einzelnen Schriftstücke ansteht, handelt es sich bei der Übernahme digitaler Informationen um einen Prozess, der mit einer teilweisen oder gar vollständigen Wandlung der Struktur, der Metadaten und der Datenformate einhergeht. Der im OAIS-Modell als Ingest bezeichnete Prozessabschnitt reicht vom Aushandeln der Übergabe über Aufbereitung und Verzeichnung bis hin zur Ablage an einem festen Lagerort.

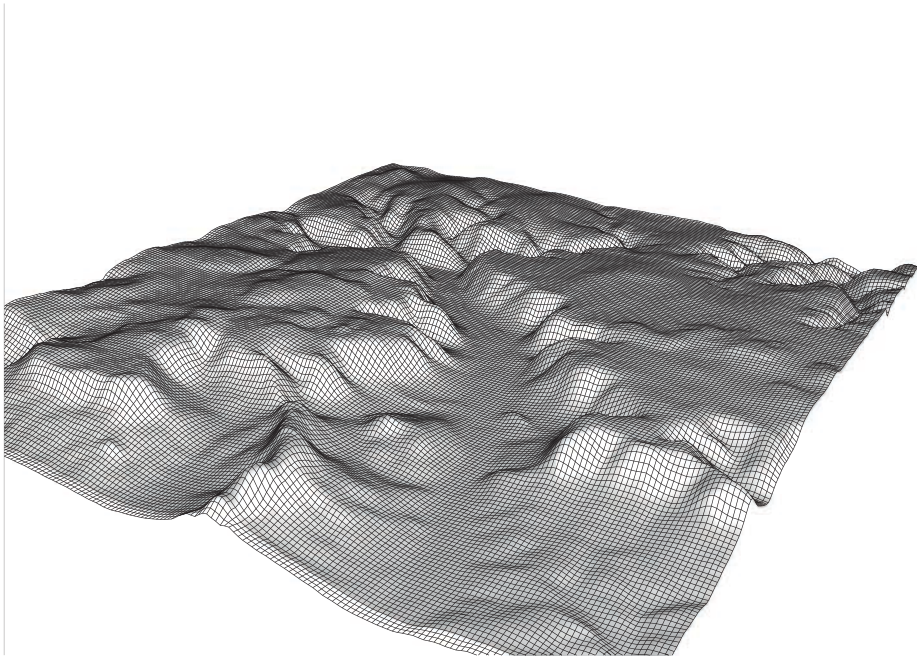


Abb. 3: Screenshot graphische Darstellung Höhenmodell 1989, Quelle: StAL EL 68 V DO 1.

Der erste Testlauf unserer AG war die Übernahme eines digitalen Höhenmodells aus dem Jahr 1989, das im LGL für die Errechnung der ersten digitalen Orthophotos gedient hatte (Abb. 3). Das damalige Höhenmodell mit seinem 100-Meter-Raster hat seinen historischen Wert heute eher als Technikdenkmal denn als Informationsquelle, ist doch das Messraster heute zwanzigmal dichter. Es geht auch nur in die Größenordnung von 100 Megabyte, was vor 25 Jahren eine Ehrfurcht heischende Datenmenge war, was aber heute kaum mehr beeindruckt.

Dann holten wir einen Abzug der sogenannten Liegenschaftskatasterdatenbank (LiKaDB) in das Landesarchiv. Diese Anwendung repräsentiert in Baden-Württemberg die automatisierte Liegenschaftskarte ALK und das automatisierte Liegenschaftsbuch ALB. Das LGL hat für das Landesarchiv den Stand des Katasters von November 2009 gleichsam eingefroren. Dies war

auch der optimale Zeitpunkt zur ersten Übernahme, denn seit 2009 werden in Baden-Württemberg Liegenschaftskarten in papierner Form nicht mehr fortgeführt.

Der Inhalt der LiKaDB wird derzeit in das neue ALKIS-Datenmodell migriert. Sobald die Migration abgeschlossen ist, wird ein weiterer Abzug des Katasters an das Landesarchiv abgegeben. Danach ist geplant, alle zehn Jahre einen weiteren Abzug an uns abzugeben. Ohne die große Unterstützung der entsprechenden Referate beim LGL wäre diese Arbeit nicht gelungen. Mit 88 Gigabyte war die Übernahme der LiKaDB die bisher größte einzelne Übernahme bei uns. Die Aktion erforderte mehrere Validierungsläufe, bis einwandfrei feststand, dass wir einen vollständigen Export im Haus hatten. Insgesamt haben das LGL und das Landesarchiv zusammen ca. 35 Personentage in das Projekt investiert.

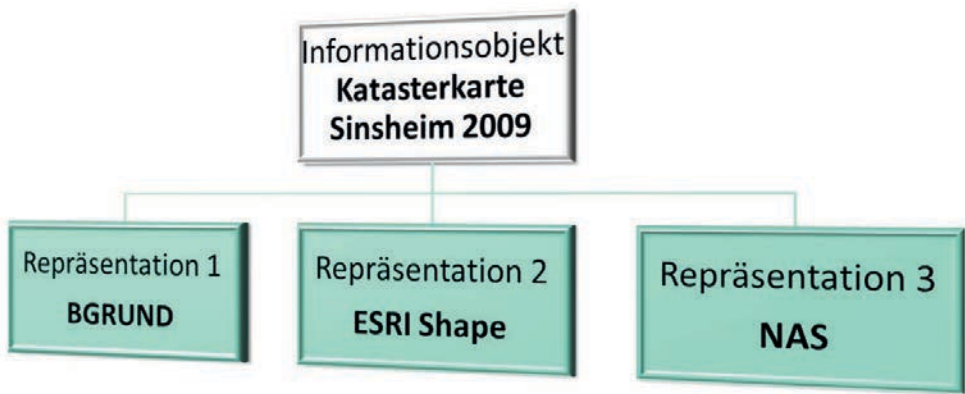


Abb. 4: Zusammenhang von Informationsobjekt und Repräsentationen.

Die abgelieferten Daten liegen in sogenannten Archival Information Packages oder AIPs vor. Die Pakete sind für jede Gemeinde einzeln verfügbar, gleichgültig ob es sich um eine kleine Landkommune oder eine Großstadt handelt. Darin liegen schlichte Dateien. Gelagert und auf Wunsch ausgeliefert wird immer das aktuell nutzbare Erhaltungsformat oder optional das ursprüngliche Datenhaltungsformat. Aktuelles Format ist bei der Liegenschaftskarte zur Zeit ESRI Shape,²⁴ das Ausgangsformat ist das landeseigene ISAM²⁵-Format BGRUND.²⁶ Bei der

²⁴ ESRI Shapefile Technical Description, 1998, <http://www.esri.com/library/whitepapers/pdfs/shapefile.pdf>.

²⁵ ISAM (Index Sequential Access Method) ist ein Ablageverfahren für Datenbankinhalte in Großrechnerarchitekturen.

²⁶ BGRUND ist ein Austauschformat der Vermessungsverwaltung Baden-Württemberg. <http://www.gismngt.de/format/bgrund.htm>.

Ablage in DIMAG²⁷ sieht man einen erheblichen Unterschied zu allen marktgängigen Enterprise Content Management Systemen. Letztere können nämlich jedes Informationsobjekt in der Regel nur einem Datenobjekt zuordnen, was zur langfristigen Erhaltung nicht ausreicht.

Stattdessen sieht unser Konzept vor, jedem Informationsobjekt mehrere Erscheinungsformen, sogenannte Repräsentationen, zuzuordnen. Tabellen wie im Automatisierten Liegenschaftsbuch werden im Format Character Separated Values (CSV) mit einer Beschreibung der einzelnen Felder und optional ebenfalls im Ursprungsformat (z.B. beim Liegenschaftsbuch WLDGE²⁸) übernommen. Die Tabellen erhalten dazu eine gründliche Beschreibung der in der Tabelle enthaltenen Attribute.

Abgerundet werden diese Pakete durch ein Dokumentationspaket. Dieses enthält:

- die Verwaltungsvorschriften und Beispielsammlungen zur Katasterführung,
- Formatspezifikationen und Zusätze zu den Shapefiles,
- Dokumentation der Formatmigration und den sogenannten Aufbereitungsbericht, der den Prozess von der Abgabe an das Landesarchiv bis zur endgültigen Einlagerung beschreibt.

Die notwendige Formatmigration aus dem BGRUND-Format nach ESRI Shape war ein gutes Beispiel für die Rolle signifikanter Eigenschaften und ihrer Prüfung bei digitaler Bestandserhaltung. Gerade die Validierung der Formatmigration von BGRUND nach ESRI Shape (Liegenschaftskarte) und von WLDGE nach CSV (Liegenschaftsbuch) hat einen Gutteil der Arbeit an der Katasterübernahme verursacht, war aber für die Glaubwürdigkeit der Unterlagen unabdingbar. Das BGRUND-Format kann Kreisbögen als Grundstücksgrenzen definieren, ESRI Shape hingegen kann Bogenformen nur durch Zerlegung in Strecken nachahmen. Die Konversion von BGRUND nach Shape erfolgte daraufhin mit der Software *BGRUND nach Shape Konverter 8.80* von Rainer Kettemann (Weil der Stadt). Die zerlegten Kreisbögen wurden begutachtet und als adäquater Ersatz eingestuft. Das zur Übernahme eingesetzte Programm IngestList des Landesarchivs erkannte darüber hinaus, dass der Konverter für die Komponente des automatisierten Liegenschaftsbuchs leere CSV-Dateien erzeugte, was wiederum eine Nachbesserung an dem zur Migration verwendeten Visual Basic Skript für MS Excel bedingte.

Gerade die Übernahme des Katasters hat wohl auf beiden Seiten, bei den Archiven wie bei der Vermessung, für Erleichterung gesorgt. Man hatte in den neunziger Jahren prophezeit, die Umstellung des Katasters und der Landesvermessung auf digitale Formen werde in den

²⁷ Vgl. den Beitrag von Christian *Keitel*: DIMAG-Kooperationen, S. 147-155 in diesem Band und Christian *Keitel* und Rolf *Lang*: DIMAG und IngestList. Übernahme und Nutzung von digitalen Unterlagen im Landesarchiv Baden-Württemberg. In: *Archivische Informationssysteme in der digitalen Welt*. Hg. von Gerald Maier und Thomas Fritz. Stuttgart 2010. S. 53–63.

²⁸ WLDGE steht für Workdatei Liegenschaftskataster Daten-Gewinnung Entschlüsselt und wird von den deutschen Vermessungsbehörden verwendet.

Archiven zu einem neuen überlieferungslosen Zeitalter²⁹ führen, doch die in Baden-Württemberg und auch in Bayern vollzogenen Übernahmen zeigen uns inzwischen, dass die düsteren Aussichten sich nicht bewahrheitet haben.

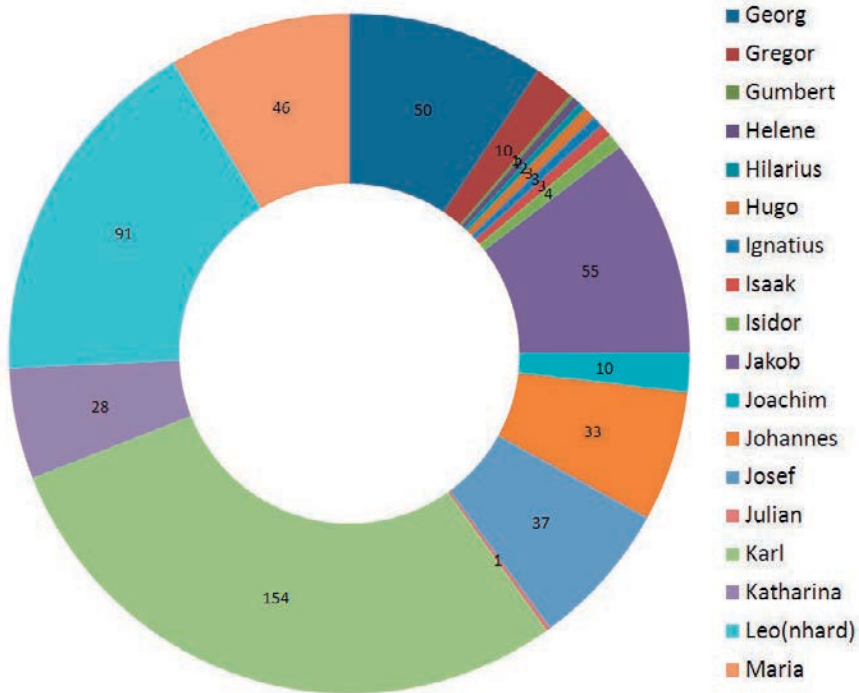


Abb. 5: Häufigkeit von Heiligennamen (G bis M) in Flurbezeichnungen Baden-Württembergs, Diagramm des Autors, Datenquelle: StAL EL 68 V DO 1303.

²⁹ Volker Buchholz: Kataster – ein Auslaufmodell in den Archiven? In: Archivierung von Unterlagen aus digitalen Systemen. Beiträge zur Tagung im Staatsarchiv Münster, 3.-4. März 1997. Hg. von Frank M. Bischoff (Veröffentlichungen der staatlichen Archive des Landes Nordrhein-Westfalen E 4). Münster 1997. S. 69–74; Gudrun Fiedler: Archivierung digitaler Katasterunterlagen. Die Fortführung eines Archivierungsmodells der niedersächsischen Staatsarchive. In: Schäfer/Bickhoff, wie Anm. 2, S. 153–162; Peter Hobeisel: Archivische Überlieferungsbildung und -sicherung von Daten des automatisierten Liegenschaftskatasters in den hessischen Katasterämtern. In: Archivierung und Zugang. Hg. von Nils Brübach (Veröffentlichungen der Archivschule Marburg, Institut für Archivwissenschaft 36). Marburg 2002. S. 37–79; Karl-Ernst Lapprian: Die digitale Flurkarte Bayerns – Entwicklung, Aufbau, Archivierung. In: Archive vor der Globalisierung? Hg. von Mechthild Black-Veltrup, Ottfried Dascher und Axel Koppetsch (Veröffentlichungen der staatlichen Archive des Landes Nordrhein-Westfalen, Reihe E, Beiträge zur Archivpraxis, H. 7). Düsseldorf 2001. S. 69–76; Udo Schäfer: Geographische Informationssysteme in der Landesverwaltung Baden-Württemberg aus archivischer Perspektive. In: Schäfer/Bickhoff (wie Anm. 2), S. 114–129.

Statt der prophezeiten Schwierigkeiten mit der Nutzung könnten Unternehmungen wahr werden, die einst als unrealistisch galten. 1997 skizzierte Hanno Vasarhelyi das Vorhaben, durch Analyse von Flurnamen aus dem Liegenschaftsbuch noch unbekanntes Wüstungen in Baden-Württemberg auf die Spur zu kommen.³⁰ Inzwischen haben wir die Daten, und nur als Beispiel für die Möglichkeiten: Wir haben sämtliche Flurbezeichnungen erhalten und man kann auszählen, welcher Heiligennamen in den Flurbezeichnungen wie häufig vorkommt (Abb. 5). Man könnte dies selbstverständlich auch auf ein Gebiet, zum Beispiel das Umfeld eines bestimmten Klosters, einschränken.³¹

Ebenso übernommen ist seit etwa einem Jahr das Digitale Landschaftsmodell in der Ausprägung des Basis-DLM in drei verschiedenen zeitlichen Ständen, nämlich von 1998, 2000 und 2010. Hier wäre es möglich gewesen, aus dem Fundus des LGL weitere Zeitschnitte, immerhin 36 Stück zwischen 1998 und 2010, zu erhalten, auch hier haben wir aus Platzgründen verzichtet.

Derzeit sind zwei weitere Übernahmen in Arbeit. Zum einen stehen die als Rasterdaten vorliegenden gescannten Topographischen Karten zur Einlagerung an, zum anderen die Datenbankinhalte des Festpunktinformationssystems auf dem Stand vor der Überführung nach AFIS. Zu der Festpunktdatenbank werden auch Scans der Karteikarten geliefert, auf denen die Festpunkte in ihrer Lage dargestellt sind.

Anschließend werden wir uns mit Orthophotos beschäftigen, die hinsichtlich der Datenmenge wohl die größte Herausforderung darstellen. Hier ist, besonders was Auflösung und die denkbaren Qualitätsverluste durch Kompression angeht, sehr genau zu überlegen. Wir werden später noch darauf zurückkommen.

Die übernommenen Daten liegen nun in unserem Aufbewahrungssystem DIMAG. Das System lässt Nachweise über die Integrität und Authentizität der erfassten Daten zu. Die Daten werden täglich auf einen zweiten Standort und wöchentlich auf einen dritten Standort kopiert, um auch im Katastrophenfall ihr Überleben sicherzustellen. DIMAG ist ausgelegt auf Formatmigrationen, das heißt: Altmodische Formate lassen sich durch neue Formate ablösen, ohne dass Daten überschrieben werden.

Außerdem wird DIMAG derzeit so eng mit unserem Verzeichnungs- und Recherchesystem verknüpft, dass auch eine automatisierte Bestellung und Nutzung über unser Online-Findmittelsystem in den Bereich des Möglichen rückt. Unser Entwicklungspartner Bayern arbeitet derzeit an einer Schnittstelle zur Erstellung solcher Nutzungspakete. Wir wollen künftigen Nutzern stets die Nutzung in einem aktuellen, verbreiteten Format ermöglichen, das den prognostizierbaren Anforderungen einer historischen Nutzung in seinen *signifikanten Eigenschaften* möglichst weitgehend entspricht. Die künftigen Nutzer nennen wir auch nach OAIS-Konvention die *designated community*.

Handelt es sich um Vektordaten, so werden diese Daten im ESRI-Shape-Format vorgehalten und abgegeben. Dazu wird auf die im Internet verfügbaren Anleitungen und Softwarelösungen

³⁰ Hanno Vasarhelyi: Die Flurnamen auf den lithographierten Flurkartenerstdrucke der württembergischen Landesvermessung 1818–1840. Ergänzendes Statement zum Vortrag von Hartmut Klüver und einige weitere Bemerkungen. In: Gemeindebeschreibungen und Ortschroniken in ihrer Bedeutung für die Landeskunde. Hg. von Eugen Reinhardt (Werkhefte der staatlichen Archivverwaltung Baden-Württemberg A 12). Stuttgart 1997. S. 191–197.

³¹ Bestellsignatur des Objekts: Staatsarchiv Ludwigsburg EL 68 V DO 1303.

hingewiesen, die eine Nutzung ermöglichen. Vor fünfzehn Jahren gab es noch große Sorgen, niemand könne sich die teure Spezialsoftware für solche Daten leisten.³² Inzwischen aber besteht ein erfreulich großes Angebot an kostenloser Software, das auch Laien einen Zugang zu solchen Daten bietet. Perspektivisch wird man eines dieser Produkte stärker in so ein Abgabepaket integrieren können. Handelt es sich um Rasterdaten, so haben wir bisherige Übernahmen im TIFF-Format vollzogen. Wir haben innerhalb der TIFF-Dateihülle auch verlustfreie Kompression mit dem Lempel-Ziff-Welch (LZW) Algorithmus zugelassen. Die Diskussion um das geeignete Format für Luftbildmaterial, sei es JPEG2000 oder TIFF mit einer verlustfreien Kompression oder JPEG2000 mit verlustbehafteter Kompression, ist noch nicht zu Ende geführt.

Archivische Bewertung der Geodaten

Archive kommen nicht umhin, gegenüber ihrem Träger und der Öffentlichkeit zu rechtfertigen, warum sie für die Erhaltung bestimmter Information Geld ausgeben. Archive leisten diese Bewertungsarbeit durch Bewertungsmodelle, das sind Dokumente, die bestimmte Themenbereiche als archiwwürdig von anderen Bereichen abgrenzen, die als nicht historisch relevant eingestuft werden.³³ Diese Bewertungsmodelle kommen beim Landesarchiv stets unter intensiver Beteiligung der betroffenen Verwaltungszweige zustande.

In der Vermessungsverwaltung ging es 2010 vor allem darum, das bereits bestehende Modell für analoge Unterlagen des Katasters und der vermessungstechnischen Arbeit des LGL³⁴ um den Bestandteil der digital vorliegenden Geodaten zu erweitern. Diese Arbeit ist heute noch nicht zu Ende, sollte aber in den nächsten Monaten erledigt sein. Es ergaben sich allgemeine Kriterien für historischen Wert, von denen einige von der klassischen archivischen Bewertungslehre abweichen und daher hier aufzuführen sind. Bei der Erstellung des Katalogs war kein Kriterium allein ausschlaggebend, sondern stets in Abwägung mit anderen Kriterien zu betrachten.

Besondere Argumente für die Archivierung

Folgende Bewertungsmotive gelten als Hinweis auf historischen Wert:

- feinsten verfügbaren Detaillierungsgrad (größter Maßstab, höchste Bodenauflösung).
- technikhistorische Meilensteine (z.B. Kataster-Datenbanken).
- Orientierungshilfen bei der Nutzung anderer Geodaten (digitalisierte topographische Karten, Orthophotos und Schummerungskarten).
- Objekte werden sehr langfristig beim LGL gebraucht.

³² Bickhoff/Rehm, wie Anm. 2, S. 140.

³³ Steve Morris: Appraisal and Selection of Geospatial Data, White Paper, prepared for Library of Congress, November 2010, http://www.digitalpreservation.gov/meetings/documents/othermeetings/AppraisalSelection_whitepaper_final.pdf

³⁴ Landesarchiv Baden-Württemberg: Vertikale und horizontale Bewertung der Unterlagen der Vermessungsämter in Baden-Württemberg, 2006, <http://www.landesarchiv-bw.de/web/46760>.

Besondere Argumente gegen die Archivierung

Unter den Argumenten, die eindeutig gegen eine dauerhafte Erhaltung sprachen, waren folgende für digitale und georeferenzierte Daten spezifisch:

- Versuchs- und Probe-Objekte, da es sich nicht um aussagekräftige Endergebnisse handeln dürfte. Ausnahme: technikhistorisch Relevantes.
- Objekte, die besondere Aufbereitungen anderer Objekte darstellen.
- Objekte, die Angaben enthalten, die sich aus Kartendaten errechnen lassen (z. B. Flächenangaben über Gebiete, die als Polygone ausgewiesen sind).
- Inhalte des betroffenen Objekts werden in Papierform durch Bibliotheken und andere Einrichtungen im vollen Umfang nutzbar erhalten (z. B. Deutsche Grundkarte 3426-5332, Freizeitkarte Kaiserstuhl).
- Objekte in allzu proprietären Dateiformaten.

Besondere Argumente für eine Übernahme in Auswahl

Weiterhin wurde eine dritte Gruppe von Eigenschaften festgestellt, die dafür sprechen, nur einen Teil der anstehenden Objekte zu übernehmen, also nur einen Bruchteil oder auch nur ein Exemplar einer Objektart:

- Berichtsobjekte. Diese sind nur erhaltenswert, sofern ein erheblicher Mehrwert gegenüber den Einzeldaten vorhanden ist oder sie sich zur Prüfung der Ausgangsdaten anbieten. Einzelne Exemplare sind aber geeignet, das Verwaltungsverfahren zu dokumentieren (Evidenzwert).
- Der Speicherbedarf des Objekts korreliert schlecht mit den Aussichten auf die Nutzungsintensität (insbesondere bei Akkumulation über mehrere Jahrzehnte).
- Rohdaten für Produkte (z. B. Luftbilder für Orthophotografien, Laserscan-Rohdaten für ein Geländemodell). Hier ist abzuwägen zwischen Vorteilen in künftigen Nutzungsfällen und den Speicherkosten der Rohdaten, die oft diejenigen des daraus erstellten Produkts übertreffen.
- Sind die Objekte im Auftrag durch einen Dienstleister erstellt worden und sind die Verwertungsrechte nach dem Urheber- und sonstigen Leistungsschutzrecht nicht übertragen worden, bestehen in der späteren Nutzung Einschränkungen. Hier ist abzuwägen, ob das Interesse an der vollständigen Erhaltung trotz solcher Einschränkungen groß genug ist.

Übernahmehäufigkeit bei laufend aktualisierten Produkten

Wenn man dann soweit ist, die Relevanz einer Objektart festgestellt zu haben, erhebt sich die Frage, ob man alle Bearbeitungsstände der Objektart erhalten muss. Als Beispiel können die vom LGL in den letzten 20 Jahren produzierten Ausgaben topographischer Kartenblätter dienen. Man hätte hier der Meinung sein können, jede Ausgabe verdiene eine Aufbewahrung in digitaler Form. Wir haben uns dagegen entschieden und drei Zeitschichten übernommen: als erste Generation die ältesten Kartenblätter, die überhaupt als digitaler Scan vorliegen, als zweite Generation alle Karten-

blätter auf dem letzten Stand bis zum Jahr 2000, und als dritte Generation alle Kartenblätter auf dem letzten Stand bis zum Jahr 2010. Als vierte Generation werden wir im Jahr 2021 alle Blätter auf dem letzten Stand bis zum Stichjahr 2020 zu uns holen. Der Abstand von 10 Jahren erschien uns allen in der Gruppe letztlich als konsensfähige Lösung für viele Arten von Überlieferung.

Was für eine dauerhafte Aufbewahrung in staatlichen Archiven sinnvoll erscheint, trifft so noch nicht auf die Verwahrung im Geschäftsprozess zu. Hier gilt es, eine Richtlinie zu finden, die dem langsam verriegelnden Rückgriff für Vermessungsaufgaben berücksichtigt. In diesem Sinne hat das bayerische Landesamt für Vermessung und Geoinformation (LVG) sich vor Kurzem eine Richtlinie für die Erhaltung von digitalen Luftbildern gegeben.³⁵ In jedem Geschäftsjahr lässt das LVG von einem Drittel der bayerischen Landesfläche aktuelle Luftbilder anfertigen. Die bayerischen Kollegen werden die jährlich anfallenden Datenmengen an Luftbildern aus der Landesbefliegung zunächst vollständig aufbewahren. Nach Ablauf von drei Jahren werden die Daten durch verlustfreie JPEG2000-Kompression auf die Hälfte ihrer bisherigen Menge reduziert. Hier findet eine erste Formatmigration statt, die Ausgangsdaten werden gelöscht. Ab dem 11. Lebensjahr, in der zweiten Lebensdekade, wird die Datenmenge nochmals geschrumpft, diesmal durch die Absenkung der Farbtiefe von 16 Bit auf 8 Bit. Das ist die zweite Formatmigration. Ab dem 21. Geburtstag der Daten, in der dritten Dekade, werden die Befliegungen eines ganzen Jahres gelöscht, sofern sie nicht auserwählt sind, auf alle Zeit die Luftbilddaten dieser Dekade zu repräsentieren. Je Jahrgang sind dann nicht mehr drei verschiedene Luftbildsätze vorhanden, sondern nur noch einer. Dieser Datenbestand wird nach Ablauf von 31 Jahren nochmals verdichtet, indem eine verlustbehaftete, jedoch für Menschenaugen kaum wahrnehmbare Kompression auf die Bilder angewandt wird. Am Ende dieses Prozesses steht also ein landesweiter, jedes Jahrzehnt repräsentierender Datenbestand.

Aus allgemeiner archivfachlicher Sicht ist eine solche oder ähnliche zeitliche Ausdünnung ein sinnvoller Weg, der die Bewertungsentscheidungen der staatlichen Archive ergänzt. Trotzdem sind hierüber Absprachen erforderlich. So wäre zu erörtern, wann eine Ablieferung an das Landesarchiv erfolgen soll. Werden die Daten erst drei Jahrzehnte nach ihrer Entstehung geliefert, so bestehen 1. Ausfallrisiken, die bei einer früheren Übertragung der Inhalte oder zumindest einer parallelen Aufbewahrung geringer wären, und 2. dürften höhere Kosten für die Bestandserhaltung anfallen, da Formatmigrationen etc. nicht auf einer einheitlichen landesweiten Basis stattfinden. Auch die Entscheidung, einen verlustbehafteten Kompressionsalgorithmus anzuwenden, wäre zunächst mit der zuständigen Archivverwaltung zu erörtern.

Robmaterial oder Produkt? Über die Bewertung von Luftbildern und Orthophotographien

Der Kontrast zwischen unseren und den Ergebnissen des LVG Bayern eignet sich auch sehr gut dazu, archivische Bewertungsentscheidungen zu erläutern. In der AG AGL wurde nämlich entschieden, zunächst Orthophotos für den langfristigen Erhalt zu übernehmen und die Ent-

³⁵ Für diese und andere Informationen bedankt sich der Verf. herzlich beim LVG, insbesondere bei Herrn Wolfgang Stößel.

scheidung über Luftbilder zurückzustellen. Orthophotos entstehen aus Luftbildern, indem die durch die Wölbung der Erdoberfläche entstehende Verzerrung herausgerechnet wird. Aus dem Rohmaterial Luftbildaufnahme wird das Produkt Orthophotographie, auf dem alle Winkel und Strecken identisch mit dem Kartenbild sind. Das LVG Bayern hingegen wird nur Luftbilder ganz langfristig aufbewahren. Die daraus abgeleiteten Orthophotos werden dort nur für zwölf Jahre vorgehalten. Wie kommt das?

Jede Entscheidung über Erhaltung und Nichterhaltung von Information bemisst sich nach dem künftigen Kundenkreis und seinen zu erwartenden Interessen. Nach intensiver Diskussion zwischen Landesarchiv und LGL bewegt sich die Nachfrage schon für aktuelles Bildmaterial regelmäßig im Bereich von Orthophotos und nur selten im Bereich von Luftbildern. Nach den Erfahrungen des Landesarchivs wird sich diese Tendenz mit zeitlichem Abstand von der Aufnahme eher verstärken. Gleichzeitig wurde klar, dass eine Neuberechnung von Orthophotos aus vorhandenem Luftbildmaterial 1. photogrammetrischen Sachverstand, 2. Informationen über die Aufnahmesensoren, 3. Wissen über das für die Entzerrung verwendete 3D-Modell und 4. spezielle Hard- und Software voraussetzt.

Das LVG Bayern traut sich eine solche Rekonstruktion von Orthophotos aus vorhandenem Luftbildmaterial jederzeit zu, auch zwanzig oder dreißig Jahre nach Erstellung der Aufnahme. Unsere Arbeitsgruppe kam hingegen zu dem Schluss, dass uns eine solche Rekonstruktion mit mehreren Jahren oder Jahrzehnten Abstand wohl nicht immer gelingen wird. Aus diesem Grund bleibt nach heutigem Stand die langfristige Erhaltung von Luftbildern beim Landesarchiv offen und wird erst nach Übernahme der Orthophotos besprochen. Das LGL Baden-Württemberg hat gleichwohl die Möglichkeit, außerhalb der im Landesarchivgesetz definierten Ziele, also zur Erfüllung der eigenen Aufgaben, einen ähnlichen Weg wie das LVG Bayern zu beschreiten.

Grundprinzipien der Nutzung

Die Nutzung historischer digitaler Geoinformation folgt etwas anderen Prinzipien als denen, die man aus modernen GIS gewohnt ist. Im GIS ist man heute blattschnittfrei, man holt sich seine Daten über das Netz aus verschiedensten Quellen und führt alles in seinem GIS-Client zusammen, der im Idealfall in einem Webbrowser läuft. Das ist ein Maß an Freiheit, das man gern für ältere Datenbestände ebenso hätte.

Doch solche Wünsche sind abzuwägen mit den Anliegen, die ein staatliches Archiv hat. Dort gelten folgende gesetzliche Anforderungen:

- Das Archivgut muss jedermann zugänglich gemacht werden.
- Archivgut wird erst nach Ablauf einer Schutzfrist von meist 30 Jahren allgemein zugänglich, zuvor kann die Schutzfrist nur im Einzelfall auf Antrag verkürzt werden.
- Archivgut wird i.d.R. durch Einsichtnahme genutzt.
- Die Abgabe von Archivgut erfolgt nur in Kopie gegen eine Gebühr für einzelne Einheiten.

Hinzu kommen finanzielle Anforderungen der staatlichen Träger. Ein Kerngeschäft der Archive im digitalen Umfeld liegt im Scannen von Akten, Amtsbüchern, Fotos und Urkunden. Hierfür sind Präsentationssysteme entwickelt worden, die zwar Karten darstellen können, aber die Fähigkeit zur Georeferenzierung nicht mitbringen. Auf der Ebene der technischen Umsetzung sind wir damit in einem schwierigen Umfeld. Man kann wohl nicht voraussetzen, dass der berühmte juristische *Jedermann*³⁶ ein GIS-System zur Hand hat. Wir müssen uns aber auf solche Nutzer einrichten, einmal da sie einen Nutzungsanspruch haben, zum anderen auch, weil einfach sehr viele unserer Nutzer Heimat- und Ortsforscher sind, also Geisteswissenschaftler mit oft geringen technischen Fertigkeiten. Kritisch zu betrachten sind die Kosten eines Geodatenservers, die für wirklich alte, selten nachgefragte Daten nicht zu rechtfertigen wären.³⁷ Gleichzeitig wollen wir aber die Nutzung der Daten in geographischen Informationssystemen weiterhin ermöglichen. Das Ergebnis ist ein Kompromiss, der sich in folgenden Prinzipien äußert:

- Die Geodaten sind in Pakete aufgeteilt, die für sich allein eingelagert, aufgefunden, genutzt und übertragen werden können. Fachleute sprechen von Submission, Archival und Dissemination Information Packages, kurz SIP, AIP und DIP. Diese Paketierung ist erforderlich, um Computer und Leitungen nicht zu überfrachten, um Gebührenabrechnungen möglich zu machen und um Bestandserhaltungsoperationen abwickeln zu können. Die Pakete sind entweder nach Verwaltungsgrenzen gebildet (z.B. Katasterkarte der Gemeinde Beispielhofen) oder nach Blattschnittmustern oder Kilometerkacheln (z.B. Blatt L6543 der Topographischen Karte 1:10.000 auf dem Stand von 1990).
- Die Geodaten enthalten immer eine Georeferenzierung. Nach außen (in Findmittelsystemen) wird stets das aktuelle Referenzierungsformat (künftig UTM/ETRS89), im Innern der Objekte kann hingegen ein veraltetes Format vorliegen. Bei anstehenden Formatmigrationen wird dieses interne Referenzierungsformat angepasst werden können.
- Sofern die Geodaten Rasterdaten (Bilddaten) sind, können sie auch ohne ein GIS genutzt werden.
- Für Raster- und Vektordaten wird man eine einfache Anleitung mitgeben, wie man eine lokale GIS-Software verwendet, um die Daten im Kontext zu betrachten.
- Bei der Verzeichnung der Geodaten wird dafür Sorge getragen, eine möglichst genaue ortsbezogene Suche auch ohne GIS-Einsatz zu ermöglichen. Das Landesarchiv nimmt hier unter den Archiven eine Vorreiterrolle ein, indem es seit Ende April 2012 mit LEO-BW eine GIS-orientierte Sicht auf die Datenbestände der beteiligten Partnerorganisationen anbietet.

³⁶ § 6 Abs. 1 LArchG Baden-Württemberg: *Jedermann, der ein berechtigtes Interesse glaubhaft macht, hat nach Maßgabe der Benutzungsordnung das Recht, das Archivgut nach Ablauf der Sperrfristen zu nutzen, [...].*

³⁷ Aus dieser Erwägung heraus wurde bei der Umsetzung der INSPIRE-Richtlinie in deutsches Recht darauf geachtet, dass nur Geodaten berücksichtigt werden, die noch in – man hat zu ergänzen – wirtschaftlicher Verwendung stehen (§ 4 Abs. 1 GeoZG).



Abb. 6: *www.leo-bw.de: Landeskunde entdecken online, ein Portal mit Kartensuche.*

Nicht nur in technischer, sondern auch in rechtlicher Hinsicht³⁸ waren Absprachen zur Nutzung zu treffen. Kein Archivkunde ist amüsiert, wenn Nutzungsgenehmigungen in einem langwierigen Beteiligungsverfahren eingeholt werden müssen. LGL und Landesarchiv haben sich daher auf ein unkompliziertes Verfahren geeinigt, nach dem die Zuständigkeit für Nutzungsgenehmigungen bestimmt wird. Die abgegebenen Geobasisdaten unterliegen der Nutzung nach Landesarchivgesetz (LArchG) entweder

- mit dem Zeitpunkt der Abgabe an das Landesarchiv, sofern das LGL hierüber sein Einverständnis erklärt hat,
- oder mit Ablauf von zehn Jahren a) nach der Entstehung des jeweiligen Objekts oder b) (bei laufend aktualisierten Fachverfahren) nach der Ausgliederung eines Zeitschnitts als Kopie aus dem aktuellen System.

Sind Geobasisdaten in der Obhut des Landesarchivs, ohne schon nach LArchG nutzbar zu sein, so werden Nutzungsanfragen vom LGL genehmigt und abgerechnet. Erfolgt die Nutzung nach LArchG, so ist allein das Landesarchiv zuständig.

³⁸ Rechtsgrundlagen vgl. <http://www.landesarchiv-bw.de/web/46788>.

Wie geht es weiter?

Die Arbeit in Baden-Württemberg ist bereits auf gutem Weg, Teilprojekte sind auf der Zielgeraden angekommen. Die AG AGL setzt ihre Arbeit in folgenden Bereichen fort:

- Um die Ablieferung von Geodaten von der Vermessung an das Landesarchiv zu erleichtern, sind geeignete Transferpakete für Geodaten zu definieren und Übertragungswege für Metadaten zu schaffen.
- Das bestehende Bewertungsmodell für papierne Unterlagen der Vermessung aus dem Jahr 2006 ist um die digitalen Bestandteile zu erweitern und zu konkretisieren.
- Es bedarf einer Übersicht, welche Papierunterlagen der Vermessungsverwaltung mit dem Übergang auf die AAA-Systemlandschaft entbehrlich werden und ob diese insgesamt oder in Auswahl archivwürdig sind.
- Gedankenaustausch und Zusammenarbeit mit anderen Verwaltungen (Europa, Bund, Kommunen) sind fortzusetzen.
- Darüber hinaus soll die archivische Bewertung in vorhandene Infrastruktur- und Klassifikationssysteme integriert werden.
- Auch verbesserte Rahmenbedingungen für die Nutzung der Archivpakete könnten realisiert werden.
- Mit dem Übergang auf die Normbasierte Austauschschnittstelle der AAA-Architektur wird auch darüber zu reden sein, wie die bisher im ESRI Shape Format übernommenen Vektorobjekte in ein zeitgemäßeres Format auf GML-Basis überführt werden sollen. Beim Open Geospatial Consortium ist bereits ein Arbeitsprozess im Gange, der auf die Definition eines archivischen Profils für die Geographic Markup Language GML abzielt.
- Themen der Zwischenarchivierung bleiben außen vor. Zwischenarchivierung sind Dienste für Geodaten, die noch für die Geschäftsprozesse gebraucht werden, über deren historischen Wert noch nicht entschieden wurde. Hierüber wird man eines Tages ergebnisoffen diskutieren.

Bayern und Hessen arbeiten bereits in ähnlicher Weise an der Umsetzung wie Baden-Württemberg. Einige andere Bundesländer und selbst einige leistungsfähige Stadtverwaltungen wollen, sobald ihre digitalen Archivsysteme im Produktivbetrieb stehen, diesen Beispielen folgen. Man kann also sagen, das Zitat vom Anfang dieses Texts ist ungültig geworden. Inzwischen arbeiten einige Archive weltweit mit digitalen Geobasisdaten, erhalten sie und machen sie nutzbar. Dabei entstehen Kosten und Aufwand, die aber finanzierbar sind. Man kann davon ausgehen, dass gemeinsame Interessen zwischen Vermessungsverwaltung, staatlichen Archiven und Nutzern eine kontinuierliche Überlieferung der Kartographie und des Vermessungswesens auch im 21. Jahrhundert sicherstellen werden.

Formaterkennung und Formatvalidierung

Theorie und Praxis

Von CLAIRE RÖTHLISBERGER-JOURDAN

In der Praxis reicht es nicht aus, dass ein Archiv definiert hat, welche Formate es für die digitale Archivierung akzeptiert. Vielmehr muss bei jedem Ingest überprüft werden, ob diese Vorgaben tatsächlich eingehalten werden; sonst handelt sich das Archiv womöglich ein Risiko für die Bestandserhaltung ein. Diese Überprüfung vollzieht sich in zwei Schritten: zunächst in einer automatischen Formaterkennung, danach in der Validierung der gesamten Datei.

Die Formaterkennung identifiziert das Format einer Datei bis zu einer bestimmten, gewünschten Granularität. Sie stützt sich dabei auf das Vorhandensein besonders charakteristischer Eigenschaften; in der Regel sind dies bestimmte Bytesequenzen innerhalb der Datei. Die Formatvalidierung überprüft, ob eine Datei der Spezifikation ihres Formats entspricht. Dabei muss jede einzelne der in der Formatspezifikation verlangten Eigenschaften überprüft werden. Nur wenn alle Eigenschaften erfüllt sind, ist die Datei valide.

Für die Formaterkennung existieren mehrere, gut etablierte Datenbanken und Werkzeuge. Die Formatvalidierung ist technisch ungleich komplexer; entsprechend ist hier das Toolangebot beschränkt, insbesondere für Formate ausserhalb des Mainstreams.

Grundlagen

Die KOST

Die Koordinationsstelle für die dauerhafte Archivierung elektronischer Unterlagen (KOST) ist ein Gemeinschaftsunternehmen des Schweizerischen Bundesarchivs, des Landesarchivs des Fürstentums Liechtensteins, 24 kantonalen und 5 kommunalen Archiven der Schweiz. Sie hat den Auftrag, ihre Träger bei der Archivierung elektronischer Unterlagen zu unterstützen. Dazu erarbeitet sie unter anderem Standards und Richtlinien als Grundlage, stellt Tools und Dienstleistungen zur Lösung konkreter Probleme und Arbeitsschritte zur Verfügung und dokumentiert in Studien und Kolloquien den Wissensstand zu einzelnen Themen. Zudem vermittelt die KOST das zusammenfliessende Wissen den beteiligten Archiven in verschiedenen Veranstaltungen.¹

¹ Siehe zu den Grundlagen und Produkten der KOST den Beitrag von Georg *Büchler* in diesem Band sowie weitere Arbeiten von demselben, Martin *Kaiser* und Christian *Engster* in früheren Tagungsbänden des Arbeitskreises.

Der Katalog archivischer Dateiformate (KaD)

Einleitung

Der Katalog archivischer Dateiformate (KaD) ist eines der ersten Produkte der KOST. Er wurde in seiner ersten Version 2007 erarbeitet, in Antwort auf ein häufig geäußertes Desiderat vieler KOST-Trägerarchive. Dabei verfolgt er zwei Ziele: Erstens zeigt er auf, welche Formate nach heutigem Kenntnisstand theoretisch archivtauglich sind und als Zielformate für die Migration oder Konversion dienen können. Zweitens dient er im Kontakt mit der Verwaltung als Referenz dafür, welche Formate aus archivischer Sicht im aktiven Lifecycle verwendet (und entsprechend empfohlen) werden können.

Die aktuelle, zweite Version des KaD datiert aus dem Jahr 2009.

Die dritte Version wird 2012 erstellt und soll im Frühjahr 2013 publiziert werden. Neben einer generellen Aktualisierung wird der Katalog insbesondere in den Bereichen Video-, Text- und Bildformate aufdatiert sein.

Formatkategorien

Für den Katalog archivischer Dateiformate sind die folgenden abstrakten Formatkategorien relevant: Textdaten, Bilddaten, Audiodaten, Videodaten und strukturierte Daten (Tabellenkalkulation, Datenbanken). Hingegen sind Programmdateien für Archive nicht relevant, da diese keine Software archivieren.

Analyse und Bewertung

Die im Katalog enthaltenen Formate wurden aus verschiedenen Blickwinkeln oder Sichten analysiert (Abb.1):

1. Eine Bewertung anhand archivfachlicher Kriterien legt offen, in welchem Mass ein Format die Anforderungen von Archiven an die Archivtauglichkeit erfüllt und welche Risiken bei seiner Verwendung zu beachten sind. Dazu wurde ein Katalog von sechs unterschiedlich gewichteten Kriterien erarbeitet.
2. Eine Best-Practice-Analyse hält fest, wie jedes Format in der Archivwelt beurteilt und in den Verwaltungen angewendet wird. Diese Sicht lässt sich in zwei weitere Kriterien übersetzen, Best Practice und Perspektive.
3. Eine Klassifizierung der Formate erlaubt es, unterschiedliche Bewertungen gemäss den ersten beiden Sichten zu verstehen, und trägt zum Entscheid über eine Empfehlung bei. Es wird unterschieden zwischen altbekannten, weit verbreiteten Formaten, die sich auf Grund ihrer Stabilität für die Archivierung eignen; neuen Formaten, die zweifellos grosse Verbreitung erlangen werden; und potentiellen Formaten, bei deren Design die Archivtauglichkeit eine besondere Rolle gespielt hat, deren Zukunft aber noch nicht absehbar ist.

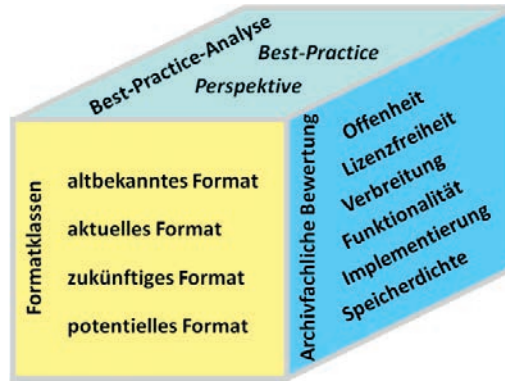


Abb. 1: Kriterien zur Analyse und Bewertung (KaD Version 1 und 2).

Die Bewertung der einzelnen Formate wird zusammengefasst in einer Empfehlung der KOST pro Formatkategorie.

Umsetzung im einzelnen Archiv

Der KaD ist nicht mehr als eine Richtlinie. Der Entscheid, welche Formate für die digitale Archivierung akzeptiert werden, obliegt dem einzelnen Archiv. Dieser Entscheid kann von Archiv zu Archiv unterschiedlich sein, weil die im KaD abgebildeten archivfachlichen Empfehlungen technischen, wirtschaftlichen oder politischen Einflüssen unterliegen.

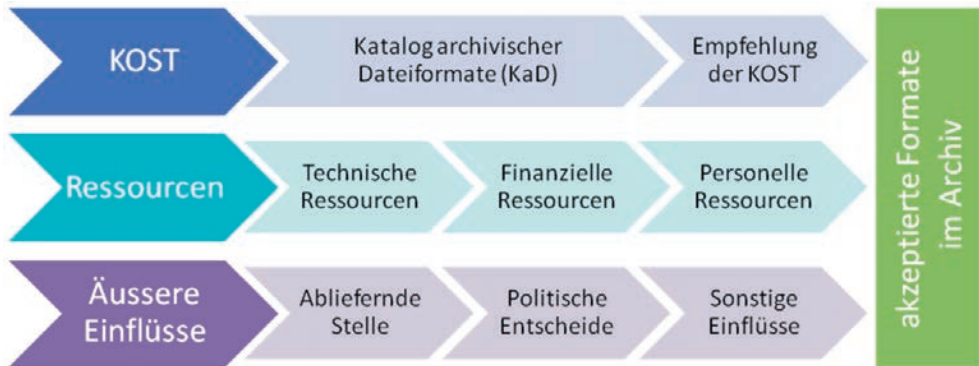


Abb. 2: Umsetzungseinflüsse im Bereich akzeptierter Formate.

Einführung in Formaterkennung und Formatvalidierung

Verifizierung der akzeptierten Formate

In der Praxis reicht es nicht aus, dass ein Archiv definiert hat, welche Formate es für die digitale Archivierung akzeptiert. Die Einhaltung dieses Entscheides muss vor jedem Ingest überprüft werden. Diese Überprüfung vollzieht sich in zwei Schritten: zunächst in einer automatischen Formaterkennung, danach in der Validierung der gesamten Datei.

Unterschied zwischen Formaterkennung und -validierung

Grundlagen

Die Formaterkennung identifiziert das Format einer Datei bis zu einer bestimmten, gewünschten Granularität. Sie stützt sich dabei auf das Vorhandensein besonders charakteristischer Eigenschaften. In der Regel sind dies bestimmte Bytesequenzen innerhalb der Datei.

Die Formatvalidierung überprüft, ob eine Datei der Spezifikation ihres Formats entspricht. Dabei muss jede einzelne der in der Formatspezifikation verlangten Eigenschaften überprüft werden. Nur wenn alle Eigenschaften erfüllt sind, ist die Datei valide.

Wenn für ein Format kein entsprechender Validator existiert, ist einzig eine näherungsweise Validierung möglich durch das Öffnen der Dateien in einem Viewer. Da Viewer in der Regel aber eine gewisse Fehlertoleranz haben, gibt allein die Tatsache, dass eine Datei ohne Fehlermeldung angezeigt werden kann, noch keine Gewissheit über ihre Validität.

Einführendes Beispiel

Eine Datei im Format PDF/A-1b enthält spezielle, charakteristische Bytesequenzen (Abb. 3). Sie beginnt mit den Zeichen %PDF-1 und enthält an einer beliebigen Stelle den RDF-Metadaten-eintrag, der sie als PDF/A, Version 1, Konformanzlevel B ausweist. Diese Bytesequenzen sind ersichtlich, wenn eine PDF/A-Datei mit einem normalen Texteditor geöffnet wird:

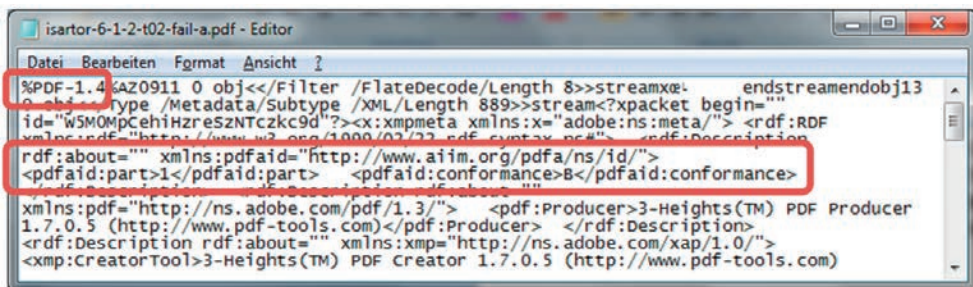


Abb. 3: Charakteristische Bytesequenzen von PDF/A-1b.

