

Bewertungskriterium Standardformat? Die Auswirkungen der Format- und Schnittstellenproblematik auf die Aussonderung und die Auswertbarkeit elektronischer Unterlagen im Digitalen Archiv

Dr. Michael Puchta M.A.
(Generaldirektion der Staatlichen Archive Bayerns)



Abgabe- und Archivformate – ein Dauerbrenner

Brandenburgisches Landeshauptarchiv setzt auf PDF/A bei der Langzeitarchivierung

Vorgabe des Transferformates war der gesuchte Königsweg

Seit 2009 werden elektronische Normen als PDF/A-Daten archiviert. Der pdfaPilot der callas software GmbH stellt dabei sicher, dass ausschließlich Daten in das Archiv gelangen, die dem PDF/A-Format als ISO-Standard entsprechen. Das Projekt hat aufgrund seiner ganzheitlichen Abbildung des Prozesses Pilotcharakter. Andere Bundesländer können von den Erfahrungen profitieren.

Das Brandenburgische Landeshauptarchiv (BLHA) ist das zentrale staatliche Archiv des Landes Brandenburg und zuständig für das Archivgut aller Stellen des Landes sowie ihrer Rechts- und Funktionsvorgänger.

Als die Bekörde 2009 das Projekt startete, betraf man auszunag eine „große Wiese“. Bis zu diesem Zeitpunkt erfolgte die Befreiung papiergebundener Akten konventionell durch einen Boten. Um einer etwaigen Flut an verschickten zu erwartenden Postkassen zu begegnen, wurden Mittel und Wege zur Konsolidierung und Konsolidierung gesucht und gefunden. „Die Herangehensweise, wie wir die Dokumente nachhaltig archivieren können, ohne in einer Flut von unterschiedlichen Formaten zu ertrinken“, erinnert sich Jörg Homburg, zuständiger Projektleiter des Brandenburgischen Landeshauptarchivs. „Als Königsweg erwies sich die Vorgabe und Festlegung eines Transferformates, ein weiterer Weg war wegen der Personal-Kapazität und der räumlichen Gegebenheiten denkbar gewesen.“

Das Land Brandenburg hat als erstes Bundesland mit der Elektronischen Normverkündung (EL.NORM) einen durchgehenden Prozess von der Ausfertigung über die Verkündung bis zur Langzeitpeicherung realisiert. Dabei haben sich die Anforderungen an elektronische Dateien und der Umgang damit sowie die Übergabe dieser Dateien an das Brandenburgische Landeshauptarchiv (BLHA) maßgeblich verändert.



Blick über Potsdam

So gelang es, die Zahl der zu bewerkstellenden Formate massiv zu reduzieren. (Quelle: Brandenburgisches Landeshauptarchiv)

auswachen Schriftgutverwaltung und Langzeitverfügbarkeit digitaler Archivdaten (Quelle: INADITION AG)

den über Dokumente elektronisch im PDF/A-Format zur Bewertung auf Archivverträglichkeit, Übernahme und An-

Konformität überprüft. Dazu hat man sich auf verschiedene Versuchsanordnungen unterschiedliche Selbstverständnisse angeschaut und stellt hierbei unter anderem auf die callas software aus Berlin. Das Unternehmen liefert bereits seit 1997 Prüftechnologien, war außerdem für Adobe und ist auch Hersteller des Preflight-Moduls, das seit der Acrobat Version 8 die Konvertierung nach PDF/A-1 und deren Prüfung ermöglicht. Dasselben Prüftechnologie hat callas auch in seine eigene Validierungssoftware, den pdfaPilot, integriert.

„Acrobat hat in unserem Bundesland großen Zuspruch, und da der pdfaPilot auf derselben Technologie basiert, haben wir uns mit der Software näher auseinandergesetzt.“ Schließlich wurde der für das BLHA im Projekt EL.NORM zuständige Partner, die H&T Greenline GmbH, be-

Arbeitgeber von Jörg Homburg. „Wir haben uns aus Sicherheitsgründen für beide Varianten entschieden, um bei Fehlermeldungen nach der Validierung über die Server-Version nachschauen zu können“, so Homburg.

Dabei prüft die Serverversion alle eingehenden Dokumente. Die Dokumenten, die Fehler aufweisen, prüft Jörg Homburg dann noch einmal stichprobenartig. Sicher werden die Ergebnisse beider Varianten zu 100 Prozent synchron.

