

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.

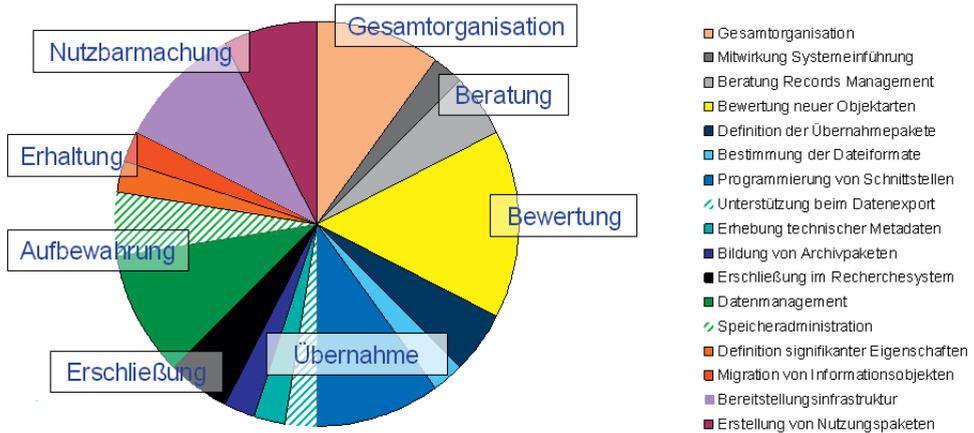


Abb. 1: Anteile am Zeit- und Personalaufwand des Digitalen Archivs Hessen (Schätzwerte für die Jahre 2010 und 2011).⁵

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 Zusatz- aufwand 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.

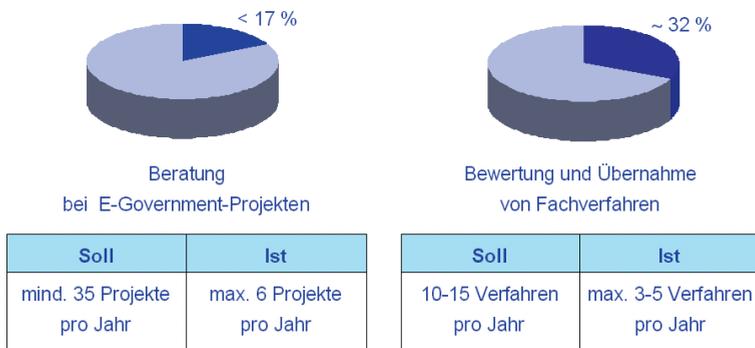


Abb. 2: Zielerreichung angesichts vorhandener Personalkapazitäten (Angaben für das Digitale Archiv Hessen 2010/2011)

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 Anbietetung 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 Anbietungspflicht 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 Anbietung 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 (Anbietung 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 vernachlässigbare

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.