Wenn diese Elemente fehlen, erkennt eine Formaterkennungs-Software eine Datei nicht als PDF/A-1b, auch wenn die Dateiendung .pdf ist. Umgekehrt erkennt sie eine Textdatei fälschlicherweise als PDF/A-1 (Abb. 4), wenn diese Bytesequenzen vorhanden sind. (Erst beim Versuch, eine solche Datei mit einem PDF-Viewer zu öffnen (Abb. 5), oder bei der PDF/A-Validierung wird der Fehler offenbar.)

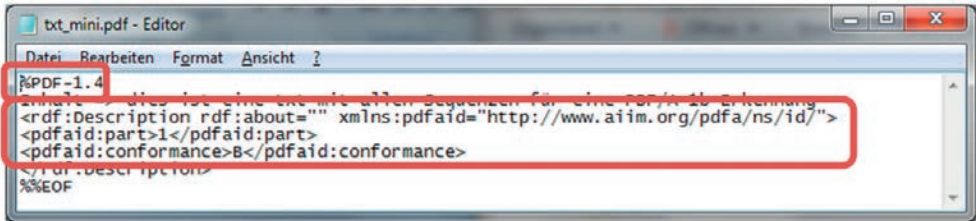


Abb. 4: Textdatei mit enthaltener PDF/A-Sequenz.

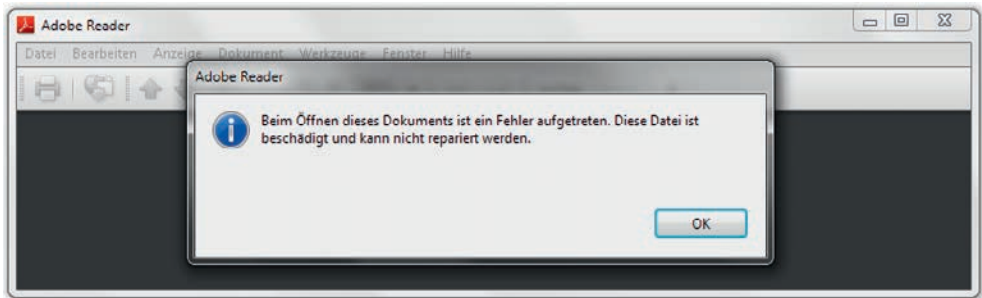


Abb. 5: Viewer-Fehlermeldung bei korrupten PDF-Dateien.

Dass das Öffnen einer PDF-Datei mit einem PDF-Viewer nicht ausreicht, zeigt das folgende Beispiel: Es handelt sich hier um eine PDF-Datei, welche die erwähnten Bytesequenzen enthält, die aber mit einem Passwort geschützt ist. Dies verletzt die Spezifikation von PDF/A. Beim Öffnen im Adobe Reader wird die Datei, gestützt einzig auf die integrierte Formaterkennung, fälschlicherweise als PDF/A identifiziert.

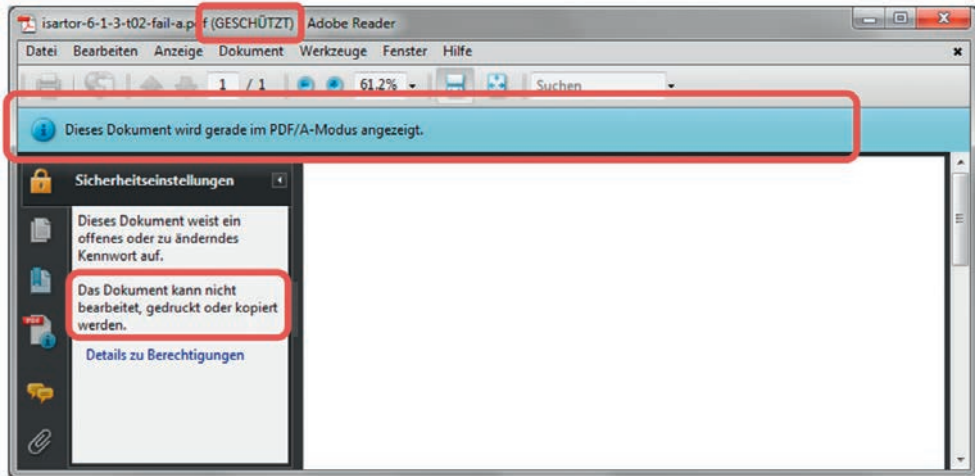


Abb. 6: Grenzen des PDF-Viewers.

Die reine Formaterkennung geht hier also fehl. Einzig ein PDF/A-Validator kontrolliert alle Eigenschaften und gewährleistet, dass es sich um ein valides PDF/A-1b handelt, das auch in Zukunft korrekt vermittelt respektive konvertiert werden kann.

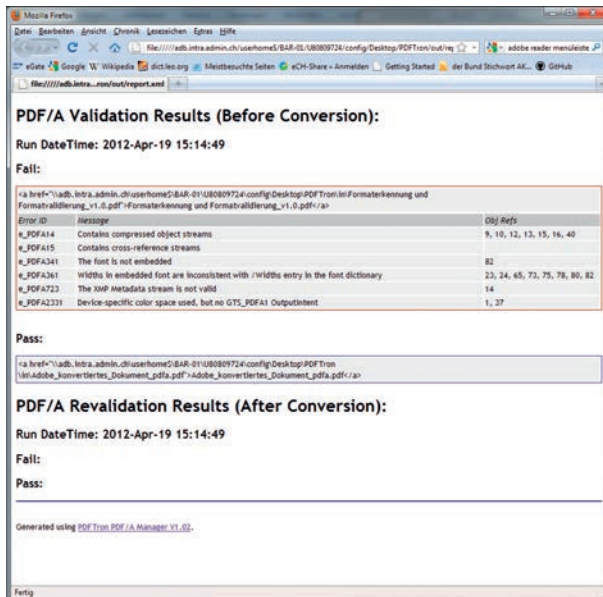


Abb. 7: Validierungsergebnisse zweier PDF-Dateien.

Nutzen und Notwendigkeit der Formaterkennung

Angesichts der aufgezeigten Grenzen und der Unschärfe der Formaterkennung, und weil immer eine anschließende Validierung nötig ist, stellt sich die Frage, weshalb es überhaupt eine Formaterkennung braucht.

Ihre Hauptrolle ist die einer Triage. Spätestens wenn die Menge der zu kontrollierenden Daten groß ist, sind die Archive auf maschinelle Unterstützung angewiesen. Anhand der Formaterkennung kann dann automatisch der korrekte Nachfolgeprozess angesprochen werden. In der nachfolgenden Grafik wird zum Beispiel anhand der Formaterkennung die Datei Test.pdf entweder zur PDF/A-Validierung, zur Konvertierung PDF zu PDF/A oder, als weitere Möglichkeit, zur Konvertierung DOC zu PDF/A weitergegeben und verarbeitet.

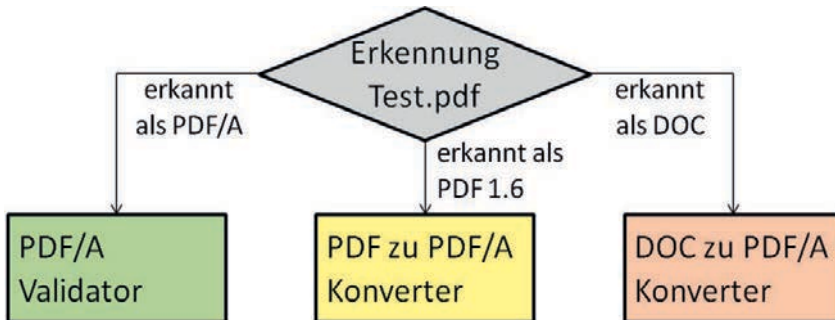


Abb. 8: Formaterkennung im Bereich der automatischen Triage.

Formaterkennung

Anforderungen an die Formaterkennung

Damit ein Tool zur Formaterkennung in den Standardprozessen der Archive verwendet werden kann, sollte es die folgenden Anforderungen erfüllen:

- Batchtauglichkeit
- Installation ohne Administratorenrechte
- Schnelle und qualitätsvolle Erkennung
- Öffentlich zugängliche sowie editierbare Liste mit den Merkmalen der einzelnen Formate

Batchtauglichkeit

Die Batchtauglichkeit ist die Grundvoraussetzung, damit die Formaterkennung auch wirklich in den Standardprozess eingebunden werden kann. Zudem müssen die Resultate in einer definierten Form ausgegeben werden, damit der Nachfolgeprozess korrekt angestoßen werden kann.

Inbetriebnahme

Wenn immer möglich soll die Software ohne Administratorenrechte installiert werden können oder eine Installation überhaupt nicht notwendig sein. Dies ist vor allem in der Testphase eine wesentliche Erleichterung.

Qualität und Geschwindigkeit

Grundsätzlich ist die Qualität höher einzustufen als die Geschwindigkeit. Die Tests, welche die KOST durchgeführt hat, basierten immer auf der Einstellung für die höchst mögliche Qualität mit entsprechendem negativen Einfluss auf die Geschwindigkeit.

Bei der Qualität ist auch die gewünschte Granularität wesentlich. Aus der Sicht der KOST sollte die Granularität genügend fein sein, um zum Beispiel mindestens eine normale PDF-Datei (z.B. PDF-1.4 oder PDF-1.7) von einer PDF/A-Datei (z.B. PDF/A-1b oder PDF/A-2u) unterschieden zu können. Die Granularität ist auch abhängig von der verwendeten Format-Datenbank.

Format-Datenbank

Die Merkmale der einzelnen Formate sollten auf einer öffentlich zugänglichen sowie editierbaren Liste beziehungsweise Datenbank basieren. Weil die Anzahl der potenziell zu identifizierenden Formate und Formatversionen sehr hoch und die Identifikation deshalb relativ komplex ist, ist es unabdingbar, dass diese Datenbank von möglichst vielen Akteuren benützt und gepflegt wird. Die in der Archivwelt am weitesten verbreitete Datenbank, welche diese Kriterien erfüllt, ist die PRONOM-Datenbank des Englischen Nationalarchivs (*The National Archives*, TNA).

Die vier bekanntesten Formaterkennungstools

Im Sinn einer Marktübersicht werden in der Folge die vier bekanntesten Formaterkennungssoftwares für die digitale Archivierung kurz beschrieben:

DROID DROID (Digital Record Object Identification) ist ein Software-Tool entwickelt durch TNA, um die automatisierte Identifikation von Dateiformaten anhand ihrer PRONOM-Datenbank durchzuführen zu können. DROID ist eine plattformunabhängige Java-Anwendung und kann sowohl von einer grafischen Benutzerschnittstelle (Graphical User Interface, GUI) als auch von einer Konsole respektive einer auf Zeichen basierenden Benutzerschnittstelle (Command Line Interface, CLI) aufgerufen werden. DROID ist kostenlos und steht unter der BSD-Lizenz für DROID v4.0. Es kann von <http://sourceforge.net/projects/droid/> heruntergeladen werden.²

² Alle Links wurden am 4.10.2012 überprüft.

- File** File ist ein UNIX-Programm zur Formaterkennung einer Datei. Die *magic numbers*, welche für die Erkennung verwendet werden, sind in einer ASCII-Textdatei namens *Magic* angegeben. File ist kostenlos und kann unter anderem von <ftp://ftp.astron.com/pub/file/> heruntergeladen werden.
- Fido** Fido (Format Identification for Digital Objects) ist ein CLI-Tool für Windows und Linux zur Identifikation von Dateiformaten. Es wurde von der Open Planets Foundation (OPF) zur einfachen Integration in automatisierte Prozesse entwickelt. Fido verwendet die PRONOM Signaturen, übersetzt sie jedoch in reguläre Ausdrücke und wendet diese direkt an. Fido ist frei erhältlich unter der Apache 2.0 Lizenz und kann von <http://github.com/openplanets/fido/downloads> heruntergeladen werden.
- Tika** Apache Tika™ Toolkit erkennt und extrahiert Metadaten und strukturierte Text-Inhalte aus verschiedenen Dokumenten in einer der vorhandenen Parser-Bibliotheken. Für die Formaterkennung wird der *Mime Magic* Detection Teil verwendet.
- Tika ist ein Projekt der Apache Software Foundation unter der Apache License, Version 2.0. Die neueste Tika Version kann von <http://tika.apache.org/download.html> kostenlos heruntergeladen werden.

Grobbewertung der Formaterkennungssoftware

Asger Blekinge hat drei der erwähnten Tools für das Scape-Projekt getestet. Die Testresultate sind auf dem Blog der Open Planets Foundation publiziert.³ Sie werden hier kurz zusammengefasst und ergänzt (Tab. 1). Dabei muss jedoch beachtet werden, dass die gewünschte Granularität und die verwendeten Einstellungen die Resultate stark beeinflussen.

³ Asger *Blekinge*: Identification tools, an evaluation. <http://www.openplanetsfoundation.org/blogs/2012-02-23-identification-tools-evaluation>.

Name	Stärke	Neutral	Schwäche
DROID	Inbetriebnahme PRONOM-DB Qualität		Batchtauglichkeit Geschwindigkeit
File	Batchtauglichkeit Inbetriebnahme Geschwindigkeit		Qualität Andere DB
Fido	Batchtauglichkeit PRONOM-DB Qualität	Inbetriebnahme	Geschwindigkeit
Tika	Batchtauglichkeit Inbetriebnahme		
Geschwindigkeit	Qualität Andere DB		

Tab. 1: Grobbewertung der Formaterkennungssoftware.

Analyse der verwendeten Formatdatenbanken

Die vier Formaterkennungstools verwenden drei verschiedene Quellen für die Erkennung einzelner Dateien. Damit sich jeder ein besseres Bild machen kann, werden die einzelnen Quellen anhand des Eintrags für PDF 1.4 erläutert. In den Fällen, in welchen diese Granularität nicht gegeben ist, wurde der PDF-Eintrag verwendet.

1. PRONOM-Datenbank

Die PRONOM-Datenbank ist von TNA erstellt worden und wird laufend durch die weltweiten Benutzer und Softwarehersteller aktualisiert und ergänzt. Die PRONOM-Datenbank wird sowohl von DROID als auch von Fido verwendet. Die Datenbank enthält rund 860 Formateinträge.

Nachfolgend ein Auszug aus dem Eintrag von PDF 1.4:

Name	Acrobat PDF 1.4 - Portable Document Format			
Version	1.4			
Other names	PDF (1.4)			
Identifiers	MIME: application/pdf Apple Uniform Type Identifier: com.adobe.pdf PUID: fmt/18			
External signatures	File extension: pdf			
Internal signatures	Name	PDF 1.4		
	Description	Header and footer		
	Byte sequences	Position type	Absolute from BOF	
		Offset	0	
		Byte order		
		Value	255044462D312E34	
		Position type	Absolute from EOF	
		Offset	0	
		Byte order		
Value		2525454F(46 460A 460D 460D0A 460D00)		

Abb. 9: PDF-1.4-Eintrag aus der PRONOM-Datenbank.

PRONOM gibt neben dem verbreiteten MIME-Type (application/pdf) auch einen eigenen Identifikator heraus, den Persistent Unique Identifier (PUID, im Beispiel fmt/18) heraus, welcher die entsprechende Granularität aufweist.

Unter Byte sequences steht, dass das PDF 1.4 mit %PDF-1.4 (Hexadezimale Schreibweise 255044462D312E34) beginnen muss und mit %%EOF oder %EOF. oder %EOF.. (Hexadezimale Schreibweise 2525454F46, 2525454F460A, 2525454F460D, 2525454F460D0A oder 2525454F460D00) endet.

2. Magic-Datensammlung

In der File-Applikation befindet sich auch die Magic-Datensammlung. Diese enthält rund 230 Dateien, welche teilweise mehrere Formatidentifikationen enthalten, jedoch nur einen einzelnen allgemeinen Eintrag zu PDF.

Nachfolgend der Eintrag zu PDF:

```

#-----
# $File: pdf,v 1.6 2009/09/19 16:28:11 christos Exp $
# pdf: file(1) magic for Portable Document Format
#
0  string      %PDF-      PDF document
!:mime  application/pdf
>5  byte       x          \b, version %c
>7  byte       x          \b.%c

```

Abb. 10: PDF-Eintrag aus der File Magic-Datensammlung.

In der Identifikationszeile (0 string %PDF- PDF document) steht, dass ein PDF mit dem Text %PDF- beginnen muss und dass bei einer erfolgreichen Erkennung *PDF document* ausgegeben werden soll. File gibt auch den MIME-Type (application/pdf) aus.

3. Tika-mimetypes.xml

In Tika org.apache.tika.detect.MagicDetector wird Tika-mimetypes.xml zur Verfügung gestellt. Tika-mimetypes.xml enthält rund 170 Dateneinträge, wovon nur ein einzelner allgemeiner Eintrag zu PDF definiert ist.

Nachfolgend der Eintrag zu PDF:

```

<mime-type type="application/pdf">
  <acronym>PDF</acronym>
  <comment>Portable Document Format</comment>
  <magic priority="50">
    <match value="%PDF-" type="string" offset="0" />
  </magic>
  <glob pattern="*.pdf" />
  <alias type="application/x-pdf" />
</mime-type>

```

Abb. 11: PDF-Eintrag aus Tika-mimetypes.xml.

Die Tika-Einträge bauen auf den MIME-Type Einträgen (z.B. application/pdf) auf.

In der Identifikationszeile (<match value="%PDF-" type="string" offset="0" />) steht, dass ein PDF mit dem Text %PDF- beginnen muss.

4. Fazit zu den Formatdatenbanken

Die Qualität der Formaterkennung durch File und Tika ist ungenügend, weil die Granularität nicht fein genug ist. Entsprechend wurden diese beiden Programme keiner Detailbewertung unterzogen. DROID und Fido hingegen wurden detailliert getestet im Hinblick auf ihre Qualität und Geschwindigkeit (als untergeordnetes Kriterium). Für den Test wurde ein Sample von 664 Dateien verwendet sowie (zur besseren Vergleichbarkeit der beiden Tools) einheitlich das DROID_SignatureFile_V55.

Detailbewertung von Fido 1.0.0

1. Qualität und Geschwindigkeit

Fido benötigte für die Erkennung mit einer theoretischen Buffergröße von 10GB weniger als 6 Minuten. Fido wendet die PRONOM-Datenbank korrekt an.

Wird die Standard-Buffergröße von 128KB verwendet, benötigt Fido zwar weniger als 20 Sekunden, erkennt aber 44 Dateien nicht korrekt, weil nicht alle Bytes der grossen Dateien eingelesen wurden und sich jeweils mindestens eine wesentliche Bytesequenz erst am Schluss der Datei befunden hätte. Im Test waren insbesondere PDF/A- und SIARD-Dateien davon betroffen. Entsprechend ist eine solche Reduzierung der Buffergröße für einen hohen Zeitgewinn in Anbetracht des Qualitätsverlustes nicht zu empfehlen.

Fido bietet im Gegensatz zu DROID jedoch die Möglichkeit, die Anzahl der für die Formaterkennung verwendeten PUIDs einzugrenzen, sei dies durch das Ausschliessen von gewissen PUIDs oder durch das Definieren zu verwendender PUIDs. In einem zweiten Test wurden einzig 10 PUIDs für die Erkennung ausgewählt, wiederum bei einer theoretischen Buffergröße von 10GB. Dabei benötigte Fido weniger als 40 Sekunden für die Erkennung. Bei den erkannten Formaten wurden durch Fido 43 PDF-Dateien anhand der Dateieindung fälschlicherweise als PDF/A erkannt. Abgesehen von dieser Kinderkrankheit von Fido wurden alle gewünschten PUIDs richtig erkannt. Sobald dieser Fehler behoben ist, ist diese Form von Eingrenzung sehr zu empfehlen, wenn man grundsätzlich mit einer kleinen Anzahl archivtauglicher Formate operiert. Der kleine Rest von dabei allenfalls nicht erkannten Dateien kann dann in einem zweiten Durchgang ohne Einschränkung der PUIDs durch Fido identifiziert werden.

2. Batchtauglichkeit

Fido ist ein Konsolenprogramm und erfüllt die Batchtauglichkeit zu 100%. Alle Einstellungen bis hin zum Format der Ausgabe werden über den Erkennungsbefehl Fido übergeben.

3. Inbetriebnahme

Für die Installation von Fido benötigt man grundsätzlich keine Administratorenrechte. Da Fido jedoch ein Python-Programm ist, muss Python auf dem Rechner vorhanden sein. Es besteht die Möglichkeit, Python 2.7 als Portable-Applikation zu verwenden.

Detaillbewertung von DROID 6.0.1

1. Qualität und Geschwindigkeit

DROID benötigte für die Erkennung mit einer theoretischen uneingeschränkten Buffergröße circa 1 Minute. Das Resultat der Erkennung entspricht den erwarteten Resultaten anhand der PRONOM-Datenbank und ist identisch mit jenem von Fido.

Wird die Buffergröße von 128KB verwendet, benötigt DROID nur noch 5 Sekunden, erkennt jedoch die gleichen 44 Dateien wie Fido nicht korrekt. Entsprechend ist eine solche Reduzierung der Buffergröße für einen hohen Zeitgewinn in Anbetracht des Qualitätsverlustes nicht zu empfehlen.

2. Batchtauglichkeit

DROID überzeugt insbesondere bei der Anwendung des GUI. Im Gegensatz dazu ist die Konsolenversion sehr unbefriedigend; es erwies sich im Rahmen der Tests insbesondere als unmöglich, einen brauchbaren Output zu generieren. Da DROID ein quelloffenes Java-Programm ist, wäre es theoretisch möglich, DROID direkt in eine Java-Applikation einzubinden. Leider konnte die KOST DROID Version 6 als Java-API (Application Programming Interface) nicht zum Laufen bringen. Eine Einbindung der Version 5 war zwar möglich, dabei wurden jedoch sehr grosse Dateien trotz uneingeschränkter Buffergröße teilweise nicht korrekt erkannt.

3. Inbetriebnahme

Für die Installation von DROID werden keine Administratorenrechte benötigt.

Fazit zur Formaterkennungssoftware

Ihr Umfang und ihre Granularität sprechen für die Verwendung der PRONOM-Datenbank für die Formaterkennung. Es ist für die Archive empfehlenswerter, sich am Ausbau und an der Verbesserung von PRONOM zu beteiligen, als eigene Formaterkennungs-Datenbanken und -Werkzeuge aufzubauen.⁴ Als Softwaretools kommen folglich DROID oder Fido in Frage. Beide sind im Gegensatz zu Tika und File deutlich langsamer, was aber grösstenteils auf die Granularität der Erkennung respektive die Grösse von PRONOM zurückzuführen ist.

⁴ In diesem Sinn ist Steffen *Bachmann* und Katharina *Ernst* zu widersprechen, welche eine gewichtete Formaterkennung in einem Modulrahmen postulieren (Formaterkennung – Ziele, Herausforderungen, Lösungsansätze. In: Auf dem Weg zum digitalen Archiv. Stand und Perspektiven von Projekten zur Archivierung digitaler Unterlagen. 15. Tagung des Arbeitskreises *Archivierung von Unterlagen aus digitalen Systemen* am 2. und 3. März 2011 in Schwerin. Hg. von Matthias *Manke* (Veröffentlichungen des Landeshauptarchivs Schwerin). Schwerin 2012. Die von ihnen geschilderten Probleme bei der Formaterkennung beruhen auf Eigenheiten oder Fehlern von PRONOM, beispielsweise im Verzicht auf eine Erkennung der einzelnen TIFF-Versionen und ihrer Ersetzung durch einen generischen TIFF-Eintrag (PUID fmt/353). Für Fehler in PRONOM sollen wenn möglich Korrekturvorschläge erarbeitet und eingereicht werden, um die Qualität der Datenbank zu erhöhen. Selbstverständlich ist zu beachten, dass immer die aktuelle PRONOM-Version respektive das neueste *Signature File* von DROID verwendet wird.

Die Stärken und Schwächen von DROID und Fido werden aus der Detailbewertung deutlich. Wer die Formaterkennung nur ab und zu einsetzen will, ist sicherlich mit dem GUI von DROID bestens bedient. Für den Einsatz in einem maschinellen Prozess empfiehlt sich die Verwendung von Fido, weil hier keine Abstriche bei der Batchtauglichkeit gemacht werden müssen. Die Geschwindigkeit sollte bei der Bewertung eine untergeordnete Rolle spielen. Für Fido spricht weiterhin, dass das Programm aktuell mindestens durch 3 Personen aus verschiedenen Archiven entwickelt wird und sein Quellcode viel übersichtlicher ist, währenddem die kontinuierliche Weiterentwicklung von DROID durch diverse Personalwechsel bei TNA erschwert wurde.

Formatvalidierung

Anforderungen an die Formatvalidierung

Damit ein Tool zur Formatvalidierung in den Standardprozessen der Archive verwendet werden kann, sollte es folgende Anforderungen erfüllen:

- Batchtauglichkeit
- Installation ohne Administratorenrechte
- Schnelle und qualitative Validierung
- Nachvollziehbarkeit / Überprüfung der Validierung

Die Anforderungen an die Formatvalidierung sind beinahe identisch mit jenen an die Formaterkennung. Bei der Qualität ist nicht nur massgebend, ob die Datei valid oder invalid ist, sondern auch, dass die Fehlermeldung dem effektiven Fehler entspricht.

Die Nachvollziehbarkeit der Validierung muss gegeben sein. Ist dies nicht möglich, weil zum Beispiel die Software nicht quelloffen ist, muss die Validierung mit Hilfe von invaliden Test-Dateien für alle Anforderungen einzeln überprüft werden.

Validierungssoftware

Aufgrund der Anforderungen an die Validierung wird in der Regel ein spezialisiertes Tool pro Format benötigt.

PDF/A

Die Validierung von PDF/A ist sehr komplex und nicht einfach realisierbar. Weil zudem die Nachfrage sehr gross ist, existiert ein Markt für PDF/A-Validatoren und sind quelloffene Tools nicht erhältlich. Die Überprüfung der Qualität der Validatoren wurde durch die KOST durchgeführt und basiert auf den Grundlagen von PDFlib. Die Resultate sind in einer Studie zu den

PDF/A-Konvertern zusammengefasst und auf der KOST-Website⁵ veröffentlicht.

Nachfolgend sind die 4 bekanntesten Validatoren⁶ in einer Produktübersicht aus den Jahren 2009–2010 zusammengefasst (Tab. 2):

PDF/A Validatoren	Adobe: Adobe Acrobat 9.1	Intarsys: PDF/A Live	PDF Tools: 3Heights PDF Validator Shell	PDFTron: PDF/A Manager
Geschwindigkeit & Robustheit: Sehr gut = <1 Minute & ohne Absturz Gut = 1 - 5 Minuten & ohne Absturz Ausreichend = >5 Minuten & ohne Absturz Mangelhaft = Absturz	Ausreichend	Gut	Sehr gut	Sehr gut
Genauigkeit: ⁷ Sehr gut = Mittelwert >=95% Gut = Mittelwert 90% - 94% Ausreichend = Mittelwert 75 - 89% Mangelhaft = Mittelwert <75% Isartor testsuite (non-conforming) 6.1 File structure 31x 97% 6.2 Graphics 47x 100% 6.3 Fonts 28x 100% 6.4 Transparency 6x 100% 6.5 Annotations 25x 96% 6.6 Actions 37x 100% 6.7 Metadata 27x 100% 6.9 Interactive Forms 3x 100% Other non-conforming ISO 19005 violations 9x 89% XMP 2004 violations 5x 20% PDF 1.4 violations 8x 38% Conforming Real world 34x 88% PDFlib samples 8x 100% Advanced XMP 16x 100% Mittelwert (conforming / non-conforming) 91%	Gut	Sehr gut	Gut	Sehr gut
Getestete Version:	9.1.0	5.0.4	1.8.32.1	1.00 (CLI)
Tester / Testjahr:	PDFlib / 2009	KOST / 2010	PDFlib / 2009	KOST / 2010
Bemerkungen:	Die Version 9.0 ist bei der Geschwindigkeit und Robustheit mangelhaft.	Test mit einer neueren Version nochmals komplett durchgeführt.		Report ist sehr übersichtlich (Kompakt mit guter Fehlermeldung).

Tab. 2: Auszug der Produktübersicht aus der KOST-Studie.

⁵ PDF/A-Validatoren. Studie der KOST. Bern 2010. http://kost-ceco.ch/cms/index.php?pdfa_validatoren_de.

⁶ In der Studie der KOST wurde ebenfalls der pdfaPilot von Callas untersucht. Da dieses Produkt in Adobe Acrobat unter dem Namen Preflight Validator enthalten ist, und da sich die Ergebnisse dieser beiden Validatoren nicht wesentlich unterscheiden, wurde für eine bessere Lesbarkeit pdfaPilot von Callas in der Übersicht weggelassen. Die komplette Produktübersicht ist in der Studie ersichtlich.

⁷ Bei der Genauigkeit ist nicht nur das Ergebnis (bestanden / durchgefallen) wesentlich, zusätzlich muss die Fehlermeldung mindestens einen realen Fehler beschreiben.

Bei der Validierung von PDF/A muss beachtet werden, dass PDF/A zwar als ISO-Standard genormt ist, jedoch im Standard nur aufgelistet wird, welche einzelnen Funktionen der zugrunde liegenden PDF-Version obligatorisch, empfohlen, eingeschränkt oder verboten sind. Diese werden vereinzelt in den Details unterschiedlich interpretiert. Zudem sind alle Dokumente, die den PDF/A Standard zusammen definieren, sehr umfangreich und sehr technisch. Dies führt dazu dass in einigen Fällen unterschiedliche Validierungsergebnisse existieren.

TIFF

Für die Validierung von TIFF-Dateien ist der KOST einzig der JHOVE-Validator bekannt. JHOVE (JSTOR/Harvard Object Validation Environment) ist quelloffen und gratis verfügbar. Eine unabhängige Überprüfung der Validierungsqualität anhand des Quellcodes ist der KOST nicht bekannt.

JPEG2000

Für die Validierung von JPEG2000-Dateien ist der KOST nebst dem JHOVE-Validator auch der jpylyzer von OPF bekannt. Beide Validatoren sind quelloffen und gratis verfügbar. Unabhängige Überprüfungen der Validierungsqualität anhand der Quellcodes sind der KOST nicht bekannt.

WAV

Für die Validierung von WAV-Dateien ist der KOST einzig der JHOVE-Validator bekannt. JHOVE ist quelloffen und gratis verfügbar. Eine unabhängige Überprüfung der Validierungsqualität anhand des Quellcodes ist der KOST nicht bekannt.

SIARD

Für die Validierung von SIARD-Dateien existierte bis 2012 kein Validator. Deshalb entwickelt die KOST gegenwärtig die Anwendung SIARD-Val (Validator für SIARD-Dateien), die unter der GPL3-Lizenz und gratis zur Verfügung gestellt wird. Geplant ist der Beta-Release für den Herbst 2012. Die Qualität der Validierung muss entsprechend anschliessend durch eine andere Institution anhand des Quellcodes überprüft werden.

Fazit zu Formatvalidierungssoftware

Nicht nur die Formatvalidierung selber, sondern auch die Qualitätssicherung der Validierungstools ist sehr aufwändig. Für viele Validatoren existieren deshalb noch keine unabhängigen Überprüfungen der Qualität ihrer Resultate. Diese Lücken sollten durch die Archivgemeinschaft geschlossen werden. Die KOST plant deshalb, ihre Arbeit an und ihre Untersuchungen von Validierungstools fortzusetzen. Welche Formate als nächste in den Fokus genommen werden, ist noch offen.

AV-Unterlagen für DIMAG

Von PETER BOHL und JOHANNES RENZ

Der Beitrag handelt von der Entstehung des AV-Archivs beim Hauptstaatsarchiv Stuttgart, seinem derzeitigen Aufgabenspektrum und davon, wie in ihrem Bestand gefährdete audiovisuelle Aufzeichnungen auf DAT-Bändern von ihren Trägern gelöst und in das Erhaltungssystem DIMAG überführt werden. Er schließt mit weiteren Planungen für audiovisuelle Unterlagen.

Das Audiovisuelle Archiv (AV-Archiv) – eine kurze Geschichte

Funktionen und Aufgaben des AV-Archivs

Das Audiovisuelle Archiv des Landesarchivs Baden-Württemberg ist ein Spezialarchiv, das organisatorisch dem Hauptstaatsarchiv Stuttgart angegliedert ist. Seine Zuständigkeit liegt in den Bereichen Übernahme und Verwahrung von audiovisuellen Unterlagen aller Abteilungen des Landesarchivs, die im Rahmen von behördlichen Ablieferungen und der Ergänzungsdokumentation (archivwürdige Unterlagen aus nichtstaatlichen Registraturen bzw. Archiven und Nachlässen) einkommen. Die wichtigsten Funktionen und Aufgaben sind die Bestandserhaltung, Digitalisierung und Erschließung der Ton- und Filmdokumente. Als ein weiterer wichtiger Tätigkeitsbereich des Audiovisuellen Archivs hat sich über die Jahre die Aufgabe des Dienstleisters für kirchliche und kommunale Archive, Wirtschafts- und Universitätsarchive herauskristallisiert, die in der Regel weder die technische Ausstattung noch die finanziellen Mittel besitzen, um dieses nicht alltägliche Archivgut sachgerecht aufzubereiten, geschweige denn das Fachpersonal dafür haben. Nachgefragt werden diese Quellen sowohl von privaten als auch von kommerziellen Nutzern, vor allem vor dem Hintergrund, dass die Bestände des Audiovisuellen Archivs über Internet-Recherchen zu fast 90 Prozent nutzbar sind. Ein leidiges Problem ist immer die Zahl der Mitarbeiter, die trotz Zunahme der Aufgaben seit Jahren mit einer Tontechnikerin und einem Dokumentar stagniert.



Abb. 1: Unterschiedliche Tonbandgeräte zum Abspielen verschiedener Tonbänder mit unterschiedlichen Bandlaufgeschwindigkeiten und Spulengrößen.

Die Ursprünge des AV-Archivs reichen in das Jahr 1988 zurück. Der Intendant des Süddeutschen Rundfunks (SDR), Hans Bausch, trat an die Landesregierung heran wegen der dauerhaften Sicherung der Sendungen des Fernsehens und Rundfunks. Er war gelernter Historiker und arbeitete lange Jahre als Zeitungsjournalist und wurde im Jahr 1959 zum Intendanten des SDR gewählt. Zur Zeit seines Amtsantritts fand in der Rundfunkanstalt eine große Löschkaktion von Tonbändern statt, um Kosten zu sparen und die Bänder erneut nutzen zu können. Möglicherweise war diese Löschkaktion einer der Auslöser für den Wunsch des scheidenden Intendanten, die historisch bedeutsame Rundfunküberlieferung durch einen Dritten dauerhaft bewahren zu lassen. Ein Jahr später schlossen das Land Baden-Württemberg, vertreten von der Landesarchivdirektion, und der SDR einen ausgesprochen restriktiven Vertrag zur Überlieferungsbildung mit genauen Angaben zu Sendetitel und Sendezeit der zum Mitschneiden zugelassenen Sendungen.¹ Das Audiovisuelle Archiv (AV-Archiv) begann mit zwei Tontechnikerinnen, einem Dokumentar und dem Leiter des AV-Archivs im Jahr 1989 seine Arbeit. Von 1990 bis 1995 wurden fast ausschließlich Sendungen des SDR mitgeschnitten und erschlossen, die von politischer, gesellschaftlicher oder historischer Bedeutung waren.

Im Jahr 1995 unter der neuen Leitung des AV-Archivs wurde der Vertrag mit dem Südwestfunk (SWF) geschlossen, der dann auch Sendungen dieses Senders mit einschloss. Im gleichen Jahr wurden in Kooperation mit dem SDR Tondokumente aus der Zeit zwischen 1945 und 1950 im AV-Archiv auf Digital Audio Tapes (DAT) kopiert und der Nutzung zur Verfügung gestellt.

¹ Vertrag zwischen dem Süddeutschen Rundfunk und dem Land Baden-Württemberg über die Verwahrung, Erhaltung, Erschließung und Nutzung von Hörfunk- und Fernsehproduktionen des Süddeutschen Rundfunks vom 17. November 1988, in: Archivrecht in Baden-Württemberg: Texte, Materialien, Erläuterungen. Bearb. von Hermann *Bannasch*. Unter Mitw. von Andreas *Matsch*. Mit einer Einführung in das Landesarchivgesetz von Gregor Richter. Stuttgart 1990, S. 49–55.

Der Landtag von Baden-Württemberg gab in der gleichen Zeit seine Tonbänder der mitgeschnittenen Landtagssitzungen ab 1982 an das AV-Archiv ab, diese Bänder wurden ebenfalls auf DAT kopiert. In den folgenden Jahren überließ der SDR weitere Bänder mit Mitschnitten der Sendungen Politischer Wochenbericht und Chronik der Woche aus den Jahren ab 1959 bis in die 1980er Jahre zum Kopieren auf DAT dem AV-Archiv. Die Originalbänder und eine DAT-Kopie gingen an den SDR zurück.

Seit 1996 wurde der Anteil der vom SDR, SWF und – ab 1998 nach der Fusion der beiden Sender zum SWR – vom SWR mitgeschnittenen Sendungen sukzessive reduziert, da der SDR bzw. dann der SWR seine Sendungen nicht rigoros löschte, sondern weiter als Sendervermögen aufbewahrte und auch zur Nutzung durch Externe zur Verfügung stellte. Gleichzeitig nahm der Anteil der von Behörden, Verbänden, Vereinen und kommunalen Archiven und Privatleuten dem AV-Archiv übergebenen AV-Materialien sehr stark zu. Das AV-Archiv entwickelte sich in den folgenden Jahren zum Dienstleister für nichtstaatliche Archive und kopiert und digitalisiert inzwischen AV-Medien für andere staatliche und nichtstaatliche Archive.

Ab 2001 digitalisiert das AV-Archiv konsequent alle einschlägigen Datenträgerformate (Film: Normal- und Super-8 Filme, 16 mm Filme, U-Matic-Filme, VCR-Videokassetten, Betacam-Kassetten, SVHS- und VHS-Kassetten). Allein 35 mm Kinofilmformate müssen aufgrund der fehlenden technischen Ausstattung und dem geringen Anfall an externe Dienstleister zur Digitalisierung übergeben werden. Ab 2011 werden die vorhandenen DAT-Kassetten, da diese Technik und das Format nicht mehr unterstützt werden, zu WAVE-Dateien umgewandelt.

Technik des AV-Archivs: Tondokumente

Das AV-Archiv verwahrt und bearbeitet unterschiedliche Originaltonträger, darunter befinden sich Studiotonbänder, Tonbänder mit unterschiedlichen Laufgeschwindigkeiten und Spulengrößen. Daneben werden auch Schallplatten und vor allem Kompaktkassetten technisch bearbeitet und digitalisiert. An Abspielgeräten müssen aufgrund der unterschiedlichen Tonträger auch die entsprechenden Studiobandmaschinen, Tonbandgeräte der unterschiedlichen Hersteller (z.B. Revox, Uher), Kompaktkassettenbandgerät von Studer und auch ein Studioplattenspieler bereitgehalten und gewartet werden. Das Digital Audio Tape (DAT) wurde ursprünglich als Tonträger für die längerfristige und qualitativ ausreichende Form zur Archivierung dieser Dokumente genutzt. Doch gerade im Bereich der audiovisuellen Medien vollziehen sich bei der Technik und den Trägermedien rasche Veränderungen. Seit 2007, als sich abzeichnete, dass DAT-Aufnahme- und Abspielgeräte vom Markt verschwinden, wurden Tonaufzeichnungen auf CD-ROM nach dem Format Audio-CD (CD-DA) erstellt und zugleich auf einem NAS-Server im WAVE-Format abgelegt.

Technik des AV-Archivs: Filmdokumente

Wie bei den Tondokumenten gilt für Filmdokumente das Gleiche. Unterschiedliche Originalfilmträger, 8 mm Filme (Normal 8, Super 8), 16 mm Filme, 35 mm Filme (Kinoformat), U-Matic, VCR (Videoformat der Firmen Philips und Grundig ca. 1970-1985), VHS (ab 1985), SVHS,

Betacam, DVDs werden dem AV-Archiv übergeben, die dann mit adäquaten, der Technik der Originale angepassten Abspiel- und Kopiergeräten bearbeitet werden müssen. Die benutzten Filmprojektoren wurden technisch umgerüstet, um die Filme digitalisieren zu können. Ebenso mussten entsprechende VCR-Videorekorder, U-Matic, VHS- und SVHS-Videorekorder, Kombigeräte (S)VHS-DVD-Rekorder und DVD-Rekorder beschafft und gewartet werden. Bis zum Jahr 2001 wurden als archivischer Filmträger SVHS-Bänder genutzt, von denen qualitativ bessere Kopien gezogen werden konnten, als von den gewöhnlichen VHS-Kassetten. Seitdem werden im AV-Archiv DVDs im VOB (MPEG)-Format genutzt, eine weitere Kopie wird auf NAS-Servern im MPEG- und AVI-Format abgelegt. Die Daten auf den NAS-Servern werden wiederum auf externen USB-Festplatten gesichert.

Umfang der AV-Medien

Das Audiovisuelle Archiv verwahrt (Stand März 2012):

- Mitschnitte von Rundfunk- und Fernsehsendungen: 1.310 DAT- und 1.230 SVHS-Kassetten.
- AV-Medien von Behörden, Dienststellen und Landtag: 2.000 DAT-Kassetten, 1.250 SVHS/VHS-Kassetten und 1.050 Kompaktkassetten
- aus Nachlässen, Vereins- und Verbandsarchiven kirchlichen und kommunalen Archiven, Wirtschaftsarchiv: 845 DAT, 120 SVHS/VHS, und 3.250 Kompaktkassetten

Insgesamt werden im AV-Archiv Ton- und Bildträger im Umfang von ca. 16.000 Stunden verwahrt, die mit ca. 40.000 Verzeichnungseinheiten erschlossen sind.



Abb. 2: Tonbänder einer Ablieferung des Landeskirchlichen Archivs Stuttgart.

Aufbereitung von AV-Unterlagen für den digitalen Massenspeicher DIMAG²

Ausgangslage

Digitale Audiotapes (DATs) sind als Tonträger in den letzten Jahren obsolet geworden, d. h., sie können nicht mehr neu im Handel erworben werden. Das gleiche gilt für Abspielgeräte, welche nur noch, und das naturgemäß in abnehmender Stückzahl, gebraucht erhältlich sind. Bei defekten Geräten dürfte der Kostenaufwand für Reparaturen, wenn überhaupt möglich, zunehmend teurer werden. Dazu kommt, dass DAT-Aufzeichnungen, obwohl bereits in digitaler Technologie erstellt, nicht in einem Dateiformat vorliegen und daher nur durch Abspielen ausgelesen werden können. Die Bitfolgen der Tonspur auf dem Band gleichen weitgehend dem WAVE-Format, die Umwandlung in eine echte Datei muss aber hardwareseitig erfolgen und ist eine Sache, die immer weniger Spezialisten beherrschen.



Abb. 3: DAT-Band (Foto: Yaoleilei via Wikimedia Commons³).



Abb. 4: DAT-Abspielgerät (Foto: Tammy2 via Wikimedia Commons⁴).

² Vgl. den Beitrag von Christian Keitel: DIMAG-Kooperationen, S. 147-155 in diesem Band und Christian Keitel, Rolf Lang: DIMAG und IngestList. Übernahme und Nutzung von digitalen Unterlagen im Landesarchiv Baden-Württemberg. In: Gerald Maier und Thomas Fritsch: Archivische Informationssysteme in der digitalen Welt (Werkhefte der Staatlichen Archivverwaltung A 23). Stuttgart 2010. S. 53–63.

³ U.a. verlinkt im Wikipedia-Artikel: http://de.wikipedia.org/wiki/Digital_Audio_Tape.

⁴ U.a. verlinkt im Wikipedia-Artikel: http://de.wikipedia.org/wiki/Digital_Audio_Tape.

Die archivischen DAT-Bestände, insbesondere die audiovisuellen Ablieferungen des baden-württembergischen Landtags,⁵ sind meist bereits in der Datenbank ScopeArchiv erschlossen, d. h. die Datenstruktur ist in relativ engen Grenzen vorgegeben. Dementsprechend sind die Findmittel in der Regel bereits online verfügbar und sollten in ihrer Struktur erhalten bleiben.

Vorüberlegungen

Da der Bitstream der DATs zwar digital, der Tonträger selbst aber als analoges Medium vorliegt, scheidet eine Emulationsstrategie von vorn herein aus, denn Bandlaufwerke für Computer sind inzwischen durch den Umstieg auf externe Festplatten und (NAS-) Server als Sicherungsmedien ebenfalls obsolet geworden. Originalbänder und DAT-Kassetten können zum einen selbst nicht erhalten werden, zum anderen werden in Zukunft auch keine Abspielmöglichkeiten mehr gegeben sein. Die im Audiovisuellen Archiv beim Hauptstaatsarchiv Stuttgart vorhandenen DAT-Bänder müssen also in jedem Fall digitalisiert werden, das heißt, es wird eine Migrationsstrategie angewandt.

Digitalisierung

In einem ersten Arbeitsgang werden die überwiegend in Mono-Qualität aufgenommenen Mitschnitte von Landtagssitzungen der 8.-11. Legislaturperiode (1982-1995), die bisher auf DAT-Bändern vorliegen, von einem Dienstleister durch ein softwaregestütztes Abtasten der Tonsignale in das unkomprimierte Standardformat WAVE⁶ überführt. Dieses wurde von der Firma Microsoft entwickelt und ist bereits seit den neunziger Jahren im Einsatz. Hierfür wurde vom Dienstleister eine dedizierte Digitalisierungsstrecke, besteht aus der Einspielstation und der Signalstrecke, aufgebaut. Eine WAVE-Datei entspricht einer Titelaufnahme der archivischen Fachdatenbank scopeArchiv⁷. In den meisten Fällen können die DAT-Bänder 1:1 in das WAVE-Dateien umgewandelt werden. In denjenigen Fällen, in denen mehrere Landtagssitzungen auf einer DAT vorlagen, ist eine Teilung der Dateien notwendig. Für den Digitalisierungsprozess waren eine Samplingrate von 48 kHz und eine Bittiefe von 16 bit durch den DAT-Standard vorgegeben. Beim Übergang in das Dateiformat blieb die Authentizität der Aufzeichnung gewahrt. Etwaige Qualitätsmängel wie z. B. dumpfe Tonqualität, falsche Real-Geschwindigkeiten, unterschiedliche Lautstärkesteuerung, Tonaussetzer, Störgeräusche, Rauschen, Ausblendungen und Verzerrungen wurden daher bei der Digitalisierung nicht behoben. Zusätzlich zu den eigentlichen WAVE-Dateien werden jeweils eine Metadaten-datei und einer Prüfsumme (Hashwert) je Digitalisat erzeugt.

⁵ Bestand HStAS R 10, <http://www.landesarchiv-bw.de/plink/?f=1-3400>.

⁶ Zum Format RIFF-WAVE http://de.wikipedia.org/wiki/RIFF_WAVE.

⁷ <http://www.scope.ch/de/scopeArchiv/SystemArchitecture.aspx>.

Validierung und Import in DIMAG

Die Digitalisate werden vom Dienstleister auf einer externen Festplatte in einer Ordnerstruktur geliefert, wobei die Ablage von jeweils einer WAVE-, einer Metadaten- und einer Prüfsummendatei in einem Container erfolgt. Die Digitalisate müssen anschließend stichprobenartig überprüft werden. Hierfür wird u. a. das vom Landesarchiv Baden-Württemberg entwickelte Tool Ingest List verwendet. Es folgt der Entwurf eines Archival Information Packages (AIP) für den digitalen Massenspeicher (DIMAG) unter Zuordnung der Digitalisate zu den archivisch gewachsenen Beständen und unter Beibehaltung von Archivsignaturen und persistenten Identifikatoren. Zusätzlich wird noch das Metadatenmodell des Landesarchivs Baden-Württemberg⁸ in Anlehnung an die international anerkannten Standards PREMIS,⁹ METS,¹⁰ IASA TC 3/4¹¹ etc. um den Bereich für AV-Unterlagen erweitert. Erst dann kann der eigentliche Import in DIMAG erfolgen, der je nach Struktur und Größe der Primär- und Metadaten automatisiert durchgeführt werden kann oder ein händischer Upload einzelner WAVE-Dateien erfolgen muss. Die DIMAG-Struktur muss sich in jedem Fall an der Archivtektonik orientieren. Liegen bereits Erschließungsdaten vor, sollte das AIP an diese angepasst werden; ist dies nicht der Fall, muss das AIP so strukturiert werden, dass dies auch in der Erschließungsdatenbank ScopeArchiv und im Online-Findmittelsystem OLF21¹² darstellbar ist.

⁸ http://www.landesarchiv-bw.de/sixcms/media.php/120/48392/konzeption_metadaten10.28354.pdf.

⁹ Preservation Metadata: Implementation Strategies, <http://www.loc.gov/standards/premis>.

¹⁰ Metadata Encoding and Transmission Standard, <http://www.loc.gov/standards/mets>.

¹¹ International Association of Sound and Audiovisual Archives, http://www.iasa-web.org/de/downloads/publications/TC03_German.pdf bzw. <http://www.iasa-web.org/de/tc04/audio-preservation>.