Gibt der pdfaPilot bei der Prüfung nach PDF/A-Konformität Fehlermeldungen an, wird die Datei für die Archivierung abgelehnt und der Absender über die Ursachen informiert. Oft fehlen Schriftarten, in Dokumenten oder Farbprofile bei digitalisierten großen Karten, die eine langfristige Reproduzierbarkeit gefährden. „Der pdfaPilot protokolliert den Validierungsprozess sehr detailliert und seine Fehlermeldungen sind verständlich“, erläutert der Projektleiter. „So können wir dem entsprechenden Mitarbeiter noch genau sagen, wie das Problem zu lösen ist. Die Qualität der eingehenden PDF/A-Dokumente konnte im Laufe der Zeit sehr stark verbessert werden.“

Das BLHA setzt auf den Normvoll PDF/A-1a, der sowohl die standortige Visuelle Reproduzierbarkeit als auch Abbildbarkeit von Text nach Unicode und die inhaltliche Strukturierung des Dokuments vorschreibt.

„Da die Ausfertigung der Gesetze nur noch elektronisch, also ohne ein entsprechendes Papierderivat, erfolgt, muss sichergestellt sein, dass die Gesetzesdokumente so archiviert werden, dass sie über Jahre hinweg stark zu lesen wie das ursprüngliche Dokument“, so Homburg. „Darüber hinaus sind wir mit PDF/A-1a bestrebt, darauf zu achten, wenn es darum geht, die Inhalte barrierefrei zur Verfügung zu stellen.“

Ausblick

Darüber prüft das Land Brandenburg, ob es die erforderlichen Konformitätsprüfungen, die die Ministerien für die PDF/A-Ermittlung aus-

Abgabe- und Archivformate – Auswahl an Empfehlungen aus dem dt. Sprachraum

- **DOMEA-Organisationskonzept 2.0: „Technische Aspekte der Archivierung elektronischer Akten“ (2004)**
- **Empfehlung des AK "Elektronische Archivierung" zur Verwendung von Textformaten bei der Archivierung elektronischer Unterlagen (2006)**
- **Praktische Lösungsansätze zur Archivierung digitaler Unterlagen des AK "Elektronische Archivierung" (2007)**
- **Katalog archivischer Dateiformate (KaD) der KOST (2010/2013)**
- **BSI Technische Richtlinie 03125: „Beweiswerterhaltung kryptographisch signierter Dokumente - Anlage TR-ESOR-F: Formate und Protokolle“ (2011)**
- **„SAGA-Modul Technische Spezifikationen“ (2011)**

Organisatorisch-theoretische Konzepte zum Datenaustausch zwischen Provenienzbildner und Archiv

- **„Producer-Archive Interface Methodology Abstract Standard“ - PAIMAS (ISO 20652:2006)**
- **nestor-Leitfaden „Wege ins Archiv - Ein Leitfaden für die Informationsübernahme in das digitale Langzeitarchiv“ (2008)**

Die Auswirkungen der Format- und Schnittstellenfrage auf ausgewählte archivische Kernaufgaben

- 1. Wie beeinflusst die Format- und Schnittstellenfrage die archivische Bewertung?**
- 2. Wie verändert die Format- und Schnittstellenfrage die Übernahme digitaler Daten?**
- 3. Was für Folgen hat die Format- und Schnittstellenfrage auf die Auswertbarkeit der elektronischen Unterlagen im Digitalen Archiv?**

Wie beeinflusst die Format- und Schnittstellenfrage die archivische Bewertung?

Bewertungsdiskussion des 19. und 20. Jahrhunderts: Von Inhalt und Herkunft der Unterlagen abstrahierende Kriterien wie Lesbarkeit und äußere Beschaffenheit der Unterlagen sind kein Entscheidungskriterium

VS.

Drohender Informationsverlust und Unlesbarkeit digitaler Archivalen, die nicht in wenigen Standardformaten ausgesondert und archiviert werden können => Forderung nach Berücksichtigung der Format- und Schnittstellenfrage bei der Bewertungsentscheidung

Wie beeinflusst die Format- und Schnittstellenfrage die archivische Bewertung?