¹² <http://www.landesarchiv-bw.de/web/46042>.

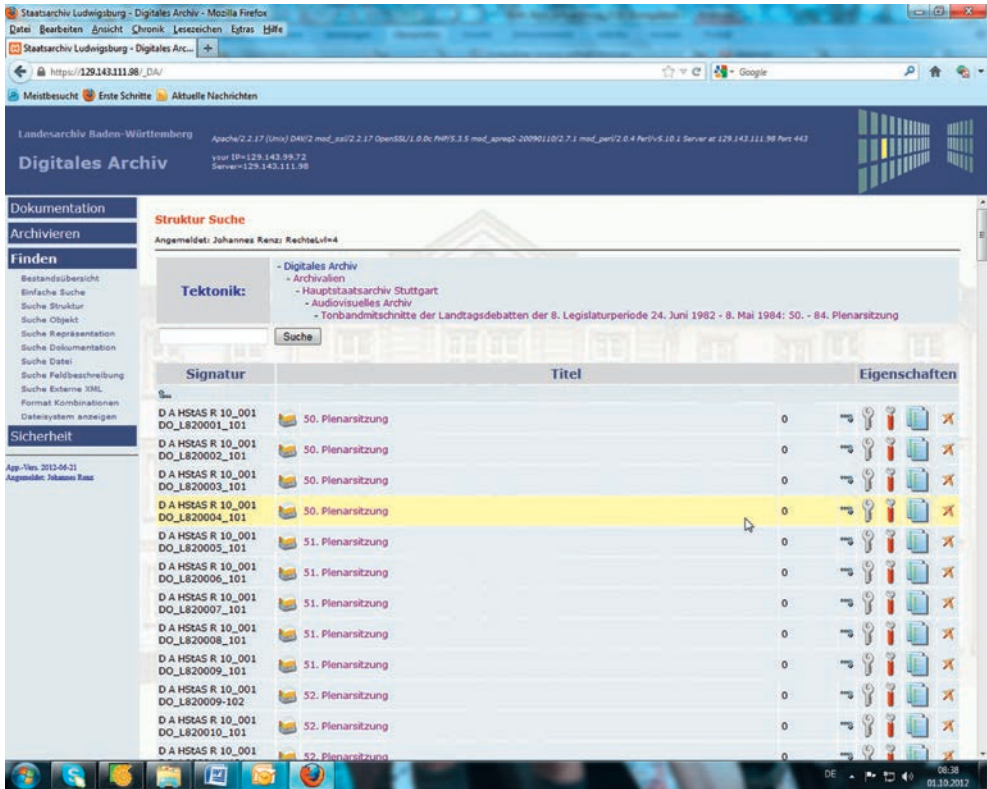


Abb. 5: Bestand R 10/001 (Landtagsitzungen 1982-1984) in DIMAG.

Nutzen des Projekts und zukünftige Möglichkeiten

Mit dem Import in DIMAG können bisher analog und/oder auf obsoleten Datenträgern vorhandene audiovisuelle Unterlagen dauerhaft gesichert werden. Nach einer Prioritätenliste (*Roadmap*) soll bis zum Jahr 2016 die Digitalisierung aller physisch gefährdeten Bild- und Tonträger – neben den DATs sind dies auch Kompakt-, VHS-/SVHS- und VCR-Kassetten – erfolgt sein.

Die bei der Digitalisierung entstandenen Dateien gelten künftig als Ersatzoriginal, was letztlich die Voraussetzung für eine physische Vernichtung des obsolet gewordenen Trägermaterials darstellt. Trotzdem ist dies freilich nur der Beginn einer dauerhaften digitalen Bestandershaltungsstrategie. Künftige Migrationen werden dem Archiv auch hier nicht erspart bleiben – die Voraussetzungen dafür wurden jedoch durch die Digitalisierung geschaffen. Wichtig ist auch hier letztlich die Sicherung der Information, nicht des Datenträgers.

Positive Nebeneffekte ergeben sich insbesondere aus Sicht des Nutzers: So können aus den Masterdateien für die archivische Benutzung komprimierte Kopien (z. B. MP3s, MP4s oder Flashvideos) *on demand* zur Verfügung gestellt werden. Dabei kann es sich entweder um eine

Gesamtdatei (z. B. eine Landtagssitzung) oder aber um einen Ausschnitt (z. B. einen Tagesordnungspunkt oder einen bestimmter Wortbeitrag) handeln. Die Reproduktionskosten können auf Grund des geringeren Aufwands sowohl für das Audiovisuelle Archiv als auch für den Benutzer gesenkt werden. Auch eine Bereitstellung im Internet, entweder eingebunden in ein Online-Findbuch oder für die Präsentation eines archivalischen Highlights, z. B. als *Archivale des Monats*¹³ oder Bestandteil einer virtuellen Ausstellung, ist jetzt möglich.

¹³ <http://www.landesarchiv-bw.de/web/53598>.

Erfahrungen bei der Übernahme von digitalen Unterlagen aller Art im Staatsarchiv St.Gallen – und kritischer Rückblick

Von FELIX STADLER

Sechs Jahre ist es her, seit im Staatsarchiv St.Gallen (StASG) die Abteilung *Elektronische Archivierung* geschaffen wurde. 2006 begann damit der bewusst schrittweise gewählte Aufbau der digitalen Langzeitarchivierung (dLZA)¹, mit dem Ziel, eine Lösung für die dLZA zu implementieren. Dieses dürfte bis Ende 2012 erreicht sein. Einige zentrale Zwischenschritte auf dem Weg zur produktiven dLZA werden im ersten Teil aufgegriffen, wobei der Situationsanalyse, der Strategie und dem Leitbild sowie den Pilotversuchen in den Unterkapiteln besondere Erwähnung zukommt. Nach dieser einleitenden Berichterstattung folgen im zweiten Teil die Arbeitshypothesen der letzten Jahre und die schließlich gewonnenen Erkenntnisse und Prognosen. Sie bilden den Schwerpunkt dieses Artikels.

Aufbau und Einführung der digitalen Langzeitarchivierung

2008 wurde das *Kompetenzzentrum (digitales) Aktenmanagement und Langzeitarchivierung* (CC DALA) gegründet. Dieses Gremium vereint unter der Leitung des StASG Vertreter von Anspruchsgruppen des StASG und bündelt die lokalen Kenntnisse aus den Bereichen (digitales) Aktenmanagement², Langzeitarchivierung, E-Government, IT, Recht und Politik. Vertreter der kantonalen Verwaltung und der Gemeinden bringen die Anliegen der Aktenbildner ein.³

2009 wurden Richtlinien bezüglich *archivtauglicher Dateiformate* sowie zu den *Anforderungen an ECM-Systeme und die Aussonderung (generell)* erarbeitet. Diese Richtlinien sind als Weisungen im kantonalen Informatikhandbuch enthalten, womit ihnen die für die erfolgreiche Implementierung nötige Verbindlichkeit zukommt. Im selben Jahr wurde auch eine Marktanalyse betreffend die möglichen Lösungen für die dLZA erstellt, um diesbezüglich einen besseren Überblick zu erhalten. Zudem wurde versucht, den Lebenszyklus von digitalen Unterlagen⁴ zu definieren.

¹ Im StASG hat sich die Bezeichnung *Digitale Langzeitarchivierung* etabliert. Wir verwenden bewusst *Langzeit*, weil erfahrungsgemäss in der IT unter Archivierung häufig nicht die dauerhafte Aufbewahrung verstanden wird.

² In Klammern steht *digitales* daher, weil auch das analoge Aktenmanagement gemeint ist und sich das digitale Aktenmanagement in der kantonalen Verwaltung von St.Gallen noch nicht so stark etabliert hat.

³ Für weiterführende Informationen: <http://www.staatsarchiv.sg.ch/home/ccdala.html>. Alle Links wurden am 19. September 2012 überprüft.

⁴ Mit *Unterlagen* werden sämtliche Arten von Dokumenten bezeichnet, die bei den Aktenbildnern im Rahmen ihrer Aufgabenerfüllung entstehen. Es kann sich dabei um Schriftgut, Bilder, Fotos, Pläne, Tondokumente etc. handeln.

Situationsanalyse

Bei der 2010 durchgeführten Situationsanalyse ging es darum, das Aussonderungspotenzial zu konkretisieren. Bis dahin war jeweils lediglich eher zufällig bekannt geworden, bei welchen Aktenbildnern Aussonderungspotenzial oder -bedarf für digitale Unterlagen vorhanden war. Unter Berücksichtigung der archivseitig vorhandenen Ressourcen war es nicht möglich, eine vollumfängliche Analyse durchzuführen. Es ist anzunehmen, dass sich mit einer lückenlosen Analyse den zurecht befürchteten Informationsverlusten zwar entgegenwirken ließe. Doch selbst mit einer lückenlos-feinmaschigen Analyse könnte ein solcher Verlust nicht ausgeschlossen werden. Insofern erschien dieses Vorgehen realistisch und vertretbar.

Gesamthaft gesehen, lassen sich die Resultate der Situationsanalyse durchaus als repräsentativ beurteilen. Demnach zeigte sich zu diesem Zeitpunkt kein dringender Aussonderungsbedarf. Hingegen konnte ein kurz- und mittelfristiger Bedarf eindeutig erkannt werden. Als Konsequenz daraus wurde der Bedarf nach einer Lösung zur dLZA deutlich bestätigt.

Die Durchführung der Situationsanalyse ermöglichte es zudem, mit den Aktenbildnern ins Gespräch über die dLZA zu kommen und bei den Gesprächspartnern dadurch eine gewisse Sensibilisierung für das Thema zu erreichen.

Strategie und Leitbild

Die Resultate der Situationsanalyse waren auch für die anschließende Erarbeitung einer gesamthaften Strategie für die dLZA bedeutend. Diese Strategie enthält die längerfristigen Ziele des StASG im Bereich der dLZA. Sowohl für das StASG selbst als auch für die Aktenbildner und die Archivbenutzer sollen Nachvollziehbarkeit, Transparenz und Vertrauen geschaffen und nachhaltig gewährleistet werden.

Um die Bildung von Verantwortungsbewusstsein bei allen Beteiligten zu ermöglichen, muss die Strategie in allgemein verständlicher Form präsentiert werden. Mit einem kurzen thematischen Einstieg sowie dem Leitbild, welches gewissermaßen die externe bzw. öffentliche Version der Strategie darstellt, sollte es möglich sein, die Zielgruppen adressatengerecht zu erreichen.⁵

Die Strategie soll zudem auch als Grundlage dienen zur Implementierung der Lösung für die dLZA nach OAIS⁶.

Pilotversuche

In den vergangenen Jahren wurden mehrere Pilotversuche durchgeführt. Dabei kamen unterschiedliche Lösungen für die dLZA zum Einsatz. Letztlich ging es immer auch darum, Ingest durchzuführen und somit die Übernahme von Unterlagen konkret zu testen. Folglich waren stets auch Aktenbildner und weitere benötigte Beteiligte, wie externe IT-Dienstleister etc., invol-

⁵ Einstieg und Leitbild sind auf der Website des StASG zugänglich: http://www.staatsarchiv.sg.ch/home/behoerden_und_verwaltung/digitale_langzeitarchivierung.html.

⁶ Open Archival Information System, ISO 14721.

viert. So ergab sich im Rahmen dieser Pilotversuche die Möglichkeit, wertvolle Erfahrungen zu machen – sowohl aus technischer, als auch aus organisatorischer Sicht.

Die Geschehnisse rund um einen dieser Pilotversuche lohnt es sich an dieser Stelle kurz zu schildern. Es bot sich der Ingest von Medienmitteilungen des Parlaments, der Regierung, der Verwaltung und der Gerichte an. Die Medienmitteilungen werden zentral durch die kantonale Kommunikationsabteilung auf der Webseite des Kantons⁷ veröffentlicht. Das StASG hat sie vor einiger Zeit als integral archivwürdig bewertet. Die Möglichkeit der effektiven Übernahme dieser Medienmitteilungen ergab sich schließlich ziemlich unerwartet, sie stellte sich dann aber bald als idealer Pilotversuch heraus.

Unerwartet war das Übernahmeangebot deshalb, weil nach der Situationsanalyse vorerst nicht damit gerechnet wurde, dass hier Aussonderungsbedarf besteht. Dieser Umstand beunruhigte insofern, als bewusst wurde, dass vermutlich auch andere Aktenbildner mit Aussonderungsbedarf, zumindest vorerst, vergessen gingen.

Ideal war das Angebot deshalb, weil die Medienmitteilungen öffentlich sind und dadurch keiner Schutzfrist unterliegen. Innerhalb des Pilotversuchs hätte somit auch der Zugriff – und damit der gesamte Lebenszyklus – getestet werden können.

Zusammen mit den verschiedenen Beteiligten starteten die fachlichen Verhandlungen zur Übernahme. Anwesend waren Vertreter der Aktenbildner, des kantonalen Dienstes für Informatikplanung (als verwaltungsinterner IT-Koordinator und produktverantwortliche Stelle für die Webseite), des externen IT-Dienstleisters, der den Betrieb des CMS⁸ sicherstellt, sowie des externen IT-Dienstleisters der Lösung für die dLZA. Mit dem StASG zusammen waren es somit Vertreter von fünf Parteien. Leider gelang es nicht, dass alle Parteien die benötigten Ressourcen zum gewünschten Zeitpunkt bereitstellen konnten, um den Pilotversuch zeitnah durchzuführen. Zur Diskussion um die notwendigen finanziellen Ressourcen kam es somit gar nicht.

Im Rahmen dieses Pilotversuchs wurden stattdessen Versichertenverzeichnisse der kantonalen Gebäudeversicherung als Extrakt aus deren Datenbank übernommen.⁹

Lösung dLZA im Staatsarchiv St.Gallen

2011 wurde die Evaluation für die Lösung dLZA im StASG durchgeführt, deren Implementierung momentan erfolgt. Als Archivinformationssystem bleiben die entsprechenden Module von *scopeArchiv*TM der Firma scope solutions ag im Einsatz. Dazu kommt *scopeOAI*TM für die dLZA. Als Repository wird *FedoraCommons* eingesetzt.¹⁰

⁷ Vgl. <http://www.sg.ch/>.

⁸ Zum Einsatz kommt *CQ5* der Schweizer Firma Day Software, seit 2010 im Besitz der Firma Adobe Systems.

⁹ Vgl. hierzu den Aufsatz von Martin *Lüthi*: Gebäudeversicherung: Ein Versichertenverzeichnis früher und heute. In: Das neue Handwerk. Digitales Arbeiten in kleinen und mittleren Archiven. Vorträge des 72. Südwestdeutschen Archivtags am 22. und 23. Juni 2012 in Bad Bergzabern. Hg. von Kai *Naumann* und Peter *Müller*. Stuttgart 2013 S. 81–84.

¹⁰ scope solutions ag: <http://www.scope.ch>. FedoraCommons: <http://fedora-commons.org/>.

Erkenntnisse und Prognosen

Die ersten sechs Jahre dLZA im StASG waren ereignisreich und boten die Gelegenheit, viele Erfahrungen zu sammeln. Beim kritischen Rückblick und beim Reflektieren wurden Erklärungen gesucht und zumindest teilweise auch gefunden, konnten Erkenntnisse gewonnen und daraus wiederum Prognosen abgeleitet werden. Diese lassen sich in die vier Bereiche *Organisation*, *SIP-Erstellung*¹¹, *Pragmatismus* und *Öffentlichkeitsarbeit* gliedern.

Organisation

Bei Übernahmen analoger Unterlagen sind normalerweise zwei Parteien beteiligt. So gesehen handelt es sich bei einer einzelnen Übernahme um ein Geschäft, das zwei Parteien gemeinsam abwickeln, nämlich der Aktenbildner und das Archiv. Grundsätzlich anders sieht es bei der Übernahme digitaler Unterlagen aus. Nebst dem Aktenbildner und dem Archiv können ggf. zudem die IT-Koordinationsstelle sowie externe IT-Dienstleister seitens des Quellsystems¹², des Rechenzentrums oder der Lösung für die dLZA (Zielsystem) beteiligt sein.

Bei einer Verdoppelung oder sogar Verdreifachung der Anzahl an Beteiligten liegt es auf der Hand, dass die gesamte Koordination und Organisation komplexer wird. Im Grunde genommen bleibt es aber eine Angelegenheit zwischen Aktenbildner und Archiv, bei dem weitere Parteien nur beteiligt sind. Das Archiv tut gut daran, die Federführung des Geschäfts bzw. die Organisation zu übernehmen, um die Wahrscheinlichkeit zu erhöhen, dass die Übernahme im Sinne des Archivs vollzogen wird.

Es ist zu erwarten, dass Übernahmen und dadurch auch deren Organisation an Komplexität verlieren, sobald sich eine Übernahme aus einem Quellsystem wiederholt und dadurch gewissermaßen zur Routine-Angelegenheit wird.

Im Gespräch, in der Diskussion oder bei der Lektüre zum Thema dLZA besteht die Gefahr, den Eindruck zu bekommen, dass die Herausforderungen der dLZA mehrheitlich oder sogar ausschließlich in den technischen Aspekten liegen. Deren Bedeutung ist unbestritten. Weil die Organisation von Übernahmen aber komplex ist, wäre es fatal, die Frage nach der Organisation zu vernachlässigen. Mehr noch: Die Klärung organisatorischer Fragen ist gesamthaft gesehen wichtiger und sollte erfolgen, bevor technische Aspekte behandelt werden. Gut organisiert ist beinahe gewonnen, unzureichend organisiert ist beinahe gescheitert.

An dieser Stelle sei ein Beispiel aus der Praxis erwähnt: Ein Aktenbildner führt Massenakten, die er auf Grund des Gesetzes jeweils nach wenigen Jahren löschen muss. Gemäß einer Bewertungsvereinbarung mit dem Archiv behält der Aktenbildner jeweils eine kleine Auswahl an Akten zurück, um sie später dem Archiv im Rahmen einer größeren Ablieferung zu übergeben. Neu werden die Massenakten ausschließlich digital geführt und jeweils nach Ablauf der gesetzlichen Aufbewahrungsfrist von der Oberbehörde gelöscht. Die digitalen

¹¹ Submission Information Packages bzw. Übernahmepakete, gemäss OAIS.

¹² Damit ist beim Aktenbildner eingesetzte Software gemeint.

Massenakten können somit gar nicht mehr zurückbehalten werden, wie das bei den analogen geschieht.

Historisch gewachsen und vom StASG gefordert, haben sich Abstände von mehreren Jahren zwischen den einzelnen Übernahmen etabliert. Das genannte Beispiel zeigt, dass es bei digitalen Unterlagen in gewissen Fällen nötig sein wird, Übernahmen in kürzeren Abständen als bisher zu tätigen. Dies wird jedoch den Vorteil haben, dass sich die Ablieferungen tatsächlich zu Routine-Angelegenheiten entwickeln werden. Die unterschiedliche Praxis bei der Übernahme analoger und digitaler Unterlagen beeinflusst die Organisation ebenfalls.

Bestimmt sind auch neue rechtliche Grundlagen bzw. verbindliche Vorgaben nötig, die sozusagen das normative Gerüst für die dLZA bilden sollen. Die früheren Bestimmungen waren lediglich auf die Archive ausgerichtet und stammen aus einer Zeit, als der dLZA und dem gesamten Lebenszyklus der Akten noch keine Beachtung geschenkt wurde. Dies war zumindest im Kanton St. Gallen die Ausgangslage. Daher entstand ein zeitgemäßes *Gesetz über Aktenführung und Archivierung*, welches seit 2011 in Kraft ist¹³. Spezifischere Vorgaben für die dLZA sind geplant und werden als dienlich im Interesse der Aktenbildner sowie des StASG betrachtet. Diesbezüglich steht u.a. noch die Wahl der Erlassform (Verordnung oder Richtlinien) aus.

Die bisher erwähnten organisatorischen Aspekte haben grundsätzlich allgemeine Gültigkeit. Im Gegensatz dazu gibt es einen Aspekt, der vermutlich spezifisch für das StASG gilt. Der Kanton St.Gallen hat ungefähr 0,5 Mio. Einwohner. Der Kanton Zürich mit ungefähr 1,4 Mio. und das Land Baden-Württemberg mit ungefähr 10,8 Mio. Einwohnern sind drei- resp. zwanzigmal grösser als der Kanton St.Gallen.¹⁴ Die deutschen Länder und Schweizer Kantone gehören jedoch derselben politischen Ebene an, womit sie diesbezüglich vergleichbar sind. Die kantonalen Archive, in der Schweiz mehrheitlich als Staatsarchive bezeichnet, entsprechen in Deutschland von der Größe her allerdings eher kommunalen Archiven, als Landesarchiven.

Die Kantone haben, unabhängig von ihrer Größe¹⁵, ähnliche und gleiche Kompetenzen bzw. Aufgaben (beispielsweise Archivieren). Es erstaunt daher nicht, dass die Kantone zur Erfüllung ihrer Aufgaben teilweise dieselben Systeme¹⁶ einsetzen. Ein kleiner Kanton hat aber grundsätzlich nicht weniger Systeme im Einsatz, als ein großer. Die Regionalisierung von Aufgaben sowie politische Fusionen finden aber eher auf der politischen Ebene der Gemeinden statt, als bei den Kantonen¹⁷.

In direkter Abhängigkeit zur Größe eines Kantons stehen hingegen die Anzahl der Verwaltungsmitarbeitenden sowie der Aktenumfang, welcher bei der Verwaltungstätigkeit entsteht. St.Gallen ist gemessen an der Einwohnerzahl der fünfgrößte Kanton (von 26), wobei der größte

¹³ Vgl. <http://www.gallex.ch/gallex/1/fs147.1.html>.

¹⁴ Die in diesem Absatz erwähnten Einwohnerzahlen sind den Artikeln *Baden-Württemberg*, *Kanton (Schweiz)*, *Kanton St.Gallen* und *Krینau* aus Wikipedia entnommen, vgl. <http://de.wikipedia.org/wiki/Wikipedia:Hauptseite>.

¹⁵ Der einwohnerstärkste Kanton der Schweiz hat ungefähr 1,4 Mio. (Zürich), der einwohnerschwächste ungefähr 16.000 Einwohner (Appenzell Innerrhoden).

¹⁶ In diesem Zusammenhang ist damit jegliche Art von eingesetzter Software gemeint.

¹⁷ Die einwohnerstärkste Gemeinde des Kantons St.Gallen hat ungefähr 72.000 (Stadt St.Gallen), die einwohnerschwächste ungefähr 250 Einwohner (Krinau).

Kanton (Zürich) dreimal grösser ist. Der Kanton St.Gallen ist genügend groß, um eine eigene Lösung für die dLZA zu haben, aber gleichzeitig zu klein, als dass heute schon ausgebildete Fachleute aus den Bereichen Aktenführung und Archivierung in der Verwaltung (außerhalb des Archivs) arbeiten. Die Kontaktpersonen in der Verwaltung sind daher nicht ausgewiesene und erfahrene Aktenführungs- und Archivierungs-Fachleute, was die Zusammenarbeit generell nicht vereinfacht oder effizienter macht. Insbesondere bei kleineren Verwaltungen dürfte die Wahrscheinlichkeit gering sein, dass direkt bei den Aktenbildnern Fachleute mit den nötigen Kenntnissen und Kompetenzen verfügbar sind, sodass bei Übernahmen auf die Beteiligung externer IT-Dienstleister verzichtet werden könnte.

Es muss also davon ausgegangen werden, dass, bezogen auf die Übernahme digitaler Unterlagen und die dLZA im Allgemeinen, das StASG deutlich mehr Aufwand als größere Archive betreiben muss, um dasselbe Ergebnis zu erreichen¹⁸. Die Verwaltungsgröße des Kantons St.Gallen ist demnach kritisch bzgl. der Umsetzung der dLZA.

SIP-Erstellung

SIP gibt es sinngemäß auch bei der analogen Archivierung bzw. der Übernahme von analogen Unterlagen. Nebst den Unterlagen selbst, den Primärdaten, sind die dazugehörenden Metadaten ebenso Teil einer Ablieferung bzw. eines analogen SIP. So gesehen gibt es eine SIP-Erstellung auch im analogen Bereich. Nach vollzogener Bewertung umfasst die SIP-Erstellung nebst der SIP-Spezifizierung auch das Verpacken der Unterlagen nach archivischen Vorgaben, was ein wesentlicher, weil aufwändiger, Arbeitsschritt ist. Gegenüber der Spezifizierung ist das Verpacken intellektuell wenig herausfordernd und wird häufig von Auszubildenden oder Aushilfskräften ausgeführt.

Im Vergleich zur SIP-Erstellung im analogen Bereich ist anzunehmen, dass sich die SIP-Erstellung im digitalen Bereich als deutlich anspruchsvoller herausstellt. Das gilt einerseits im Besonderen für die SIP-Spezifizierung (Organisation und technische Aspekte). Die Anzahl verschiedener SIP wird ziemlich groß sein, womit es auch immer wieder zu weiteren SIP-Spezifizierungen kommen wird. Andererseits gilt dies auch für die übrigen Arbeitsschritte der SIP-Erstellung, die bedeutend komplexer sein werden, als die entsprechenden Arbeitsschritte im analogen Bereich.

Sobald eine SIP-Erstellung zur Routine-Angelegenheit wird, kann auch Personal mit geringeren fachlichen Kenntnissen für die SIP-Erstellung eingesetzt werden.

Geht es um die Zuständigkeit für die SIP-Erstellung, heißt es erfahrungsgemäß oft, diese liege bei den Aktenbildnern. Dieser Eindruck entlarvt sich allerdings als zu pauschal und undifferenziert. Wie im vorigen Unterkapitel erwähnt, ist die Verwaltung des Kantons St.Gallen vergleichsweise klein. In größeren Verwaltungen dürfte es von den personellen Ressourcen und den damit verbundenen Kenntnissen des betreffenden Personals her eher möglich sein, die SIP-Erstellung erfolgreich auszulagern, sodass die Verwaltung bzw. die Aktenbildner dafür zuständig sind. In kleineren Verwaltungen sind diese Voraussetzungen kaum gegeben. Darüber hinaus unterscheidet

¹⁸ Ein konkretes Beispiel dazu folgt im Unterkapitel SIP-Erstellung.

sich die Anzahl der eingesetzten Systeme bei unterschiedlich großen, aber gleichartigen Verwaltungen kaum. Dieser Umstand unterstreicht, dass die Auslagerung bei kleineren Verwaltungen weder realistisch, noch sinnvoll ist. Daraus lässt sich eine indirekte Korrelation zwischen der Verwaltungsgröße und der Größe des Archiv-Anteils an der SIP-Erstellung begründen. Es ist daher anzunehmen, dass im StASG der eigene Anteil an der SIP-Erstellung beträchtlich sein wird.

Dass Archive kleinerer Verwaltungen die SIP-Erstellung auslagern können, ist bei Routine-Angelegenheiten sowie dort, wo es sich um Massenakten handelt, am ehesten denkbar.

Pragmatismus

Ist die Größe einer Verwaltung kritisch, lässt sich das auch auf das Archiv ableiten. Die Ausführungen in den vorherigen Unterkapiteln implizieren, dass Archive kritischer Größe häufig pragmatisch handeln müssen. Letztlich läuft es darauf hinaus, pragmatisch im vertretbaren Rahmen zu sein. So wird es bekanntlich auch in der analogen Archivierung praktiziert. Der Prüfung der Vertretbarkeit kommt eine zentrale Bedeutung zu, denn sie hat bei sämtlichen Entscheiden und bewusst zu geschehen.

Das StASG ist wegen seiner kritischen Größe gezwungen, ausgesprochen pragmatisch zu handeln. Dabei wird in Kauf genommen, dass nicht immer der mustergültige Ansatz gewählt werden kann, sondern eine andere Variante gewählt werden muss. Flexibilität und auch Kompromissbereitschaft sind gefragt. In der Folge werden einige dieser pragmatischen Handlungsweisen beschrieben.

Wie bereits im vorigen Unterkapitel angetönt, wird erwartet, dass der Anteil des StASG an der SIP-Erstellung beträchtlich sein wird. Darin sind durchaus auch positive Aspekte zu sehen. So ist es bspw. möglich, intensiv Erfahrungen mit der SIP-Erstellung zu sammeln und damit möglichst viele Erkenntnisse für zukünftige Übernahmen zu gewinnen.

Eine zentrale Konvertierungsplattform für PDF/A konnte in der kantonalen Verwaltung bis heute nicht eingeführt werden. Obschon PDF/A als archivtaugliches Dateiformat verlangt wird, scheint es wenig zielführend zu sein, von Beginn an kategorisch darauf zu bestehen, dass von den Aktenbildnern ausschließlich validierte PDF/A-Dateien übergeben werden. Wenn ein Aktenbildner die Anforderungen nicht erfüllen kann, besteht seitens des StASG die Bereitschaft, die Konvertierung selbst vorzunehmen.

Es ist unbestritten, dass mit den Aktenbildnern jeweils Vereinbarungen betreffend die Übernahme getroffen werden müssen. Damit sind Abmachungen gemeint, die über die SIP-Spezifizierung hinaus gehen, wie z.B. Absprachen bezüglich des eigentlichen Transfers der Daten. Es bedarf einer Klärung darüber, wie, auf welchem Speicher, auf welchem Weg etc. die Daten zum Archiv gelangen.

Wie bereits oben erwähnt, wurden im Rahmen eines Pilotversuchs Versichertenverzeichnisse der kantonalen Gebäudeversicherungsanstalt als Extrakt aus der dort verwendeten Datenbank übernommen. Die Übernahme erfolgte in einigen Punkten noch nicht nach den Vorstellungen des StASG entsprechend. Die Absprache betraf lediglich die Bewertung bzw. die Auswahl der zu übernehmenden Datenfelder aus der Datenbank.

Die Daten wurden nicht als SIP übernommen, sondern es handelte sich dabei um eine einzige Datei. Die SIP-Erstellung erfolgte im Anschluss an die Übernahme durch das StASG.

Der Transfer der ersten Daten erfolgte per interner Post, das heißt physisch. Die Datei war auf einer CD-R gespeichert und die Ablieferung wurde damit nicht so dokumentiert, wie es bei der Übernahme analoger Unterlagen sonst üblich ist, das heißt mit einer Angebotsliste, einer Ablieferungsbestätigung und den übrigen Formalitäten. Als ein Jahr später der nächste Extrakt abgeliefert wurde, überwog dennoch die Freude über die unaufgefordert erfolgte Ablieferung die ernüchternde Erkenntnis, dass sich an der etwas achtlosen Art und Weise der Ablieferung nichts geändert hatte. Ein solches Vorgehen soll mittel- und längerfristig nicht zur Praxis werden, sondern wird voraussichtlich nur noch bis zur Inbetriebnahme der Lösung für die dLZA toleriert.

Öffentlichkeitsarbeit

Aktenführung, Aussonderung und Archivierung werden die Aktenbildner vermutlich nie als ihre Kerngeschäfte betrachten, was sie auch gar nicht müssen. Und dennoch ist es zentral, dass die Aktenbildner diese Bereiche ernst nehmen und die hierfür notwendige Sensibilität erlangen. So handelt es sich bei der Aktenführung um eine alltägliche Tätigkeit, der deutlich mehr Gewicht zukommen muss, als es bisher teilweise der Fall ist. Um Übernahmen digitaler Unterlagen möglichst reibungslos durchführen zu können, ist es nötig und zentral, die Kooperation respektive die Bereitschaft der Aktenbildner dazu zu verbessern.

Fragt sich, wie das erreicht werden kann beziehungsweise was Archive unternehmen können, um besser auf ihre Anliegen aufmerksam zu machen, damit sich bei den Aktenbildnern als konstantes Nebengeschäft etabliert, was zur Optimierung der Aussonderung und Archivierung beiträgt. In der Frage wird auf die Antwort bereits angespielt. Es geht darum, Aufmerksamkeit zu erregen, Werbung zu machen und Öffentlichkeitsarbeit zu betreiben. Die Öffentlichkeitsarbeit soll auf die Aktenbildner und überhaupt auf sämtliche Beteiligte (auch externe IT-Dienstleister etc.) ausgerichtet sein. Es muss das Ziel sein, die Herausforderungen der dLZA gemeinsam meistern zu können. Obschon auch die Öffentlichkeitsarbeit nicht Kerngeschäft ist, muss sie wohl bedacht und wirkungsvoll erfolgen.

Im Bewusstsein darum und angesichts der gegebenen Möglichkeiten entsteht im StASG allmählich eine bunte Palette an einzelnen konkreten Maßnahmen:

- Dass Wert auf die Auftritte im Web und im Intranet gelegt wird, versteht sich von selbst. Dort sind Informationen und verschiedene Dokumente verfügbar.
- Online verfügbar ist ein kurzer Einstieg ins Thema dLZA, sowie das Leitbild, welches der externen Version der Strategie entspricht und die dLZA detaillierter behandelt.
- Dazu kommen Richtlinien bezüglich archivtauglicher Dateiformate sowie der Anforderungen an ECM-Systeme und die Aussonderung (generell). Ebenfalls ist ein Glossar mit Begriffen zur Aktenführung und Archivierung zugänglich, womit versucht wird, Fachbegriffe verständlich zu umschreiben.
- Sobald die Lösung der dLZA in Betrieb ist und erste Übernahmen vollzogen sein werden, ist eine Informationsveranstaltung geplant, um allen Beteiligten darüber mündlich zu berichten. Dies soll ergänzend zu schriftlichen Informationen eine Vermittlung in direkter Form sein, um unmittelbar miteinander in Kontakt zu kommen.

Erfahrungsgemäß muss seitens der Archive einiges unternommen werden, um die Eintrittshürde ins Archiv zu senken. Das zeigt sich bei den Bürgern wie bei den Aktenbildnern gleichermaßen, und sowohl im analogen, als auch im digitalen Bereich. Im Rahmen von Ablieferungsverfahren die Aktenbildner beziehungsweise die Mitarbeitenden der entsprechenden Aktenbildner zu Führungen ins StASG einzuladen, hat sich inzwischen zu einer Tradition entwickelt. Was dadurch ermöglicht wird, kann sinngemäß vielleicht als haptische Wahrnehmung des Archivs bezeichnet werden. Es ergibt sich die Möglichkeit, Archivgut zu präsentieren (teilweise sogar themenspezifisch zu den Aktenbildnern) sowie gleichzeitig über archivische Themen und Tätigkeiten zu berichten. Diese Führungen bewähren sich und haben eine positive Auswirkung auf die Verfahren respektive auf deren Ergebnisse.

Die Erfahrung zeigt, dass mit der Präsentation von Archivgut sowie der Berichterstattung archivischer Themen und Tätigkeiten durchaus Werbung in eigener Sache gemacht werden kann. Dazu eignet sich auch die Präsentation von Archivgut in kleinen Einheiten, regelmäßig erweitert auf der eigenen Webseite¹⁹. Auf der Startseite des Intranets erscheint jeweils auch ein Teaser, der auf neue solcher Beiträge hinweist, womit viele Mitarbeitende der Verwaltung darauf aufmerksam gemacht werden.

Ein archivisches Thema, welches Potenzial hat und sich in hervorragender Weise eignet, um damit Öffentlichkeitsarbeit zu betreiben, ist die Webarchivierung. Nach entsprechenden ersten erfolgreichen Tests kam man zur Auffassung, dass genau dieser Bereich bei der Öffentlichkeit des StASG durchaus Interesse wecken könnte. Web und Internet sind vielen Leuten intuitiv nahe, weil sie sich täglich damit befassen und darin bewegen.

Schlussbemerkungen

Diese Prognosen sind weder als Orakel, noch als Worst-Case-Szenario zu verstehen. Obschon einem vieles im Bereich der dLZA manchmal unberechenbar vorkommen mag, wird erwartet, dass diese Prognosen realistisch sind und mit ziemlich großer Wahrscheinlichkeit auch eintreffen werden. Die Vorbereitungen auf das, was die Zukunft bringt, richten sich jedenfalls an diesen Überlegungen aus.

Ob das Erwartete eintrifft, wird sich zu einem späteren Zeitpunkt zeigen. Es ist im Interesse des StASG, sich kontinuierlich zu verbessern. Grundlage dazu sind regelmäßige kritische Rückblicke.

¹⁹ Was andernorts häufig als *Archivalie des Monats* oder ähnlich bezeichnet wird, heisst im StASG ‚Aufgefallen im Staatsarchiv‘. Vgl. <http://www.staatsarchiv.sg.ch/home/publikationen/aufgefallen.html>.

Der Nachfolger des DOMEA®-Konzepts

Das Organisationskonzept elektronische Verwaltungsarbeit und seine Auswirkungen auf die elektronische Archivierung

Von STEFFEN SCHWALM

Das DOMEA-Konzept bildete mehr als 10 Jahre den zentralen Standard zur Einführung der elektronischen Aktenführung und Vorgangsbearbeitung. Trotz der umfassenden Bekanntheit und zahlreichen Projekte hat sich die E-Akte und elektronische Bearbeitung bislang nicht flächendeckend durchgesetzt. Das Organisationskonzept elektronische Verwaltungsarbeit löst das DOMEA-Konzept ab und geht mit der Bedarfsorientierung hinsichtlich der möglichen elektronischen Prozessunterstützung neue Wege. Dies bedeutet in der Konsequenz jedoch eine Diversifizierung der Verfahren und Unterlagen. Das zentrale DMS/VBS für eine Behörde oder alle Behörden eines Landes mit E-Akte und IT-gestützter Vorgangsbearbeitung ist zukünftig nur eine von zahlreichen Möglichkeiten. Der Aufsatz führt in das neue Organisationskonzept elektronische Verwaltungsarbeit ein und zeigt mögliche Auswirkungen auf die elektronische Archivierung.

Das DOMEA®-Konzept im Rückblick

Bedeutung des DOMEA®-Konzepts

Das DOMEA®-Konzept galt als das maßgebliche und umfassendste Kompendium zur elektronischen Aktenführung und Vorgangsbearbeitung in der öffentlichen Verwaltung. Trotz des Fokus auf die Bundesverwaltung gewann das DOMEA®-Konzept den Status als der bedeutendste Quasi-Standard im Kontext im öffentlichen Bereich. Es bildete die wesentliche Grundlage der Mehrheit der Einführungsprojekte in Bund, Ländern und Kommunen.

Aufbauend auf den Grundprinzipien des Verwaltungshandelns beschrieb das Organisationskonzept die Anforderungen an den elektronischen Geschäftsgang vom Posteingang über die Bearbeitung bis zum Postausgang. Ergänzt wurden die Ausführungen durch grundlegende Einführungszenarien.

Das DOMEA®-Organisationskonzept 2.1 wurde ergänzt durch zahlreiche Ergänzungsmodulare zu spezifischen Themen so u.a.:

- elektronische Archivierung
- Virtuelle Poststelle oder
- Scannen

Besondere Bedeutung, insbesondere für die Ausschreibung und Auswahl konkreter Produkte, gewann der Anforderungskatalog sowie die hierauf basierende Zertifizierung von Produkten durch die KBSt im Bundesministerium des Innern. Dieser wurde, im Gegensatz zum Organisationskonzept, 2005 jedoch nicht überarbeitet. Vielmehr wurden die Kriterien nur formal umstrukturiert. Abb. 1 verdeutlicht die verschiedenen Komponenten des DOMEA®-Konzepts und deren inhaltliche Beziehungen.



Abb. 1: Komponenten des DOMEA®-Konzepts.

Einführungsstand der E-Akte und Vorgangsbearbeitung in der öffentlichen Verwaltung

In den letzten zehn bis fünfzehn Jahren wurden auf allen Verwaltungsebenen umfassende Projekte zur Einführung der elektronischen Aktenführung und IT-gestützten Vorgangsbearbeitung begonnen. Trotz Erfolgen bei der Automatisierung strukturierter Prozesse kann von einer flächendeckenden Einführung elektronischer Akten und medienbruchfreier elektronischer Geschäftsprozesse derzeit nicht gesprochen werden. Mag in einzelnen Behörden die E-Akte und Vorgangsbearbeitung durchgängig umgesetzt sein, so verharret eine Vielzahl von Einführungsprojekten in den Pilotbereichen.

Medienbrüche, hybride Akten mit Teilen auf Laufwerken, Mailpostfächern, in Papierform oder Fachverfahren sind noch immer Standard in vielen Behörden auf Bundes-, Landes- und kommunaler Ebene. So wurde nicht zuletzt im Bericht des Bundesrechnungshofs im Jahr 2010 kritisch angemerkt, dass die Schriftgutverwaltung der Bundesbehörden dringend zu verbessern sei, so insbesondere aufgrund mangelnder Vollständigkeit und Nachvollziehbarkeit – Aspekte, die mit dem DOMEA®-Konzept und dessen praktischer Umsetzung durch Einführung von E-Akte und Vorgangsbearbeitung überwunden werden sollten.

Ziel der häufig landes- und kommunenweit eingerichteten Projekte war es zumeist, einheitliche Lösungen für eine ganze Landes- oder Kommunalverwaltung zu schaffen, um so einen effizienten Ressourceneinsatz zu ermöglichen und Lizenzkosten zu sparen. Hinzu kamen Überlegungen, mit einer einheitlichen Aktenstruktur und Prozessabbildung sowie vor allem einem einheitlichen Verfahren die Einführung des DM/VBS zu erleichtern.

Eine umfängliche elektronische Umsetzung der in der Verwaltung in der Regel vorherrschenden teilstrukturierten Prozesse konnte bislang nicht erreicht werden. Die anfängliche Euphorie ist in vielen Fällen der Ernüchterung gewichen.

Die Gründe liegen unter anderem in folgenden Punkten:

- Unterschätzung des organisatorischen Anteils an Konzeption und Einführung (DMS/VBS-Projekt sind zu mindestens 70% Organisationsprojekte)
- Technikfokussierung der Projekte
- unangepasste elektronische Abbildung der papierbasierten Bearbeitung.
- mangelnde Anpassung der organisatorischen Rahmenbedingungen auf die elektronische Aktenführung und Bearbeitung
- mangelnde Leitungsunterstützung
- Unterschätzung der notwendigen Aufwände im Veränderungsmanagement und entsprechender Unterstützung der Mitarbeiter
- Schwächen in der Softwareergonomie der DOMEA®-zertifizierten Produkte.

Hinzu kamen inhaltliche Schwächen des DOMEA®-Konzept als maßgeblichem Standard, auf die im Folgenden eingegangen wird.

Schwächen des DOMEA®-Konzepts

Rückwirkend betrachtet, beinhaltet das DOMEA®-Konzept einige inhaltliche Schwächen, die den beschriebenen Einführungsstand von DMS/VBS in der öffentlichen Verwaltung beeinflussen. Hierzu zählen vordergründig:

Fokussierung auf eine Ziellösung

Das DOMEA®-Konzept fokussierte in allen Einführungszenarien auf die flächendeckende Einführung der elektronischen Vorgangsbearbeitung. Für strukturierte Prozesse, also solche, die immer nach demselben Muster, mit faktisch denselben Schritten ablaufen, birgt eine Prozessautomatisierung mittels Workflow ein erhebliches Optimierungspotenzial. So können immer wiederkehrende Prozessschritte elektronisch schnell und effizient umgesetzt werden. Im Gegensatz dazu zeichnen sich teilstrukturierte Prozesse wie sie besonders in der planenden Verwaltung üblich sind dadurch aus, dass die Prozessschritte nicht vorhersehbar, sondern für jeden Geschäftsvorfall unterschiedlich sein können. Diese lassen sich erfahrungsgemäß per E-Mail unaufwendig elektronisch unterstützen. Dagegen bedeutet der regelmäßige Aufbau von Geschäftsgangmustern im DMS/VBS je Geschäftsvorfall zur Prozessautomatisierung via Vorgangsbearbeitung

eine größere Umstellung der Arbeitsweise mit dem damit verbundenen Aufwand zum begleitenden Veränderungsmanagement.

Nicht in jedem Fall stellt die elektronische Vorgangsbearbeitung die optimale Lösung zur elektronischen Verwaltungsarbeit dar. In vielen Fällen ist es zunächst ausreichend, die elektronische Akte einzuführen und so der ohnehin bereits erfolgenden elektronischen Bearbeitung, nämlich per E-Mail, die rechts- und beweissichere Ablage in Form der E-Akte hinzuzufügen, in der die E-Mail dann abgelegt werden kann. Damit wird die verfassungsmäßig verbriefte ordnungsgemäße Schriftgutverwaltung wiederhergestellt und gleichzeitig die Arbeitsweise der Bearbeiter nur begrenzt verändert. Der Geschäftsgang selbst wird über die Veraktung der E-Mail nachgewiesen. Dies ermöglicht einen schnelleren Projekterfolg und ist in vielen Fällen zunächst ausreichend. Hierauf basierend kann dann gezielt ausgewählt werden, für welche Prozesse eine Automatisierung mittels elektronischer Vorgangsbearbeitung sinnvoll und ressourceneffizient umsetzbar ist. Mit der bisherigen Stufe 2 im Stufenkonzept (Standard) nach dem DOMEA[®]-Konzept, konnte dies so nicht erreicht werden, da das DOMEA[®]-Konzept hier die Bearbeitung jedoch vordergründig in Papierform vorsah, was jedoch in Gegensatz zur Verwaltungsrealität stand.

Hinzu kommt eine zunehmend interaktive Bearbeitung in Projekten und Gremien, wo eine auf die sequentielle Bearbeitung ausgerichtete Vorgangsbearbeitung nur begrenzt dienlich ist. Vielmehr werden in diesen Bereichen Collaborationswerkzeuge eingesetzt und damit häufig Parallelstrukturen zum auf dem Papier führenden DMS/VBS aufgebaut.

Mangelnde Aktualität

Die letzte Aktualisierung des DOMEA[®]-Organisationskonzepts erfolgte im Jahr 2005. Der Anforderungskatalog ist seit 1999 faktisch unverändert. Er wurde 2005 nur umgruppiert, jedoch nicht inhaltlich an die technische Entwicklung angepasst.

Vor diesem Hintergrund beinhaltete das DOMEA[®]-Konzept zuletzt kaum oder keine Ausführung zu aktuellen Themen und Fragestellungen. Dabei handelte es sich zum einen um rechtliche und fachliche Aspekte wie z.B. rechtliche Fragen:

- ersetzendes Scannen,
- Beweiswert elektronischer Unterlagen,
- Formerfordernisse,

oder fachliche Fragen wie:

- Prozessabbildung per E-Mail,
- Collaboration sowohl in linienübergreifender Projekt- und als auch in Gremienarbeit,
- informelle Abstimmungen,
- Wissensmanagement mit Web 2.0-Werkzeugen.

Zum anderen blieben technische Entwicklungen ausgespart. Hervorzuheben ist hier vor allem den Übergang vom Betrieb dezentraler Verfahren in den einzelnen Behörden zum Auf-

bau zentraler E-Government-Dienstplattformen betrieben durch öffentliche IT-Dienstleister. Dabei werden IT-Verfahren innerhalb der Infrastruktur eines oder mehrerer Rechenzentren zu Diensten zusammengefasst und von mehreren Behörden in getrennten Mandanten und Datenbeständen genutzt. Dies ermöglicht die Bündelung der personellen und finanziellen Ressourcen und so ein hohes Maß an Standardisierung. Die Folge ist ein effizienter IT-Betrieb und eine Erleichterung der elektronischen Verwaltungsarbeit, da nicht jede Behörde das einzelne Verfahren selbst aufbauen muss. Öffentliche IT-Dienstleister mit entsprechender Dienststruktur bestehen auf Bundes-, Landes- und kommunaler Ebene (DLZ-IT im Bund, Landes-IT-Dienstleister oder kommunale Rechenzentren).

Daneben entsprach auch die ursprüngliche Idee des DOMEA®-Konzepts, das DMS/VBS als führendes System zu betrachten, in dem der Bearbeiter die Mehrheit seiner Aufgaben bearbeitet und in das die verschiedenen Fachverfahren integriert werden, nicht mehr der technischen Realität. Vielmehr wird das DMS/VBS als ein Basisdienst betrachtet, der es ermöglicht, vollständige elektronische Akten zu bilden und Workflows zu unterstützen. Dabei ist es unerheblich, ob der Bearbeiter direkt im DMS/VBS arbeitet oder visuell im Fachverfahren und dieses im Hintergrund Funktionen des DMS/VBS nutzt, so z.B. die Aktenbildung, ohne dass dem Anwender das DMS/VBS tatsächlich physisch sichtbar gemacht wird. Die nachstehende Grafik verdeutlicht diesen Ansatz.

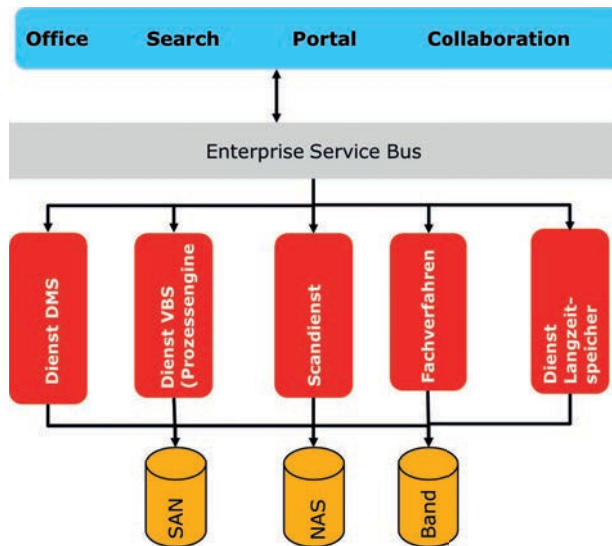


Abb. 2: DMS innerhalb einer IT-Dienststruktur.

Fokussierung auf die Produktausschreibung

Das DOMEA[®]-Organisationskonzept 2.1 beschrieb umfassend die fachlich-organisatorischen Anforderungen an die elektronische Aktenführung und Vorgangsbearbeitung. Im Fokus der Projekte in den Behörden stand erfahrungsgemäß jedoch vordergründig der DOMEA[®]-Anforderungskatalog. Dieser enthielt generisch funktionale Anforderungen an DMS/VBS. Häufig wurde der Katalog nahezu unverändert in eine Leistungsbeschreibung für die Produktausschreibung übernommen und hierzu ein Produkt beschafft, mit dem Verständnis, wenn dieses den Anforderungskatalog erfüllt, dann ist die Einführung der E-Akte und Vorgangsbearbeitung auch sehr einfach möglich.

Was dabei deutlich unterschätzt wurde, war der organisatorische Aufwand. Die Einführung von E-Akte und Vorgangsbearbeitung ist zu mindestens 70 % ein Organisationsprojekt und erfordert erfahrungsgemäß eine umfassende Fachkonzeption der künftig elektronischen Schriftgutverwaltung, des Geschäftsgangs sowie der jeweiligen Geschäftsprozesse anhand der konkreten Rahmenbedingungen der jeweiligen Behörde oder Organisationseinheit. Vielfach liegen in den Behörden vor Einführung eines DMS/VBS nur begrenzt vollständige Akten vor. Vielmehr dient die Einführung der E-Akte auch der Wiederherstellung der ordnungsgemäßen Schriftgutverwaltung. Dem muss das Projekt Rechnung tragen.

Im Vordergrund sollte daher der konkrete Bedarf der Behörde, deren rechtlich-organisatorischer und technischer Rahmen stehen. Hinzu kommen ein ständiges begleitendes Veränderungsmanagement und vor allem die Anpassung des eigenen Regelungskanons an die elektronische Bearbeitung. Diese Aspekte wurden häufig unterschätzt, so dass zwar ein Produkt beschafft wurde, die Einführung jedoch über wenige Pilotbereiche nicht hinaus kam. Darüber hinaus war das Zertifizierungsverfahren der KBSt zunehmend umstritten. Hintergrund war, dass es eine unabhängig von der konkreten Einsatzumgebung reine Produktzertifizierung darstellte. Teilweise wurden Produktfunktionen zertifiziert, die nur für das Zertifizierungsverfahren vorhanden, für das am Markt angebotene Produkt jedoch noch gar nicht verfügbar waren.

Mit Blick auf den begrenzten Einführungsstand der elektronischen Akte sowie elektronischer Geschäftsprozesse in den Behörden und die Schwächen des DOMEA[®]-Konzepts erfolgte dessen grundlegende Überarbeitung. Im Ergebnis entstand das Organisationskonzept elektronische Verwaltungsarbeit, das im Folgenden vorgestellt wird. Das Konzept wurde vom Bundesministerium des Innern unter maßgeblicher Unterstützung durch die BearingPoint GmbH, so u.a. dem Autor dieses Beitrags, erarbeitet.

Organisationskonzept elektronische Verwaltungsarbeit

Abgrenzung zum DOMEA[®]-Konzept

Im Gegensatz zum DOMEA[®]-Konzept fokussiert das Organisationskonzept elektronische Verwaltungsarbeit (Organisationskonzept eVA) auf die bedarfsgerechte Einführung elektronische Geschäftsprozesse. Ob es sich dabei um eine Prozessunterstützung per E-Mail oder

elektronischer Vorgangsbearbeitung handelt oder ob collaborative Elemente genutzt werden, ist ausschließlich bedarfsabhängig. Das Organisationskonzept eVA greift dabei die rechtlichen, fachlichen und technischen Entwicklungen sowie die Projekterfahrungen in den Behörden unmittelbar auf. Im Vordergrund steht dabei die organisatorische Umsetzung der elektronischen Verwaltungsarbeit. Unter elektronischer Verwaltungsarbeit wird dabei die

- elektronische Schriftgutverwaltung (E-Akte) einschließlich der elektronischen Langzeit-speicherung und Aussonderung sowie
- elektronische Prozessunterstützung durch:
- elektronische Vorgangsbearbeitung,
- elektronische Zusammenarbeit und
- elektronische Fachverfahren verstanden.

Das Organisationskonzept elektronische Verwaltungsarbeit enthält keinen eigenständigen Anforderungskatalog an Produkte. Auf eine Produktzertifizierung wird zukünftig verzichtet, um den Fokus auf die Entwicklung bedarfsgerechter Konzepte und die fachlich-organisatorischen Aspekte zu lenken. Die grundlegenden Bausteine zur organisatorischen Umsetzung der elektronischen Verwaltungsarbeit enthalten jedoch in kompakter Form jeweils funktionale Basisanforderungen an entsprechende IT-Verfahren.

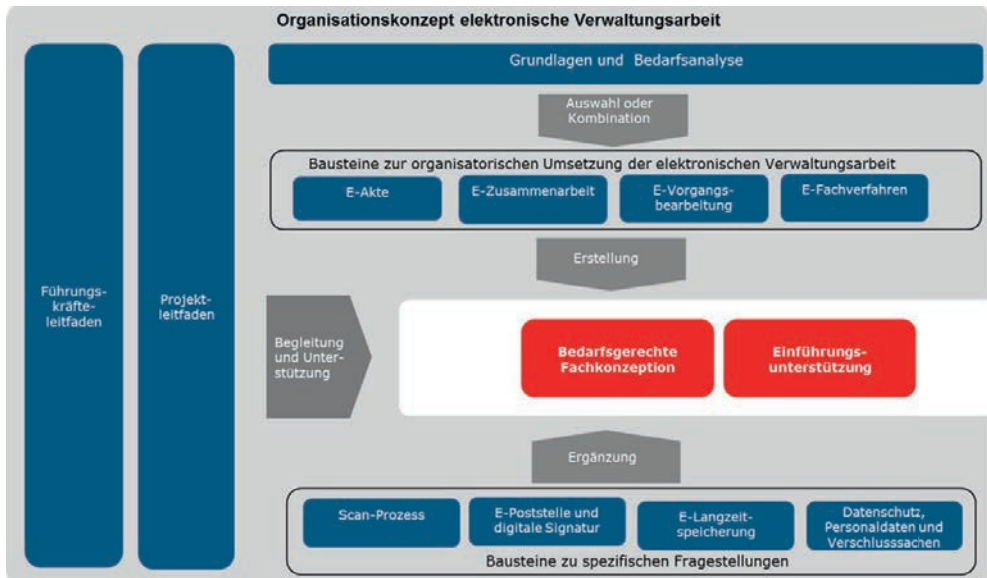


Abb. 3: Aufbau des Organisationskonzepts elektronische Verwaltungsarbeit.

Aufbau und Bausteine

Die Abb. 3¹ zeigt den Aufbau des Organisationskonzepts elektronische Verwaltungsarbeit im Überblick. Den Einstieg in das Konzept bildet das Dokument *Grundlagen und Bedarfsanalyse*. Dieser beschreibt zunächst den grundlegenden rechtlichen Rahmen der elektronischen Verwaltungsarbeit, so insbesondere

- die Regelgebundenheit des Verwaltungshandelns und
- das Prinzip der Aktenmäßigkeit.

Hierauf aufbauend erfolgt eine kritische Analyse des Status quo der elektronischen Verwaltungsarbeit sowie des DOMEA[®]-Konzepts sowie die Darstellung relevanten gesellschaftlichen, fachlichen und technischen Trends. Diese bilden den Rahmen für die Zielsetzung und Zielgruppen des Konzepts. Das Organisationskonzept elektronische Verwaltungsarbeit soll Behörden dabei unterstützen die rechtlichen, fachlichen und technischen Anforderungen an die bestmögliche elektronische Unterstützung ihrer Aufgabenerfüllung zu identifizieren und hierauf basierend bedarfsgerechte SOLL-Konzepte für die Umsetzung entwickeln zu können. Zielgruppe sind Projektleiter und Mitarbeiter in Projekten zur elektronischen Verwaltungsarbeit. Für Führungskräfte besteht mit dem Führungskräfteleitfaden ein eigenes Dokument.

Die wichtigste Rolle des Dokuments *Grundlagen und Bedarfsanalyse* besteht in der grundlegenden Darstellung der Bausteine und Leitfäden des Organisationskonzepts. Für die organisatorische Umsetzung der elektronischen Verwaltungsarbeit liegen folgende zentrale Bausteine vor:

- E-Akte
- E-Vorgangsbearbeitung
- E-Zusammenarbeit
- E-Fachverfahren.

Der Baustein *E-Akte* beschreibt dabei die rechtlichen und fachlichen Anforderungen an die elektronische Schriftgutverwaltung. Das in Art. 19 Abs. 4 und 20 Abs. 3 Grundgesetz bestimmte Prinzip der Aktenmäßigkeit stellt den wichtigsten formalen Rahmen behördlicher Aufgabenerfüllung dar. Dieses ist sowohl bei papierbasierter als auch bei elektronischer Bearbeitung, in Form vollständiger Papierakten oder vollständiger elektronischer Akten umzusetzen. Der Baustein E-Akte bildet damit faktisch die Basis der weiteren Bausteine. Diese beschreiben die verschiedenen Möglichkeiten, wie die Geschäftsprozesse der öffentlichen Verwaltung elektronisch unterstützt werden können. Abb. 4 verdeutlicht dies beispielhaft.

¹ Vgl. auch: <http://www.verwaltung-innovativ.de>.

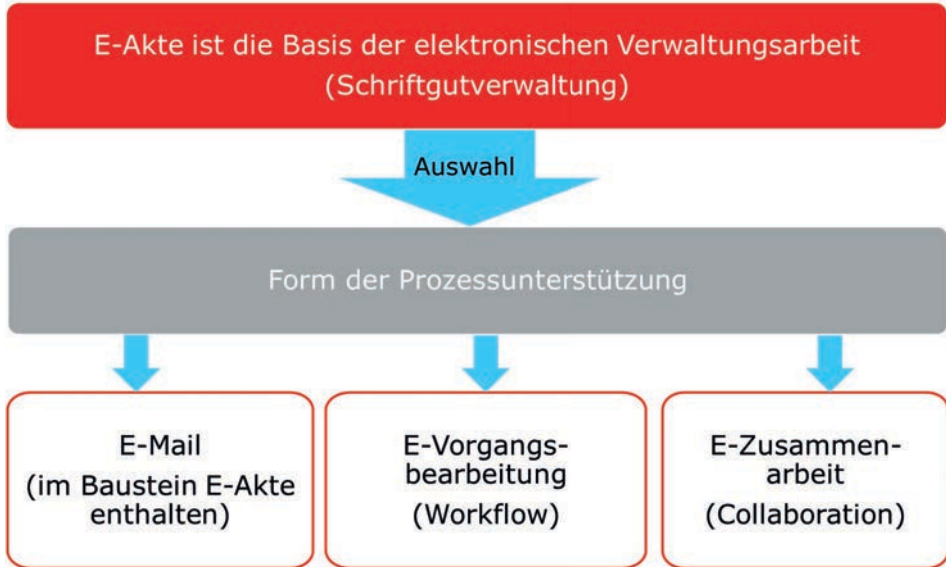


Abb. 4: Beziehung zwischen den einzelnen Bausteinen.

Neben der elektronischen Schriftgutverwaltung enthält der Baustein E-Akte eine Darstellung, wie der behördliche Geschäftsgang per E-Mail unterstützt werden kann. Dies ist eine Form der elektronischen Prozessunterstützung wie sie das Organisationskonzept eVA vorsieht.

Im Dokument *Grundlagen und Bedarfsanalyse* sind darüber hinaus beispielhaft Kriterien enthalten, anhand derer die Behörden ihre Prozesse analysieren können, um zu entscheiden, welche Form der elektronischen Prozessunterstützung sich jeweils anbietet. Ziel des Organisationskonzepts ist es, im Gegensatz zum DOMEA®-Konzept, nicht in einer Behörde alle Prozesse mit elektronischer Vorgangsbearbeitung zu unterstützen. Im Gegenteil, Ziel ist es die Prozesse bedarfsgerecht elektronisch abzubilden, so dass es möglich ist, dass in einer Behörde z.B. die E-Mail die Basis elektronischer Geschäftsgänge bildet, einzelne strukturierte Prozesse per Vorgangsbearbeitung und die bestehenden Projektgruppen und Gremien mittels collaborativer Elemente der E-Zusammenarbeit elektronisch unterstützt werden. Abb. 5 verdeutlicht diese Diversifizierung der Prozesse und Unterlagen.

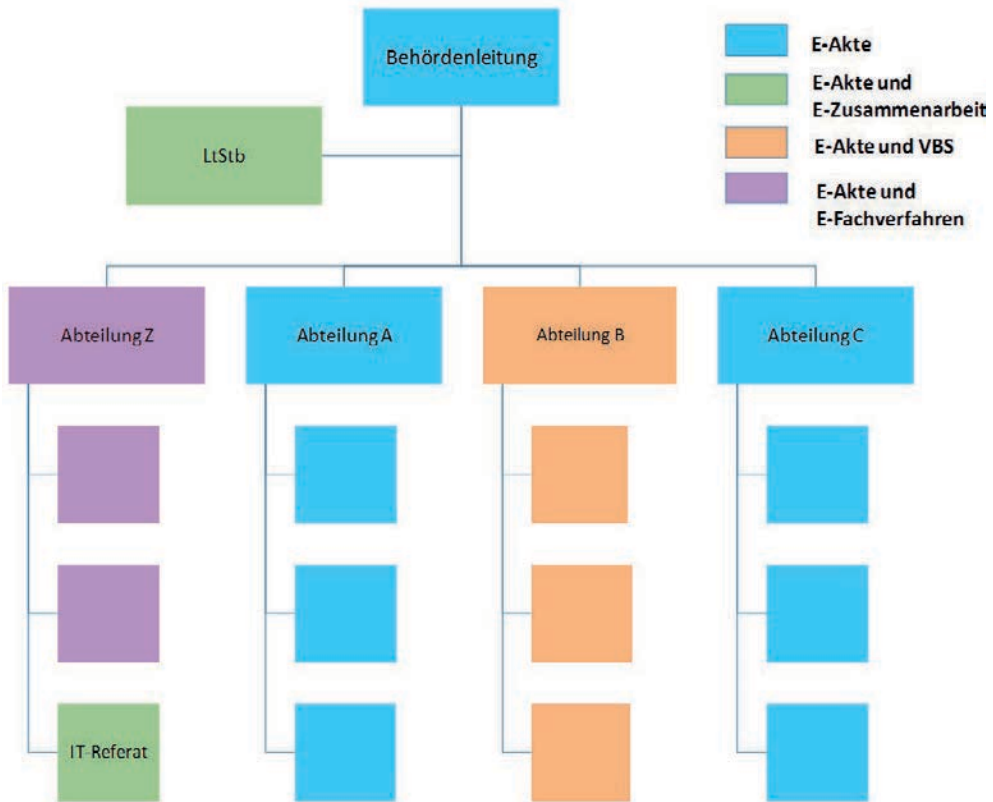


Abb. 5: Diversifizierung der Prozessabbildung.

Entscheidend dabei ist, dass die entstehenden Dokumente in der E-Akte abgelegt werden. Eine Prozessunterstützung per E-Mail, Vorgangsbearbeitung, E-Zusammenarbeit oder Fachverfahren ist damit nur mit Vorhandensein der E-Akte möglich.

Die *E-Vorgangsbearbeitung* beschreibt die Prozessunterstützung mit einem Workflowsystem wie es aus dem DOMEA®-Konzept bekannt ist. Der Baustein *E-Zusammenarbeit* greift dagegen die zunehmende Nutzung collaborativer Arbeitsweisen auf. In Konzentration auf dokumentbezogene Werkzeuge (Teamräume, Blog, Wiki etc.) wird zum einen beispielhaft beschrieben wie mit Hilfe collaborativer Elemente ein behördlicher Geschäftsgang in Projektgruppen und Gremien abgebildet werden kann und zum anderen der Aufbau eines behördlichen Wissensmanagements sowie einer anwendungsübergreifenden Recherche dargestellt.

Allen Bausteinen gemein ist die Beschreibung der notwendigen organisatorischen Regelungsbedarfe bei der Umsetzung in den Behörden.

Ergänzt werden diese grundlegenden Bausteine durch einen Projektleitfaden, der die Vorgehensweisen und Methoden zur Umsetzung eines Projekts zur elektronischen Verwaltungsarbeit darstellt. Hervorzuheben sind hier insbesondere

- das Vorgehensmodell zur bedarfsgerechten Einführung der elektronischen Verwaltungsarbeit
- die umfassenden Ausführungen zum Qualitäts-, Risiko- und Veränderungsmanagement in entsprechenden Projekten.

Um die Führungskräfte gezielt zu unterstützen enthält das Organisationskonzept einen eigenen Führungskräfteleitfaden.

Derzeit in Entwicklung befinden sich ergänzende Bausteine zu spezifischen Fragestellungen so:

- Scan-Prozesse,
- E-Poststelle und digitale Signatur,
- E-Langzeitspeicherung

Künftig hinzugefügt werden sollen die Bausteine

- E-Fachverfahren, was besonders für die Kommunen und Fachbehörden wichtig sein dürfte
- Datenschutz, Personaldaten und Verschlusssachen.

Weitere Ergänzungsbausteine, die im DOMEA®-Konzept zuletzt ein kaum überschaubares Ausmaß annahmen, sind nicht geplant.

Rahmenbedingungen des Organisationskonzepts eVA

Das Organisationskonzept eVA wurde in Zusammenhang mit der IT-Strategie des Bundes und der hier bestehenden inhaltlichen Fokussierung auf serviceorientierte Architekturen entwickelt. Es kann als organisatorische Umsetzung einer diensteorientierten IT bezeichnet werden. Bedarfsgerecht wendet die Behörde faktisch die E-Akte und hierzu entweder E-Mail, Vorgangsbearbeitung oder E-Zusammenarbeit an. Ebenso wie sie technisch z.B. einen DMS-Dienst nutzt und diesen um die Nutzung der Collaborationsplattform, des Signaturdienstes und des Scandienstes des jeweiligen öffentlichen IT-Dienstleisters ergänzt. Die Dienstenutzung gründet organisatorisch auf Service-Level-Agreements zwischen Behörde und IT-Dienstleister.

Rechtlich und fachlich greift das Organisationskonzept insbesondere die Arbeiten zum E-Government-Gesetz des Bundes und begleitende Entwicklungen auf. So werden z.B. auch Fragen zum ersetzenden Scannen oder dem Beweiswert elektronischer Unterlagen im Baustein *E-Akte* aufgegriffen.

Bezug zu korrespondierenden Standards und Normen

In das Organisationskonzept eVA flossen die aktuellen geltenden Standards und Normen sowie Gutachten und Best Practices unmittelbar ein. Hierzu zählen z. B.

- DIN ISO-15489: Schriftgutverwaltung
- Manual zur DIN ISO-15489 des für die Normung in Schriftgutverwaltung und Records Management zuständigen Arbeitskreises Schriftgutverwaltung im DIN NABD 15
- Planungen und Überlegungen zum E-Government-Gesetz des Bundes
- Ergebnisse des Provet-Projekts der Universität Kassel
- Rechtsgutachten zum ersetzenden Scannen
- Empfehlungen des AfO im Bundesministerium des Innern
- Grundsatzpapier der AG IT-gestützte Verwaltungsarbeit des KoopA-ADV zur Aktenrelevanz
- Leitfäden und Best Practices des Bundesarchivs
- Regelungen und Richtlinien aus Bundes- und Landesbehörden
- SAGA 5.0²
- IT-Rahmenarchitektur des Bundes
- BSI TR-03125: Beweiserhaltung kryptographisch signierter Dokumente.
- DIN 31644: Aufbau und Unterhalt digitaler Langzeitarchive (Entwurfssfassung)
- Erweiterungsmodul des DOMEA[®]-Konzepts zur elektronischen Archivierung
- Empfehlung zur Durchführung von Wirtschaftlichkeitsbetrachtungen in der Bundesverwaltung (WiBe) 4.1
- DIN-Fachbericht 158:2009-09
- Unterlage für Ausschreibung und Bewertung von IT-Leistungen V (UfAB V)

Auswirkungen auf die elektronische Archivierung

Diversifizierung von Prozessen und Unterlagen

Das neue Organisationskonzept elektronische Verwaltungsarbeit folgt einem modularen Aufbau. Auf Grundlage der elektronischen Schriftgutverwaltung werden die Geschäftsprozesse bedarfsgerecht elektronisch unterstützt. Dies führt im Ergebnis dazu, dass in einer Behörde nicht mehr alle Organisationseinheiten einzig und allein die elektronische Vorgangsbearbeitung nutzen, sondern vielmehr die E-Mail-Unterstützung die Basis bildet und einzelne strukturierte Prozesse per Workflow, Projektgruppen dagegen mit Hilfe der E-Zusammenarbeit ihre Prozesse elektronisch abbilden. Entscheidungskriterium bildet einzig und allein die rechtliche und fachli-

² SAGA ist ein umfangreiches Dokument, das Standards, Technologien, Architekturen und Methoden für E-Government-Anwendungen in der Bundesverwaltung qualifiziert, evaluiert und klassifiziert sowie Empfehlungen für deren Anwendung gibt.

che Anforderung der jeweiligen Organisationseinheit. Die öffentliche Verwaltung ist gemäß dem Prinzip der Aktenmäßigkeit gezwungen, ihre Entscheidungsprozesse transparent und nachvollziehbar für Dritte zu dokumentieren. Sie ist dabei jedoch frei, ob die Prozessschritte jedoch anhand der von E-Mails, automatisiert erzeugten Bearbeitungs- und Protokollinformationen eines Vorgangsbearbeitungssystems oder mit Hilfe collaborativer Funktionen vollständig nachgewiesen werden. Wichtig ist nur, dass der Nachweis in vollständigen elektronischen Akten erfolgt, so dass das Prinzip der Aktenmäßigkeit gewahrt bleibt.

Diese Abkehr vom eher monolithischen Ansatz des DOMEA®-Konzept, der auf die Einführung der workflowgetriebenen IT-gestützten Vorgangsbearbeitung fokussierte, soll die Einführung durchgehend elektronischer und damit medienbruchfreier Prozesse erleichtern. Im Ergebnis führt dies jedoch zu einer Diversifizierung sowohl der Arbeitsweise als auch der entstehenden Unterlagen. Die Heterogenität der Daten in den elektronischen Akten und Vorgängen steigt. Dies erhöht in der Folge die Komplexität hinsichtlich der anbieterpflichtigen Unterlagen und damit auch den notwendigen fachlichen wie technischen Aufwand bei der Übernahme ins digitale Archiv.

Dienstorientierung zum effektiven Ressourceneinsatz

Auf allen Verwaltungsebenen erfolgt seit einigen Jahren die Zentralisierung der IT in öffentlichen Rechenzentren, die von zentralen, ressort- und landesweiten oder kommunenübergreifenden öffentlichen IT-Dienstleistern betrieben werden. Dies entspricht zudem der IT-Strategie in Bund, Ländern und Kommunen, die von einer serviceorientierten Architektur ausgehen. Dabei werden zunächst Basiskomponenten der elektronischen Verwaltungsarbeit wie z.B.

- DMS
- Vorgangsbearbeitungssystem/Workflowengines
- Signatur- und Zeitstempelkomponenten/Virtuelle Poststelle
- Scanverfahren
- übergreifende Recherchertools
- Fachverfahren

als generische Dienste in komplexen E-Government-Dienstplattformen aufgebaut und behördenübergreifend genutzt. Speziell bei DMS/VBS und Fachverfahren erfolgt dabei eine Trennung der Datenbestände in einzelnen Mandanten, so dass die Sicherheitsanforderungen gewahrt bleiben. Ziel ist es, gemeinsame Infrastrukturen zu schaffen und so zum einen, einen effizienten IT-Betrieb zu ermöglichen, zum anderen die Umsetzung behördenübergreifender elektronischer Prozesse zu erleichtern, da alle Behörden eines Ressorts, eines Landes oder einer Kommune in einer gemeinsamen IT-Infrastruktur eines öffentlichen IT-Dienstleisters arbeiten.

Im Kontext sinkender Haushaltsmittel und Personalressourcen müssen die Behörden und damit auch die Archive die notwendigen Technologien (z.B. DMS-Fachverfahren, Prozessengines, Konvertierungstools, Authentifizierungslösungen, Archivsysteme, Speicher- und Netzinfrastrukturen oder Retrievaltools) zur Aufgabenerfüllung nicht mehr selbst aufbauen, integrieren

und betreiben. Im Gegenteil, es werden nur jeweils die Dienste auf Basis von Service-Level-Agreements genutzt, die auch tatsächlich benötigt werden. Zudem sind die notwendigen Anforderungen an die IT-Sicherheit nach IT-Grundschutz des BSI nicht selbst zu gewährleisten, sondern vielmehr vom IT-Dienstleister, für eine Vielzahl von Verfahren. Damit kann z.B. der Aufbau eines digitalen Archivs spürbar erleichtert werden. Abb. 6 zeigt beispielhaft ein solches diensteorientiert aufgebautes digitales Archiv.

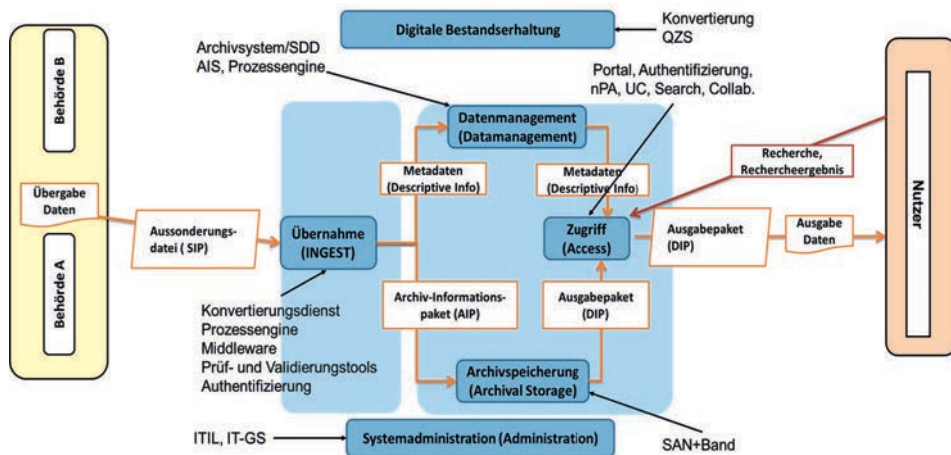


Abb. 6: Aufbau eines digitalen Archivs auf Basis von IT-Diensten.

Andererseits erhöht sich so auch die Komplexität der Verfahren, in denen potenziell anbieterpflichtige Unterlagen entstehen. Zudem muss das Archiv in die Lage versetzt werden, die eigenen archivfachlichen und technischen Anforderungen gegenüber dem IT-Dienstleister zu vertreten. Dies erfordert einen umfangreichen Wissensaufbau, den Kap. 4 näher beschreibt.

Steigende Standardisierung

Die technische Realisierung übergreifender E-Government-Dienste und hierauf basierend gemeinsamer Infrastrukturen sowie durchgängig elektronischer Prozesse erfordert ein hohes Maß an Standardisierung. Dies gilt sowohl aus technischer Sicht hinsichtlich Schnittstellen, Dateiformaten, Speichertechnologien oder Prozessketten als auch organisatorischer und fachlicher Sicht. Die Umsetzung einer bedarfsgerechten elektronischen Verwaltungsarbeit erfordert die Rezeption von Standards zum einen in der Schriftgutverwaltung, zum anderen in der Prozessanalyse und schlussendlich Prozessabwicklung. Ein behördenspezifischer Bebauungsplan zur elektronischen Verwaltungsarbeit, der verschiedene Formen elektronischer Prozessunterstützung zulässt, bedingt die Definition einer Basis, eines Standards, der für alle Organisationseinheiten anzuwenden ist, so z.B. Nutzung der E-Mail. Um eine vollständige Aktenführung sicherzustellen

ist es darüber hinaus empfehlenswert auf gemeinsamen Regeln basierende Ablagestrukturen zu nutzen, was einer innerbehördlichen Standardisierung gleichkommt. Zudem wären die bestehenden organisatorischen Rahmenbedingungen entsprechend anzupassen. Bedarfsgerechte elektronische Verwaltungsarbeit ist faktisch nur mit organisatorischen, fachlichen und technischen Standards denkbar.

Mit Blick auf die geltenden Aufbewahrungsfristen zwischen 2 und 110 Jahren oder dauernd, ist Standardisierung zudem die Grundlage einer langfristigen wie beweissicheren Langzeitspeicherung. Angesichts der kurzen Innovationszyklen der IT, so ist erfahrungsgemäß ca. alle 10 Jahre eine Datenmigration durchzuführen, gewinnen zudem Lösungen zur wirtschaftlichen Aufbewahrung elektronischer Unterlagen exponentiell an Bedeutung.

Die Realisierung dezentraler, verfahrens- und herstelleregebundener Lösungen ist im Kontext zentraler E-Government-Dienstplattformen sowie der damit verbundenen Diversifizierung von Systemen und Daten wirtschaftlich nicht umsetzbar. Die Entwicklung einer vertrauenswürdigen Aufbewahrung kann aus rechtlichen wie finanziellen Gründen als essentiell bezeichnet werden.

Entsprechend den geltenden rechtlichen und fachlichen Rahmenbedingungen muss eine beweissichere Aufbewahrung, die das Prinzip der Aktenmäßigkeit langfristig sicherstellt, die

- Authentizität,
- Integrität,
- Verlässlichkeit (Vollständigkeit und Nachvollziehbarkeit),
- Verkehrsfähigkeit,
- Verfügbarkeit,
- Lesbarkeit

der Unterlagen bis zum Ablauf der Aufbewahrungsfristen gewährleisten.³ Für diesen Aufgabenkomplex wurden in den letzten Jahren entsprechende Standards und Normen etabliert, so z.B.

- ISO-14721:2012 (OAIS-Modell),
- DIN-31644: 2012-4 Information und Dokumentation – Kriterien für vertrauenswürdige digitale Langzeitarchive
- TR-03125 des BSI,⁴
- TransiDoc,
- ISO-19005-1:2005 bzw. ISO-19005-2:2011 (PDF/A).

³ Vgl. Alexander *Rofsnagel*: Langfristige Aufbewahrung elektronischer Dokumente. Anforderungen und Trends. Baden-Baden 2007 i.V.m. ISO-15489:2001. Information and Documentation Records Management, Genf 2001 sowie Stefanie *Fischer-Dieskau*: Das elektronisch signierte Dokument als Mittel zur Beweissicherung. Baden-Baden 2006.

⁴ BSI TR-03125: Beweiswerterhaltung kryptographisch signierter Dokumente. Version 1.1, Bundesamt für Sicherheit in der Informationstechnik, Bonn 2011.

Eine beweisichere Aufbewahrung erfordert demnach die Definition und Einführung definierter und dokumentierter Prozesse zur Gestaltung, Ablage, Abruf definierter und damit formalisierter, validierter und vor allem selbsttragender Datenpakete, die die elektronischen Unterlagen im Aktenzusammenhang einschließlich der notwendigen Daten für den Nachweis der o.g. Eigenschaften gegenüber Prüfbehörden und Gerichten in langzeitspeicherfähiger Form enthalten. Hierfür ist eine vollständige System- und Datenträgerunabhängigkeit der Langzeitspeicherung sehr empfehlenswert. Dadurch wird zum einen eine vertrauenswürdige Aufbewahrung ermöglicht, da zum einen die Authentizität, Integrität etc. durch die Eigenschaften des Datenpakets, das die aufbewahrungspflichtigen Unterlagen enthält, selbst sichergestellt und nachweisbar gehalten werden. Zum anderen wird eine langfristige Abhängigkeit vom Hersteller des IT-Verfahrens oder des Speichersystems vermieden. So bedeutet die Ablage auf einem WORM-Speicher, dass die Unterlagen in herstellerabhängiger Form abgelegt werden. Ohne das originale Speichersystem sind die Daten faktisch nicht lesbar. Zudem bedeutet die gem. §§ 6 Sig G und 17 SigV zur Beweissicherung notwendige Übersignierung signierter Unterlagen⁵ bei WORM-Speichern die Duplizierung der gespeicherten Daten, da ein Schreiben auf dem Datenträger aufgrund der WORM-Funktionalität faktisch ausgeschlossen ist. Die öffentliche Verwaltung bewahrt die Unterlagen selbst auf, nicht die IT-Verfahren und Datenträger, weshalb zunehmend die Langzeitspeicherung in selbsttragenden Datenobjekten erfolgt. Mit den bestehenden Standards und Normen, so insbesondere dem OAIS-Modell und der TR-03125 des BSI wird dies auch in beweisicherer Form ermöglicht. So verweist auch der Entwurf des E-Government-Gesetzes hinsichtlich der Beweissicherheit elektronischer Unterlagen explizit auf die TR-03125.

Diese Erfordernisse werden aktuell in konkreten Projekten auf allen Verwaltungsebenen (z.B. Bundesministerium für Gesundheit, Land Mecklenburg-Vorpommern) umgesetzt und zentrale Langzeitspeicher als Dienste aufgebaut, um eine übergreifende und damit wirtschaftliche Nutzung zu ermöglichen, zumal die beschriebenen Anforderungen an die Langzeitspeicherung für alle aufbewahrungspflichtigen und insofern auch anbietungspflichtigen Unterlagen gelten. Ein solcher zentraler Langzeitspeicher verringert zudem den Aufwand für Aussonderung und Übernahme, da

1. die potenziell archivwürdigen Unterlagen bereits beim Übergang in den Langzeitspeicher in eine langzeitspeicherfähige, standardisierte Form, also ein selbsttragendes Datenpaket zu überführen sind, um deren langfristigen Erhalt bis zur Aussonderung sicherzustellen.
2. die Aussonderung nur noch aus einem Verfahren, dem zentralen Langzeitspeicher erfolgt.

In diesem Punkt führt die zunehmende Standardisierung in der laufenden Bearbeitung sowie der Langzeitspeicherung zu einer Erleichterung in der elektronischen Archivierung.

⁵ Vgl. *Rafsnagel*, wie Anm. 3, i.V.m. Fischer-Dieskau, wie Anm. 3 sowie Jan Skrobotz: *Das elektronische Verwaltungsverfahren. Die elektronische Signatur im E-Government*, Berlin 2005.

Steigende Heterogenität der Daten

Sowohl die Dienstorientierung der IT als auch der modulare Ansatz des Organisationskonzepts elektronische Verwaltungsarbeit dienen dem Ausbau durchgängig elektronischer Prozesse. Dies unterstützt zum einen die stetig steigende Menge aufbewahrungspflichtiger und damit anbietungspflichtiger Unterlagen sowie zum anderen deren Heterogenität.

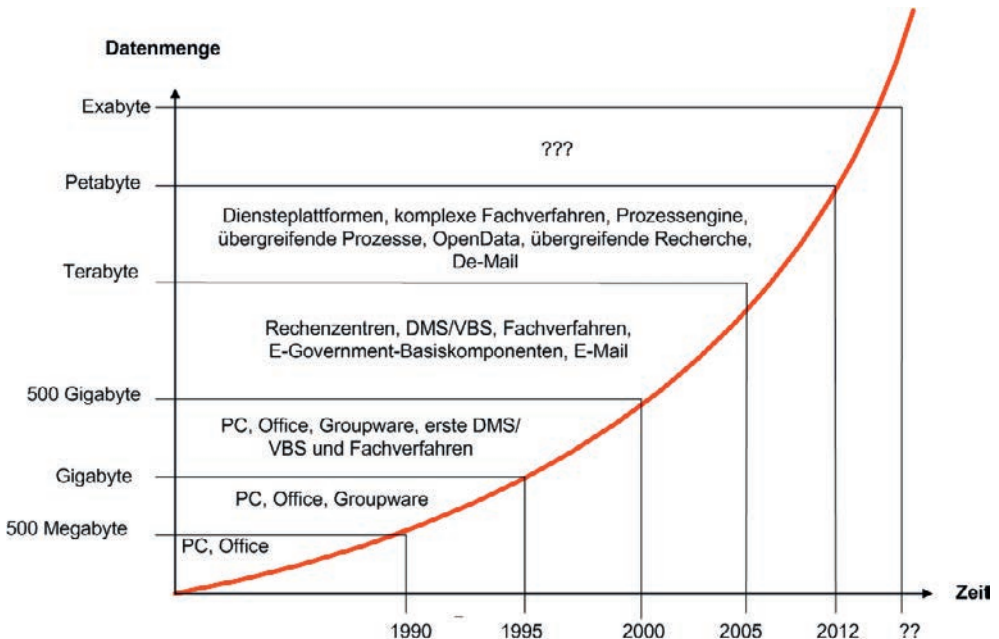


Abb. 7: Exponentiell steigende heterogene Menge an datenschutzpflichtigen Daten.

Diese Aspekte wären bei Lösungen zur Bewertung, Übernahme und Archivierung der elektronischen Unterlagen in allen öffentlichen Archiven zu berücksichtigen.

Nutzung der Synergien zwischen Langzeitspeicherung und Archivierung

Langzeitspeicherung und Archivierung können als zwei Seiten einer Medaille bezeichnet werden. Die Anforderungen sind vergleichbar. In beiden Fällen gilt es, die

- Authentizität,
- Integrität,
- Verlässlichkeit (Nachvollziehbarkeit),
- Verkehrsfähigkeit,

- Verfügbarkeit,
- Lesbarkeit

der elektronischen Unterlagen zu sichern. Mit Blick auf die kurzen Innovationszyklen der IT⁶ ist es erfahrungsgemäß unerheblich, ob die Daten für 10, 20 und 100 Jahre oder dauerhaft aufbewahrt werden. In beiden Fällen gilt es, selbsttragende Datenpakete in system- und datenträgerunabhängiger Form zu erzeugen, um deren langfristige Nutzbarkeit unter Wahrung der genannten Anforderungen zu gewährleisten. Unterschiede lassen sich zum einen hinsichtlich der in der Langzeitspeicherung bestehenden unmittelbaren Rechtswirkung identifizieren, die in der Archivierung nicht mehr gegeben ist. Zum anderen ist der Aufbewahrungszweck zu nennen – Beweissicherung in der Langzeitspeicherung, Überlieferungssicherung in der elektronischen Archivierung. Die fachlichen Anforderungen sind jedoch vergleichbar. Auch ist in beiden Fällen ein hohes Maß an Standardisierung:

- definierte und dokumentierte Prozesse
- Definition, Erzeugung, Ablage, Abruf formalisierter und validierter zur langfristigen Aufbewahrung geeigneter Datenpakete

sowie Maßnahmen und Bestimmung von Verantwortlichkeiten zur langfristigen Datenerhaltung notwendig. Die eingesetzten Technologien und relevanten Standards sind vielfach gleich. In beiden Fällen ist OAIS die empfehlenswerte Richtschnur zur Konzeption und Umsetzung des Verfahrens. Vordergründig gestaltet sich die Anwendung differenziert. Dies betrifft vordergründig die genutzten Funktionen der jeweiligen IT-Dienste, so z.B. des eingesetzten INGEST-Tools zur Ausgestaltung der Datenpakete. Die notwendigen Komponenten und Prozesse (Konvertierungsdienst, Prozessengine, Archivsystem, Archival Storage etc.) sind nicht nur vergleichbar, sondern im Grundsatz faktisch gleich. Lediglich in deren Anwendung bestehen Unterschiede, so z.B. hinsichtlich der Frage welche beschreibenden und technischen Metadaten erhoben werden oder die Entfernung qualifizierter elektronischer Signaturen in der elektronischen Archivierung. Selbsttragende Datenpakete sind in beiden Fällen notwendig, faktisch verändert sich nur deren Gestaltung.

Insofern lassen sich zwischen Langzeitspeicherung und Archivierung Synergieeffekte in Aufbau und Betrieb nutzen, in dem der Aufbau paralleler Infrastrukturen für ähnliche und teilweise gleiche Abläufe vermieden wird. Hierfür ist es allerdings notwendig, die Herangehensweisen in beiden Fällen zu kennen und dementsprechend pragmatisch verbinden zu können. Langzeitspeicherung und Archivierung können insofern als zwei Seiten einer Medaille betrachtet werden.

⁶ Vgl. das Kapitel zur steigenden Heterogenität der Daten weiter unten.

Notwendiger Ausbau der Behördenberatung

Die steigende Flexibilität in Konzeption und Umsetzung der elektronischen Verwaltungsarbeit durch die neuen Möglichkeiten elektronischer Prozessunterstützung des Organisationskonzepts eVA erleichtert die Einführung medienbruchfreier Geschäftsprozesse. Dem Bedarf der einzelnen Organisationseinheit kann durch die jeweils geeignete Form

- E-Mail
- E-Vorgangsbearbeitung
- E-Zusammenarbeit

gezielt entsprochen werden. Gleichzeitig erhöht sich damit jedoch auch die fachliche Komplexität der Projekte. So ist neben der vollständigen E-Akte vor allem eine gemeinsame Basis zu finden, um zwischen den Organisationseinheiten, sofern mehrere Formen elektronischer Prozessunterstützung genutzt werden, die Unterlagen austauschen zu können bei gleichzeitig vollständigem Nachweis der Bearbeitungsschritte in der E-Akte. Hier entsteht ein hohes Maß an Unterstützungsbedarf seitens der anbietungspflichtigen Stellen, der seitens der Archive faktisch nur durch einen Ausbau der Behördenberatung bedient werden kann. Dies ist umso wichtiger, als dass die Diversifizierung der Prozesse und damit auch der Unterlagen die Gefahr birgt, die Bedeutung einer ordnungsgemäßen Schriftgutverwaltung zu unterschätzen. Dies betrifft in der Regel Institutionen, in denen Collaborationswerkzeuge genutzt werden. Hier gilt es eine klare Trennung zwischen formalem Geschäftsgang und informellem Handeln festzuschreiben, um die ordnungsgemäße Veraktung zu gewährleisten. Gerade bei der Schaffung der organisatorischen Regeln und der Umsetzung der Schriftgutverwaltung dürfte der Beratungsbedarf im Zuge der neuen Möglichkeiten elektronischer Verwaltungsarbeit ansteigen.

Hinzu kommt die Frage der langfristigen Aufbewahrung. Die Behörden stehen vor der Herausforderung, die steigende Menge aufbewahrungspflichtiger Unterlagen und vor allem deren Heterogenität langfristig beweissicher aufzubewahren. Bislang bestehen hierzu nur punktuell Erfahrungen. Die Kernaufgabe der Archive liegt, im Gegensatz zu den anbietungspflichtigen Stellen, in der dauerhaften Aufbewahrung von Unterlagen unter Wahrung der im vorigen Abschnitt genannten Anforderungen. Mit einer Konzentration auf die eigenen Kernaufgaben, durch die Nutzung von IT-Dienstleistern für Aufbau und Betrieb des digitalen Archivs unter fachlicher Kontrolle des zuständigen Archiv selbst eröffnet sich die Chance, die Behördenberatung auch hinsichtlich der Aufbewahrung und Erhaltung digitaler Unterlagen auszubauen.

Als erfahrungsgemäß vertrauenswürdiger Aufbewahrungsort sowie Ansprechpartner sind die Archive geradezu prädestiniert, eine solche Beratungsfunktion wahrzunehmen. Angesichts der Unsicherheiten in der langfristigen beweissicheren Aufbewahrung elektronischer Unterlagen in den Behörden kann das Archiv so zum internen Wissensdienstleister über die Schriftgutverwaltung hinaus auch im Kontext Langzeitspeicherung werden. Positive Nebeneffekte sind z.B.:

- spürbarer Mehrwert für den Archivträger und damit verbunden
- eine Qualitätsverbesserung in Schriftgutverwaltung und Langzeitspeicherung für den Archivträger sowie

- die Nutzung gemeinsamer Standards und technischer Komponenten in Langzeitspeicherung und Archivierung (in Mandantentrennung) und damit:
 - Erleichterung des Aufbaus digitaler Archive sowie
 - Reduktion der Komplexität in der Aussonderung.

Kooperation und Vernetzung

Die steigende fachliche und technische Komplexität im Zuge der Dienstorientierung der IT sowie des Organisationskonzepts eVA erscheint für kleinere und mittlere Archive, die in der Regel die Mehrheit der Archive darstellen, angesichts der erfahrungsgemäß begrenzten personellen wie finanziellen Ressourcen nur bedingt zu bewältigen. Vor diesem Hintergrund ist eine archivübergreifende Kooperation und Zusammenarbeit empfehlenswert. Dies ermöglicht sowohl die Bündelung der Kompetenzen als auch die Chance, bestehende Erfahrungen nachnutzen zu können. Erfahrungen in der elektronischen Archivierung besitzen derzeit vorwiegend das Bundesarchiv sowie Landesarchive.

Der Ansatz des Landesarchivs Baden-Württemberg, verschiedene Kooperationsmodelle anzubieten und so die Kräfte zu bündeln und die Zusammenarbeit zum gegenseitigen Nutzen auszubauen, ist ein zentrales Beispiel einer solchen Kooperation. Im Kontext von Haushaltskonsolidierung und sinkenden personellen Ressourcen kommt der archivspartenübergreifenden Kooperation und Vernetzung zur elektronischen Archivierung grundlegende Bedeutung zu. Wie auf dem Archivwissenschaftlichen Kolloquium in Marburg im November 2011 diskutiert, steht zu überlegen, ob in Deutschland eine Koordinierungsstelle zur elektronischen Archivierung äquivalent der KOST in der Schweiz eine Lösung darstellen kann. Diese *deutsche KOST* könnte durch die Entwicklung von Best Practices, Standards und beispielhaften Frameworks kleinere und mittlere Archive systematisch unterstützen sowie Vernetzung und Kooperation fördern. Damit könnte der verdienstvolle Arbeitskreis *Archivierung von Unterlagen aus digitalen Systemen* als Gremium zum Wissens- und Erfahrungsaustausch gezielt untersetzt werden. Dabei erscheint es empfehlenswert, konsequenterweise die Zusammenarbeit mit der Schweizerischen KOST zu suchen, um von den Erfahrungen der Kollegen in Aufbau und Umsetzung einer solchen Koordinierungsstelle zu profitieren.

Kompetenzerweiterung als Chance und Notwendigkeit

Die sich aus dem neuen Organisationskonzept elektronische Verwaltungsarbeit und den darin aufgegriffenen organisatorischen und technischen Entwicklungen im E-Government ergebenden Herausforderungen und Chancen bedingen eine fundierte Erweiterung der archivischen Fachkompetenz seitens der öffentlichen Archive. Um daneben die Produkte und Leistungen öffentlicher IT-Dienstleister entsprechend der archivfachlichen Rahmenbedingungen nutzen zu können, ist es essentiell, auf Augenhöhe kommunizieren zu können. Ziel sollte es sein, das Spannungsfeld zwischen dem technischen Standardisierungsinteresse des IT-Dienstleisters und der damit verbundenen begrenzten Bereitschaft zur Umsetzung spezieller Lösungen, die in der

Form bislang in der bestehenden Infrastruktur nicht existieren, auf der einen Seite und den berechtigten fachliche Anforderungen des Archiv auf der anderen Seite, aufzulösen.

Der Wissensaufbau betrifft dabei vor allem folgende Themenkomplexe:

Fachlich-organisatorisch	technisch
<ul style="list-style-type: none"> • Datenschutz • Datensicherheit • Rechtlich-organisatorische Regeln • Service Level Management • E-Government • Prozessmanagement • Wissensmanagement 	<ul style="list-style-type: none"> • IT-Sicherheit • IT-Grundschutzkataloge BSI • Verschlüsselung • E-Government-Dienste, SOA • Storage, Schnittstellen • Preservation Planning • Systemintegration und -design

Tabelle 1: notwendiger Wissensaufbau der öffentlichen Archive.

Wie auf der Podiumsdiskussion auf dem Archivwissenschaftlichen Kolloquium 2011 diskutiert, wäre diese Kompetenzerweiterung auch in der archivfachlichen Ausbildung zu berücksichtigen. Darüber hinaus gilt es archivintern, für die neuen Kompetenzen Verantwortlichkeiten zu benennen und organisatorisch zu verorten, um handlungsfähig zu sein.

Dies erfordert zudem die Einstellung geeigneten Fachpersonals, das die oben genannten Kompetenzen und Erfahrungen im Kontext E-Government, IT-Dienste, IT-Infrastruktur sowie den rechtlich-organisatorischen und fachlichen Rahmenbedingungen etc. aufweisen kann. In diesem Zusammenhang bewegen sich die Archive in enger Konkurrenz mit der gesamten öffentlichen Verwaltung sowie der Privatwirtschaft. Auch in den Behörden selbst sowie den Unternehmen ist der Bedarf nach entsprechend qualifiziertem Personal außerordentlich hoch, was die aktuelle Stellensituation unterstreicht. Beim *Kampf um die Köpfe* spielt besonders in der öffentlichen Verwaltung die Attraktivität des Arbeitsplatzes eine erhebliche Rolle.

Dabei lassen sich neben den klassischen Anreizen wie die relative Arbeitsplatzsicherheit des öffentlichen Dienstes beispielhaft nennen:

- Entscheidungs- und Gestaltungsspielräume,
- flexible Aufstiegsmöglichkeiten,
- Technische Ausstattung des Arbeitsplatzes
 - elektronische Bearbeitung,
 - Nutzung kollaborativer Werkzeuge oder
- die Präsentation nach außen.

Das Organisationskonzept elektronische Verwaltungsarbeit und die darin aufgegriffenen rechtlichen, fachlichen und technischen Entwicklungen stellen für die elektronische Archi-

vierung Chancen zur Vereinfachung dar. Sei es die Nutzung von Diensten zum Aufbau des digitalen Archivs, die damit verbundene Möglichkeit zur Konzentration auf die eigentlichen Kernaufgaben, die Positionierung als interner Wissensdienstleister für die anbieterpflichtigen Stellen im Sinne eines Ausbaus der Behördenberatung mit den damit verbundenen Vorteilen oder die Nutzung der Synergien zwischen Langzeitspeicherung und Archivierung – die elektronische Archivierung kann künftig potenziell einfacher und effizienter realisiert werden. Voraussetzungen hierfür sind allerdings der beschriebene Wissensaufbau sowie eine Bündelung der Kompetenzen und Erfahrungen durch eine verstärkte, archivspartenübergreifende Vernetzung und Kooperation. Die Idee einer deutschen KOST kann dabei zu einem zentralen Erfolgsfaktor für eine elektronische Archivierung für alle Archivsparten werden. Nicht zuletzt gilt es, die Stellung der Archive als vertrauenswürdige Partner der öffentlichen Verwaltung in der beweissicheren elektronischen Verwaltungsarbeit auszubauen und hierfür das notwendige Fachpersonal zu gewinnen.

Das Organisationskonzept eVA ist eine Chance für Behörden und Archive gleichermaßen – nutzen wir sie!

Archivierung digitaler Unterlagen im Staatsarchiv Basel-Stadt Rückblick auf den Infrastrukturaufbau und erste Erfahrungen

Von LAMBERT KANSY und MARKUS LOCH

Der Beitrag beschreibt rückblickend den Aufbau der Infrastruktur für die digitale Archivierung im Staatsarchiv Basel-Stadt in den Jahren 2010–2012 sowie erste Erfahrungen mit den realisierten Werkzeugen. Die Darstellung fokussiert auf die Implementierung einer standardisierten Ablieferungsschnittstelle (SIP), die Neudefinition der Struktur der Archivpakete (AIP) sowie die Entwicklung des digitalen Magazins auf technischer Ebene. Abschließend werden mögliche Ausbauschritte skizziert.

Einleitung

Das Staatsarchiv Basel-Stadt führte von 2008 bis 2011 das Projekt Informatisierung III durch, das sich neben der Digitalisierung von Archivgut und der Integration des Archivinformationssystems in die kantonale Systemlandschaft mit dem Aufbau von Infrastrukturen und Geschäftsprozessen zur Archivierung digitaler Unterlagen befasste. 2010 wurde der damalige Projektstand an der 14. Tagung des Arbeitskreises für die Archivierung von Unterlagen aus digitalen Systemen vorgestellt.¹ Der nachstehende Beitrag beschreibt die seither erreichten Ergebnisse und stellt die gemachten Erfahrungen vor, bevor er mit einem Ausblick auf die weiteren Entwicklungslinien schließt.

Ergebnisse im Überblick

Im zweiten Quartal 2010 wurden die in der ersten Etappe des Projekts entwickelten neuen Komponenten des Archivinformationssystems, das Modul Ingest und der Ingestserver sowie die Anbindung des Repositorysystems FedoraCommons, im Rahmen eines Releasewechsels von

¹ Vgl. Lambert Kansy: Archivierung digitaler Unterlagen im Staatsarchiv Basel-Stadt. Rückblick auf den Infrastrukturaufbau und erste Erfahrungen. In: Neue Entwicklungen und Erfahrungen im Bereich der digitalen Archivierung: von der Behördenberatung zum Digitalen Archiv. 14. Tagung des Arbeitskreises *Archivierung von Unterlagen aus digitalen Systemen* vom 1. und 2. März 2010 in München. Hg von Susanne Wolf, München 2010, S. 5562, http://www.staatsarchiv.sg.ch/home/auds/14/_jcr_content/Par/downloadlist/DownloadListPar/download_6.ocFile/Text%20Kansy.pdf. Alle Links in diesem Beitrag wurden am 20. September 2012 überprüft.

scopeArchiv eingeführt. Mit dem damit erreichten Stand konnten 2010 und 2011 im Rahmen eines Pilotversuchs in zwei Ablieferungen erstmals originär digitale Unterlagen aus dem System ARAP (elektronisches Polizeijournal der Kantonspolizei Basel-Stadt) übernommen werden.² Die dabei gesammelten – geringen – Erfahrungen sind in den Ausbau der Ingestwerkzeuge eingeflossen. Sie betreffen die Handhabung der Werkzeuge sowie die Kommunikation zwischen Aktenbildner resp. abliefernder Stelle und dem Staatsarchiv. Darüber hinaus wurden grundsätzliche Weiterentwicklungswünsche festgehalten, zum Beispiel der Aufbau einer generalisiert anwendbaren Ablieferungsschnittstelle.

Insgesamt wurden acht Arbeitspakete sehr unterschiedlichen Umfangs definiert:

1. Festlegung eines Persistent Identifier (PID) für die Archival Information Packages (AIP)
2. Ausbau der Protokollierung des Ingestprozesses
3. Verknüpfung zwischen der Ingest- und der Ablieferungsverwaltung
4. Betrieb der FedoraCommons-Datenbank auf einer Oracle-Datenbank statt wie bisher auf einer PostgreSQL³-Datenbank
5. Implementierung einer Standardschnittstelle für SIP und Anpassung der AIP-Definition
6. Verbesserter DIP-Viewer als Alternative zum Standard-Disseminator von FedoraCommons
7. Schaffung einer Möglichkeit mit hoher *usability*, um innerhalb des SIP browsen zu können, zum Zwecke der stichprobenhaften Kontrolle bei der manuellen Validierung des SIP
8. Realisierung eines SIP-Generators.

Unabhängig von diesem Ausbau der Infrastruktur wurden die 2009 definierten Geschäftsprozesse – von der Bewertung über Anbieten und Abliefern bis zur Erschließung und Magazinierung – weiter ausgearbeitet und in Form eines Organisationshandbuchs mit der bestehenden Aufbauorganisation abgestimmt. Dabei wurde an der Zerteilung der prospektiven Bewertung festgehalten: Die Makrobewertung hat die Bewertung der Aufgaben des Aktenbildners zum Gegenstand, die Mikrobewertung stellt eine Bewertung der angebotenen Unterlagen resp. des aktenbildenden Systems dar. Im Rahmen des Übernahmeprozesses findet bei vorliegender prospektiver Bewertung lediglich noch eine Kontrolle des Angebots und der Ablieferung statt, mit der die Übereinstimmung von zu übergebenden Unterlagen mit in der Bewertung festgelegten Kriterien überprüft wird.

Im Folgenden wird das zentrale Arbeitspaket AP5 genauer dargestellt.

² Zum Pilotprojekt der ARAP-Übernahmen vgl. Lambert *Kansy*: Pilot-Übernahme elektronischer Unterlagen durch das Staatsarchiv Basel-Stadt, 2011, <http://www.archiversum.com/wp-content/uploads/2011/07/Kansy-02.pdf>.

³ PostgreSQL (oft kürzer als Postgres bezeichnet) ist ein Open Source Datenbankmanagementsystem.

Implementierung einer Standard-Schnittstelle für Übernahmepakete (SIP)

Bei der Entwicklung der Ingestwerkzeuge hat das Staatsarchiv Basel-Stadt als Pilotübernahme die Ablieferung von Unterlagen aus dem System ARAP der Kantonspolizei definiert. Dafür wurde ein spezifisches SIP definiert, das auf dem XML-Schema des Übernahme-Assistenten von scope solutions basiert. Dieses Übernahmepaket wird bis auf Weiteres bei der Ablieferung aus ARAP verwendet. Von Beginn an wurde jedoch diese Pilotlösung nicht als Basis für eine durch verschiedene Quellsysteme zu beliefernde SIP-Schnittstelle verstanden. Die Standard-Ablieferungsschnittstelle sollte auf bestehenden Standards aufbauen, um die Chancen zu vergrößern, dass die Quellsysteme der Aktenbildner diese Übergabeschnittstelle bedienen können. Ziel ist es, die Anzahl der durch das Archiv zur Verfügung gestellten Ablieferungsschnittstellen möglichst gering zu halten, um den Bewirtschaftungsaufwand langfristig überhaupt bewältigen zu können.

Ende 2009 stellte das Schweizerische Bundesarchiv eine von ihm entwickelte Ablieferungsschnittstelle vor.⁴ Eine eingehende Überprüfung dieser SIP-Definition durch die KOST im Rahmen ihres Projekts bento ergab 2010, dass nur wenige Modifikationen notwendig waren, um die Schnittstelle auch für andere Archive nutzbar zu machen. So wurde die Spezifikation um die Möglichkeit ergänzt, zusätzliche Metadaten, die nicht in den Elementen von metadata.xsd abgebildet werden können, dennoch in strukturierter und damit auch maschinell verarbeitbarer Form im SIP abliefern zu können. Im Rahmen des bento-Projekts wurden zwei Wege definiert: zum einen die Definition von Key-Value-Paaren und zum anderen die Integration von Elementen, die in fremden XML-Schemata definiert sind. Auf diese Weise können beliebige weitere Metadaten in der Metadatendatei metadata.xml des SIP übernommen werden.⁵ Die SIP-Spezifikation des Bundesarchivs, ergänzt um die Erweiterung der Key-Value-Paare, soll 2012 als eCH-Standard verabschiedet werden.⁶ Aufgrund der sich abzeichnenden Akzeptanz dieser Ablieferungsschnittstelle erscheint deren Implementierung als Standard-Ablieferungsschnittstelle gerechtfertigt.

Neudefinition der Metadatenstruktur der Archivpakete (AIP)

Die Neudefinition des Standard-SIP konnte sich auf die Implementation eines – in Entwicklung befindlichen – Standards abstützen. Bei der Bildung der Archivpakete (AIP) hingegen wurde mit EAD ein bestehender und verbreiteter Metadatenstandard durch das – ebenfalls im Rahmen des KOST-Projekts bento entwickelte – XML-Schema xIsadg ersetzt.⁷ Grund hierfür ist die mehrdeutige Abbildbarkeit von ISAD(G)-Metadaten in EAD und die Komplexität dieses Standards, die auf seine ursprüngliche Zielsetzung (Aufbereitung gedruckter Findbücher für die Präsentation im Internet) zurückgeht. Beides macht EAD aus Sicht des Staatsarchivs nicht zum idealen Standard für die Abbildung von beschreibenden archivischen Metadaten innerhalb

⁴ Siehe <http://www.bar.admin.ch/themen/00876/00877/>.

⁵ Zum Projekt bento der KOST siehe http://kost-ceco.ch/cms/index.php?bento_de.

⁶ Zum Verein eGovernment Schweiz eCH siehe <http://www.ech.ch>.

⁷ Zur Spezifikation von xIsadg siehe http://kost-ceco.ch/cms/index.php?bento_de.

eines AIP, die aufgrund der Granularität der AIPs auf den Verzeichnisstufen Dossier und Dokument (sowie allenfalls Serie) keinesfalls Umfang und Struktur eines Bestand beschreibenden Findbuchs haben.

Das xIsadg-Schema basiert auf dem im Projekt eDavid des Stadtarchivs Antwerpen entwickelten XML-Schema isad.xsd und erweitert es vor allem hinsichtlich der Möglichkeit hierarchischer Schachtelung von Verzeichniseinheiten.⁸ Zudem wurde die Definition einiger ISA-D(G)-Felder im XML-Schema angepasst. Auch die Frage der Redundanzvermeidung wurde in xIsadg umzusetzen versucht, indem Vererbungsregeln definiert werden, die freilich nicht durch das Schema validiert werden können. Es wurde in xIsadg die Möglichkeit der Übernahme beliebiger zusätzlicher Metadaten in Form von key-value-Paaren geschaffen, ähnlich wie bei der Erweiterung der SIP-Spezifikation. Zusätzlich können hier auch fremde XML-Schemata integriert werden. Zwecks Referenzierbarkeit von Primärdaten, weiteren Metadaten Dateien sowie Archivpaketen wurden entsprechende Elemente definiert:

- `aipReference` – für die Referenzierung von Archivpaketen
- `primaryDataLocator` – für die Referenzierung von Primärdaten
- `secondaryDataLocator` – für die Referenzierung von Primärdaten über deren Beschreibung in anderen XML-Dateien.

Während die beiden letzten Referenzierungsmöglichkeiten für die Verwendung innerhalb eines AIP geschaffen wurden, wird mit dem Verweis auf ein AIP die Nutzung von xIsadg für den Austausch von beschreibenden Metadaten zwischen Archivinformationssystemen möglich. Die für die Nutzung von xIsadg innerhalb eines AIP notwendigen Anpassungen wurden vom Staatsarchiv Basel-Stadt gemeinsam mit der Geschäftsstelle der KOST definiert.

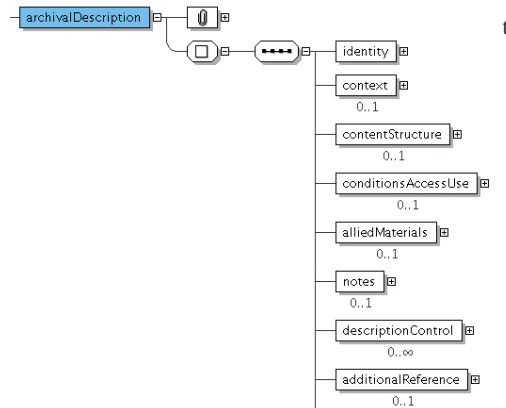


Abb. 1: Grundstruktur von xIsadg.

⁸ Zum Projekt eDavid siehe <http://www.edavid.be/eng/index.php>.

Die übrige Struktur des AIP wurde unverändert belassen. Es ist als JAR-Container definiert, der Metadaten und Primärdaten in drei Verzeichnissen getrennt enthält. Die Metadaten bestehen aus zwei Dateien im Verzeichnis metadaten: den technischen und administrativen Metadaten gemäß dem PREMIS-Standard, Version 2.1, und den beschreibenden Metadaten gemäß dem xIsadg-Schema.

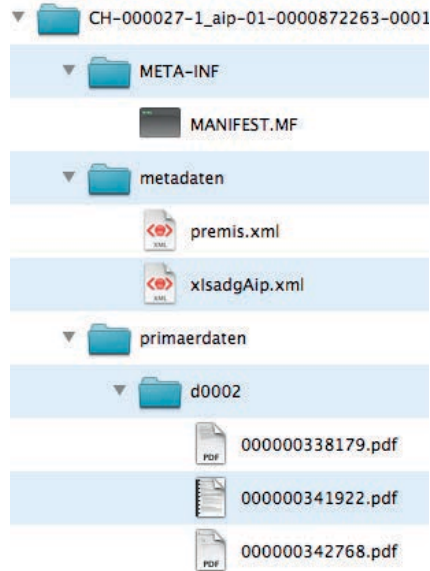


Abb. 2: AIP-Struktur.

- Das Verzeichnis META-INF enthält mit der Datei MANIFEST.MF Metadaten über das AIP und dessen Verzeichnis- und Dateiinhalte.
- Das Verzeichnis primaerdaten enthält in beliebiger Schachtelung in Verzeichnissen die Primärdaten.
- Der Zugriff auf die Primärdaten erfolgt von den beschreibenden Informationen in der Datei xIsadgAip.xml aus über eine xPointer-Verknüpfung mittels der Datei premis.xml.

Die JAR-Datei des Archivpakets erhält als Dateinamen den Persistent Identifier (PID), der mit dem ISIL-Code des Staatsarchivs beginnt, und nachfolgend die Art des Informationspakets angibt. Die weiteren Abschnitte des PID setzen sich aus der Kennung des Archivinformationssystems zusammen, gefolgt von der dort vorhandenen ID für die mit dem AIP referenzierte Verzeichnungseinheit sowie einer Versionsnummer des AIP.

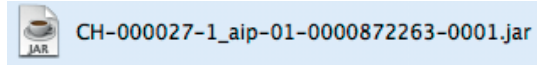


Abb. 3: Beispiel für PID des AIP.

Infrastrukturausbau

Parallel zu dem Ausbau der Ingestwerkzeuge erfolgte ein Ausbau der Infrastruktur des digitalen Magazins (digiMag).

Im Rahmen des Projekts Informatisierung III wurde somit die vom Staatsarchiv Basel-Stadt im Bereich der Archivinformatik eingesetzte IT-Infrastruktur deutlich ausgebaut. Bestand zuvor das Archivinformationssystem scopeArchiv aus einem Datenbankserver, einem Applikationsserver für den Online-Archivkatalog sowie den lokalen Client-Anwendungen, so wurden die folgenden drei- virtualisierten – Server im Rahmen des Projekts zusätzlich beschafft:

- FTP-Server für die Entgegennahme der Ablieferungspakete (SIP)
- scopeIngest-Application-Server (mit IIS-Webserver für den SIP-Browser)
- createBento-Webserver (Apache/Tomcat-Webserver für den SIP-Generator).

Daneben wurde auch das digitale Magazin, sowohl in Hinblick auf seine Speicherkapazität als auch in Hinblick auf seine Architektur, massiv ausgebaut. Die neuen Komponenten des digiMag sind:

- FedoraCommons-DB-Server
- FedoraCommons-Application-Server (mit Apache/Tomcat)

Mit Blick auf das OAIS-Referenzmodell deckt die Infrastruktur (Hardware und Software) des Staatsarchivs Basel-Stadt nach Abschluss des Projekts die Prozesse Ingest, Storage, Data Management und – wenn auch nur in rudimentärer Form – Access ab. Gegliedert nach Software-Schichten ergibt sich die nachstehende Übersicht.

Infrastrukturlayer	OAIS-Prozess	StABS-Prozess	Anwendung/Lösung
Archival Information System	Access, Data Management	Erschließung, Benutzung	scopeArchiv, scopeQuery
Digital Archiving Software	Ingest, Storage	Übernahme, Magazinierung	FTP-Server, scopeIngest und FedoraCommons
Storage Management Software	Storage	Magazinierung	FedoraCommons, Primärspeicher
Storage Hardware	Storage	Magazinierung	Primärspeicher und arcun

Tabelle 1: Infrastruktur-Systematik

Ausbau des digitalen Magazins

Mit Beginn des Informatisierungsprojekts Info III gab es im Staatsarchiv Basel Stadt ein Netzwerklaufwerk (Protokoll: SMB/CIFS) zur Speicherung von Dateien. Dieses bestand aus einer handelsüblichen NAS-Speicherbox mit 5 Festplatten, die als RAID 5 konfiguriert waren und netto 3,5 TB Platz enthielten. Dieses Laufwerk war an allen Desktop Rechnern des Archivs gemountet. Ein kleiner Kreis von Archivmitarbeitern war mit Schreibrechten ausgestattet und konnte dort Digitalisate und digitales Archivgut in Form von einzelnen Dateien ablegen. Die übrigen Mitarbeitenden verfügten nur über Lesezugriff. Zur Sicherung der Daten stand ein weiteres gleichartiges Laufwerk als Spiegel zur Verfügung, auf das der Verzeichnis-Ast des ersten Laufwerkes repliziert wurde. Auf diesen Spiegel konnte ausschließlich der Systemadministrator zugreifen; er entsprach also in etwa einem Schattenarchiv (dark archive).

Beim Replizieren wurde überprüft, ob Dateien auf dem Spiegel verwaist waren, weil die Ursprungsdatei auf dem Urbild gelöscht oder verändert worden war. Falls das vorkam, wurde der Administrator über eine E-Mail informiert, und es war seine Aufgabe, dieser Abweichung nachzugehen: Durch Nachfragen bei den entsprechenden Archivmitarbeitern musste er herausfinden, ob die Löschung der Datei bewusst oder aus Versehen geschah.

Mit dieser einfachen Bewirtschaftung wurde eine sehr hohe Sicherheit gegen Ausfall der Dateien erzielt. Insbesondere das Risiko einer versehentlichen Löschung wurde so vermindert. Ein Backup hilft im Falle einer versehentlichen Löschung nur dann, wenn der Irrtum innerhalb der Backup-Frist bemerkt wird.

Als die Datenmenge im Laufe des Informatisierungsprojekts Info III zunahm, war schnell klar: Es wird ein größerer Speicher benötigt. Die neue Lösung sollte aber nicht laufwerkbasierend sein, sondern in einer Netzwerkschnittstelle bestehen, der die Dateien übergeben werden (Repositoryum). Ein solches Repositoryum funktioniert ähnlich wie eine Garderobe mit Garderobenmarken:⁹ Man bringt ein Kleidungsstück – in unserem Fall eine Datei – und erhält dafür eine Garderobenmarke. Dem entspricht in unserem Fall eine Zeichenkette als Persistent Identifier. Diese PID wird im Archiv Informations System abgelegt. Möchte man die Datei wieder haben, geht man (übers Netzwerk) zum Repositoryum, übergibt die PID und erhält dafür die Datei, so wie man an der Garderobe sein Kleidungsstück gegen Vorlage der Garderobenmarke zurück erhält.

Um es vorwegzunehmen: Die produktive Umgebung arbeitet immer noch (März 2012) mit einer laufwerkbasierenden Lösung. Allerdings ist das Laufwerk (der SMB-Server) auf einen größeren Host mit mehr Festplatten umgezogen. Das Testsystem befindet sich auf demselben Host und ist bereits auf Repositoryum-Betrieb umgestellt. Als entsprechende Repositoryum Software dienen FedoraCommons und eine Oracle Datenbank. Die Umstellung des Produktivsystems wird im zweiten Quartal 2012 erfolgen.

Die Vorbereitung des Umzugs vom alten System mit den beiden NAS-Speicherboxen auf ein neues, größeres System begann mit der Evaluierung eines neuen Speichers samt Rechner

⁹ Diese von Kai *Naumann* an der Tagung des AK AUdS 2012 gebrauchte Analogie haben wir bei der Ausarbeitung des Vortrags gerne aufgegriffen, weil sie den Sachverhalt anschaulich schildert.

(Host). Die zugrunde liegenden drei Komponenten Betriebssystem – Volume Manager – Dateisystem standen im Fokus der Evaluierung.

Die Anforderungen an das neue System lauteten grob umrissen wie folgt:

1. Eigener Host mit der Möglichkeit, Shellscrippts auszuführen
2. Speichervolumen deutlich grösser als 10 TB und weiter ausbaufähig
3. Möglichst preiswert (<1000 SFr/TB in der Anschaffung, weitere 1000 SFr/TB und Jahr im Betrieb)
4. Garantie der Bitstream Preservation möglichst schon im Betriebssystem
5. Thin Provisioning¹⁰ einfach und elegant möglich, da mehrere Virtuelle Systeme den Speicher nutzen
6. Gutes Zusammenspiel der Hard- und Softwarekomponenten.

Keine dieser Vorgaben war für sich alleine zwingend erforderlich; ein Plus im einen Bereich konnte ein Minus in einem anderen Bereich kompensieren. Es gab auch Parameter, die das Disk-Subsystem nicht zu haben brauchte, die allerdings oft bei Offerten als Vorteil angepriesen werden:

1. Mandantenfähigkeit (das Archiv ist und bleibt einziger Mandant des Disk-Subsystems)
2. Geschwindigkeit (wegen der höheren Ausfallwahrscheinlichkeit der Komponenten eher störend, deshalb SATA-Festplatten statt SAS).

Das neue System sollte die beiden Netzwerklaufwerke ablösen, und Digitalisate und originär-digitales Archivgut sollten über eine Netzwerkschnittstelle eingeliefert werden können. Es war aber absehbar, dass diese Repositorium-Funktionalität nicht rechtzeitig fertiggestellt werden konnte. Die bereits laufenden Digitalisierungsaktivitäten brachten zudem große Mengen an Digitalisaten, die vorläufig auf einem Netzwerklaufwerk zwischengelagert werden mussten. Wegen dieser Übergangszeit musste das anzuschaffende System das SMB-Protokoll beherrschen, eine Bedingung, die einfach erfüllt werden konnte, da eine gut funktionierende OpenSource Implementierung des SMB Protokoll (SAMBA) existiert. Diese kam entsprechend zum Einsatz. Das Umkopieren kann später stattfinden. Dabei werden die Dateien vom Netzwerklaufwerk geholt, mit Metadaten versehen, in ein SIP verpackt und so ins digitale Magazin eingeschleust. Damit ist sichergestellt, dass die Digitalisate korrekt im Archiv-Informationen-System verzeichnet sind und dass keine unverzeichneten Digitalisate auf dem Repositorium liegen.

¹⁰ Zum Begriff *Thin Provisioning* weiter unten mehr.

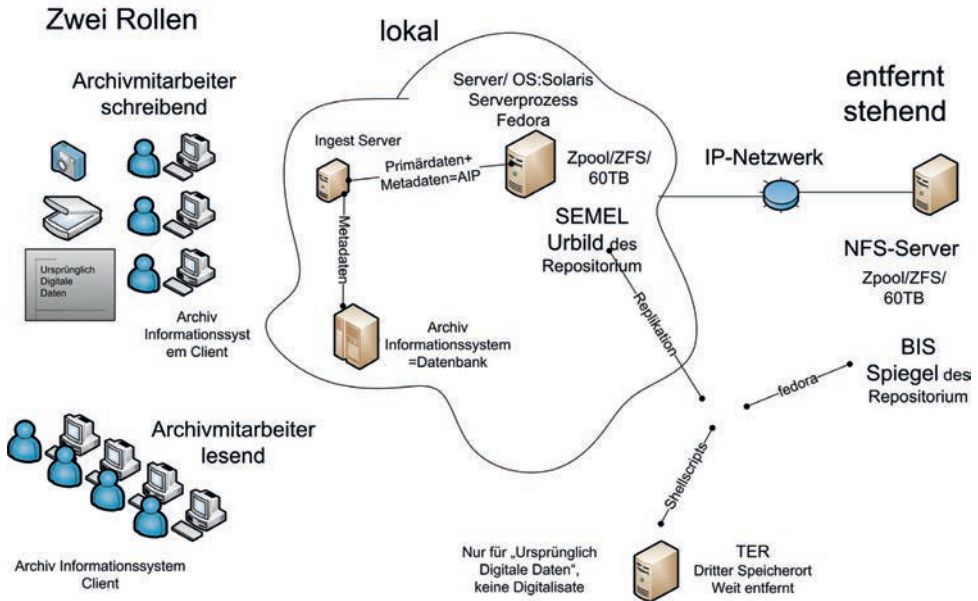


Abb. 4: Repository-Architektur des digiMag.

Eine gute *Thin Provisioning* Fähigkeit des Gesamtsystems war unter anderem wegen dieses später geplanten Umkopierens wichtig. Hierbei wird frei werdender Platz auf der einen Seite (Netzwerklaufwerk) der anderen Seite (Repositoryum) zur Verfügung gestellt. So können beispielsweise auf einem Speicher mit 20 TB verfügbarem Platz, wovon 16 TB mit Daten belegt sind, diese 16 TB umkopiert werden, ohne dass vorübergehend ein zweiter Speicher mit mindestens der Größe 16 TB benötigt wird.

Aus Kostengründen ausgeschlossen waren von vornherein große Hardware RAIDs für SAN-Netzwerke, wie zum Beispiel EMC, Hitachi oder Netapp. Diese Systeme sind sowohl in der Anschaffung wie im Betrieb viel zu teuer (mindestens um den Faktor 5), und der Vorteil der Mandantenfähigkeit bringt keinerlei Gewinn für ein Archiv. Ebenfalls überflüssig ist die hohe Geschwindigkeit dieser Systeme. Die Vorgabe des Speichervolumens auf über 10 TB schloss den Einsatz von preisgünstigen Speicherboxen aus dem SOHO Umfeld aus. 2009 hatten die größten erhältlichen Festplatten 1 TB und die größten SOHO Speicherboxen acht Einschübe, sodass damit netto eine Speicherkapazität von 6,5 Terabyte – bei Einsatz von RAID 5 – zu erzielen gewesen wäre.

Die Lösung sollte von wenigen Herstellern mit aufeinander abgestimmten Komponenten bezogen werden können. Im Idealfall ist das ein einzelner Hersteller, der alles – Hardware und Software – liefert. Gerade beim Einsatz von RAID Systemen ist das korrekte Zusammenspiel von Volume Manager, Treibersoftware und Firmware von Disk und Controller unerlässlich. Die häufigsten Bugs finden sich dort, ohne dass man als Anwender bei zwei Lieferanten einen davon haftbar machen könnte. Die Hersteller schieben im Zweifelsfall die Schuld hin und her.

Der lokale Informatikdienstleister des Kantons wurde aufgefordert, ein Angebot abzugeben, und offerierte ein großes Hardware RAID mit SAN-Anschluss. Die Kosten dafür waren allerdings für das Staatsarchiv Basel-Stadt untragbar hoch (43.000 SFr/TB). Angeschafft wurde letztendlich ein kleiner SUN sparc Rechner mit einem JBOD Disk-Subsystem. Dies brachte in Kombination mit zfs als Volume Manager und Dateisystem die maximale Erfüllung der vorgegebenen Kennwerte. Mit 24* 1 TB Festplatten wurde ein Netto-Speichervolumen von 20 Terabyte erzielt.

zfs brachte als Filesystem und Volume Manager eine Reihe von Vorteilen mit, die anderen Filesystemen fehlen, aber gerade in einer Archivlandschaft nützlich sind. Das sind:

1. Scrubbing: Die Möglichkeit, auf Betriebssystemebene alle Daten auszulesen und so die Lesbarkeit der Daten auf der Festplatte periodisch sicherzustellen.
2. Self healing: Wird beim Lesen ein Fehler detektiert, so werden die Daten der betroffenen Festplatte umgehend auf andere Platten kopiert.
3. Hashtrees: Beim Schreiben einer Datei wird ein Hashwert gebildet, der bei allen zukünftigen Lesevorgängen im Hintergrund überprüft wird. Wird ein Fehler entdeckt, so wird er falls möglich korrigiert, ansonsten wird eine Fehlermeldung an die Applikation geschickt. Beim Scrubbing wird der Hashwert ebenfalls geprüft.
4. Thin Provisioning: zfs bietet mit dem Konzept von Disk-Pools ein sehr elegantes Thin Provisioning ohne kompliziertes Umbauen eines RAID. Filesysteme können regelmäßig vergrößert und neu aufgeteilt werden und so in den freien Disk-Pool hineinwachsen.

All diese Kennwerte kennzeichnen aus unserer Sicht zfs gegenüber anderen Filesystemen und Volume Managern als überlegenes System. Als wichtigstes Merkmal seien die Hashtrees genannt, die kein anderes Filesystem anbietet. Damit kann auf einfache Art und Weise die von DIN 31644 unter 5.3.2 geforderte Integrität (K15 Unversehrtheit) der Datenströme sichergestellt werden.

Rückblick und erste Erfahrungen

Zu konstatieren ist, dass einige anfangs gesetzte Ziele des Teilprojekts 2 Digitale Archivierung nicht erreicht werden konnten. Verantwortlich hierfür sind zu einem kleineren Teil Veränderungen im Projektverlauf selbst; zum größeren Teil jedoch die Konkurrenz um die gleichen Personalressourcen durch andere Projekte. So mussten aus Termingründen Projekte wie die Einführung der Online-Bestellung auf Mitte 2011 und das Teilprojekt Digitalisierung von Archivgut zeitweise gegenüber dem hier dargestellten Projekt priorisiert werden. Daraus resultierten entsprechende Verzögerungen bei der Projektabwicklung. Es zeigte sich überdies, dass die Ausbaumaßnahmen einen größeren konzeptionellen Aufwand mit sich brachten als anfänglich geplant. Die Entwicklung verlief daher stärker iterativ als geplant, in engem Kontakt zwischen scope solutions als Auftragnehmer und dem Staatsarchiv als Auftraggeber.

Nicht umgesetzt wurde die 2010 skizzierte Anwendung *Bewerten und Anbieten*, mit der die vorarchivischen Geschäftsprozesse der Bewertung und des Anbietens abgewickelt werden sollten.

Grund für den Verzicht auf die Realisierung waren knappe finanzielle Ressourcen sowie zahlreiche offene Punkte in Bezug auf die Form der Umsetzung (Anwendung mit lokalem Client oder webbasierte Anwendung) und die Integration mit dem Archivinformationssystem. Auch wurden die Informationsangebote über die SIP-Schnittstelle und generell den Ansatz des Staatsarchivs Basel-Stadt bei der Archivierung digitaler Unterlagen aus Ressourcenmangel zurückgestellt. Im Bereich des Storage wurde die Anbindung von arcun als Teil des digitalen Magazins sowie der Umbau desselben zur Repository-Lösung nicht im Projektrahmen durchgeführt. Diese Arbeiten werden 2012 nachgeholt.

Trotz der (noch) nicht umgesetzten Ziele kann die um Projektauftrag 2008 formulierte Zielsetzung *Aufbau von Prozessen und Bereitstellung der erforderlichen Infrastruktur zur Archivierung digital erzeugter und archivierter Unterlagen der kantonalen Verwaltung* als erreicht gelten. Mit Abschluss des Projekts verfügt das Staatsarchiv Basel-Stadt über eine funktionsfähige Infrastruktur für die Archivierung digitaler Unterlagen, die 2012 produktiv gestellt wird. Zudem sind die archivischen Geschäftsprozesse von der Bewertung bis zur Magazinierung auf die digitale Archivierung vorbereitet und mit der Aufbauorganisation abgestimmt worden. Die geschaffene Infrastruktur stellt eine Erweiterung der bereits bestehenden Systeme des Archivinformationssystems und des digitalen Magazins dar. Auch bedeutet die Implementierung der KOSTbento-SIP-Schnittstelle und die Verwendung von xIsadg für die beschreibenden Metadaten des AIP eine Orientierung an Standards resp. die Nutzung von allgemein zugänglichen Spezifikationen und senkt so in beiden Fällen die Risiken von kompletten Eigenentwicklungen.

Im Bereich des digitalen Magazins konnte die dauerhafte bitstream preservation als Anforderung umgesetzt werden. Somit ist hier die notwendige Basis vorhanden, auf der Maßnahmen der digitalen Bestandserhaltung (digital preservation) aufgebaut werden können.

Aus der Projektmanagementsicht handelte es sich um ein ambitioniertes Vorhaben, in dem zahlreiche IT-Komponenten erst einmal entwickelt und anschließend zu einem Gesamtsystem integriert werden mussten. Es zeigte sich bereits zu Projektbeginn, dass es im Sinne der Minimierung von Aufwand und Risiko sinnvoll war, auf der vorhandenen Infrastruktur aufzubauen und das Archivinformationssystem scopeArchiv als führendes System beizubehalten und zu erweitern. Somit setzte das Staatsarchiv weiter auf die Nutzung einer bewährten Standardsoftware und in Teilbereichen der Entwicklung auch auf die Nutzung von offenen Standards. Zugleich resultierten daraus aber im Projektverlauf inhaltliche und terminliche Abhängigkeiten von Fahrplänen des Herstellers und archivübergreifenden Arbeiten im Rahmen des KOSTbento-Projekts.

In Bezug auf die einzelnen Komponenten wurde anfänglich die Bedeutung eines standalone-Werkzeugs zur Erzeugung schnittstellenkonformer SIP unterschätzt. Es zeigte sich im vorarchivischen Bereich seit 2009 deutlich, dass zahlreiche Aktenbildner archivwürdige digitale Unterlagen nur in Form von Dateiablagen anbieten können – sei es, weil diese in Ermangelung eines Records Management-Systems so abgelegt werden oder weil Fachapplikationen und legacy applications nur in dieser Form Unterlagen exportieren können und der Einbau einer SIP-Schnittstelle unverhältnismäßig ist. Es musste auch für diese Situation eine allgemeine Lösung gefunden werden, da die manuelle Generierung von SIP fehleranfällig und sehr aufwendig ist. Die Alternative hätte in der Öffnung des Ingestprozesses dahin gehend bestanden, dass beliebige Formen der Ablieferung hätten verarbeitet werden müssen. Dieser Ansatz führt jedoch dazu, dass das

Archiv Aufgaben der Aktenbildner übernimmt und die Anforderungen an die Übernahmewerkzeuge in Bezug auf deren Flexibilität in massiver Weise gesteigert werden, da dann jede Übernahme digitaler Unterlagen aus einem Quellsystem faktisch auf die Realisierung einer eigenen SIP-Schnittstelle hinausläufe. Dies ist sowohl mit Blick auf den zu leistenden Aufwand im Archiv als auch die Zuständigkeiten der Beteiligten im Übernahmeprozess abzulehnen. Auch führt eine Vielzahl von Übernahmeschnittstellen zu einem hohen Aufwand bei der Bewirtschaftung derselben im Verlauf des Lebenszyklus der Quellsysteme wie auch der Übernahmewerkzeuge des Archivs. Die Entwicklung eines webbasierten SIP-Generators war die Antwort hierauf, die nunmehr die Flexibilität des Archivs bei Übernahmen massiv erhöht hat.

Über alles betrachtet stehen Prozesse und Infrastruktur in einer ersten Version zur Verfügung und es gilt nun, damit weitere Erfahrungen im produktiven Einsatz zu sammeln, die zur Weiterentwicklung und Ausbauten führen werden.

Ausblick

Unter Einbeziehung der nicht realisierten Ziele und der bisherigen Erfahrungen sind in folgenden Bereichen weitere Ausbau-, respektive Entwicklungsschritte in Planung begriffen:

- Im Bereich Übernahme wird der bestehende ARAP-Ingest soweit angepasst werden, dass die gleiche AIP-Struktur generiert wird, wie im neu definierten Standard-Ingest (bento-Ingest). Sollte sich das Quellsystem ARAP in eine Richtung weiterentwickeln, die größere Anpassungen bei der SIP-Schnittstelle erforderlich macht, ist der Einsatz des SIP-Generators zu prüfen.
- Generell ist der Ingest-Workflow zu flexibilisieren. So soll der Schnittpunkt des SIP bei der Generierung der Archivpakete flexibel werden und zusätzlich zur heutigen Situation nicht nur auf Dossierstufe, sondern auch auf Bestands-, Serien- und Dokumentstufe erfolgen können. Auch muss die Einfügeposition der abgelieferten Unterlagen in die Archivtekonik zwischen den Verzeichnissstufen Bestand und Dokument frei wählbar sein.
- Schließlich ist der Ingest-Workflow so weit zu modifizieren, dass er auch für die dauerhafte Speicherung von Digitalisaten genutzt werden kann.

Der Umbau des digitalen Magazins zur Version 2.0, d.h. die Anbindung von arcun an FedoraCommons und den Wechsel des Zugriffs vom Netzwerklauferwerk zum Repository-Zugriff erfolgt im Anschluss an den Abschluss dieses Projekts. Auch wird die Nutzung von Zeitsignaturen und Checksummen zur Sicherstellung der Anforderungen Integrität und Authentizität angedacht. In betrieblicher Hinsicht wird die Minimierung des Stromverbrauchs durch den Einsatz von MAID-Systemen und damit eine Senkung der Betriebskosten sowie die Erhöhung der Ausfallsicherheit bei Systemstillständen angestrebt. Mittel- bis langfristig ist schließlich der Gesamtbetrieb des digitalen Magazins auf die zentralen Informatikdienste zu übertragen, um archivintern Ressourcen für die arbeitsintensive und nicht delegierbare Kernaufgabe der digitalen Bestandserhaltung freizusetzen.

Der Aufbau von Komponenten für die digitale Bestandserhaltung und die Schaffung eines digitalen Lesesaals, mit dem die im O AIS vorgesehenen Zugangs- und Benutzungsfunktionen umgesetzt werden, bildet für die nächsten Jahre den Kernbereich der konzeptionellen Weiterentwicklung.

Der nestor-Leitfaden zur Digitalen Bestandserhaltung und seine Folgen für die Archive¹

Von CHRISTIAN KEITEL

Längst hat die Literaturproduktion zur digitalen Archivierung schwindelerregende Höhen erreicht, es vergeht kaum ein Jahr, in dem nicht ein neuer, noch wichtigerer Standard vorgestellt und der Aufmerksamkeit der Fachöffentlichkeit anempfohlen wird. Es ist daher nur legitim zu fragen:

Bedarf es wirklich eines neuen Standards zur digitalen Bestandserhaltung? Warum sollten sich Archivarinnen und Archivare mit dem Leitfaden zur digitalen Bestandserhaltung beschäftigen?

Die maßgeblichen Konzepte für die digitale Archivierung wurden bislang zum überwiegenden Teil nicht von den klassischen Archiven entwickelt.² OAIS, PREMIS und die Ergebnisse der Projekte CEDARS (Significant Properties) und PLANETS wurden maßgeblich von Mitarbeitern anderer Einrichtungen, den neuen digitalen Archiven verfasst. Das Performance Model wurde zwar vom australischen Nationalarchiv veröffentlicht, aber bislang nur wenig von den klassischen Archiven rezipiert.

Die vornehme Zurückhaltung der klassischen Archive bei der Entwicklung neuer Konzeptionen zur Bewältigung der digitalen Archivierung hat natürlich Folgen. Ein Archiv, das sich heute aufmacht, um digitale Unterlagen zu übernehmen, ist gezwungen, die an anderer Stelle entwickelten Konzepte zu verwenden. Dass sich dabei Bruchstellen zu den hergebrachten Konzepten auftun, ist geradezu zwangsläufig. Wie vermitteln wir zwischen der in PREMIS abgebildeten Trennung von Informationsobjekt (intellectual entity) und Repräsentation einerseits und der noch in ISAD(G) kodifizierten unverbrüchlichen physischen und logischen Einheit des Archivals? Wie können wir die Stufenerschließung nach ISAD(G) vereinbaren mit dem für OAIS grundsätzlichen Konzept des Archivierungspakets (AIP)? Und daran anschließend: Werden wir in einigen Jahren gezwungen sein, auf eine adäquate Abbildung von Tektonik und Klassifikation in unseren Archivierungssystemen verzichten zu müssen, da es nur noch Systeme auf dem Markt gibt, die den bisherigen konzeptionellen Vorgaben der neuen digitalen Archive entsprechen und für den viel größeren Markt der Bibliotheken entwickelt wurden?

Auf dem zentralen Gebiet der digitalen Bestandserhaltung hat die internationale Fachgemeinschaft derer, die sich mit digitaler Archivierung im weiteren Sinne beschäftigen, bislang wenigstens fünf fundamentale Ansätze entwickelt:

¹ In Erinnerung an Ernst *Fuchs*, Professor der Theologie, Lehrer der Hermeneutik und Großvater des Verfassers.

² Christian *Keitel*, Archivwissenschaft zwischen Marginalisierung und Neubeginn. In: *Archivar* 64 (2011), S. 33–37.

- OAIS sieht im Preservation Planning einen von sechs zentralen Funktionsbereichen und ordnet diesem vier wesentliche Aufgaben zu: Erhaltungsstrategien und Standards entwickeln, Paketmodelle und Migrationspläne entwickeln, vorgesehene Zielgruppe beobachten und Technologie beobachten.³
- Ebenfalls von OAIS stammt das Konzept der Designated Community. Digitale Archive sollten ihr gesamtes Handeln an den anzunehmenden Interessen ihrer künftigen Nutzer ausrichten. So unbestritten dieses Konzept seit der ersten Veröffentlichung im Jahr 2002 auch ist, es gibt doch recht wenige Überlegungen, wie dies in die Praxis umgesetzt werden könnte.
- 1999 schon stellte das CEDARS Projekt seine Überlegungen zu den Significant Properties vor.⁴ Durch den von Margret Hedstrom und Christopher Cal Lee 2002 auf dem DLM-Forum in Barcelona gehaltenen Vortrag fand das Konzept breite Resonanz in einschlägigen Fachkreisen und 2008 im Inspect Projekt eine erste konzeptionelle Umsetzung.⁵
- Dass im Kern die von Menschen wahrnehmbare Information erhalten werden sollte, hat bereits OAIS festgehalten. Deutlicher formuliert wird es im ebenfalls 2002 erschienenen Performance Model der National Archives of Australia.⁶
- Schließlich zählt auch der 2005 veröffentlichte PREMIS-Standard mit Begriffen wie *Intellectual Entity* und *Representation* zur Gruppe der Konzepte, die für die Erhaltung digitaler Objekte einschlägig sind.⁷

Die fünf aufgeführten Konzepte wurden mit unterschiedlicher Intensität von der internationalen Fachdiskussion diskutiert. Dennoch konnte bislang kein übergreifendes Modell vorgelegt werden, das diese unterschiedlichen Ansätze unter einem Dach vereinigt hätte.⁸ Man bediente sich vielmehr aus dieser Fundgrube unterschiedlichster Vorstellungen manchmal hier und manchmal da und hoffte darauf, dass sich auf die Dauer schon ein einheitliches Vorgehen bei der Erhaltung digitaler Objekte entwickeln würde.

Digitale Archivierung ist derweilen dabei, von der zwar Angst machenden, aber doch in fernerer Zukunft liegenden Bedrohung zu einer unmittelbar anstehenden und ganz konkreten Aufgabe zu werden. Nun müssen die staatlichen Archive die digitalen Unterlagen bewerten und

³ Referenzmodell für ein Offenes Archiv-Informationssystem, Deutsche Übersetzung der nestor-Arbeitsgruppe OAIS-Übersetzung / Terminologie, nestor-materialien 16, Frankfurt/Main 2012

⁴ Vgl. Kelly Russell: Digital Preservation and the Cedars Project Experience, <http://www.worldcat.org/arcviewer/1/OCC/2007/09/28/0000073852/viewer/file507.html>. Alle Abrufe wurden am 17.8.2012 überprüft.

⁵ Margaret Hedstrom und Christopher Cal Lee: Significant properties of digital objects: definitions, applications, implications. Proceedings of the DLM-Forum 2002, INSAR Supplement VII, S. 218–227. Zum Inspect-Projekt s. <http://www.significantproperties.org.uk/>.

⁶ Helen Heslop, Simon Davis und Andrew Wilson: An approach to the preservation of digital records. Canberra 2002, http://www.naa.gov.au/Images/An-approach-Green-Paper_tcm16-47161.pdf.

⁷ Premis [Preservation Metadata: Implementation Strategies]: Version 2.1. (Ohne Ort) 2011, <http://www.loc.gov/standards/premis/version-2-1-announcement.html>.

⁸ Auch das PLANETS Projekt hat seinen Schwerpunkt darauf gelegt, ein Verfahren zur Bestimmung der zum gegenwärtigen Zeitpunkt am besten geeigneten Formate zu entwickeln, wodurch die Kohärenz der heutigen mit den künftigen Erhaltungsbemühungen aber nicht gesichert wird.

übernehmen, für die sie gemäß den Archivgesetzen bereits seit über zwanzig Jahren zuständig sind. Für die kommunalen und kirchlichen Archive, die Archive der freien Wirtschaft, des Rundfunks und aus anderen Bereichen kann ähnliches gesagt werden. Die Aufgabe der digitalen Archivierung steht allorten auf der Tagesordnung. Und mit dem Einstieg in die digitale Archivierung stellt sich erneut die Frage, welche Konzepte dabei aufzugreifen sind.

Gerade auf dem Gebiet der digitalen Bestandserhaltung ergeben diese Konzepte ein vielstimmiges und schwer auflösbares Bild. Wie kann beispielsweise zwischen der am Objekt haftenden Performance und den unabhängig vom Objekt formulierten Wünschen künftiger Nutzer vermittelt werden? Die einfachsten Lösungen sind nun, einfach zu beginnen und darauf zu hoffen, dass sich die offenen Fragen mit wachsender Erfahrung quasi von selbst klären. Das mit diesem Weg verbundene Risiko liegt zunächst darin, dass die klassischen Archive gezwungen sein könnten, Konzepte einzusetzen, die von anderen Einrichtungen für ganz andere Kontexte entwickelt wurden. Es gibt aber auch ganz praktische Schwierigkeiten, die es nahelegen, diesen Weg nicht zu favorisieren. Da ist zunächst die Frage eines in sich stimmigen kohärenten Vorgehens. Bei der digitalen Bestandserhaltung müssen die Archive immer wieder Entscheidungen treffen: Welches Dateiformat, welches Zeichenformat soll für die nächste Migrationsstufe gewählt werden? Auf was kann verzichtet werden, wenn nicht alles erhalten werden kann? Wie kann ein Archiv gewährleisten, dass die Nutzungsmöglichkeiten, die ein Archivar bei der archivischen Bewertung gesehen hat, auch nach Jahrhunderten noch bestehen? Wie gewährleisten wir also über viele Jahrzehnte einen einheitlichen, möglichst widerspruchsfreien Prozess bei der Bestandserhaltung? Dabei unterscheidet sich die digitale Bestandserhaltung durch die zahlreichen, regelmäßig wiederkehrenden Entscheidungen sehr deutlich von der Bestandserhaltung von Papier und Pergament. An die Stelle der passiven Bestandserhaltung tritt eine aktive Bestandserhaltung.

Die Kette der Entscheidungen beginnt spätestens mit der Bewertung der digitalen Unterlagen durch den Archivar.⁹ Sie beginnt also bereits vor der Übernahme. Damit ist den digitalen Archiven die Frage der Kohärenz in der digitalen Bestandserhaltung von Beginn an eingeschrieben. Nun ist es vielleicht denkbar, dass sich die Archivare eines digitalen Archivs bestimmte Gedanken zur Erhaltung ihrer digitalen Archivalien machen, die sie dann schriftlich oder (wahrscheinlicher) mündlich an ihre Nachfolger tradieren. Dies mag mit einigen wenigen Objekten noch funktionieren. Es wird aber nicht mehr funktionieren, wenn Millionen digitaler Archivalien jeweils adäquat und kohärent erhalten werden sollen. Für eine solche Aufgabe sollte dann doch ein einheitliches Konzept vorliegen, an dem sich die digitale Bestandserhaltung orientieren kann. Darüber hinaus wird es Menschen nicht mehr möglich sein, die anstehenden Aufgaben alle selbstständig umzusetzen. Die Umsetzung der getroffenen Entscheidungen muss daher an Maschinen (Computer) übertragen werden. Computer können dies nur vornehmen, wenn sie eindeutige und damit programmierbare Anweisungen bekommen, was ebenfalls von der Existenz eines Konzepts für die digitale Bestandserhaltung abhängig ist.

⁹ Denkbar ist es auch, diese Kette bereits mit der Einführung der digitalen Systeme beginnen zu lassen, in denen die digitalen Unterlagen dann später entstehen sollen. Da Archivare hierauf aber in aller Regel und entgegen verschiedentlich geäußerten Hoffnungen nur einen sehr beschränkten Einfluss haben, soll dieser Aspekt hier nicht weiter verfolgt werden.

Neben der schieren Menge ist es auch noch aus einem weiteren Grund ratsam, sich bereits heute auf die Suche nach einem einheitlichen Konzept zur digitalen Bestandserhaltung zu machen. Jede Entscheidung in der digitalen Bestandserhaltung birgt auch die Gefahr in sich, die Inhalte des Archivals irreversibel zu verändern. Auch besteht die Möglichkeit, die aus Gründen der Bestandserhaltung notwendigen Veränderungen nicht ausreichend zu dokumentieren, so dass zu einem späteren Zeitpunkt nicht mehr gesagt werden kann, ob noch ein authentisches Archival vorliegt. Da also die Authentizität der Archivalien bereits ganz zu Beginn der digitalen Archivierung durch die Erhaltungsmaßnahmen gefährdet werden kann, bedarf es klarer Vorstellungen, an was sich die Authentizität der digitalen Archivalien bemessen und wie diese auch erhalten werden kann.

Zusammenfassend lässt sich daher feststellen, dass es für digitale Archive gute Gründe gibt, bereits jetzt Ausschau nach einem einheitlichen Konzept zur digitalen Bestandserhaltung zu halten. Im Oktober 2011 hat die nestor-AG Digitale Bestandserhaltung nun einen *Leitfaden zur digitalen Bestandserhaltung* zur öffentlichen Kommentierung veröffentlicht, der erstmals ein Modell zur Integration der oben kurz angesprochenen fünf Konzepte anbietet. Der Leitfaden wurde von Archivaren, Bibliothekaren und Mitarbeitern anderer Einrichtungen verfasst. Durch die interdisziplinäre Erarbeitung stellt er einen Kompromiss zwischen den Ansichten seiner Mitglieder dar. Als Rahmenstandard ist er zugleich von seinem Anspruch her für verschiedene Umsetzungen offen. Damit stellt sich die Frage, wie er in den klassischen Archiven umgesetzt werden kann. Eine Antwort hierauf kann nur eine breitere Diskussion unter den klassischen Archiven geben. Auch besteht für diese noch die Möglichkeit, ihre Vorstellungen bereits frühzeitig in die Entwicklung des Konzepts einzubringen und ebenso rechtzeitig die Auswirkungen des Konzepts auf die bisherigen Standards und Prozesse zu prüfen.

Darüber hinaus gibt es zwei weitere Gründe für die klassischen Archive, sich jetzt ganz konkret mit dem Leitfaden zur digitalen Bestandserhaltung zu beschäftigen. Zum einen hilft die Beschäftigung mit einem übergreifenden Modell wie dem Leitfaden zur digitalen Bestandserhaltung jedem einzelnen Archiv, die eigenen Anforderungen zu konkretisieren. Zweitens sind viele Archive gerade erst dabei, tastend ihren Weg in die digitale Archivierung zu finden. Die hierfür notwendigen Aufwände können beträchtlich reduziert werden, wenn sich die klassischen Archive wenigstens in Grundzügen auf ein einheitliches Modell verständigen, wie sie die digitale Bestandserhaltung verstehen und umsetzen möchten. Die daraufhin entwickelten Vorstellungen und Softwaretools können dann auch von anderen Archiven eingesetzt werden. Es spart also auch Geld, wenn man sich rechtzeitig mit diesen konzeptionellen Fragen beschäftigt.

Der Leitfaden zur digitalen Bestandserhaltung beschreibt die Prozesse, die bei der Verarbeitung der gespeicherten Daten anstehen und schlägt zweitens vor, wie diese klassifiziert und auf eine systematische Art und Weise erhalten werden können. Es geht also um Verarbeitungs- und Erhaltungsprozesse.

Verarbeitungsprozesse

Zwischen den nackten unverarbeiteten Dateien, die auf einem Datenträger gespeichert sind und der Feststellung eines Archivars, diese oder jene Informationen sollten archiviert werden, liegen verschiedene Verarbeitungsprozesse, die im Leitfaden erstmals in einem Zusammenhang dargestellt werden. Dabei erfolgt die Verarbeitung zunächst durch eine Maschine (den Computer), dann durch den Menschen selbst. Diese Verarbeitung beginnt als Wahrnehmungsprozess. Mit der Darstellung der Stationen dieses Prozesses wird deutlich, welche Eingriffsmöglichkeiten und -notwendigkeiten bestehen. Folgende Stationen lassen sich unterscheiden:

- Die physischen Daten: Sie müssen vom Archiv auf jeden Fall erhalten werden, um die nachfolgenden Verarbeitungsschritte erst möglich zu machen.
- Die verarbeitende Maschine, also Hard- und Software des Computers. Dieser Bereich entwickelt sich aus Sicht der Archive mit geradezu naturgesetzlicher Zwangsläufigkeit. Entsprechend gering sind die Einwirkungsmöglichkeiten der Archive. Stattdessen müssen sie die von den Herstellern angebotenen Systeme als gegeben hinnehmen.
- Die Performance, also die Ausgabe der verarbeiteten Daten auf einem Monitor, durch einen Lautsprecher oder ein Eingabegerät für Computerspiele. Nun erst werden die Inhalte für einen Menschen über seine Sinne wahrnehmbar. Digitale Archivierung versucht, die Fähigkeit zu erhalten, ähnliche Performances auch in Zukunft reproduzieren zu können.
- Die wahrgenommenen Sinneseindrücke werden vom Rezipienten verarbeitet. Dieser integriert die Inhalte in sein bereits bestehendes Vorverständnis¹⁰ und kann sie dadurch verstehen. Es entsteht ein Informationsobjekt.

Welche Vorteile ergeben sich durch eine derartige kleinteilige Darstellungsweise?

- Sie definiert die Aufgabe der Archive näher. Egal ob das Archiv nun eine Migrations- oder eine Emulationsstrategie verfolgt: Es strebt danach, seine Fähigkeit zu erhalten, eine der ersten Performance ähnliche aktuelle Performance herstellen zu können. Die erste Performance ist in diesem Zusammenhang diejenige, die ein Archivar bei der archivischen Bewertung vorfand.
- Die Objekte an den beiden Enden des Prozesses, also Repräsentationen und Informationsobjekte, können klar voneinander unterschieden und archivisch angesprochen (bearbeitet) werden.
- Durch die Unterscheidung zwischen Performance und Informationsobjekt kann der unvermeidliche subjektive Anteil des Archivars genauer bestimmt werden.
- Die Unterscheidung zwischen Performance und Informationsobjekt ermöglicht es, auch die nicht sinnlich erfassbaren Elemente der Informationsobjekte konzeptionell zu greifen. Von einer Datenbank lassen sich beispielsweise in der Performance nur sehr geringe Teile sinnlich wahrnehmen. Erst durch die intellektuelle Durchdringung der mit

¹⁰ Damit hält die Hermeneutik, die bislang eher in Literaturwissenschaft, Philosophie und Theologie diskutiert wurde, Einzug in die archivfachlichen Diskurse. Vgl. Hans-Georg *Gadamer*: *Wahrheit und Methode*. Tübingen 1960 und Ernst *Fuchs*: *Hermeneutik*, Stuttgart/Bad Cannstatt 1954.

der Datenbank einhergehenden Möglichkeiten bekommt der Archivar eine Vorstellung davon, welches Objekt zur archivischen Bewertung und mithin zur Archivierung ansteht.

- Die Performance wird zu dem Ort, an dem die erhaltenswerten Eigenschaften vor und nach einer Migration gemessen und verglichen werden können. Hier müssen die signifikanten Eigenschaften benannt werden. Damit wird auch klar, dass ein Großteil der Tools, die derzeit Eigenschaften von digitalen Objekten erheben, sich nur auf die Datenebene bezieht. Entsprechend deutlich wird der noch ausstehende Forschungs- und Entwicklungsbedarf.
- Migrations- und Emulationsstrategie können in einem Modell dargestellt und damit in ihren Leistungen verglichen werden.

Der Prozess grenzt aber nicht nur ab, er integriert auch. Denn jede Änderung der Daten im Zuge der Migrationsstrategie oder der Software im Zuge der Emulationsstrategie führt zu einer anderen Performance. Dem Wahrnehmungsprozess schließen sich Erhaltungsprozesse an, die letztlich wiederum zu einer Änderung der Daten und/oder der Software führen.

Erhaltungsprozesse

Einzelne digitale Archivalien können wahrscheinlich ohne größere konzeptionelle Vorstellungen eine Zeit lang erhalten werden. Probleme entstehen erst, wenn die verschiedenen Erhaltungsbestrebungen über einen sehr langen Zeitraum konsistent gehalten werden müssen, wenn also der Nachfolger des heutigen Archivars im Objekt immer noch das sehen sollte, das der zuerst mit der Erhaltung betraute Archivar in ihm sah. Dieser Nachfolger benötigt dafür Hinweise, wie denn die ursprüngliche Performance aussah. Probleme können aber auch früher auftreten. Wenn nämlich nicht nur eine Handvoll digitaler Objekte, sondern Hunderttausende und Millionen dieser Objekte erhalten werden sollen. Nun soll der Computer wissen, wie er den heutigen Archivar zu unterstützen hat. Beide, der Nachfolger und auch der als Assistent auftretende Computer, benötigen eindeutige Angaben, um ihre Arbeit vornehmen zu können.

Um effizient vorzugehen und eine von Anfang an einheitliche Bestandserhaltung zu gewährleisten, sollten die Hinweise bereits während der Bewertung erfasst werden. Dabei ist es sowohl zu ihrer leichteren Erfassung als auch für eine später möglichst effiziente Verarbeitung naheliegend, die Hinweise in strukturierter Form zu erfassen. Es bedarf also einer Klassifikation dieser Hinweise.

Von welchen Ausgangspunkten können die Hinweise nun abgeleitet werden? Da sind zunächst natürlich die Objektarten selbst. Mit ihren Eigenschaften bestimmen sie den Rahmen, in dem künftige Nutzer im Archiv Informationen finden und verarbeiten können. Diese Objekte müssen, wie oben ausgeführt, durch eine Kette aktiver Entscheidungen des Archivars erhalten werden. Zumeist gibt es verschiedene Möglichkeiten der Erhaltung. Es muss also eine Entscheidung getroffen werden, die sich nicht eindeutig aus den Eigenschaften des Objekts ableiten lässt. Folglich müssen noch andere, außerhalb der Objekte liegende Kriterien hinzugezogen werden, um die Hinweise zur Erhaltung definieren zu können. Limitierende Faktoren sind hier zunächst die finanziellen Möglichkeiten des Archivs oder der interessierten Fachgemeinde. Ebenso wichtig sind die Erwartungen, die die künftigen Nutzer voraussichtlich an die Archivalien richten

werden. Es gilt also, diese Hinweise durch Berücksichtigung der von den Archivalien mitgebrachten Eigenschaften, der finanziellen Möglichkeiten und der anzunehmenden Erwartungen künftiger Nutzer strukturiert zu erfassen.

Der Leitfaden zur digitalen Bestandserhaltung sieht zunächst eine grobe Klassifikation digitaler Unterlagen vor. In diesem ersten Schritt werden anhand der in der Performance aufscheinenden Eigenschaften erste Gruppen gebildet. Diese Gruppen können mit dem Leitfaden zur digitalen Bestandserhaltung als *Informationstypen* bezeichnet werden. Beispielsweise könnten Akten, schwach strukturierte Daten, Fachverfahren, Webseiten und Audio- bzw. Videoobjekte unterschieden werden. Denkbar wäre es aber auch, die Fachverfahren in einem weiteren Sinne in Statistikverfahren einerseits und Fachverfahren im engeren Sinne zu unterscheiden. Auch könnten aus letzteren noch die geographischen Informationen (landläufig Geodaten) ausgegliedert werden, die wiederum in *Geodaten der Vermessungsverwaltung* und *Geodaten aus Fachverwaltungen* oder in *Geobasisdaten* und *Geofachdaten* aufgeteilt werden könnten. Diese Aufteilungen geben eine grobe erste Einteilung vor. Es wäre wünschenswert, wenn hierüber eine breitere Diskussion zwischen den Archiven einsetzen könnte. Zugleich sind diese Einteilungen auch abhängig von den von den einzelnen Archiven verfolgten Zielen. Sie können daher per definitionem nicht ein für alle Mal für alle Archive verbindlich festgelegt werden.

Die Festlegung der Informationstypen macht das Feld der zu erhaltenden Objekte überschaubarer. Sie genügt aber nicht, um die notwendigen Hinweise geben zu können. Am deutlichsten wird dies bei den Eigenschaften. Archive können aus verschiedenen Gründen nicht alle Eigenschaften übernommener Objekte erhalten. Zunächst ist dies in aller Regel schon aus finanziellen Gründen nicht möglich. Hinzu kommt noch die fehlende Dokumentation heute verbreiteter Dateiformate, wie dies beispielsweise sehr schön bei den Microsoft Office Formaten beobachtet werden kann. Selbst wenn also unbegrenzte Mittel zur Verfügung stünden, könnte daher nicht alles erhalten werden, da nicht alle Eigenschaften der Dateiformate dokumentiert sind. Es muss also ausgewählt werden. Die Teilmenge der zu erhaltenden Eigenschaften sind die signifikanten Eigenschaften. Sie müssen in einem zweiten Schritt ermittelt werden.

Die signifikanten Eigenschaften können nur durch einen bewussten Auswahlprozess festgelegt werden. Jede Entscheidung für die Eigenschaften A, B und C ist häufig eine Entscheidung gegen die Eigenschaften D, E und F. Hier kommen nun im verstärkten Maße die Faktoren zum Tragen, die außerhalb der einzelnen Archivalien liegen. Was ist finanzierbar? Was ist ganz praktisch umsetzbar? Was ist vor allem im Interesse der künftigen Nutzer?

Gerade die Interessen der künftigen Nutzer lassen sich nicht ganz einfach benennen. Wer kann schon in die Zukunft sehen? Manche Annahmen sind nicht möglich: Wir wissen nicht, ob in Zukunft ein Familienforscher Vorfahren mit dem Namen Schmidt, Müller oder Mayer sucht. Andere Annahmen können dagegen mit größerer Wahrscheinlichkeit getroffen werden. Die nestor AG digitale Bestandserhaltung geht davon aus, dass alle Nutzungen digitaler Archivalien sich letztlich auf eine von vier Grundformen zurückführen lassen:

- Betrachtung oder Wahrnehmung des Gesamtobjekts (z.B. Lesen eines Romans oder Ansehen eines Films)
- Suche nach einer spezifischen Information (z.B. über die Volltextsuche)
- Einlesen des Archivals in eine externe Software zur Weiterverarbeitung
- Abspielen des Archivals (nur bei Software, z.B. bei Computerspielen)

Bei einem Archiv können nun sehr unterschiedliche Informationsobjekte vom selben Informationstyp ankommen. Zumeist dürften für die Informationsobjekte eines Informationstyps unterschiedliche Nutzungsszenarien vorstellbar sein. Für die Informationsobjekte des Typs Text ist es z.B. denkbar, dass manche vor allem als ganzes Objekt genutzt werden sollen (z.B. ein typographisch aufwändig gestaltetes Gedicht), während andere außerdem noch auf einzelne Zeichenfolgen hin durchsuchbar sein sollten. Es ist daher notwendig, die Informationstypen weiter zu unterteilen. Dabei bietet es sich an, alle Informationsobjekte mit denselben signifikanten Eigenschaften in eine gemeinsame Gruppe zu bringen, da sie aufgrund ihrer signifikanten Eigenschaften auf dieselbe Art und Weise erhalten werden können. Diese Gruppe nennt der Leitfaden zur digitalen Bestandserhaltung *Erhaltungsgruppe*.

Die signifikanten Eigenschaften können als Bedingungen verstanden werden, die von künftigen Dateiformaten erfüllt werden müssen. Um noch einmal das soeben genannte Beispiel des Informationstyps Text aufzugreifen, könnten die Objekte, die nur als Gesamtobjekt genutzt werden (das Gedicht), auch als Bild archiviert werden, während sehr umfangreiche Strafverfahrensakten vermutlich nur unter Zuhilfenahme der Volltextsuche sinnvoll genutzt werden können. Diese Objekte stellen daher an die künftigen Dateiformate andere Anforderungen als das digitale Gedicht.

	Reihenfolge der Zeichen	Verarbeitbarkeit	Auszeichnung	Adressierbarkeit	Textebenen	Mögliche Formate (2012)	Mögliche Formate (2100)	Mögliche Formate (2200)
Text lesen	x					TIFF
Information suchen	x	x				PDF/A
Text weiterverarbeiten	x	x	x	x	x	XML

Tabelle 1: Mögliche Erhaltungsgruppen für Text, Annahme: Eine Erhaltungsgruppe pro Nutzungsziel. Links die Nutzungsziele, oben die möglichen signifikanten Eigenschaften bzw. rechts davon die daraus ableitbaren Dateiformate.

	Reihenfolge der Zeichen	Verarbeitbarkeit	Auszeichnung	Adressierbarkeit	Textebenen	Mögliche Formate (2012)	Mögliche Formate (2100)	Mögliche Formate (2200)
Text lesen	x					XML
Information suchen	x	x						
Text weiterverarbeiten	x	x	x	x	x			

Tabelle 2: Mögliche Erhaltungsgruppen für Text, Annahme: Eine Erhaltungsgruppe für drei Nutzungsziele. Links die Nutzungsziele, oben die möglichen signifikanten Eigenschaften bzw. rechts davon die daraus ableitbaren Dateiformate.

Die signifikanten Eigenschaften werden mit der Bewertung festgelegt. Sie gelten für die gesamte Lebenszeit des Archivals. Wenn die Dokumente zu einem Strafverfahren in einer zeichenbasierten (und damit im Volltext durchsuchbaren) Form übernommen wurden, sollte diese Durchsuchbarkeit auch nach der zweiten, dritten und vierten Dateiformatsmigration noch gegeben sein. Ähnliches könnte von der Verteilung der Farben und Formen in einem Bild, der Zahl der Datensätze in einer Tabelle und der Länge einer Audiodatei gesagt werden. Nach jeder Migration müssen diese Werte neu erhoben und mit den übernommenen Werten verglichen werden. Wenn sie voneinander abweichen, ist das gewählte Dateiformat nicht geeignet, um die definierten signifikanten Eigenschaften zu unterstützen. Die signifikanten Eigenschaften stellen damit den Maßstab, vor dem sich alle künftigen Dateiformate bewähren müssen. Dabei ist es von Vorteil, dass dieser Maßstab nur einmal (ganz am Anfang) festgelegt werden muss.

Nach der Festlegung der signifikanten Eigenschaften sollte das Archiv noch bestimmen, welche dieser Eigenschaften absolut und welche nur zu bestimmten (definierten) Graden erfüllt werden müssen. Die Zahl der Datensätze einer Tabelle muss sicherlich vollständig erhalten werden. Dagegen sind im Farbraum eines Bildes durchaus Abweichungen denkbar.

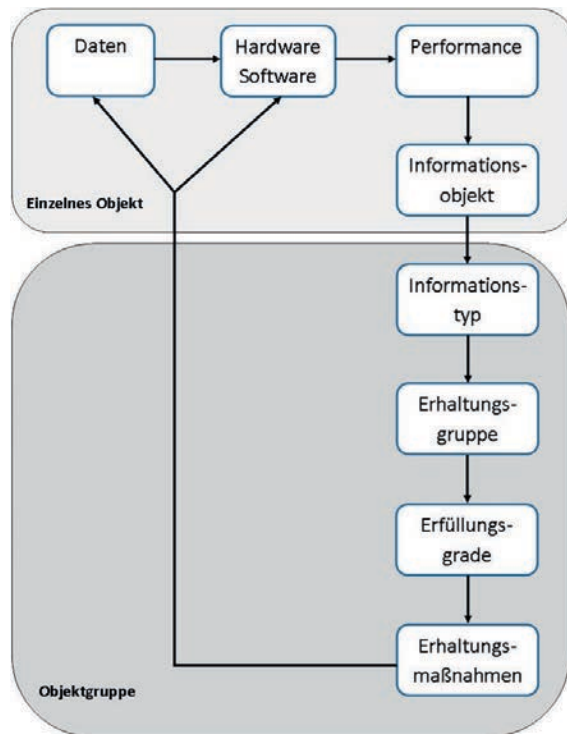


Abb. 1: Prozessfolge bei der digitalen Bestandserhaltung: Das einzelne Objekt wird wahrgenommen, für die Objektgruppe werden Festlegungen getroffen.

Neben den signifikanten Eigenschaften stellt der Leitfaden zur digitalen Bestandserhaltung noch weitere Bereiche vor, die bei der Erhaltung digitaler Unterlagen wichtig sind. Er beschreibt die von OAIIS erstmals definierte Representation Information und grenzt sie von den signifikanten Eigenschaften ab¹¹. Er beschreibt konkret, wie sich sowohl die Migrationsstrategie als auch die Emulationsstrategie mit dem definierten Instrumentarium umsetzen lässt. Und er enthält Hinweise zu den von OAIIS vorgesehenen Aufgaben zur Beobachtung der Nutzergruppen und der technischen Entwicklungen.

Authentizität

Die signifikanten Eigenschaften ermöglichen es, bei verschiedenen Prozessen der digitalen Archivierung Entscheidungen zu priorisieren. Dateiformate, die alle Eigenschaften des Objekts unterstützen, sind jenen vorzuziehen, die nur die signifikanten Eigenschaften unterstützen. Es ist wichtiger, die signifikanten Eigenschaften eines übernommenen Objekts zu sichern, als die übrigen Eigenschaften dieses Objekts. Solche Dateiformate, die alle signifikanten Eigenschaften unterstützen, sind jenen vorzuziehen, die nur manche unterstützen et cetera.

Die signifikanten Eigenschaften stellen also die Messlatte dar, an der sich die Erhaltungsmaßnahmen bewähren müssen. Es ist daher nur konsequent, an ihnen auch die Anforderungen im Bereich der Authentizität festzumachen. Die Werte, die bei der Überprüfung der signifikanten Eigenschaften eines Informationsobjekts erhoben wurden, werden vom digitalen Archiv zusammen mit den anderen Metadaten dieses Objekts gespeichert. Im Normalfall dürfte dies bei fast allen digitalen Archiven das Dateiformat XML sein. Diese in XML gespeicherten Informationen können relativ leicht in ein künftiges Nachfolgeformat übertragen werden. Wir wissen daher noch in fernerer Zukunft, dass auch in der dritten Repräsentation eines Audioobjekts dessen Länge 7 Minuten betrug. Wir wissen dies, obwohl die Primärdaten dieser dritten Repräsentation von den dann gängigen Computersystemen schon längst nicht mehr gelesen werden können. Damit geben die zu den signifikanten Eigenschaften erhobenen Daten einen lückenlosen Aufschluss darüber, wie dieses Objekt erhalten wurde. Abgesehen von dieser Nachweisfunktion definieren die signifikanten Eigenschaften aber auch den Rahmen dessen, was das digitale Archiv von dem Informationsobjekt auf jeden Fall erhalten wollte. Die signifikanten Eigenschaften stellen daher auch für den Nachweis der Authentizität die zentrale Messlatte.

¹¹ Im Umfeld des PLANETS Projekts wurde diskutiert, inwieweit sich diese beiden Metadaten Gruppen voneinander unterscheiden lassen. Representation Information umfasst die für eine konkrete Nutzung erforderlichen Metadaten. Signifikante Eigenschaften stellen dagegen die Norm, gegenüber der sich alle künftigen Performances bewähren müssen.

Resümee

Der Leitfaden zur digitalen Bestandserhaltung entwirft ein konsistentes Modell zur digitalen Archivierung. Die geschilderten Prozesse sollen eine Verarbeitung auch sehr großer Mengen digitaler Archivalien ermöglichen. Mit der Bewertung setzen sie bereits sehr früh im Lebenszyklus digitaler Objekte ein. Es liegt daher im Interesse der Archive, sich mit dem Leitfaden auseinanderzusetzen und eine eigenständige Position oder auch eine konkrete Umsetzung zu entwickeln.

nestor-Workshop
„Brauchen wir Koordinierungsstellen
für die digitale Archivierung?“

Warum ist Kooperation bei der digitalen Archivierung unumgänglich?

Von CHRISTIAN KEITEL

Strategische Vorträge zur digitalen Archivierung enthalten in aller Regel wenigstens zwei Grundbestandteile. Zum einen wird das anscheinend unverzichtbare Schaubild zu den Funktionsbereichen digitaler Archive aus dem OAIS-Standard gezeigt. Zweitens wird allerorten versichert, dass digitale Archivierung nur in Kooperation gelingen könne. Weshalb dies so sein soll, wird dann für gewöhnlich nicht ausgeführt. Dieser Frage soll hier etwas eingehender nachgegangen werden. Die Untersuchung der Schaubildverwendungsfrequenz muss anderen Forschungsvorhaben vorbehalten bleiben.

Aufrufe zur Kooperation

Bereits 1991 wurde in der Europäischen Gemeinschaft eine Expertengruppe beauftragt, Fragen der digitalen Archivierung im Allgemeinen und der Kooperation im Speziellen nachzugehen. 1994 konnte diese Gruppe den Bericht *Archives in the European Union* vorlegen, in dem ein erhöhter Koordinations- und Kooperationsbedarf bei der digitalen Archivierung konstatiert wurde. Die Europäische Union förderte daher mit dem DLM-Forum eine Reihe grundlegender internationaler Tagungen. Es ist bezeichnend, dass gleich bei der ersten Tagung, die vom 18. bis 20. Dezember 1996 in Brüssel stattfand, sich drei von 12 Themen bei den Parallelveranstaltungen mit Fragen der Zusammenarbeit und der Kooperation beschäftigten.¹ Einige Jahre später förderte die Europäische Union drei- oder vierjährige Projekte wie Erpanet,² Planets³ und Digital Preser-

¹ Grenzüberschreitende Zusammenarbeit: Bedingungen und Durchführung; Kooperation europaweit und auf interdisziplinärer Ebene – Handlungsmöglichkeiten; Kooperation im Rahmen der Europäischen Union und die globale Informationsgesellschaft. Vorträge und Ergebnisse des DLM-Forums über elektronische Aufzeichnungen, INSAR Beilage II (1997). Die Eröffnungs- und Schlussansprachen werden in der Begleitpublikation programmatisch eingeleitet: „Das Forum richtet sein Augenmerk auf die Möglichkeiten einer verbesserten Zusammenarbeit zwischen den Mitgliedsstaaten und auf der Ebene der Europäischen Union im Hinblick auf Verwaltung, Lagerung, Konservierung und Retrieval elektronischer Daten.“

² Im Mission Statement von Erpanet wird festgehalten: „ERPANET will not directly carry out new research to develop such tools, but it will create a coherent platform for proactive co-operation, collaboration, exchange and dissemination of research results and experience in the preservation of digital objects.“ <http://www.erpanet.org/about.php>. Alle Webseiten wurden am 17.8.2012 überprüft.

³ <http://www.planets-project.eu/>. Ziel von PLANETS Projekts war es, praktische Hilfsmittel und Software zur digitalen Archivierung zu entwickeln.

vation Europe,⁴ die sich alle dem Gedanken der Kooperation verpflichtet fühlten. 2010 lud die Kommission dann Vertreter der drei Initiativen ein, die Kriterienkataloge für vertrauenswürdige digitale Archive verfasst hatten.⁵ Ergebnis war ein Memorandum of Understanding zwischen diesen drei Initiativen.⁶ Seit 1991 hat die Europäische Union daher immer wieder Kooperationen gefördert und auch eingefordert.

Aber auch außerhalb Europas fanden die Überlegungen zur Kooperation Wiederhall. Die von der Vollversammlung der UNESCO 2003 verabschiedete *Charter on the Preservation of Digital Heritage* enthält ebenfalls die Verpflichtung, Kooperationen zu fördern.⁷ Darüber hinaus wird die Einrichtung spezieller Stellen überlegt, die in den einzelnen Mitgliedsstaaten die Zusammenarbeit in der digitalen Archivierung fördern sollten:

Member States may wish to designate one or more agencies to take coordinating responsibility for the preservation of the digital heritage, and to make available necessary resources. The sharing of tasks and responsibilities may be based on existing roles and expertise.

Eine derartige Stelle war bereits ein Jahr zuvor in Deutschland im Rahmen des nestor-Projekts eingerichtet worden. Die zentrale Idee von nestor ist es, die verschiedenen mit Fragen der digitalen Archivierung befassten Fachleute zusammenzubringen und Informationen zum Thema frei auszutauschen. Obwohl die Projektförderung 2008 auslief, konnte nestor – nun als Kooperationsprojekt – seinen Anspruch als zentrale Informationsplattform zu dem Thema in Deutschland auch in den nächsten Jahren weiter ausbauen.⁸

Vergleichbare Gedanken können auch auf der anderen Erdhalbkugel nachverfolgt werden. Die 2004 gegründete Australasian Digital Recordkeeping Initiative (ADRI) geht über die digitale Archivierung im engeren Sinne hinaus. Aber auch hier wird ein gemeinsames Vorgehen angestrebt. 2006 brachte das australische Nationalarchiv den Kooperationsgedanken auf den Punkt:

*Digital archiving is not cheap, and we do not have the luxury of time to allow each domain to develop the levels of expertise required, when others have already established them within the collective group. We cannot afford to 'reinvent the wheel'. So collaborative work is the way of the future ...*⁹

Allein die Produkte, die vom nestor-Projekt oder von ADRI vorgelegt werden konnten, belegen, dass Kooperation in der Tat ein Schlüssel zur digitalen Archivierung ist. Sie belegen aber nicht, weshalb dies so ist. Auch bleibt die Frage offen, ob Kooperation auch für die deutschen (klassischen) Archive den zentralen Ansatz zur Bewältigung der digitalen Archivierung darstellen kann.

⁴ <http://www.digitalpreservationeurope.eu/>. „DPE addresses the need to improve coordination, cooperation and consistency in current activities to secure effective preservation of digital materials.”

⁵ DIN 31644, ISO 16363 und Data Seal of Approval.

⁶ http://datasealofapproval.org/sites/default/files/20100709_020_signed%20MoU%20to%20create%20a%20European%20Framework%20for%20Audit%20and%20Certification%20of%20Digital%20Repositories.pdf.

⁷ http://portal.unesco.org/en/ev.php-URL_ID=17721&URL_DO=DO_TOPIC&URL_SECTION=201.html.

⁸ <http://www.langzeitarchivierung.de>.

⁹ National Archives of Australia, Digital archiving in the 21st century. Archives Domain discussion paper, September 2006, <http://www.caara.org.au/wp-content/uploads/2010/03/DigitalArchiving21C.pdf>.

Die am Eingang gestellte Frage soll nun in diesem Sinne konkretisiert werden. Dabei soll zunächst nach dem Anlass zur digitalen Archivierung, dann nach den zu überwindenden Problemen und schließlich nach möglichen Lösungsmöglichkeiten gefragt werden. Erst danach wird es möglich sein, die Notwendigkeit zu Kooperationen in die fachlichen Notwendigkeiten einzuordnen.

Digitale Unterlagen

Gibt es überhaupt Anlass, die Frage nach den Kooperationen zu stellen? Die Legitimität dieser Frage ergibt sich aus verschiedenen Feststellungen:

- Lochkarten wurden in Deutschland seit 1910 eingesetzt. Unter den ersten Anwendern waren die Statistischen Ämter in Stuttgart, Karlsruhe und Dresden. Damit können wir heute auf eine über einhundertjährige Tradition der maschinengestützten Datenverarbeitung zurückblicken.
- Computer werden in den staatlichen Verwaltungen seit den 50er Jahren eingesetzt. So wurde 1956 in der Flurbereinigungsdirektion in Bamberg ein Zuse Z 11 eingesetzt, ein Jahr später wurde vom Landesamt für Flurbereinigung und Siedlung (LFS) Baden-Württemberg eine Geodätische Rechenstelle eingerichtet und der Rechenstanzer IBM 604 angemietet.¹⁰
- Fachverfahren sind seit den 1970er Jahren verbreitet. 1974 erklärte ein Vertreter der baden-württembergischen Landesverwaltung, dass seinerzeit nahezu kein Verwaltungszweig mehr ohne Computer auskomme¹¹.
- 1980 begann der Siegeszug des Personal Computers und damit die Möglichkeit für jeden einzelnen Sachbearbeiter, seine Unterlagen selbständig elektronisch abzulegen.
- 1995 setzte sich das Internet in Deutschland durch. Nicht wenige anbieterpflichtige Stellen legen einen Teil ihrer Unterlagen nur noch im Internet ab. Da sich Internetseiten verhältnismäßig schnell ändern, müssen die Archive hier bereits nach kurzen Zeiträumen die Seite sichern, sofern sie an den enthaltenen Informationen Interesse haben.

Es lässt sich daher feststellen, dass es in sehr vielen Bereichen der anbieterpflichtigen Stellen archivreife digitale Unterlagen gibt. Aber auch in den Archiven liegen seit Jahrzehnten unzählige Fotos, Audio- und Videounterlagen, die nur durch eine rechtzeitige Digitalisierung erhalten werden können. Da der Verfall der originalen Datenträger nicht aufgehalten werden kann, werden in wenigen Jahren nur noch die Digitalisate verwendbar sein, wenn sie denn rechtzeitig erstellt worden sind. Sie sind daher den genuin digitalen Unterlagen gleichzustellen. Sowohl bei den genuin digitalen Unterlagen als auch bei den Archivalien, die nur über Digitalisierung gesichert werden können, drängt also die Zeit, wenn größere Überlieferungslücken vermieden werden sollen.

¹⁰ 50 Jahre Datenverarbeitung in der Flurbereinigungsverwaltung Baden-Württemberg 1957–2007. Dokumentation (Schriftenreihe des Landesamts für Flurneuordnung Heft 17). Kornwestheim [ca. 2007].

¹¹ StAL EL 17 I Bü 162a, Tagungsprotokoll S. 16.

Digitale Archive

Diese Einschätzung der Lage scheint inzwischen von einer ganzen Reihe von Archiven geteilt zu werden. Ablesbar ist das sprunghaft gestiegene Interesse an digitaler Archivierung schon an den Teilnehmerzahlen der nestor-Veranstaltungen zur digitalen Archivierung. Bis etwa 2008 konnten auf allgemeinen Informationsveranstaltungen zum Thema im Schnitt etwa 40 bis 60 Teilnehmer begrüßt werden.¹² Zum ersten nestor-Praktikertag im Hauptstaatsarchiv Stuttgart am 30. November 2010 kamen dann über 150 Teilnehmer, obwohl kein einziger Vortrag besondere Neuigkeiten versprach. Auch die beiden nächsten nestor-Praktikertage 2011 und 2012 wurden von über 100 Interessierten besucht. Während der Verfasser dieses Artikels bis etwa 2008/2009 im Durchschnitt einmal in der Woche um Rat in Fragen der digitalen Archivierung gefragt wurde, kann er nun pro Arbeitstag mit mindestens einem Anruf rechnen.

Dennoch verläuft der Aufbau digitaler Archive nur schleppend. Hierfür können mindestens drei Gründe angeführt werden: Kosten, zwischenarchivische Abstimmungsbedarfe und innerarchivische Änderungsbedarfe.

Zu den Kosten

Allein die Quantität der zu archivierenden digitalen Objekte wird die allermeisten digitalen Archive binnen kürzester Zeit vor erhebliche Probleme stellen. Einige wenige digitale Archivalien können noch problemlos ohne Archivierungssystem und mit herkömmlichen PCs und Datenträgern verwaltet werden. Mehrere Faktoren lassen aber nun die Menge der zu verwaltenden Daten und Dateien schnell ansteigen. Zunächst bestehen digitale Objekte häufig aus mehreren Dateien. Dann müssen sie durch eine größere Zahl von Metadaten zutreffend beschrieben werden. Die Zuordnung von Metadaten und Primärdaten ist selbstverständlich dauerhaft zu gewährleisten. Schließlich findet die vielzitierte Explosion der Menge der digitalen Daten auch bei den anbieterpflichtigen Stellen statt. Diese verschiedenen Effekte führen dazu, dass die digitalen Archive aller Voraussicht nach in sehr kurzer Zeit Millionen digitaler Dateien so verwalten müssen, dass a) nichts verloren geht, b) die zugehörigen Metadaten mit ihrem Bezug zu den Primärdaten erhalten werden, c) die mit der Verwaltung einhergehenden Prozesse beschrieben werden und d) nur die vorgesehenen Archivmitarbeiterinnen und Archivmitarbeiter die Primärdaten und Metadaten archivisch erschließen, aufbereiten und einsehen können. Die schiere Menge erfordert also eingespielte Prozesse und Hilfsmittel.

Die Umsetzung der digitalen Archivierung wird ferner erschwert durch die Komplexität, die tendenziell sowohl durch die Zahl der zur Übernahme anstehenden Objekte als auch durch die Zahl der umzusetzenden Standards ansteigt. Ein Beispiel aus dem Bereich der Standards mag dies verdeutlichen. Ingest ist im Lebenslauf digitaler Objekte diejenige Phase, in der die digitalen Objekte übernommen und aufbereitet werden. In dem maßgeblichen Standard Open Archival

¹² Zur Situation 2008 s. Christian Keitel: Elektronische Archivierung in Deutschland. Eine Bestandsaufnahme. In: Für die Zukunft sichern! Bestandserhaltung analoger und digitaler Unterlagen, 78. Deutscher Archivtag 2008 in Erfurt (Tagungsdokumentation zum Deutschen Archivtag Bd. 13), Fulda 2009. S. 115–128.

Information System (OAIS = ISO 14721) ist Ingest eine von insgesamt sechs zentralen Funktionseinheiten bei der digitalen Archivierung.¹³ Diese Funktionseinheit wird gegen Ende des Standards unterteilt in insgesamt fünf Unterfunktionen. Mit PAIMAS hat nun die Gruppe der OAIS entwickelnden Luft- und Raumfahrtorganisationen einen Substandard zu OAIS entwickelt, der außerdem noch die vorangehenden Schritte der Bewertung umfasst.¹⁴ PAIMAS besteht aus etwa 30 Phasen, die allerdings in 88 Teilschritte unterteilt werden. Im Laufe dieser Differenzierung werden zwar die einzelnen Aufgaben genauer beschrieben, sie verbleiben aber auf einem sehr abstrakten Niveau. Welches Archiv ist nun in der Lage, ein Sechstel aller Funktionsbereiche der digitalen Archivierung mit insgesamt 88 Schritten abzubilden?¹⁵

Sowohl die Menge der zu archivierenden Dateien und Daten als auch die Komplexität von Daten und Standards legen es nahe, möglichst viele Arbeiten von Computern verrichten zu lassen. Eine geeignete Antwort stellen hier die verschiedenen Systeme zur Archivierung digitaler Unterlagen. Nur stellt sich dadurch die Kostenfrage eben auf einem anderen Feld. Sowohl die Eigenentwicklung eines Archivierungssystems als auch deren Kauf bei einem kommerziellen Anbieter erfordern den Einsatz von deutlich sechsstelligen Beträgen. Die Menge der Dateien und Daten, die Komplexität der Standards und auch die Archivierungssysteme selbst sind daher Faktoren, die die Kosten deutlich in die Höhe treiben.

Zwischenarchivische Abstimmungsbedarfe

Digitale Archive sind in einem noch höheren Maße als ihre konventionellen Vorgänger darauf angewiesen, ihre Anforderungen untereinander abzustimmen. Dies gilt zunächst für die ablieferungspflichtigen Stellen. Viele digitale Unterlagen entstehen in speziellen IT-Systemen. Genannt werden können hier beispielsweise Dokumentenmanagementsysteme, Vorgangsbearbeitungssysteme oder ganz allgemein Fachverfahren. Ein Zugriff auf diese Unterlagen erfolgt also zunächst über das System. Da das System aber nicht archiviert werden kann, müssen die Unterlagen aus dem System exportiert werden, es bedarf also einer Exportschnittstelle. Diese Systeme werden in aller Regel in zahlreichen anbietungspflichtigen Stellen eingesetzt, weshalb zumeist pro System ganz unterschiedliche Archive betroffen sind. Die Entwicklung derartiger Schnittstellen ist aufwändig. Es kann daher nicht sein, dass jedes Archiv eine eigene Schnittstelle von *ihrer* anbietungspflichtigen Stelle einfordert. Stattdessen werden die Stellen, die ein spezielles IT-Verfahren einsetzen, darauf Wert legen, dass sie zusammen nur eine Schnittstelle finanzieren müssen. Zwar entstehen auf Archivseite hier keine weiteren Kosten (zumindest sollte es so sein). Es ist aber notwendig, sich mit den anderen Archiven über eine gemeinsame Schnittstellendefinition auszutauschen. Vergleichbare Abstimmungen waren bei konventionellen Unterlagen nicht erforderlich.

¹³ Vgl. Referenzmodell für ein Offenes Archiv-Informationssystem. Deutsche Übersetzung, bearbeitet von der der nestor-Arbeitsgruppe OAIS-Übersetzung/ Terminologie, Frankfurt/Main 2012.

¹⁴ ISO 20652: Space Data and Information Transfer Systems – Producer-Archive Interface – Methodology Abstract Standard.

¹⁵ Christian Keitel: Ways to Deal with Complexity. Beitrag zur iPRES-Konferenz an der British Library, 2008, http://www.bl.uk/ipres2008/presentations_day2/45_Keitel.pdf.

Aber auch auf Seiten der Nutzer ist es notwendig, dass sich die Archive besser abstimmen. Zunächst haben die deutschen Archive noch lange keine Übereinkunft erzielt, wie sich die verschiedenen Erscheinungsformen digitaler Archivalien so abbilden lassen, dass sie den Bedürfnissen ihrer Benutzer Genüge tun. Es gibt beispielsweise weder Konsens darüber, das Repräsentationenmodell generell anzuwenden¹⁶ noch über dessen unterschiedliche Ausgestaltungsmöglichkeiten. Solange aber jedes Archiv seine digitalen Archivalien anders als die Kollegen vom Nachbararchiv nachweist, ist nicht mit einer verstärkten Nachfrage durch die Nutzer zu rechnen. Und wenn die Nutzer dann in die Archive kommen sollten, dürften sie dort in jedem Archiv wieder ein anderes Nutzungspaket antreffen. Das bedeutet, die Nutzer müssen sich in jedem einzelnen Fall neu in die spezielle Logik der Nutzungspakete des Archive X, Y oder Z einarbeiten. Im Kehrschluss erscheint es daher auch bei den Nutzungspaketen dringend erforderlich, dass sich die Archive auf einheitliche Standards verständigen, um archivübergreifende Nutzungsvorhaben nicht unnötig zu erschweren. Sowohl die abgebenden Stellen als auch die Nutzer haben daher ein erhebliches Interesse an einem einheitlichen Auftreten der Archive. Die sich daraus ergebenden Abstimmungsbedarfe sollten nicht unterschätzt werden.

Innerarchivische Änderungsbedarfe

An allererster Stelle sind aber die Archive durch den Paradigmenwechsel zur digitalen Archivierung herausgefordert, sowohl ihre Grundsätze als auch ihre Prozesse zu überdenken. Handelt es sich überhaupt um so große Änderungen, dass zu Recht von Paradigmenwechsel gesprochen werden kann? Auf einer abstrakten Ebene kann darauf hingewiesen werden, dass die Archive bisher Datenträger archiviert haben. Jeder einzelne übernommene Datenträger hatte die Aura des Originals. Bei digitalen Unterlagen lassen sich Datenträger nicht mehr dauerhaft erhalten. Es geht daher darum, die übernommene Information, also etwas Nicht-Stoffliches, zu erhalten durch alle Änderungen der Datenträger und Dateiformate, Hard- und Software hinaus.¹⁷ Diese auf abstrakter Ebene vielleicht noch nicht sehr bedrohliche Aussicht ändert sich sehr schnell, wenn die Änderungsbedarfe konkret auf die einzelnen Phasen im Lebenszyklus der digitalen Unterlagen bezogen aufgeführt werden. Hier einige Beispiele:

In der Bewertung digitaler Unterlagen wird es auch in Zukunft darum gehen, ob die Unterlagen archivwürdig sind. Außerdem muss aber überlegt werden, in welcher Form sie für künftige Nutzer am ehesten von Interesse sind. Sollte bei einem graphisch aufwändig gestalteten künstlerischen Text eher die Möglichkeit zur Volltextsuche oder das Erscheinungsbild erhalten werden? Daran anknüpfend ist zu fragen, welche signifikanten Eigenschaften durch alle Migrationen

¹⁶ Christian Keitel: Das Repräsentationenmodell des Landesarchivs Baden-Württemberg. In: Neue Entwicklungen und Erfahrungen im Bereich der digitalen Archivierung. Von der Behördenberatung zum Digitalen Archiv. 14. Tagung des Arbeitskreises „Archivierung von Unterlagen aus digitalen Systemen“. Hg von Susanne Wolf. München 2010, S. 69–82. Jetzt auch unter http://www.staatsarchiv.sg.ch/home/auds/14/_jcr_content/Par/downloadlist/DownloadListPar/download_8.ocFile/Text%20Keitel.pdf.

¹⁷ Vgl. die Definition von „langfristig“ in Referenzmodell für ein Offenes Archiv-Informations-System, wie Anm. 13, S. 12.

hinweg auf jeden Fall erhalten werden sollten. Nutzungsziele und signifikante Eigenschaften bringen zwei Fragestellungen mit sich, die sich in dieser Form bislang bei konventionellen Unterlagen nicht gestellt haben.

Die Archive können digitale Unterlagen wie bislang ihre konventionellen Vorgänger auf tragbaren Datenträgern übernehmen. Handelt es sich um wiederbeschreibbare Datenträger, entstehen neue Anforderungen an die spätere Glaubwürdigkeit dieser Unterlagen. Erfolgt die Übernahme über eine gesicherte Netzverbindung, muss gegen das Ausspähen vertraulicher Informationen Vorsorge getroffen werden.

An die Stelle der Entmetallisierung und des Einbettens in säurefreie Pallien und Archivboxen treten nun Migrationen in archivtaugliche Dateiformate und die Anreicherung der Metadaten um weitere Informationen, die für die Bestandserhaltung und spätere Nutzung von Belang sind.

Die Archivierung digital gespeicherter Informationen bedeutet, dass die physischen Grundlagen (Dateien, Dateiformate, Computer, Software) ständig ausgetauscht werden müssen, während die Information sich möglichst wenig verändern sollte. Aus der bislang anzustrebenden passiven Bestandserhaltung wird so eine aktive Bestandserhaltung. Immer wieder müssen zum Erhalt Entscheidungen getroffen werden.

Sogar bei der Nutzung ändert sich vieles. Durch die Unterscheidung von Archivierungs- und Nutzungspaketen ist es möglich, verschiedene Archivalien als ein benutzbares Objekt auszugeben. Ebenso ist es aber auch möglich, verschiedenartige Benutzungsobjekte aus einem Archivierungspaket zu formen. Beispielsweise könnten aus den 9,9 Millionen Datensätzen der Volkszählung 1970 alle Datensätze ausgegeben werden oder nur die Datensätze der Bewohner von Esslingen oder nur die Datensätze aller Schornsteinfeger oder eben nur die Schornsteinfeger von Esslingen.

Bei allen Phasen im Lebenszyklus der digitalen Archivalien muss daher jedes einzelne Archiv, seine bisherigen Prozesse und Grundsätze überdenken. Einerseits ist jedes Archiv für sich herausgefordert. Auch muss jedes Archiv für sich entscheiden, wie es auf diese Herausforderungen reagiert. Zugleich ist es aber so, dass alle anderen Archive vor denselben Fragen stehen. Was aber bislang fehlt sind Orte, an denen sich Archive über diese Fragen intensiv austauschen können und an denen sie auch eine entsprechende Unterstützung hierfür bekommen.

Zusammenfassend kann daher festgestellt werden, dass die mit der digitalen Archivierung einhergehenden Kosten, Abstimmungs- und Änderungsbedarfe die Archive vor sehr große Herausforderungen stellen.

Kooperationen

Obwohl also der Problemdruck groß ist, sind einfache Lösungen nicht zu sehen, da die publizierten Standards sich auf einem sehr abstrakten Niveau bewegen und nicht schematisch übertragen lassen. Ein zusätzliches Problem entsteht dadurch, dass bei den einzelnen Archiven sehr unterschiedliche Ausgangsvoraussetzungen gelten, es also auch nicht die eine für alle Archive sinnvolle Lösung geben kann. Dennoch gibt es natürlich ganze Gruppen von Archiven, die von vergleichbaren Voraussetzungen ausgehen könnten.

Es ist kaum denkbar, dass ein Archiv auf diese Herausforderungen im Alleingang die passenden Antworten entwickeln kann. In dieser Situation ist es besonders misslich, dass deutsche Archive bei der digitalen Archivierung bislang kaum in der Lage sind, sich die nötigen Konzepte durch gemeinsames praktisches Arbeiten an konkreten Problemen zu erarbeiten. Gefragt werden sollte daher weniger nach dem Ob einer Kooperation und mehr nach dem Wie, also nach deren Umsetzung.

Auf dem Gebiet der digitalen Archivierung hat das nestor-Projekt bereits eindrücklich nachgewiesen, wie viel durch Kooperationen erreicht werden kann. In nestor kooperieren Archive, Bibliotheken, Museen und andere bei der Entwicklung von Konzepten zur digitalen Archivierung. Nestor hat verschiedene Ergebnisse vorgelegt. Am meisten nachgefragt und zitiert werden aber die Ergebnisse, die in interdisziplinär besetzten nestor Arbeitsgruppen erarbeitet wurden.¹⁸

Diese Arbeitsgruppenergebnisse stehen sowohl für die Stärke als auch für die Schwäche von nestor. Zunächst sind es Texte, die durch einen längeren Diskurs unter den AG-Teilnehmern gegangen sind und danach auch der Fachöffentlichkeit zur Kommentierung vorgelegt wurden. Die Texte sind also sehr genau durchdacht und entsprechend fundiert. Zum anderen müssen sich diese Texte auf einem sehr abstrakten Niveau bewegen, da sie für alle möglichen Anwendungsbereiche Gültigkeit besitzen sollen. Damit kommen ihnen dieselben Probleme zu, die auch für die oben erwähnten Standards gelten.

Wer setzt diese Konzepte nun in praktisches Handeln um? Dabei sind die konkreten Anforderungen der Archive dann doch etwas anders als diejenigen der Bibliotheken und Museen. Die Bibliotheken haben auf diese Herausforderung mit dem Projekt dp4lib reagiert.¹⁹ In der deutschen Archivwelt gibt es – abgesehen von einzelnen punktuellen Absprachen – noch keinen vergleichbaren Ansatz. Zu fragen ist daher, wie zwischen den konkreten Bedürfnissen des einzelnen Archivs und den abstrakten Konzepten von nestor und anderen einschlägigen Standards eine Brücke geschlagen werden kann. Dabei handelt es sich nicht nur um ein nebensächliches Problem. Tatsächlich dürfte die digitale Archivierung für die meisten Archive zu den Pflichtaufgaben zählen. Auf Dauer wird es sich keines dieser Archive erlauben können, seine Pflichtaufgaben zu vernachlässigen, da es ansonsten in den Augen des Archivträgers seine Existenzberechtigung aufs Spiel setzen würde.

Dass digitale Archivierung nur in Kooperation gelingen kann, scheint klar. Wahrscheinlich sind Kooperationsangebote auf verschiedenen Ebenen nötig, die sich je nach Aufgaben- und Themenbereich unterscheiden. Wie aber diese Angebote aussehen sollten, darüber sollte noch eingehend gesprochen werden.

¹⁸ Genannt werden können beispielsweise der Kriterienkatalog vertrauenswürdige digitale Langzeitarchive, hg. von der nestor-Arbeitsgruppe Vertrauenswürdige Archive – Zertifizierung (nestor-Materialien 8), Version 2, Frankfurt/Main 2008; Wege ins Archiv – Ein Leitfaden für die Informationsübernahme in das digitale Langzeitarchiv, hg. von der nestor-Arbeitsgruppe Standards für Metadaten, Transfer von Objekten in digitale Langzeitarchive und Objektzugriff, 2009 oder Leitfaden zur digitalen Bestandserhaltung. Vorgehensmodell und Umsetzung, Version 1.0, verfasst und herausgegeben von der nestor-Arbeitsgruppe Digitale Bestandserhaltung, Frankfurt/Main 2011. Alle Publikationen von nestor können von der Webseite <http://www.langzeitarchivierung.de> heruntergeladen werden.

¹⁹ <http://dp4lib.langzeitarchivierung.de/>.

Die KOST: Sieben Jahre Good Practice

Von GEORG BÜCHLER

Im November 2004 nahm die schweizerische Koordinationsstelle für die dauerhafte Archivierung elektronischer Unterlagen KOST in Bern ihre Arbeit auf. Die in den über sieben Jahren ihres Betriebs gesammelten Erfahrungen können für die Diskussion, ob und in welcher Form in Deutschland Koordinierungsstellen für die digitale Archivierung geschaffen werden sollen, als Input und Anschauungsmaterial dienen. Die in beiden Ländern unterschiedlichen Rahmenbedingungen sind dabei immer zu beachten. Der vorliegende Artikel unternimmt es, Organisation und Produkte der KOST zu beschreiben und abschließend einige Erkenntnisse zu formulieren. Der Vortragscharakter wurde dabei weitgehend beibehalten.¹

Form

Strategiestudie

In der Diskussion um digitale Archivierung zu Beginn unseres Jahrhunderts war die Forderung nach Zusammenarbeit und Kooperation ein Gemeinplatz. Viele geografisch oder thematisch definierte Communities von Gedächtnisinstitutionen machten sich in einer Grundsatzdiskussion über mögliche Formen der Zusammenarbeit diese Forderung zu eigen. In der schweizerischen Archivlandschaft war es die sogenannte Strategiestudie,² welche 2002 frühere Überlegungen bündelte und konkrete Vorgehensweisen anregte. Im Rahmen ihrer Empfehlungen forderte sie unter dem Titel *Koordinations- und Beratungsstelle* insbesondere: *Auch nach dem Abschluss der Studie und der Arbeiten des Lenkungsausschusses und des Projektteams soll ein Organ bestehen, das die weiteren Aktivitäten koordiniert und den Archiven beratend zur Seite stehen kann.*³ Sie präziserte, dass die *Einrichtung der Koordinations- und Beratungsstelle [...] das zentrale Element der kurzfristigen Empfehlungen* sei und empfahl eine *Gesamtdotierung von ca. 150–200 Stellenprozenten.*⁴

¹ Die dem Referat zugrundeliegende Präsentation ist online verfügbar unter <http://prezi.com/kyiwenbgmovc/die-kost-sieben-jahre-good-practice/>. [Alle Links wurden am 25.09.2012 überprüft].

² Thomas Schärli et al.: Gesamtschweizerische Strategie zur dauerhaften Archivierung von Unterlagen aus elektronischen Systemen (Strategiestudie), Basel 2002. <http://www.vsa-aas.org/de/aktivitaet/archivdirektoren/strategiestudie/>.

³ Schärli, wie Anm. 2, S. 157.

⁴ Schärli, wie Anm. 2, S. 157.

Errichtung der KOST

Für die Erarbeitung der Strategiestudie hatten der Verband schweizerischer Archivarinnen und Archivare VSA und die Konferenz der leitenden Archivarinnen und Archivare auf Kantons- und Bundesebene sowie des Fürstentums Liechtenstein (KLA CH/FL, jetzt Archivdirektorenkonferenz, ADK) zusammengearbeitet. Für die Umsetzung der zentralen Empfehlung beanspruchte die ADK die Federführung.⁵ Sie unterstrich damit, dass sie die größeren – also die kantonalen und die nationalen – Archive als prädestiniert erachtete für die Schaffung und den Aufbau einer neuen gemeinsamen und mit viel spezialisiertem Know-how ausgestatteten Institution: Fragen der digitalen Archivierung stellten sich vor zehn Jahren vor allem in größeren Häusern.

Bund, Kantone und das Fürstentum Liechtenstein schlossen dazu eine Verwaltungsvereinbarung ab.⁶ Die Wahl dieser Form der Grundlage siedelt die Zusammenarbeit nicht auf einer fachlichen, sondern auf einer politischen Ebene an. Die Involvierung der jeweiligen Regierungen verleiht der KOST ein größeres Gewicht. Die Verwaltungsvereinbarung wird ergänzt durch ein Betriebsreglement.

Organisatorische Ausgestaltung

Die Grundlagentexte spezifizieren zunächst die Strukturen. Höchstes Organ der KOST ist die Aufsichtskommission, der sämtliche Träger angehören. Diese genehmigt Budgets, Jahresrechnungen und Rechenschaftsberichte und wählt einen Steuerungsausschuss aus drei (zukünftig fünf) Mitgliedern, der die operativen Geschäfte führt und der Geschäftsstelle vorgesetzt ist. Diese ist administrativ beim Schweizerischen Bundesarchiv in Bern angesiedelt.⁷ Anders als andere Institutionen der schweizerischen Archivszene verfügt die KOST damit über besoldetes Personal und ein eigenes Budget. Die Ausschreibung der ersten zwei Stellen für einen Archivar und einen Informatiker zeigt die Absicht, den Graben zwischen den Archiven und der Welt der Informatik durch eine Fachstelle überwinden zu wollen, die beide Sprachen spricht.

Um die in der Strategiestudie empfohlene Personalausstattung zu erreichen, wurde ein Jahresbudget von rund 200 000 CHF angestrebt. Dieses wurde zur einen Hälfte zu gleichen Teilen, zur anderen entsprechend der Wohnbevölkerung unter den Trägern aufgeteilt. Daraus ergibt sich der Jahresbeitrag jedes Trägers als Summe aus einem Sockelbeitrag von 5 400 CHF (ca. 4 500 EUR) und einem variablen Beitrag von 0,018 CHF pro Einwohner. Die Jahresbeiträge werden der Bevölkerungsentwicklung und der Teuerung angepasst. Die Leistungen des Bundes werden mit der zur Verfügung gestellten Infrastruktur abgegolten.

⁵ Siehe dazu und zum Folgenden die Akten zur Gründung der KOST in CH-BAR 074, Schweizerische Archivdirektorenkonferenz.

⁶ Verwaltungsvereinbarung über die Zusammenarbeit der Schweizerischen Eidgenossenschaft mit den Kantonen und dem Fürstentum Liechtenstein bezüglich Errichtung und Betrieb einer Koordinationsstelle für die dauerhafte Archivierung elektronischer Unterlagen (KOST), beschlossen am 25.02.2003. <http://kost-ceco.ch/cms/download.php?7ad4e0458eaabf2dab886251cdd3eb5c>.

⁷ Ebenfalls diskutiert, aber verworfen, wurde eine Angliederung an die *École supérieure d'information documentaire* in Genf, also ein Fachhochschulinstitut, was dem oben geäußerten Gedanken der *Selbsthilfeorganisation* widersprochen hätte.

Diese Kostenaufteilung hat zwei wesentliche Konsequenzen. Einerseits hält sie die Belastung für die einzelnen Träger in einem verantwortbaren Rahmen. Für kleine Archive überschreitet der Jahresbeitrag nur unwesentlich den Sockel von 5 400 CHF, für die grössten Träger bewegt er sich in der Größenordnung von 30 000 CHF, bleibt also im Rahmen einer kleinen Teilzeitstelle. Andererseits bedingt der Sockelbeitrag implizit eine minimale Größe eines potentiellen Trägerarchivs. Bei der späteren Öffnung der KOST für Kommunalarchive zeigte sich, dass dieser Betrag in der Regel die für die digitale Archivierung verfügbaren Mittel kleinerer Institutionen rasch übersteigt.

Geschichtlicher Abriss und Entwicklung

Nach der Verabschiedung der Verwaltungsvereinbarung durch die ADK am 25.02.2003 dauerte es über ein Jahr, bis das vereinbarte ambitionöse Quorum erreicht war und 15 Kantone sowie der Bund diese ratifiziert hatten. Beim operativen Start im November 2004 waren 18 Staatsarchive, das Schweizerische Bundesarchiv und das Liechtensteinische Landesarchiv mit von der Partie; mit einem Budget von 203 000 CHF konnten 120 (später 140) Stellenprozent durch zwei Teilzeitmitarbeitende besetzt werden. Nach der Startphase sah eine Erweiterungsrunde in den Jahren 2007 und 2008 den Beitritt sämtlicher noch fehlender Kantone (außer Appenzell Innerrhoden und dem Tessin) sowie der drei ersten Städte. 2012 zählt die KOST 31 Träger (Schweizerische Eidgenossenschaft, Fürstentum Liechtenstein, 24 Kantone, 5 Kommunen), verfügt über ein Jahresbudget von 330 000 CHF und beschäftigt 3 Mitarbeitende, die sich 200 Stellenprozent teilen.

Inhalt

Prämissen

Aufgabe der KOST ist es nicht, die Probleme ihrer Trägerarchive zu lösen, sondern diese bei der Problemlösung zu unterstützen. Sie sieht sich heute dabei mit 31 in vielerlei Hinsicht sehr ungleichen Archiven konfrontiert. Die Anzahl Mitarbeitender variiert zwischen drei und über sechzig, darunter können sich mehrere Archivinformatiker befinden oder kein einziger. Einige Archive haben bereits umfangreiche digitale Bestände übernommen, andere haben keinerlei praktische Erfahrung mit der digitalen Archivierung.⁸ Das Angebot der KOST muss dieser Vielfalt Rechnung tragen. Es tut dies, indem es eine möglichst große Bandbreite von Theorie und Praxis abdeckt und die Träger der KOST in unterschiedlichen Arten und Intensitäten involviert.

Diese Involvierung der Träger ist für den Erfolg der KOST unabdingbar. Allein mit den beschränkten personellen und finanziellen Ressourcen der Geschäftsstelle wäre das gegenwärtig

⁸ Einen Überblick über die Situation der KOST-Träger im Bezug auf die Archivierung digitaler Unterlagen gibt die *Umfrage zum Stand der digitalen Archivierung in der Schweiz*, die die KOST im Frühjahr 2011 durchgeführt hat: <http://kost-ceco.ch/cms/index.php?id=3,280,0,0,1,0>.

tige Angebot der KOST nicht zu gewährleisten. Indem sich Mitarbeitende der Trägerarchive in vielen Projekten in diverser Art und Weise engagieren, vervielfältigen sie dieses Angebot und es ergeben sich mannigfaltige Wechselwirkungen. Die Geschäftsstelle unterstützt diese Form von Austausch und Kooperation unter anderem dadurch, dass sie in der Regel den Großteil der organisatorischen Verantwortung trägt.

Das Produktportfolio der KOST

Pilotlösungen: Zusammen mit einem oder mehreren Trägerarchiven erarbeitet die KOST Prozesse, Definitionen und Schnittstellen für eine spezifische Art von Unterlagen oder Daten und führt die Übernahme durch. Bestand das Hauptziel solcher Projekte zu Beginn der Tätigkeit der KOST darin, am konkreten Objekt Erfahrungen zu sammeln und Hemmschwellen abzubauen, dienen sie inzwischen vermehrt dazu, das Instrumentarium der KOST (Tools, Standards, Konzepte, etc.) in der Praxis anzuwenden und zu testen. Pilotlösungen sind aber auch Experimentierfelder der Zusammenarbeit mit Dritten, nämlich mit Softwareanbietern, Informatikdienstleistern und Berufsverbänden. Dies gilt insbesondere im Bereich des E-Government, wo die KOST die Federführung eines priorisierten Vorhabens zur digitalen Archivierung im Rahmen der nationalen E-Government-Strategie innehat und die Zusammenarbeit mit anderen Akteuren sucht.

Standards: Bereits zu Beginn der Tätigkeit der KOST wünschten sich ihre Träger Hilfestellungen in Form von Richtlinien und Standards. Erste solche Produkte erarbeitete die KOST im Alleingang; erwähnt sei hier an erster Stelle der *Katalog archivischer Dateiformate KaD* von 2007/08, dessen dritte Version gegenwärtig in Arbeit ist.⁹ Mit der Gründung einer Fachgruppe Digitale Archivierung im Rahmen von eCH, dem schweizerischen Verein für die Erarbeitung von Standards für das E-Government, stellte die KOST die Standardisierungsaktivitäten in ihrem Geschäftsbereich auf eine breitere und breiter akzeptierte Basis.¹⁰

Erarbeitung von Grundlagen: Es gibt in der Schweiz keine mit *nestor* vergleichbare Institution, welche wichtige Themen in Zusammenarbeit und auf einem hohen Abstraktionsniveau untersucht. Die KOST ist einer der wenigen Orte, wo gemeinschaftlich theoretische Grundlagen für die digitale Archivierung erarbeitet werden können. Dies geschieht hauptsächlich in zwei Formen. Neue, noch wenig diskutierte, aber als wichtig erkannte Themen beleuchtet die KOST im Rahmen von Kolloquien, zu welchen sie vornehmlich externe Referentinnen und Referenten einlädt und deren Ergebnisse sie publiziert. Andere Themen, die erhöhte Aufmerksamkeit verdienen, untersucht die KOST-Geschäftsstelle in der Form von Studien, die sie auf ihrer Website veröffentlicht.

⁹ <http://www.kost-ceco.ch/wiki/whelp/KaD/index.php>.

¹⁰ http://www.ech.ch/vechweb/page?p=page&site=/Gremien/Fachgruppen/digitale_archivierung.

Tools und Services: Mit dem allmählichen Fortschreiten hin zur Praxis hat die Erarbeitung von Werkzeugen stark an Bedeutung gewonnen. Erste Services der KOST beschränkten sich auf kurze Handreichungen für konkrete Aufgaben im Tagesgeschäft der Archive. Inzwischen zeichnet sich die Landschaft der Konzepte und Softwarewerkzeuge für die digitale Archivierung hinreichend genau ab, dass Lücken identifiziert und gezielt gefüllt werden können. Ein konkretes Beispiel dafür sind Software-Tools zur Implementierung von Standards. So hat die KOST für den in Vernehmlassung befindlichen eCH-Standard 0160, Archivische Ablieferungsschnittstelle (SIP), einen Validator und einen Browser entwickelt.¹¹ Weitere Validierungstools sind in Erarbeitung oder in Planung. Ein spezielles Dienstleistungsangebot ist arcun, eine Remote-Speicherlösung für digitale Archivalien.¹²

Beratung: Das individuellste Angebot der KOST steht allen Trägerarchiven gemäß ihren Bedürfnissen zur Verfügung. Es kann sich in einer kurzen Telefonauskunft erschöpfen, aber auch bis zur Begleitung von oder Mitarbeit in Projekten gehen. Je mehr die Archive sich selber mit der digitalen Archivierung befassen, desto mehr wird die Beratung durch die KOST genutzt. Es steht zu erwarten, dass dieses Angebot in den kommenden Jahren ausgebaut werden muss.

Zusammenarbeit: Die KOST arbeitet nicht allein in ihrem Elfenbeinturm, sondern sucht die Vernetzung mit anderen Initiativen: mit archivischen Arbeitsgruppen, mit *User Groups* von relevanter Software, mit Standardisierungsgruppen. Es gilt im Minimum, andere Initiativen nicht zu konkurrenzieren, idealerweise aber, sich richtig zu ergänzen. Während der Hauptteil der Zusammenarbeit innerhalb der Grenzen der Schweiz erfolgt, existieren auch international diverse Kontakte, in Deutschland vor allem in Form einer Teilnahme an den jährlichen Tagungen des Arbeitskreises *Archivierung von Unterlagen aus elektronischen Systemen* und durch die Mitarbeit in den *nestor*-Arbeitsgruppen Langzeitarchivierungsstandards, OAIS und Bestandserhaltung.

Erkenntnisse

Die Beschreibung von Form und Inhalt der KOST geben eine Vorstellung von ihrem Funktionieren. Ergänzt werden soll diese abschließend durch fünf zentrale Erkenntnisse aus siebeneinhalb Jahren Koordinations- und Beratungsarbeit, die im Folgenden thesenartig präsentiert werden.

¹¹ Zu SIP-Val, dem Validator für SIPs, siehe http://kost-ceco.ch/cms/index.php?sip-val_de. Zum SIP-Browser siehe http://kost-ceco.ch/cms/index.php?sip-browser_de.

¹² *Arcun* ist das rätoromanische Wort für *Speicher* und steht zugleich kurz für *Archival Repository for Collaborative Use*. Arcun beruht auf einer genauen Spezifikation der archivischen Anforderungen an eine Speicherlösung und setzt diese in Zusammenarbeit mit einem kommerziellen Anbieter um; siehe http://kost-ceco.ch/cms/index.php?arcun_de.

1. Der Angebotsmix ist der Schlüssel zum Erfolg

Durch ein breites Produktangebot erlaubt es die KOST allen Trägerarchiven, unabhängig von ihrer Größe und ihrem Entwicklungsstand in der digitalen Archivierung, von ihren Leistungen zu profitieren. So mag sich ein ressourcenstarkes Archiv über Monate hinweg in einem Pilotvorhaben intensiv engagieren, während es einem weniger weit fortgeschrittenen Archiv möglich ist, durch den gelegentlichen Veranstaltungsbesuch mit der aktuellen Diskussion im Kontakt zu bleiben. Zudem erlaubt die Breite des Angebots zusehends mehr Verknüpfungen: Die KOST entwickelt Tools, welche ihre Standards implementieren, testet Resultate von Kolloquien in Pilotvorhaben und anderes mehr.

2. Ein Gesamtbild konkretisiert sich

Die KOST hat als Dienstleister für Archive selber weder Archivsprengel noch Archivalien, mithin auch keine eigene Archivlösung. Die Umriss- und einzelne Elemente einer möglichen Lösung (auf verschiedenen Komplexitätsstufen) werden durch ihre Arbeit aber immer sichtbarer. Grundlage und Ordnungsprinzip dafür sind die *Minimalanforderungen an die digitale Archivierung*, ein 2009 publiziertes Richtlinienendokument, das die gemeinsam anerkannte Gesamtsicht der Problematik digitale Archivierung wiedergibt.¹³

3. Die Schwerpunkte wandern, aber wenig und langsam

Die Arbeit und der Angebotsmix der KOST blieben in den bald acht Jahren ihrer Existenz überraschend stabil. Zu beobachten sind allerdings Verschiebungen der Schwerpunkte: Zu Beginn lag ein Hauptaugenmerk auf den Pilotlösungen, die vielen Archiven einen Einstieg in die praktische Arbeit mit digitalen Archivalien ermöglichten. Nach einer zweiten Phase, welche einen Fokus auf der Erarbeitung von theoretischen Grundlagen und Standards sah, gewinnen in den letzten Monaten die Erarbeitung von Tools und die individuelle Beratung der Träger an Bedeutung. Die neuen Schwerpunkte verdrängen jedoch die alten nicht, sondern bewirken einzig eine maßvolle Gewichtsverlagerung.

4. Die Kooperation über die staatlichen Ebenen hinweg gewinnt an Bedeutung

Die verstärkte Zusammenarbeit über die staatlichen Ebenen hinweg sowie eine allgemeine Tendenz zur Verlagerung von Aufgaben (oder zumindest von Definitionsmacht) von den Kantonen weg zum Bund erfordern neue Vorgehensweisen für die Archivierung. Insbesondere nehmen die kantonal, aber nach Bundesrecht vollzogenen Aufgaben zu. In solchen Fällen erheben und bewirtschaften die Kantone beispielsweise Daten in einer vom Bund bereitgestellten und betrieb-

¹³ KOST (Hg.), *Minimalanforderungen an die digitale Archivierung*. 2009. http://kost-ceco.ch/cms/index.php?minimal_specifications_de.

benen Applikation. Datenhoheit und technischer Ansatzpunkt für die Archivierung divergieren. Die KOST ist ein idealer Akteur für die Koordination entsprechender Archivierungsprojekte.

5. Kooperation bleibt schwierig, Eigeninitiative bleibt attraktiv

Der Wert der Kooperation ist unumstritten. Man darf jedoch die Augen nicht verschließen vor der Tatsache, dass Kooperation gegenüber einer Einzelkämpferlösung weniger attraktiv sein kann. Es ist in der Regel reizvoller, ein eigenes Projekt zu führen als eine gemeinsam gefundene Lösung einfach umzusetzen. Dies gilt sowohl für die Informatikdienstleister als auch für die Archive. Die größere Autonomie und die bessere Anpassung an die je eigene Situation müssen abgewogen werden gegen die Effizienz des Mitteleinsatzes. So müssen Kooperation und Koordination in einem fortlaufenden Prozess immer wieder neu erarbeitet werden.

Erfahrungen des LWL-Archivamts bei der Koordination von Projekten zum Langzeiterhalt elektronischer Daten

Von PETER WORM

Die Herausforderungen, die mit dem langfristigen Erhalt digitaler Informationen verbunden sind, sind gerade für kleine und mittelgroße Archive, wie sie im nichtstaatlichen Bereich überwiegen, so groß, dass sie nicht im Alleingang gelöst werden können. In den folgenden Ausführungen werden nach einer kurzen Darstellung der Aufgaben des LWL-Archivamts zwei Wege exemplarisch aufgezeigt, wie in Kooperationsvorhaben auf regionaler und in archivspartenübergreifender Weise tragfähige Lösungen erarbeitet werden können.

Das LWL-Archivamt für Westfalen

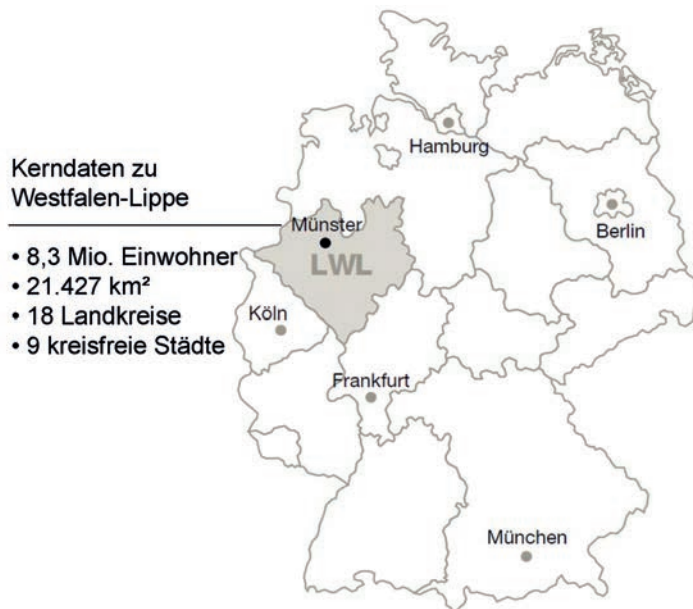


Abb. 1: Lage und Kerndaten zum Landesteil Westfalen-Lippe (Stand Dez. 2009).

Das LWL-Archivamt für Westfalen ist eine Beratungseinrichtung des Landschaftsverbands Westfalen-Lippe, einer Gebietskörperschaft oberhalb der Kreise, die umlagefinanziert Aufgaben

für die unmittelbaren (Kreise, kreisfreie Städte) und mittelbaren (kreisangehörige Städte und Gemeinden) Mitgliedskommunen übernimmt. In erster Linie sind das die Bereiche der Behindertenhilfe, -ausbildung und -versorgung und der psychiatrischen Kliniken (inkl. Maßregelvollzug), aber auch der Jugendhilfe und der Kultur (LWL-Landesmuseen und -Kulturdienste). Schon 1927 wurde vom Rechtsvorgänger, dem preußischen Provinzialverband, eine Archivberatungsstelle zur Unterstützung des nichtstaatlichen Archivwesens eingerichtet.¹

Heute gibt das Amt u.a. zweimal im Jahr die Zeitschrift *Archivpflege in Westfalen-Lippe* heraus, veranstaltet jährlich einen der größten regionalen Archivtage in Deutschland, und unterstützt die rund 250 Kommunalarchive im Landesteil Westfalen-Lippe mit Fortbildungen, einer Restaurierungswerkstatt und individueller Beratungsleistung. Es sorgt für die Erschließung und Nutzbarmachung der gut 100 Adelsarchive der Region und unterhält eine Außenstelle mit 1,5 Personalstellen im Westfälischen Wirtschaftsarchiv in Dortmund. Schließlich ist das LWL-Archivamt Archiv des Landschaftsverbands, in dessen mehr als 200 Dienststellen über 13.000 Beschäftigte arbeiten. Seit 2001 ist auch das Westfälische Literaturarchiv im Haus angesiedelt. Von den acht Facharchivarinnen und -archivaren sind sieben in der Kommunalarchivpflege tätig; für vier davon ist das das Hauptaufgabefeld.²

Im eigenen Verband versuchen wir seit über fünfzehn Jahren archivische Interessen bei IT-Projekten einzubringen und haben das Engagement seit den ersten DMS-Einführungen verstärkt.³ Wir hoffen Ende des Jahres mit einem elektronischen Langzeitarchiv in den Produktivbetrieb gehen und die ersten elektronischen Übernahmen durchführen zu können.

¹ Norbert Reimann: Die Sorge um die Archive als Aufgabe der landschaftlichen Kulturpflege in Westfalen. In: *Der Märker* 45 (1996), S. 139–153 (online abrufbar unter: www.lwl.org/waa-download/pdf-download/Sorge_um_Archive.pdf (alle Links wurden am 26.7.2012 überprüft).

² Einen Überblick über unsere Tätigkeit bietet unsere Homepage www.lwl-archivamt.de.

³ Katharina Tiemann: Überlegungen zum Umgang mit Dokumentenmanagement-Systemen beim Aufbau des Archivs des Landschaftsverbands Westfalen-Lippe. In: *Übernahme und Bewertung von kommunalem Schriftgut, Datenmanagement-Systeme (Texte und Untersuchungen zur Archivpflege 12)*. Münster 2000. S. 113–122; Brigitta Nimz: Besuch der *Arbeitsgruppe Archivierung von Unterlagen aus digitalen Systemen* im Bundesarchiv am 28.8.2001. In: *Archivpflege in Westfalen-Lippe* 55 (2001) S. 43–45; Peter Worm: Archivische Begleitung einer DMS-Auswahl und -Einführung bei den Kommunalen Versorgungskassen für Westfalen-Lippe. In: *Internationale Archivsymposien in Brauweiler (D) (2005), Trier (D) (2006) und Hasselt (B) (2007)*. *Annalen (Miscellanea Archivistica. Studia 176)*. Brüssel 2008. S. 193–205; zuletzt den Stand zusammenfassend: Katharina Tiemann und Peter Worm: Zwischen Domea-Anspruch und kommunaler Wirklichkeit. Werkstattbericht aus der Einführung eines DMS im Landschaftsverband Westfalen-Lippe. In: *Auf dem Weg zum digitalen Archiv. Stand und Perspektiven von Projekten zur Archivierung digitaler Unterlagen*. Hg von Matthias Manke (Veröffentlichungen des Landeshauptarchivs Schwerin). Schwerin 2012 S. 60–67.

Das Modell regionaler Arbeitskreise

Als Aufbruchssignale, sich den archivischen Herausforderungen, die aus der digitalen Verwaltung erwachsen, in der breiten Masse der Kommunalarchive anzunehmen, können zwei Positionspapiere gesehen werden, die die drei archivischen Arbeitskreise des Nordrhein-Westfälischen Landkreistags (AKKA), des Städtetags (ASGA) und des Nordrhein-Westfälischen Städte- und Gemeindebunds (ARGE) 2005 (*Archivische Aufgaben im digitalen Zeitalter*) und 2008 (*Handreichung zur Organisation der digitalen Archivierung*) als Empfehlungen der kommunalen Spitzenverbänden des Landes auf den Weg brachten.⁴ Verbunden waren diese Aufrufe mit dem Wunsch der Archive an die Archivberatungsstellen, mit dem notwendigen Know-how ausgestattet zu werden. Um diesem Wunsch gerecht zu werden, führten Dr. Gläser (ehem. LVR – Archivberatungs- und Fortbildungszentrum, jetzt INFORA) und ich eine Umfrage bei den rund 450 Kommunalarchiven in NRW durch, die die organisatorische Einbindung des Archivs, den Wissensstand der Archivarinnen und Archivare und den IT-Einsatz in der Trägerkommune zum Thema hatte.⁵ Während die Kenntnisse im Bereich der klassischen Schriftgutverwaltung von mehr als 80% der Antwortenden zumindest mit *ausreichend* eingeschätzt wurden, lag diese Selbsteinschätzung im Bereich der inhaltlichen Fragen der elektronischen Archivierung nur bei gut 30% vor, in technischen Fragen sahen sich nur 25% mit ausreichenden Kenntnissen ausgestattet. Die Kontakte des Archivs zur Verwaltungsspitze beurteilten die meisten Archive als zumindest ausreichend gut, während die allermeisten mit der lokalen EDV-Abteilung gut auskamen. Schlechte oder gar keine Kontakte gab es zum zuständigen Kommunalen Rechenzentrum: nur ca. 25% schätzten ihr Verhältnis zum Rechenzentrum mit ausreichend oder besser ein. Auffällig korreliert dieses Ergebnis mit der Frage, ob die Archive bei der Einführung von neuer Software beteiligt würden. Das konnten nur 28% der Befragten bejahen.

Die Verwaltungsorganisation und -praxis stellt sich im Bereich des IT-Einsatzes in den meisten kreisangehörigen Städten und Gemeinden wie folgt dar: Mit dem Gesetz über kommunale Gemeinschaftsarbeit von 1979 wurden in NRW die Grundlagen gelegt, auf denen in der Folgezeit zumeist in der Rechtsform des Zweckverbands eine Vielzahl von kommunalen Rechenzentren gegründet wurden. Häufig orientierte man sich an Kreisgrenzen, manchmal gingen Rechenzentrumsgebiete aber auch weit darüber hinaus. In Entwicklungsgemeinschaften wurden Softwares zunächst selbst entwickelt, später auch beschafft oder zusammen mit den Softwarefirmen auf die Gegebenheiten vor Ort angepasst. Auch der Ausbau von Netzinfrastruktur und die Ermöglichung des Datenaustauschs waren und sind zentrale Aufgabenbereiche kommunaler

⁴ Archivische Aufgaben im digitalen Zeitalter. In: *Archivpflege in Westfalen-Lippe* 63 (2005) S. 55–57. Handreichung zur Organisation der digitalen Archivierung. In: *Archivpflege in Westfalen-Lippe* 68 (2008) S. 34–36.

⁵ Florian Gläser und Peter Worm: Ergebnisse der Umfrage der Archivberatungsstellen zum Fortbildungsbedarf im Bereich *Archivierung elektronischer Unterlagen*. In: *Archivpflege in Westfalen-Lippe* 70 (2009) S. 50–59.

Rechenzentren.⁶ Die IT-Stellen vor Ort kümmern sich um die Bereitstellung von Hard- und Standardsoftware und unterstützen die Sachbearbeiter als *First-Level-Support*. Die Entscheidung, welche Software gerade in Kernbereichen kommunaler Verwaltung angeschafft wird, und z. B. welche Schnittstellen und Funktionalitäten sie mitbringen muss, wird in den Rechenzentren getroffen, selten in den IT-Stellen oder Fachämtern vor Ort.⁷

Diese Organisationsform bedeutete für unsere Arbeit als Archivberatungsstelle, dass wir, um archivische Interessen wirkungsvoll vertreten zu können, den Kontakt zu den Rechenzentren aufbauen und verstetigen mussten. Es entstand seit dem Jahr 2005 das Modell regionaler Arbeitskreise,⁸ an dem drei Gruppen mit bestimmten Kernkompetenzen und Zielrichtungen teilnehmen:

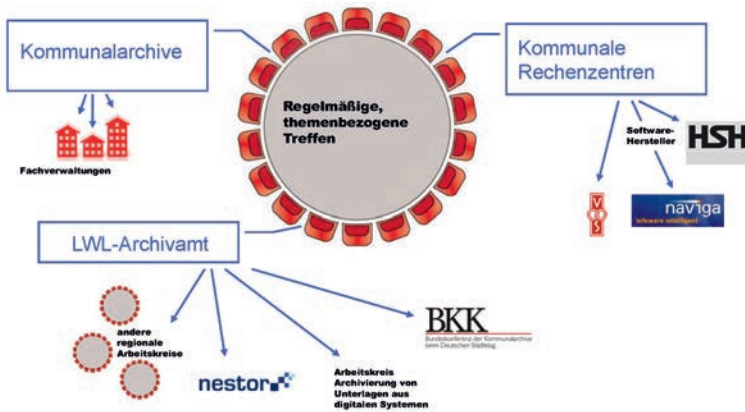


Abb. 2: Schematische Darstellung zur Zusammensetzung der regionalen IT-Arbeitskreise.

⁶ Hubert Kochjohann und Peter Worm: Sicherung elektronischer Daten im Verbund. Die Zusammenarbeit von Kommunalarchiven und kommunalen Rechenzentren. In: *Archivpflege in Westfalen-Lippe* 71 (2010) S. 42–45, bes. S. 42–43.

⁷ Etwas anders stellt sich die Situation in den Großstädten dar, hier fallen die beschriebenen Funktionen von Rechenzentrum und IT-Stelle in eins. Daneben gibt es auch kleinere Städte und Gemeinden, die ohne Rechenzentrum arbeiten (v.a. im Kreis Unna, im Ennepe-Ruhr-Kreis), oder bei denen das Rechenzentrum nur eine Art *Beschaffungsverbund* ist, aber keinen Serverbetrieb leistet (z. B. die KAAW im nördlichen und westlichen Münsterland).

⁸ Vorbild-Charakter hatte der *Facharbeitskreis Archivwesen* beim KRZN Niederrhein, vgl. Bert Thissen: Die Arbeit des Facharbeitskreises Archivwesen beim Kommunalen Rechenzentrum Niederrhein. Ein Erfahrungsbericht. In: *Handlungsstrategien für Kommunalarchive im digitalen Zeitalter. Beiträge zu einem Workshop im Rathaus Oberhausen, 14.12.05*. Hg. von Norbert Reimann (Texte und Untersuchungen zur Archivpflege 19). Münster 2006. S. 19–24; erste Vorstellung der Arbeitsweise der regionalen Arbeitskreise bei den Rechenzentren: Peter Worm: Elektronische Unterlagen der kommunalen Verwaltungen – ein Fall fürs Archiv? In: *Archivpflege in Westfalen-Lippe* 66 (2007) S. 39–41 und Eckhard Möller: *koop-owl.arch – Die Zusammenarbeit von Archiven und IT-Dienstleistern in Ostwestfalen-Lippe bei der Bewertung von Fachanwendungen*. Ebd. S. 41–44.

1. Die interessierten Kommunalarchive einer Region – sie haben die Bewertungshoheit über die Unterlagen ihrer Kommune, bringen die Kontakte zu ihren Fachämtern und die Erfahrung aus dem analogen Bereich mit. Die Veranstaltungen bieten ihnen die Möglichkeit, sich dem Thema *Elektronische Unterlagen* anzunähern und Know-how in der praktischen Beschäftigung damit aufzubauen.
2. Die kommunalen Rechenzentren einer Region – sie betreiben die großen Fachverfahren für ihre Mitgliedskommunen und sichern unter dem Schlagwort *Archivierung* die in ihnen enthaltenen Daten. Die Rechenzentren erweitern als potenzielle Betreiber von elektronischen Langzeitarchiven ihre Dienstleistungspalette und können sich in den Arbeitskreisen mit den archivischen Fachanforderungen vertraut machen. Sie vermitteln diese Anforderungen ggf. auch an die Herstellerfirmen von Fachsoftwares weiter.
3. Als Vertreter des LWL-Archivamts organisiere ich die Treffen mit und vermittele zwischen den Gesprächspartnern. Darüber hinaus sehe ich meine Aufgabe im Wissenstransfer hin zu Arbeitsgruppen im archivischen Bereich wie dem IT-Ausschuss der Bundeskonferenz der Kommunalarchive beim Deutschen Städtetag oder den Archivsparten übergreifenden Arbeitskreis *Archivierung von Unterlagen aus digitalen Systemen*, auf Fachtagungen oder durch persönliche Kontakte sowie natürlich auch zu anderen regionalen Arbeitskreisen.

Die Vorgehensweise in den IT-Arbeitskreisen beginnt zumeist mit einer Bestandsaufnahme der in den Verwaltungen eingesetzten Fachverfahren und einer archivischen Vorprüfung, ob diese Verfahren archivische Belange berühren könnten. Das ist zum Beispiel der Fall, wenn prinzipiell archivwürdige Papierakten durch die Fachämter nicht mehr gebildet oder stark ausgedünnt werden (hybride Aktenführung). Mit dem Einsatz elektronischer Registerverfahren oder von Dokumenten-Managementsystemen kann eine Verwaltung bewusst in Richtung *eAkte* steuern, doch über *starke Fachverfahren*, mit denen Verwaltungsabläufe gesteuert werden und die die Ablage von im Verfahren erstellten oder eingescannten Dokumenten erlauben, können schleichende Übergänge zur elektronischen Aktenführung begünstigt werden. Eine besondere Dringlichkeit und Aufmerksamkeit ist geboten, wenn Wechsel der Fachverfahren bevorstehen. Bei dieser Gelegenheit erfolgt oft eine Ausdünnung des Altdatenbestands und es kann darüber hinaus zu Verlusten im Zuge der Datenmigration kommen. Die Rückfalloption *Wir heben die Sicherungsbänder des Altverfahrens natürlich weiter auf!*, mit der die Fachabteilungen beruhigt werden, bleibt aus archivischer Sicht wohl eine theoretische, da weder die Lauffähigkeit des Verfahrens auf neueren Betriebssystemen erhalten wird, noch ein kontrollierter Zugriff oder Aussonderung, ein Auslesen der Daten in archivgeeigneter Form oder gar bestandserhalterische Maßnahmen an diesen Daten möglich sind.

Nach der Vorbewertung⁹ setzt ein iterativer Prozess ein, bei dem die besonders wichtigen und besonders dringlichen Verfahren zuerst in den Blick genommen werden. Es wird überprüft,

⁹ Die Ergebnisse der bisher durchgeführten archivischen Vorbewertungen (Kreisverwaltung und Gemeinden des Kreises Gütersloh (2006), Kreisverwaltung Lippe (2006), kreisfreie Stadt Bielefeld (2006), Kreisverwaltung und Gemeinden des Kreises Soest (2010), Kreisverwaltung und Gemeinden des Märkischen Kreises (2012)) sind oft auch als *Work-in-progress* zur Nachnutzung auf den Internetseiten des LWL-Archivamts online abrufbar www.lwl.org/LWL/Kultur/Archivamt/Archiv_IT/Elektronische_Fachverfahren/.

welche Standardschnittstellen und Ausgabemöglichkeiten für Daten es gibt und ob diese sich für archivische Zwecke eignen. Gute Grundlagen bieten zum Beispiel die sog. XÖV-Schnittstellen, die für den Datenaustausch öffentlicher Stellen untereinander Verwendung finden.¹⁰ Gerade für Registerverfahren, bei denen die Inhalte entscheidend sind und weniger die äußere Form, ergibt sich mit XML-kodierten Primärdaten eine schlanke, weiter maschinell auswertbare und gut zu erhaltende Überlieferung. In Zuge der Schnittstellendefinition muss auch eine Feinbewertung der Fachverfahren stattfinden, d. h. die genaue archivische Definition, welche Inhalte gesichert werden sollen und wie die erhaltenswerte Form dieser Inhalte aussehen soll. Es schließen sich die technische Adaption der Schnittstelle und Testläufe bis hin zur Produktivsetzung an. Spätestens wenn die ersten (End-) Archivdaten fließen, muss es ein elektronisches Langzeitarchiv geben, das diese aufnehmen kann.

Derzeit gibt es drei archivische IT-Arbeitskreise, an denen wir mitarbeiten:

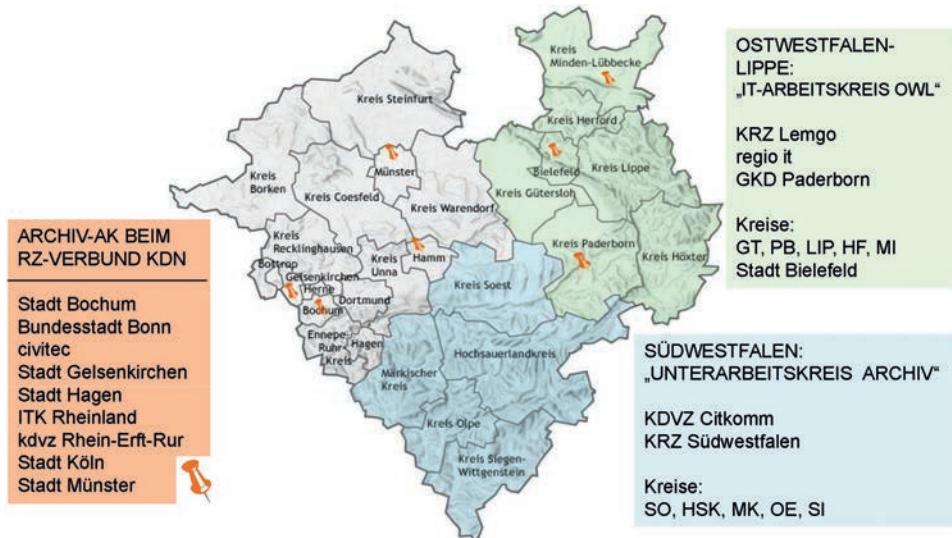


Abb. 3: Räumliche Einordnung der bestehenden regionalen IT-Arbeitskreise in Westfalen-Lippe.

Der älteste ist der IT-Arbeitskreis Ostwestfalen-Lippe, der seit 2005 regelmäßig 2–3-mal im Jahr tagt. An ihm sind als Rechenzentren die GKD Paderborn, die regio it (ehem. INFOKOM Gütersloh AöR), das KRZ Lemgo sowie Archiv-Vereine aus fast allen betroffenen Kreisen und der kreisfreien Stadt Bielefeld beteiligt. Sie sind z.T. in Entscheidungsgremien und Projektgruppen der Rechenzentren eingebunden worden, um Software-Auswahlen oder -Einführungen z. B. im DMS-Bereich archivisch begleiten zu können. Das erste Thema, dessen man

¹⁰ Eine Übersicht über die bestehenden XÖV-Standards bietet das XRepository unter der URL: www.xrepository.deutschland-online.de/.

sich angenommen hat, sind die Einwohnermelderegister.¹¹ Sie wurden bereits in den frühen 80er-Jahren von analogen Karteikarten, die pro zugezogenen oder geborenen Einwohner einer Stadt oder Gemeinde angelegt worden waren, in eine elektronische Form überführt. Auf sie geht auch der erste XÖV-Standard *XMeld* zurück. Die besondere Problematik bei den Melderegistern besteht in komplizierten Lös- und Aufbewahrungsregeln, die durch den Datenschutz vorgegeben werden. Der Hauptdatensatz muss nach Tod oder Wegzug zunächst 5 Jahre im Produktivsystem, dann noch 45 Jahre in einem gesonderten Speicherbereich aufbewahrt werden – Teildaten, wie beispielsweise die sog. Familienverkettung, d. h. der Hinweis auf die Eltern bei einem Kind, sind aber mit der Volljährigkeit des Kindes zu löschen. Vergleichbar mit der analogen Einwohnermeldekarte – und damit archivwürdig – ist der Hauptdatensatz aber nur, wenn die zuvor zu löschenden Teildaten nicht wegfallen, sondern mit ihm vor der Abgabe ins Archiv wieder verbunden werden.¹² Durch diese Rahmenbedingungen war in einigen Fällen bereits ein Datenverlust eingetreten. Es war unverzügliches Handeln der Archive gefragt, wenn dieser Überlieferungsverlust gestoppt werden sollte. Bevor die Aufwände für einen Eingriff in das Einwohnermeldeverfahren vorgenommen werden konnten, bedurfte es der Rechtsklärung, also der Abwägung von Datenschutz und archivischer Anbietungspflicht. In einer ersten Anfrage des Archivamts an den Datenschutzbeauftragten des Landes NRW wurde dem Datenschutz der Vorrang eingeräumt und die Familienverkettung als zu löschende und nicht durch die Archive zu rekonstruierende Information bewertet. Bei einer danach initiierten Anfrage des Städte- und Gemeindebunds beim NRW-Innenministerium, in der die Nachteile dieser Praxis für die Kommunen und die Bürgerinnen und Bürger dargestellt wurden, wurden die Bedenken des Datenschutzes jedoch hinten an gestellt und eine Rekonstruktion der vollen Meldedatensätze in archivischem Kontext erlaubt.¹³

Damit stand einer Kontaktaufnahme zu einem der größten Hersteller von Meldesoftware, der Firma HSH Berlin, nichts mehr im Wege. Die INFOKOM Gütersloh AöR als das Rechenzentrum, das der Hauptansprechpartner für das Verfahren in Ostwestfalen ist, holte mit der *citeq*, dem Rechenzentrum der Stadt Münster, das diese Funktion fürs Münsterland erfüllt, einen strategischen Partner mit ins Boot. Das Stadtarchiv Münster verstärkte auf archivischer Seite das Team. Gemeinsam formulierte man die Anforderungen einer XML- möglichst XMeld-basierten

¹¹ Über das Problem und die Lösung der rechtlichen Fragen berichtet Rolf-Dietrich *Müller*: Archivierung von Meldedaten – Wichtige Entscheidung des Innenministeriums NRW. In: *Archivpflege in Westfalen-Lippe* 66 (2007) S. 44.

¹² Peter *Worm*: Archivierungslösung für die Einwohnermeldedaten. In: *Archivpflege in Westfalen-Lippe* 71 (2010) S. 60–62; detailliertere Beschreibung demnächst in: *Ders.*: Zwischen melderechtlicher Löschvorschrift und archivrechtlicher Anbietungspflicht – Ansätze zur Sicherung der elektronischen Einwohnermelderegister. In: *Digitale Registraturen – digitale Archivierung. Pragmatische Lösungen für kleinere und mittlere Archive?* Beiträge zum 16. Archivwissenschaftliches Kolloquium der Archivschule Marburg. Hg. von Irmgard Christa *Becker*, Dominik *Haffner* und Karsten *Uhde* (Veröffentlichungen der Archivschule Marburg 55). Marburg 2012. S. 65–80.

¹³ Vgl. *Müller*, wie Anm. 11; der vollständige Text des Schreibens des Innenministeriums an den Städte- und Gemeindebund ist abrufbar unter: www.kommunen-in-nrw.de/mitgliederbereich/mitteilungen/detailansicht/dokument/archivierung-von-meldedaten.html.

Aussonderungsschnittstelle für die Löschmoden, die auch einen Teil der Basisdaten der Betroffenen transportieren sollte, um so eine rudimentäre Recherche und die spätere Zusammenführung von Lösch- und Restdatensatz realisieren zu können. Die Feinabstimmung und Implementierung der Schnittstelle ins Einwohnermeldeverfahren MESO benötigte mehrere Jahre (teilweise verzögert durch dringendere Vorhaben wie die Einführung des neuen Personalausweises), doch mit dem Jahreswechsel 2010/11 stand eine getestete Version für den Produktiveinsatz bereit. Ergänzt wurde die Schnittstelle durch HSH mit der Recherche- und Beauskunftungs Oberfläche *archivo*, die es erlaubt, nach Personen und Familienzusammenhängen zu suchen und im Rahmen des Archivgesetzes auch Auskünfte daraus zu erteilen.¹⁴ Die Schnittstelle wurde inzwischen durch einen Mitbewerber, die Firma AKDB (Produkt OK.EWO) implementiert; auch eine weitere Recherche- und Beauskunftungssoftware *EWO Altdaten Archiv*¹⁵ ist dazu programmiert worden. Es ist geplant, die Schnittstelle in den XÖV-Standard XMeld zu integrieren und so eine größere Herstellerunabhängigkeit und Verbindlichkeit zu erreichen.

Für den erfolgreichen Abschluss des Projekts war es notwendig, dass alle Beteiligten ihre Kompetenz und Mitspracherechte in die Waagschale warfen: Die Kommunalarchive mussten die Archivwürdigkeit der zu löschenden Daten und die generelle Zuständigkeit auch für die elektronische Überlieferung zunächst intern gegenüber den Rechenzentren, dann aber auch extern gegenüber dem Datenschutz argumentativ vertreten. Diese Erklärungen mussten gegenüber der eigenen Verwaltung fortgesetzt werden, als es um die Anschaffung des HSH-Programms *archivo* in jeder teilnehmenden Kommune ging. Bei diesen *Kämpfen* konnte das LWL-Archivamt nur unterstützend tätig sein, indem eine gemeinsame Empfehlung der beiden nordrhein-westfälischen Archivberatungsstellen an die Bürgermeisterinnen und Bürgermeister im Land versandt wurde, die auf den Handlungsbedarf bei den Einwohnermeldedaten und die vorhandene technische Lösung hinwies.¹⁶ Die Rechenzentren als Kunden der Herstellerfirma waren deren *natürliche Ansprechpartner*, die den technischen Sachverstand für die Mitarbeit an der Schnittstellendefinition mitbrachten und ihre Kontakte spielen ließen, um mit größerer Marktmacht die Anforderungen durchzusetzen.

Schließlich müssen sich Archive und deren Archivträger damit abfinden, dass für sie spezifisch angefertigte IT-Lösungen Aufwände erzeugen und somit auch Geld kosten. Während im staatlichen Bereich die Schnittstellenaufwände meist an die abgebende Stelle abgewälzt werden können, ist das bei einer Einheitsverwaltung wie einer Kommune nicht der Fall – hier macht es nur bedingt einen Unterschied, aus welcher Haushaltsstelle das Geld fließt. Die Haushaltsstelle des Archivs (so vorhanden) ist in der Regel so schmal bemessen, dass sie nicht für die Finanzierung von Lizenzkosten oder anderen anteilig umgelegten Kosten für IT-Dienstleistungen erhalten kann. Meist springt hier die IT-Abteilung einer Gemeinde oder einer Stadt mit ihrem Etat ein. Für die Zwischenarchivlösung im Einwohnermeldebereich kann mit guten Gründen auch das Bürgeramt in die Finanzierung eingebunden werden. Kernarchivische Bereiche, wie zum

¹⁴ Zur Produktbeschreibung des Herstellers www.hsh-berlin.com/modules.php?name=HSH_Content&cid=92.

¹⁵ Dazu www.afz.lvr.de/rundumsarchiv/flyer_ewoaltdatenarchivextern.pdf.

¹⁶ Neben den personalisierten Schreiben, die postalisch zugestellt wurden, ist die Empfehlung in allgemeiner Form online abrufbar unter: www.lwl.org/waa-download/pdf/Rundschreiben_Meldedaten_2010.pdf.

Beispiel ein kollektiv rechenzentrumsweit betriebenes Langzeitarchiv, benötigen eine verstetigte Finanzierung, die von konjunkturellen Schwankungen und kollektiven Sparrunden, wie sie gerade der Kulturbereich in den letzten Jahrzehnten regelmäßig erlebt hat, nicht betroffen ist.

Als nächstes Arbeitsfeld hat man sich im IT-Arbeitskreis Ostwestfalen-Lippe des Fachsystems AKDN-sozial angenommen, einer Software für Sozialämter, Jobbörsen und Jugendämter, in der v.a. die Leistungen nach SGB II und XII abgewickelt werden. Zum Teil münden die Informationen in elektronischen Fallakten in DMS-Kontexten, zum Teil besteht eine hybride Überlieferung von fallbezogenen Daten im Fachsystem und in Papierform. Das Produkt AKDN-sozial wird schwerpunktmäßig von einer eigenbetriebsähnlichen Einrichtung bei der im IT-Arbeitskreis engagierten GKD Paderborn entwickelt, so dass gute Einflussmöglichkeiten auf das Produkt bestehen.

Ein weiterer archivischer Arbeitskreis für IT-Fragen konnte 2010 in Südwestfalen für den Zuständigkeitsbereich der KDVZ Citkomm und des KDZ Westfalen-Süd eingerichtet werden (vgl. Abb. 3); er ist als Unterausschuss des Strategieausschusses verfasst und ist somit stärker institutionalisiert als sein ostwestfälisches Pendant. Bei den ersten Treffen hat man sich neben der Übernahme der *archivo-Lösung* für das Einwohnermelderegister und mit Vorüberlegungen für eine Aussonderungsschnittstelle für den Bereich des Gewerberegisters beschäftigt. Das von den meisten Städten und Gemeinden eingesetzte GERIS-Web ist eine Entwicklung der KDVZ Citkomm, auch hier bestehen also gute Einflussmöglichkeiten auf das Produkt.

Schließlich konnte beim Rechenzentrumsverbund KDN,¹⁷ in dem viele Großstädte und Kommunalverbände NRWs organisiert sind (vgl. Abb. 3), ein Arbeitskreis der Archive etabliert werden, in dem zuletzt unter anderem die Kölner Entwicklungen für ein elektronisches Langzeitarchiv vorgestellt wurden, wie es in ähnlicher Form auch beim Bundesarchiv, dem Stadtarchiv Stuttgart und im LWL-Archivamt eingesetzt wird bzw. derzeit eingerichtet wird. Auch Vorüberlegungen zur Archivierung der Katasterunterlagen aus dem bundesweit eingesetzten ALKIS wurden von den Kölner Kollegen präsentiert.

Definition von Anforderungen an die Anbietung und Aussonderung elektronischer Personenstandsregister

Das zweite Beispiel betrifft die Definition von *Anforderungen an die Anbietung und Aussonderung von elektronischen Personenstandsregistern*. Bei der Neuordnung des Personenstandswesens wurde fest-

¹⁷ Auf der Internetpräsenz des KDN (www.kdn.de) sind die folgenden Institutionen als Mitglieder aufgeführt: Die Städte Bielefeld, Bochum, Bonn, Gelsenkirchen, Hagen, Köln, Mönchengladbach, Mülheim a.d. Ruhr, Münster, Ratingen, Remscheid, Wuppertal, der Kreis Mettmann, die Rechenzentren civitec, GKD Paderborn, ITK Rheinland, kdvs Rhein-Erft-Rur, krz Minden-Ravensberg/Lippe, LVR-Info-Kom des Landschaftsverbands Rheinland, der Landschaftsverband Westfalen-Lippe und der Landeswohlfahrtsverband Hessen. Vorsitzende des Archivarbeitskreises sind Anja *Gusseke* (Stadtarchiv Münster) und Andreas *Berger* (Hist. Archiv der Stadt Köln).

gelegt, dass die Register in einem vom Fachverfahren getrennten Registerverfahren für die Zeit der Fortführung vorzuhalten und danach den nach der jeweiligen Landesregelung zuständigen Archiven anzubieten sind. Mit dem Fachverfahren – es handelt sich in nahezu 100% der Standesämter um das Produkt *Autista* – wird das Tagesgeschäft erfüllt, es leitet den Standesbeamten/die Standesbeamtin durch die Bearbeitungsschritte. Die Ergebnisse in Form der Beurkundungen werden dagegen in den Registerverfahren vorgehalten (sie ersetzen insofern ziemlich 1:1 die bisherigen Personenstandsbücher), hier gibt es derzeit drei marktgängige Verfahren: Ein Produkt des hessischen kommunalen Rechenzentrums ekom21, ein Produkt des Verlags für Standesamtswesen, der auch *Autista* betreibt, und die Firma *accenture*, die eine Lösung für landesweite Zentralregister geschaffen hat, wie sie in Bayern oder Thüringen vorgesehen sind. Gleich welches Softwareprodukt: Die Register sind nach Ende der Fortführungsfristen von den Archiven dauerhaft zu erhalten.

Bei der Frage, wer die *zuständigen Archive* sind, haben sich, wie es sich für einen föderalen Staat gehört, mehrere Verfahrensmodelle etabliert:

1. Personenstandsregister und Sicherungsregister verbleiben in kommunaler Hoheit und werden an das Kommunalarchiv bzw. an das zuständige Kreisarchiv übergeben. Bei kreisfreien Städten bleiben beide Register in städtischer Hand (z. B.: Baden-Württemberg).
2. Die Personenstandsregister übernimmt das für das Standesamt zuständige Kommunalarchiv; die Sicherungsregister werden an staatliche Archive (Staatsarchiv/Personenstandsarchiv) übergeben (z. B. Hessen oder NRW).
3. Die Personenstandsregister und Sicherungsregister gehen beide in die staatlichen Archive des jeweiligen Bundeslands (Stadtstaaten Berlin und Hamburg).

Wenn man also daran gehen will, eine Aussonderung und eine entsprechende Schnittstelle zu standardisieren, dann muss den unterschiedlichen Produktivsystemen und den unterschiedlichen potenziellen Zielsystemen und ihren organisatorischen Rahmen Rechnung getragen werden. Ein zwischen staatlichen und kommunalen Archiven abgestimmtes Konzept mit einheitlichen Vorgaben schien am erfolgsversprechendsten und langfristig tragfähigsten. Deshalb bildete sich eine ad-hoc Arbeitsgruppe mit zwei Vertretern der AG Elektronische Systeme in Justiz und Verwaltung (AG ESys)¹⁸ und einer Vertreterin und zwei Vertretern des Unterausschusses für Informationstechnologie (BKK IT-Ausschuss) der Bundeskonferenz der Kommunalarchive beim Deutschen Städtetag,¹⁹ wobei der BKK IT-Ausschuss die Federführung in den Verhandlungen übernahm.

Die Arbeitsgruppe Personenstand einigte sich zunächst auf eine erste Fassung eines Anforderungskatalogs, der alle zu vereinheitlichenden Positionen auslotete, aber bewusst die organi-

¹⁸ Ralf-Maria *Guntermann*, Landesarchiv NRW Düsseldorf / Münster und Peter *Sandner*, Hessisches Hauptstaatsarchiv Wiesbaden.

¹⁹ Miriam *Eberlein*, Stadtarchiv Heilbronn, Michael Wettengel, Stadtarchiv Ulm, Peter *Worm*, LWL-Archivamt für Westfalen sowie koordinierend der Leiter des IT-Ausschusses der BKK Robert Zink, Stadtarchiv Bamberg.

satorischen Fragen ausklammerte, die sich notwendigerweise zwischen z. B. einem Personenstandsarchiv mit vergleichsweise massenhaften Übernahmen und einem Kommunalarchiv, bei dem günstigstenfalls innerhalb des Intranet Daten verschoben werden müssen, unterscheiden. Hier bleibt Anpassungsbedarf (aber auch -spielraum), um auf die Gegebenheiten vor Ort eingehen zu können. Der Entwurf wurde zur Prüfung und Kommentierung an die BKK- und ARK-Mitglieder verteilt und die konsolidierte Fassung in ein Standardisierungsgremium eingebracht, dass die Aufgabe hatte, eine XÖV-Schnittstelle zwischen Fach- und Registerverfahren (Personenstandsregister) zu erstellen. Die Archivschnittstelle an dieses Vorhaben anzuhängen, bot zahlreiche Synergieeffekte; ein Argument, dem sich auch das zunächst skeptische Standardisierungsgremium nicht entziehen konnte. Diskutiert wurde die Frage, als wie repräsentativ und damit bindend die von uns formulierten archivischen Anforderungen angesehen werden konnten. Die Gefahr, dass Spezialanforderungen und Sonderwünsche die erwähnten Effekte zunichtemachen könnten, musste zunächst ausgeschlossen werden.

Das zweite strittige Thema war, ob eine Archivschnittstelle *schon jetzt* notwendig sei, angesichts der Tatsache, dass die Fortführungsfrist der ersten regulär geführten elektronischen Register erst nach 2040 endet. Hier half der Verweis auf Einzelbeispiele u.a. aus Hessen, bei denen Nachbearbeitungen älterer Personenstandseinträge im elektronischen System vorgenommen worden waren: Bei dieser Verwaltungspraxis wird der ursprüngliche Eintrag in Papierform komplett digital nacherfasst; die elektronische Form wird so die rechtsrelevante Form und der Papiereintrag bekommt den Status, wie ihn Sammelakten haben. Die Fristenberechnung nutzt aber das ursprüngliche Beurkundungsdatum zur Berechnung der Fortführungsfristen, so dass zumindest in Hessen schon 2013 (!) mit ersten archivreifen elektronischen Registerbeiträgen zu rechnen ist. Andere Bundesländer wie NRW sind zurückhaltender mit der Nacherfassung der Papierregister und beschränken sich v.a. auf die Heirats- und Geburtenbücher, da hier eine höhere Nachnutzbarkeit der elektronischen Einträge gegeben ist und gleichzeitig deutlich längere Fortführungsfristen die Abgabe an die archivische Zuständigkeit hinauszögern.

Ein dritter Diskussionspunkt war, ob die Archive die Register nicht in genau der Form zu übernehmen hätten, wie sie in den Standesämtern geführt worden sind. Ob also eine Umformung in langfristig stabile Formen und Formate überhaupt zulässig sei (und wenn das zu bejahen sei – wer für den Aufwand bezahlen müsse!). Nach einiger Überzeugungsarbeit, bei der v.a. das nordrhein-westfälische Archivgesetz²⁰ eine gute Argumentationshilfe darstellte, wurde das Anlie-

²⁰ ArchG NW, GV.NRW Ausgabe 2010 Nr. 11 vom 30.3.2010 Seite 183 bis 210, hier bes. § 3 (4): *Das Landesarchiv wirkt bei der Festlegung von landesweit gültigen Austauschformaten zur Archivierung elektronischer Dokumente mit. und weiter in (5): Im Rahmen seiner Zuständigkeit berät das Landesarchiv die Behörden, Gerichte und sonstigen öffentlichen Stellen des Landes bei der Verwaltung, Aufbewahrung und Sicherung ihrer Unterlagen. Die obersten Landesbehörden stellen sicher, dass die anbietenden Stellen in ihrem Geschäftsbereich die in Absatz 4 genannten Austauschformate beachten. Das gilt sowohl bei der Planung, vor der Einführung und bei wesentlichen Änderungen von IT-Systemen, die zu nach § 2 Absatz 1 i.V.m. § 4 Absatz 1 anzubietenden elektronischen Dokumenten führen. Soweit hiervon ausnahmsweise abgewichen werden soll, ist bereits vor der geplanten Nutzung anderer Formate und Techniken Einvernehmen mit dem Landesarchiv zu erzielen, um die spätere Übernahme des Archivgutes sicherzustellen. Dies entfällt, wenn Formate oder Techniken eingesetzt werden, die nach einem Verfahren nach Artikel 91 c Absatz 2 GG (Länderübergreifende Standards) abgestimmt sind.*

gen der Archive gehört und über die von uns aufgestellten Anforderungen im Standardisierungsgremium beraten. Gleichzeitig fanden flankierende Gespräche mit Vertretern des Bundesinnenministeriums und mit Vertretern des Verbands der Landesbeamten statt, um die archivischen Belange auch auf der politischen und der Ebene der Fachverwaltungen abzusichern.

Mit den Herstellern der Registerverfahren wurde darüber hinaus ein ganztägiger Workshop veranstaltet. Bei dem Treffen wurden die archivischen Vorstellungen mit den in den Verfahren tatsächlich enthaltenen und auslesbaren Daten abgeglichen und in das Konzept Vorschläge von Professor Burkhardt Renz (Technische Hochschule Mittelhessen), dem *geistigen Vater* der ersten, noch nicht XÖV-standardisierten Schnittstelle zwischen den Verfahren, eingearbeitet. Das Ergebnis wurde über mehrere Abstimmungsrunden zunächst mit den Teilnehmern des Workshops, danach mit der BKK und der ARK konsolidiert und in einer Version 2.0 auf den Internetseiten der BKK publiziert.²¹

Die Integration in den XÖV-Standard und die technische Realisierung der Schnittstelle stehen derzeit noch aus, da das Standardisierungsgremium unter Zeitdruck an der Version 1.0 der Definition von XePR arbeitete und die archivischen Anforderungen in der Kürze der Zeit nicht mehr eingearbeitet werden konnten. Nach der Abnahme der Version 1.0 ist für den Betrieb und die Weiterentwicklung des Standards nunmehr die Koordinierungsstelle für IT-Standards (KoSIT) zuständig, so dass sich unsere Ansprechpartner für die Umsetzungsarbeiten ändern.

Die Arbeit an einem Aussonderungskonzept und einer Archivschnittstelle für die Personenstandsregister hat gezeigt, dass Kooperation auf dem Gebiet der elektronischen Überlieferungsbildung anlass- und themenbezogen zwischen den Archivsparten funktioniert: Die Qualität des entstandenen Konzepts resultiert maßgeblich aus den sich gegenseitig befruchtenden Ansätzen und Vorschlägen, die von allen Mitarbeitenden eingebracht worden sind. Ebenso wertvoll waren auch die Anmerkungen von dritter Seite aus dem Kreis der Landesbeamten, der Herstellerfirmen und von Professor Renz. Darüber hinaus gehende Beratungsbedarfe gab es im bisherigen Verfahren nicht. Abschließend muss man betonen, dass die Übertragbarkeit der in der ad-hoc Arbeitsgruppe gefundenen Organisation nur bedingt gegeben ist, da eine Vermischung der archivischen Zuständigkeiten, wie sie beim Personenstandswesen besteht, die große Ausnahme darstellt.

Ergebnisse

Aus unserer Arbeit an konkreten Projekten lassen sich folgende Ergebnisse ableiten:

1. Die Definition und Durchsetzung archivischer Positionen und Anforderungen erfordert wechselnde Allianzen, da sich die Beteiligten und deren Interessen mit jedem neuen Ziel, das man mit seinem Archiv ins Auge fasst, unterscheiden. Größte Gemeinsamkeiten be-

²¹ Arbeitshilfe Anforderungen an die Anbietung und Aussonderung elektronischer Personenstandsregister und elektronisch geführter Sammelakten. Teil I: Elektronische Personenstandsregister, Status: Version 2.0 vom 2012-06-01, online abrufbar unter: www.bundeskonferenz-kommunalarchive.de/empfehlungen/Arbeitshilfe_Archivische_Anforderungen_Personenstandsregister_V2.0.pdf.

sitzen im kommunalarchivischen Bereich die Archive einer Rechenzentrumslandschaft, da sich hier die meisten Überschneidungen hinsichtlich der verwendeten Verfahren und Verfahrenswege ergeben. Auf dieser Basis lassen sich beispielsweise elektronische Langzeitarchive aufsetzen, die für kleine und mittelgroße Städte allein nicht wirtschaftlich und zuverlässig genug zu betreiben sind, oder Schnittstellen für verbundweit eingesetzte Fachverfahren durchsetzen.

2. Der regionale Bezug von auf dieser Basis eingerichteten Arbeitsgemeinschaften sichert Kontinuität und nutzt etablierte Vernetzungen der (Fach-)Verwaltungen und IT-Infrastrukturen untereinander nach. Dass das LWL-Archivamt in einer koordinierenden Funktion nicht unähnlich der KOST in der Schweiz auftreten kann, resultiert aus der jahrzehntelangen beratenden Tätigkeit, die uns bei den Mitgliedskommunen Westfalen-Lippes zu einer *festen Größe* macht. Sollte andernorts eine vergleichbare Organisation aufgebaut werden, erscheint es gerade im kommunalen Kontext sinnvoll, sie bei einer bestehenden Dachorganisation anzubinden, die Erfahrung und Anerkennung in diesen Kreisen genießt. Zu denken wäre an die kommunalen Spitzenverbände oder an die von ihnen getragene Kommunale Gemeinschaftsstelle für Verwaltungsmanagement (KGSt), die evtl. mit bestehenden oder einzurichtenden Regionalstellen arbeiten müsste. Ratschläge von einer (halb-)staatlichen Stelle können – wenn schon nicht von den Kommunalarchiven – so doch von den Archivträgern, die zumindest in NRW sehr auf ihre kommunale Selbstverwaltung bedacht sind, als Einmischung bzw. Kontrolle bewertet werden oder sogar finanzielle Forderungen nach dem Konnexitätsprinzip nach sich ziehen.
3. Punktuelle Beratungen durch externe Organisationen, die einmalig zu einer Sitzung oder einer Veranstaltung hinzukommen, können vielleicht Anregungen einbringen oder bei Spezialfragen helfen, werden aber selten dauerhafte Effekte bewirken. Vielmehr muss es mittelfristig zum Aufbau eines Know-how-Grundstocks über und Engagements für elektronische Archivierung bei den Akteuren in den Regionen kommen, das sich nach einer Entwicklungsphase (5-10 Jahre) selbst trägt und eigenständig fortgeführt wird. Beratung braucht neben Vertrauen also eine gewisse Verstetigung.
4. Eine wie auch immer geartete Beratungs- oder Koordinierungsstelle für die elektronische Überlieferung wird nicht in der Lage sein, strukturelle Defizite, die gerade im Archivwesen des ländlichen Raums bestehen, in den Griff zu bekommen. Funktionieren die archivischen Arbeitsprozesse im analogen Bereich nicht und gibt es kein halbwegs qualifiziertes Personal vor Ort, wird auch die elektronische Archivierung keine guten Ergebnisse bringen und spätestens im Bereich der Nutzung ins Leere laufen. Es sind Ansätze notwendig, die die Papier- und die Datenwelt im Blick haben
 - bei der Kommunikation mit den Registraturbildnern und den datenhaltenden Stellen,
 - bei der Bewertung und Übernahme,
 - bei der Bestandserhaltung und der Nutzung.

Brauchen wir Koordinierungsstellen für die digitale Archivierung? Zusammenfassung der Diskussion¹

Von ULRICH SCHLUDI

Angesichts der großen Herausforderungen, vor welche die digitale Archivierung die Archive stellt, sprachen sich die Teilnehmer eines nestor-Workshops Mitte Juli 2012 in Stuttgart mehrheitlich für die Einrichtung von Koordinierungsstellen zur digitalen Archivierung auch in Deutschland aus. Diese Koordinierungsstellen sollten einen nachhaltigen Aufbau von Wissen ermöglichen, spartenübergreifend und zweistufig organisiert sein und ein breites Angebot ähnlich der Schweizer KOST bieten. Die Diskussion hierüber soll nun in die Organisationen und Gremien der verschiedenen Archivsparten hineingetragen werden.

Unter der Leitfrage *Brauchen wir Koordinierungsstellen für die digitale Archivierung?* trafen sich am 18. Juli 2012 knapp 50 Archivare aus verschiedenen Archivsparten und Regionen Deutschlands, aber auch aus der Schweiz und Österreich zu einem nestor-Workshop in Stuttgart. Robert Kretschmar forderte in seiner Begrüßung dazu auf, Erfahrungen, die in der Schweiz, Westfalen oder anderswo gesammelt worden seien, zu rezipieren und die Diskussion in Deutschland auf dieser Basis voranzubringen. Ziel dieses ergebnisoffenen Diskussionsprozesses müsse es sein, zu verbesserten Strukturen der Koordination und Zusammenarbeit auf dem Gebiet der digitalen Archivierung zu gelangen.²

In seinem Einführungsreferat begründete Christian Keitel die Notwendigkeit von Koordination und Kooperation bei der digitalen Archivierung mit den hohen Kosten und dem großen Abstimmungsbedarf zwischen den Archiven. Gemeinsam überdacht werden müssten aber auch die archivischen Arbeitsprozesse, werde die digitale Archivierung doch das archivische Arbeiten in allen Arbeitsbereichen verändern.³ Georg Büchler genauso wie Peter Worm stellten in ihren Vorträgen die praktischen Erfahrungen vor, die die *Koordinierungsstelle für die dauerhafte Archivierung elektronischer Unterlagen* (KOST) in der Schweiz und das LWL-Archivamt in Westfalen in den letzten Jahren gemacht haben. Peter Worm stellte dabei die Wichtigkeit einer regionalen Verortung von Koordinierungsstellen heraus. Diese allein stelle einen

¹ Der vorliegende Bericht wurde zuerst publiziert in der Zeitschrift *Archivar* 66 (2013), S. 67–70. Die hier abgedruckte Fassung wurde um die Besprechung der drei in diesem Band abgedruckten Referate gekürzt und gleichzeitig am Ende unter Berücksichtigung der zwischenzeitlich eingetretenen Entwicklungen geringfügig erweitert; vgl. auch unten S. 316.

² Abgedruckt in diesem Band, ab S. 279.

³ Vgl. die Schriftfassung oben, S. 281–288.

nachhaltigen Erfolg der Beratungstätigkeit sicher.⁴ Georg Büchler hingegen betonte, dass die KOST ihr Angebot bewusst auf die Vielfalt ihrer Träger ausrichte. So habe sich ein Mix ganz verschiedener Angebote als der Schlüssel zum Erfolg erwiesen.⁵

In der anschließenden Diskussion dominierte zunächst die Frage, inwiefern nichtstaatliche oder nicht beteiligte Archive aus der Arbeit der KOST Nutzen ziehen können. Büchler wies in diesem Zusammenhang daraufhin, dass die Angebotspalette der KOST so breit sei, dass vieles nicht nur für die staatlichen, sondern auch für die kommunalen Archive nutzbar und hilfreich sei. Aber auch wer keinen Beitrag zur KOST leiste, könne von ihr profitieren: So würden die Arbeitsergebnisse allen Archiven über die Webseite der KOST zur Verfügung gestellt. Wer die Mitgliedschaft fachlich bewältigen könne, sei aber inzwischen ohnehin Mitglied geworden. Für die übrigen kleinen und kleinsten Archive habe die KOST keinen Auftrag. Diese Lücke füllten in der Schweiz die Kantonalarchive. Jene drängten die kleineren Gemeindearchive zur Zusammenarbeit auf Kantonalebene, und auch für die Universitätsarchive seien die Staatsarchive der Kantone zuständig.

Bezüglich des Referats von Peter Worm wiesen mehrere Teilnehmer des Workshops auf die Parallele zwischen der Schweizer KOST und der Beratungstätigkeit des LWL-Archivamts zur digitalen Archivierung hin. Westfalen-Lippe habe also schon eine KOST, so das Fazit. Da es in Deutschland aber nirgendwo anders solche Strukturen gäbe, sprächen Bedarf und Nachfrage nach dem Beratungsangebot des LWL-Archivamts gerade für Koordinierungsstellen in Deutschland. Die bereits existierenden Verbände könnten eine solche Aufgabe aufgrund ihrer Struktur nicht übernehmen.

Kritisch hinterfragt wurde aus dem Zuhörerkreis, inwiefern die im Vortrag von Peter Worm erläuterte Vorgehensweise des LWL-Archivamts bei der Übernahme elektronischer Personenstandsregister auf andere Projekte übertragen werden könne. Dass es in diesem Fall gelungen sei, unter Einbeziehung der betreffenden Arbeitskreise der Bundeskonferenz der Kommunalarchive (BKK) und der Archivreferentenkonferenz (ARK) zu einer einheitlichen Lösung zu kommen, die man dann gegenüber den Herstellern auch noch durchzusetzen vermochte, sei vor allem den guten Kontakten des Referenten und dem Ansehen des LWL-Archivamts zuzuschreiben. Kopieren könne man eine solche Vorgehensweise aber nicht. Anderswo brauche es schon die zur Diskussion stehenden Koordinierungsstellen, die sowohl regionale Schlagkraft entfalten als auch die Basis für eine bundesweite Kooperation darstellen müssten. Ohnehin aber hätten solche Projekte Auswirkungen auf das ganze Bundesgebiet – und gerade deshalb könne ein regionaler Arbeitskreis nicht allein vorangehen, sondern man benötige eine Organisation wie die Schweizer KOST.

⁴ Auch das Referat von Peter Worm findet sich in diesem Band, siehe oben, S. 297–309.

⁵ Abgedruckt oben, S. 289–295.

Die zentrale Frage: Brauchen wir Koordinierungsstellen?

In der anschließenden Generaldebatte über das Tagungsthema wurde deutlich, dass die Vertreter der versammelten Archive die zentrale Frage des Workshops fast ausnahmslos mit Ja beantworteten. Der Tenor der Workshop-Teilnehmer war hier eindeutig: Zu klein sei die Personalausstattung der allermeisten Archive in Deutschland, als dass sie das Projekt *Digitale Archivierung* allein bewältigen und das dafür nötige Know-how allein erarbeiten könnten.

Eine andere Lageeinschätzung vertraten nur wenige Anwesende, darunter das Bundesarchiv und einige Archive aus Nordrhein-Westfalen. Von dieser Seite wurde gegen eine Koordinierungsstelle vorgebracht, dass man selbst über genügend Mittel und Personal verfüge, um diese Aufgaben anzugehen, für sich selbst also keine KOST benötige. Außerdem sei die digitale Archivierung eine gesetzliche Pflichtaufgabe. Infolgedessen müsse man sich dieser Herausforderung selbst stellen. Ergänzend sei man gerne dazu bereit, Software-Entwicklungs- und Anwenderpartnerschaften anzubieten, in denen sich Archive zusammentun könnten, die vor ähnlichen Problemen stünden und ähnliche Lösungsansätze suchten. Vertreten wurde aber auch, dass für die digitale Archivierung ad hoc-Arbeitsgruppen ausreichten. Außerdem würden die Aufgaben einer Koordinierungsstelle in ihrem regionalen Bereich schon vom Archivamt für Westfalen-Lippe wahrgenommen. Für den Landesverband seien diese Bedürfnisse insofern schon abgedeckt.

Dieser Argumentation wurde entgegengehalten, dass die personelle und finanzielle Ausstattung jener Archive weit jenseits der Möglichkeiten fast aller anderen Archive in Deutschland lägen. Wenn jene also keinen Bedarf an Koordinierungsstellen sähen, so sage dies nichts über den Bedarf der übrigen deutschen Archive aus. Selbst die wenigen großen Kommunalarchive in Deutschland, die personell vergleichsweise gut aufgestellt seien, könnten das Projekt *Digitale Archivierung* nicht alleine schultern. Viel schwieriger aber sei die Lage für die große Zahl kleiner und kleinster Archive, seien es Stadt-, Gemeinde- oder Kreisarchive, Wirtschafts-, Parlaments- oder Universitätsarchive. Ihre Personalausstattung reiche häufig schon kaum dafür aus, die archivischen Fachaufgaben in der analogen Welt zu bewältigen. Würde man sie mit der neuen Herausforderung der digitalen Archivierung alleine lassen, drohten sie zu scheitern – das Ergebnis wäre ein schwarzes Loch in der Überlieferung, wie ein Vertreter der Universitätsarchive bemerkte.

Der Status einer gesetzlichen Pflichtaufgabe hindere im Übrigen nicht daran, Kooperations- und Koordinierungsgremien aufzubauen oder sich bei einer anderen Stelle Rat zu holen. Noch dazu sei nicht einmal jedes Staatsarchiv in der Lage, dem gesetzlichen Pflichtauftrag nachzukommen und die digitale Archivierung allein zu bewältigen, so die Vertreter zweier kleinerer Landesarchive, die für ihre Archive selbst Bedarf an einer KOST anmeldeten. Sie sahen sich auch außerstande, die immer mehr zunehmenden Anfragen auf Unterstützung beim Aufbau eines digitalen Archivs positiv zu beantworten. In der gegenwärtigen Finanzsituation der öffentlichen Hand sei es zudem nicht mehr vertretbar, wie früher zahlreiche Parallellösungen in Deutschland aufzubauen. Stattdessen müsse man den Weg von Kooperationslösungen beschreiten.

Die bestehenden Strukturen aber führten hier nicht weiter, wie die Vertreter mehrerer Kommunalarchive betonten. Es brauche vielmehr Personen, die sich dieser Aufgabe zu 100% und

unabhängig widmeten. Deshalb hülfe auch kein neuer Arbeitskreis. Dieser könne eben auch nur mit denselben schon einschlägig bekannten Personen besetzt werden, die aber im Alltag bereits in ganz anderen Aufgaben stünden und gar nicht die Kapazität hätten, darüber hinaus noch effektive Hilfe bei der digitalen Archivierung in anderen Archiven zu leisten.

Ad hoc-Arbeitsgruppen seien ebenso wenig eine Lösung. Eine kontinuierliche Unterstützung, die auf nachhaltige Erfolge und den Aufbau von Know-how in den einzelnen Archiven abziele, sei mit ihnen nicht zu erreichen. Bei einer solchen Mammut-Aufgabe brauche es eine dauerhafte, verstetigte Einrichtung.

Dasselbe gelte für die vorgeschlagenen Entwicklungs- und Anwenderpartnerschaften. Kleineren und kleinsten Archiven sei es weder finanziell, personell noch von ihrem Kenntnisstand her möglich, sich an solchen Software-Partnerschaften zu beteiligen und das Tempo der großen Archive mitzugehen.

Innerhalb von nestor seien solche Ziele im Übrigen auch nicht erreichbar. Der Ansatz von nestor sei gedächtnisinstitutionen-übergreifend und nicht archivspezifisch. Dazu komme, dass nestor überhaupt nicht die Absicht verfolge, Lösungswege für die Praxis zu erarbeiten. Nestor entwickle vielmehr Leitlinien, deren praktische Ausgestaltung Sache der beteiligten Bibliotheken, Museen und Archive sei. So hätten die Bibliotheken inzwischen parallel zu nestor spartenspezifische Schritte zur digitalen Archivierung unternommen, um auf diese Weise konkrete Mittel und Wege für die bibliothekarische Praxis zu finden. Als Beispiel wurde das Projekt dp4lib genannt. Genau dasselbe müssten auch die Archive tun.

Umfrage nach Archivsparten

Nach diesen ersten Stellungnahmen bat der Leiter des Workshops, Christian Keitel, die anwesenden Vertreter der verschiedenen nichtstaatlichen Archive um Stellungnahmen im Hinblick auf ihre jeweilige Archivsparte. In ihrer Antwort plädierten die anwesenden Vertreter der Kommunal-, Universitäts-, Wirtschafts- und Parlamentsarchive, die das Wort ergriffen, fast einhellig für die Einrichtung von Koordinierungsstellen. Dabei legten die Kommunalarchive Wert darauf, dass man die kommunalen (Spitzen-)Verbände in die Pläne einbeziehe. Von Seiten des Vertreters der Universitätsarchive wurde betont, dass man in dieser Archivsparte derzeit noch sehr weit von Lösungen entfernt und finanziell auch gar nicht in der Lage sei, das Projekt *Digitale Archivierung* alleine voranzutreiben. Der Weg der Universitätsarchive führe momentan geradewegs in den Datenverlust. Ähnlich großen Bedarf meldeten zwei Teilnehmer aus der Gruppe der Wirtschaftsarchive an. Universitäts- und Wirtschaftsarchive plädierten im Übrigen für Koordinierungsstellen, die archivspartenübergreifend organisiert seien. Schließlich seien auch die Probleme archivspartenübergreifend – und ihre Teilnahme sonst unter Umständen gefährdet, wenn sich staatliche und kommunale Archive in je eigenen Zirkeln zusammenschlossen.

Welche Unterstützung brauchen wir?

An diese Umfrage nach Archivsparten schloss sich eine Diskussion an, welche Unterstützung sich die deutschen Archive von Koordinierungsstellen zur digitalen Archivierung erwarten würden.

Die einzelnen Vertreter sprachen sich zunächst dafür aus, dass solche Koordinierungsstellen die Archive bei der digitalen Archivierung beraten sollten. Außerdem solle es wie in der Schweiz möglich sein, dass sich einige Archive mit einer Koordinierungsstelle zusammenschließen und gemeinsam die Lösung für ein bestimmtes Problem entwickelten (good practice); Koordinierungsstellen sollten daher auch über das Wissen verfügen, um die verschiedenen Archive, die an ein und demselben Problem arbeiten, zusammenzubringen. Sie sollten ferner eine Bündelungsfunktion wahrnehmen, an der Entwicklung von Standards arbeiten, Leitlinien erstellen, Tools entwickeln und ganz allgemein eine Informationsplattform zur digitalen Archivierung darstellen.

Spartenübergreifend und mehrstufig

Auf die Frage, ob Koordinierungsstellen eher spartenbezogen oder doch besser regional ausgerichtet sein sollten, sprachen sich viele Wortmeldungen für eine regionale und archivspartenübergreifende Struktur aus. Die Probleme der digitalen Archivierung seien schließlich spartenübergreifend dieselben und am leichtesten in regionaler Zusammenarbeit zu lösen. Gleichzeitig optierten Wortmeldungen für ein zweistufiges Modell: Eine Stelle auf einer übergeordneten Ebene könne beispielsweise Standardisierungs- und Bündelungsfunktionen wahrnehmen, während die regionalen Koordinierungsstellen Pilotlösungen entwickeln, beraten und Schulungen durchführen sollten. Die regionale Zuständigkeit dieser Koordinierungsstellen müsste allerdings nicht unbedingt mit den Bundesländern deckungsgleich sein; sie sollte sich am besten an bestehende Strukturen der Zusammenarbeit anlehnen, so ein Vorschlag.

Weitere Vorgehensweise

Angeregt wurde aus dem Teilnehmerkreis, das Thema an ARK und BKK heranzutragen und den Meinungsstand auch sonst auf allen Ebenen und in allen Gruppierungen auszuloten. Prof. Dr. Wilfried Reininghaus (Präsident des Landesarchivs Nordrhein-Westfalen und Vorsitzender der nächsten Archivreferentenkonferenz) bot an, das Thema auf die Tagesordnung der Herbsttagung der ARK Ende September in Köln zu setzen. Dr. Robert Zink (Leiter des Stadtarchivs Bamberg, stellvertretender Vorsitzender der BKK) schlug vor, gleiches für die BKK zu übernehmen. Dies stieß allgemein auf Zustimmung. ARK und BKK könnten sich dann über das weitere Vorgehen verständigen.

Sachstand zum 1.4. 2013

Zwischenzeitlich ist diese Verständigung erfolgt und haben sich die Bundeskonferenz der Kommunalarchive beim deutschen Städtetag und die Archivreferentenkonferenz gegen eine zentrale Koordinierungsstelle mit bundesweiter Zuständigkeit und für eine Koordinierung im Rahmen regionaler Verbünde für die digitale Archivierung ausgesprochen.

Autorinnen und Autoren

Dr. Peter Bobl, Archivar, ist als Referatsleiter im Hauptstaatsarchiv Stuttgart zuständig für nichtstaatliches Archivgut und leitet seit 1995 das Audiovisuelle Archiv des Landesarchivs Baden-Württemberg.

Landesarchiv Baden-Württemberg, Abt. Hauptstaatsarchiv Stuttgart
Konrad-Adenauer-Straße 4, D-70173 Stuttgart

Georg Büchler, lic. phil., ist seit 2004 als wissenschaftlicher Mitarbeiter und Berater für die Koordinationsstelle für die dauerhafte Archivierung elektronischer Unterlagen KOST tätig.

Koordinationsstelle für die Archivierung elektronischer Unterlagen KOST

c/o Schweizerisches Bundesarchiv

Archivstrasse 24, CH-3003 Bern

Dr. Katharina Ernst, Archivarin, leitet die Dienststelle „Amtliches Schriftgut“ im Stadtarchiv Stuttgart. Sie ist Obfrau des DIN Normenausschusses NABD15 zur „Schriftgutverwaltung und Langzeitverfügbarkeit digitaler Informationsobjekte“.

Landeshauptstadt Stuttgart, Kulturamt – Stadtarchiv

Bellingweg 21, D-70372 Stuttgart

Susanne Fröblich, Mag. phil., Archivarin, seit 1999 im Österreichischen Staatsarchiv, Arbeitsschwerpunkte Schriftgutübernahme, Bewertung, Skartierung und Erschließung; seit 2007 Projektkoordinatorin und Fachexpertin für das „Digitale Archiv Österreich“ sowie Koordinatorin für Bewertung und digitale Aktenübernahmen im Archiv der Republik.

Österreichisches Staatsarchiv, Abt. Archiv der Republik

Nottendorfer Gasse 2, A-1030 Wien

Jörg Homberg, Dokumentar, IT-Bereichsleiter im Brandenburgischen Landeshauptarchiv und Leiter des Projekts EL.UZA und verantwortlich für den Aufbau des digitalen Archivs (EL.ARCHIV).

Brandenburgisches Landeshauptarchiv

Zum Windmühlenberg, D-14469 Potsdam

Karsten Huth, Dokumentar, ist im Sächsischen Staatsarchiv Abt. Zentrale Aufgabe, Grundsatz für die Archivierung elektronischer Unterlagen zuständig. Im Rahmen des Projekts LeA leitet er das Teilprojekt „Elektronische Archivierung“. Er ist Mitglied des DIN Normenausschusses NABD15 zur „Schriftgutverwaltung und Langzeitverfügbarkeit digitaler Informationsobjekte“.

Sächsisches Staatsarchiv, Zentrale Aufgaben, Grundsatz

Archivstraße 14, D-01097 Dresden

Lambert Kansy, lic. phil., Archivar, arbeitet seit 2000 im Staatsarchiv Basel-Stadt, ist seit 2008 Leiter der Abteilung Informatik und zuständig für den Bereich Archivinformatik.

Staatsarchiv des Kantons Basel-Stadt
Martinsgasse 2, CH-4001 Basel

Dr. Christian Keitel, Archivar, leitet im Landesarchiv Baden-Württemberg Abt. Fachprogramme und Bildungsarbeit das Referat Überlieferungsbildung und koordiniert die digitale Archivierung im Landesarchiv Baden-Württemberg.

Landesarchiv Baden-Württemberg, Abt. Fachprogramme und Bildungsarbeit
Eugenstraße 7, D-70182 Stuttgart

Corinna Knobloch, Archivarin, 2002-2011 Tätigkeit beim Staatsarchiv Sigmaringen, seitdem im Staatsarchiv Ludwigsburg für die Übernahme und Archivierung digitaler Unterlagen von Behörden, Gerichten und sonstigen Landesstellen zuständig.

Landesarchiv Baden-Württemberg, Abt. Staatsarchiv Ludwigsburg
Arsenalplatz 3, D-71638 Ludwigsburg

Prof. Dr. Robert Kretzschmar, Archivar, Präsident des Landesarchivs Baden-Württemberg seit 2005, zuvor Leiter des Hauptstaatsarchivs Stuttgart; Arbeitsschwerpunkte u.a. Archivwissenschaft, Überlieferungsbildung und Historische Hilfswissenschaften.

Landesarchiv Baden-Württemberg
Eugenstraße 7, D-70182 Stuttgart

Rolf Lang, Dipl. Ing. (FH), ist für die Entwicklung und Programmierung von IT-Programmen zur digitalen Archivierung zuständig (DIMAG, IngestList u.a.).

Landesarchiv Baden-Württemberg, Abt. Verwaltung
Eugenstraße 7, D-70182 Stuttgart

Markus Loch, Dipl. Ing., Informatiker, ist seit 2008 im Staatsarchiv Basel-Stadt zuständig für den Aufbau und Betrieb der IT-Infrastruktur.

Staatsarchiv des Kantons Basel-Stadt
Martinsgasse 2, CH-4001 Basel

Heike Maier, Dokumentarin, ist zuständig für die digitale Langzeitarchivierung im Stadtarchiv Stuttgart.

Landeshauptstadt Stuttgart, Kulturamt – Stadtarchiv
Bellingweg 21, 70372 Stuttgart

Dr. Kai Naumann, Archivar, ist im Landesarchiv Baden-Württemberg Abt. Staatsarchiv Ludwigsburg für die Übernahme und Archivierung digitaler Unterlagen von Behörden, Gerichten und sonstigen Landesstellen zuständig.

Landesarchiv Baden-Württemberg, Abt. Staatsarchiv Ludwigsburg
Arsenalplatz 3, D-71638 Ludwigsburg

Dr. Burkhard Nolte, Archivar, ist im Sächsischen Staatsarchiv Abt. Zentrale Aufgaben, Grundsatz u. a. für Grundsatzfragen der Überlieferungsbildung und für die Archivierung elektronischer Unterlagen zuständig. Im Rahmen des Projekts LeA hat er die Gesamtprojektleitung inne.
Sächsisches Staatsarchiv, Zentrale Aufgaben, Grundsatz
Archivstraße 14, D-01097 Dresden

Johannes Renz, Archivar, ist Sachgebietsleiter im Hauptstaatsarchiv Stuttgart (Referat Sonderarchivgut/Nutzung) mit Schwerpunkt Digitales Archivgut.
Landesarchiv Baden-Württemberg, Abt. Hauptstaatsarchiv Stuttgart
Konrad-Adenauer-Straße 4, D-70173 Stuttgart

Bernhard Rieder, lic. phil., Historiker, stellvertretender Abteilungsleiter Überlieferungsbildung.
Staatsarchiv des Kantons Zürich
Winterthurerstrasse 170, CH-8057 Zürich

Claire Röhli Berger-Jourdan, Ingenieurin (FH) im Bereich der Mechatronik (Mechanik, Elektronik, Informatik), ist seit 2010 in der Koordinationsstelle für die dauerhafte Archivierung elektronischer Unterlagen (KOST) hauptsächlich für die Formate und Validierungstools verantwortlich.
Koordinationsstelle für die Archivierung elektronischer Unterlagen KOST
c/o Schweizerisches Bundesarchiv
Archivstrasse 24, CH-3003 Bern

Dr. Peter Sandner, Archivar, seit 2006 am Hessischen Hauptstaatsarchiv in Wiesbaden, ab 2010 dort Leitung der Abteilung „Archivfachliche IT / Digitales Archiv“
Hessisches Hauptstaatsarchiv
Mosbacher Straße 55, D-65187 Wiesbaden

Dr. Sigrid Schieber, Archivarin, 2005-2009 Projektstätigkeit im Landesarchiv Baden-Württemberg (Schwerpunkt Portale), Leiterin der Koordinierungsstelle Retrokonversion und des Stadtarchivs Kassel, seit November 2009 zuständig für den Aufbau des „Digitalen Archivs Hessen“ beim Hessischen Hauptstaatsarchiv in Wiesbaden.
Hessisches Hauptstaatsarchiv
Mosbacher Straße 55, D-65187 Wiesbaden

Dr. Ulrich Schludi, Archivar, 2008-2012 Leiter des Unternehmensarchivs der Ravensburger AG, seit 1. April 2012 Landesarchiv Baden-Württemberg mit den Schwerpunkten Pressestelle, Überlieferungsbildung und Nutzung.
Landesarchiv Baden-Württemberg, Abt. Fachprogramme und Bildungsarbeit
Eugenstraße 7, D-70182 Stuttgart

Dr. Christoph Schmidt, Archivar, gehört dem „Kompetenzteam Elektronische Unterlagen“ des Landesarchivs NRW an und ist für Grundsatzfragen der elektronischen Archivierung zuständig. Landesarchiv Nordrhein-Westfalen, Fachbereich Grundsätze
An den Speichern 11, D-48157 Münster

Karlheinz Schmitt, Informatiker, Projektmanager und ITIL Service Manager, seit 2009 an der Deutschen Nationalbibliothek verantwortlich für Langzeitarchivierungsdienste für Dritte, 2010-2012 Projektleitung für „Digital Preservation for libraries“ (DP4lib), momentan für den Aufbau eines Service- und Betriebsmodells einschließlich Kostenermittlung zuständig.
Deutsche Nationalbibliothek, Informationstechnik
Adickesallee 1, D-60322 Frankfurt am Main

Steffen Schwalm, Senior Consultant bei der BearingPoint GmbH mit den Schwerpunkten Schriftgutverwaltung, Compliance und Beweissicherheit elektronischer Unterlagen, Langzeitspeicherung und elektronische Archivierung sowie Standardisierung und Normung für Kunden der öffentlichen Verwaltung.
BearingPoint, Management & Technology Consultants
Kurfürstendamm 207-208, D-10719 Berlin

Felix Stadler, Informations- und Dokumentationsspezialist (FH), arbeitet im Staatsarchiv St. Gallen, Abteilung digitale Langzeitarchivierung.
Kanton St. Gallen, Departement des Innern, Amt für Kultur, Staatsarchiv
Regierungsgebäude, CH-9001 St. Gallen

Ilka Stahlberg, Archivarin, ist Abteilungsleiterin und IT-Beauftragte im Brandenburgischen Landeshauptarchiv, arbeitet schwerpunktmäßig im Bereich digitale Überlieferungsbildung und Records Management.
Brandenburgisches Landeshauptarchiv
Zum Windmühlenberg, D-14469 Potsdam

Dr. Gabriele Stüber, Archivarin, Leiterin des Zentralarchivs der Evangelischen Kirche der Pfalz, Speyer; Federführung im Arbeitskreis „Kirchenbuchportal“ des Verbandes kirchlicher Archive, Mitglied im Arbeitskreis Archivierung von Unterlagen aus digitalen Systemen seit 1997.
Zentralarchiv der Evangelischen Kirche der Pfalz
Domplatz 6, D-67346 Speyer

Dr. Peter Worm, Archivar, als Referent in der Archivberatung zuständig für die kommunalen und privaten Archive in den Kreisen Gütersloh, Herford, Soest und Warendorf und in den kreisfreien Städten Bielefeld und Münster sowie für die Querschnittsaufgabe „Elektronische Unterlagen“.
Landschaftsverband Westfalen-Lippe, LWL-Archivamt für Westfalen
Jahnstr. 26, D-48147 Münster

Mike Zuchet, M.A., Dokumentar, arbeitet im Referat Organisationsbestände des Archivs der sozialen Demokratie und ist darüber hinaus für Fragen der digitalen Langzeitarchivierung zuständig.
Archiv der sozialen Demokratie der Friedrich-Ebert-Stiftung
Godesberger Allee 149, D-53175 Bonn