Hindernisse einer Bewertung nach Format und Schnittstelle:

- 1. Ein Großteil der künftigen elektronischen Überlieferung speist sich aus Fachverfahren, deren Inhalte vielfach in schlecht oder gar nicht dokumentierten hochproprietären Formaten vorliegen**
- 2. Sehr oft Fehlen einer an den archivischen Bedürfnissen orientierten Aussonderungsschnittstelle**
- 3. Teilweise Lagerung der noch nicht archivreifen Altdaten auf externen Datenträgern ohne deren Berücksichtigung bei Migrationsmaßnahmen**

Wie beeinflusst die Format- und Schnittstellenfrage die archivische Bewertung?

Lösungsansatz

- 1. Intensive Beratungstätigkeit, um Abgabeformat und Schnittstellen den archivischen Bedürfnissen anzupassen**
- 2. Im Ausnahmefall: Übernahme archivwürdiger Daten im Ursprungsformat unter Dokumentation oder sogar Sicherung des Quellsystems => Last der Lesbarmachung/Emulation der Überlieferung wird auf den Archivbenützer übertragen**

Wie verändert die Format- und Schnittstellenfrage die Übernahme digitaler Daten?

Problemfelder in der Praxis:

1. **Bestehende Export- und Importschnittstellen der Quellsysteme sind nicht immer für die Übermittlung archivwürdiger Daten an das zuständige Archiv geeignet**
2. **Auseinanderklaffen zwischen der datenverwaltenden Stelle einerseits und der datennutzenden bzw. datenerzeugender Stelle andererseits**
3. **Erschwerte Durchsetzung rechtlicher Normen (Personal- und Kostenfaktor) im Aussonderungsgeschäft**

Beispiel für lese- und schreibberechtigte Stellen in der InVeKoS-Datenbank (sogen. Kompetenztabelle) (<http://www.zi-daten.de/Kompetenzen.html>)

Meldefristen - Mozilla Firefox

Datei Bearbeiten Ansicht Chronik Lesezeichen Extras Hilfe

Meldefristen +

www.zi-daten.de/Kompetenzen.html

W - Wikipedia (de)

Meistbesucht Schwarzes Brett Google Bayerische Landesbibli... Die staatlichen Archiv... Staatsbibliothek Bahnverbindungen Routenplaner Map24 Marburg

Kompetenzen zu Entity ZA_NURANG

| ENTITY | AKTIONNR | Code | TYP_BETR | Code-Text | BEREICH | Code |
|---|----------|------|----------|--|---------|------|
| ZA_NURANG | | | | | | |
| Anzeigen System Test Produktion exakte Suche nein ja Download | | | | | | |
| ZA_NURANG | 1 | X | 1001 | IVK - InVeKoS-Betrieb | 1 | E |
| ZA_NURANG | 1 | X | 5003 | Z3 - ZID-Landeszentralstelle Zahlungsansprüche - ZA | 8 | L |
| ZA_NURANG | 1 | X | 5311 | ZAL - Verwaltungsstelle zur Abwicklung von Zahlungsansprüchen (landesweit) | 8 | L |
| ZA_NURANG | 1 | X | 5312 | ZAR - Verwaltungsstelle zur Abwicklung von Zahlungsansprüchen (für den Regierungsbezirk) | 6 | R |
| ZA_NURANG | 1 | X | 5313 | ZAK - Verwaltungsstelle zur Abwicklung von Zahlungsansprüchen (für den Landkreis) | 2 | K |
| ZA_NURANG | 2 | R | 0 | NUL - Default-Typ (dieser Typ ist keinem Benutzer zugeordnet, bestimmt aber die Standardkompetenzen) | 1 | E |
| ZA_NURANG | 2 | R | 98 | ZDB - Zentrale Datenbank | 10 | A |
| ZA_NURANG | 2 | R | 5003 | Z3 - ZID-Landeszentralstelle Zahlungsansprüche - ZA | 8 | L |
| ZA_NURANG | 2 | R | 5013 | BZ3 - ZID-Bundeszentralstelle Zahlungsansprüche - ZA (nur Leserechte) | 10 | A |
| ZA_NURANG | 2 | R | 5311 | ZAL - Verwaltungsstelle zur Abwicklung von Zahlungsansprüchen (landesweit) | 8 | L |
| ZA_NURANG | 2 | R | 5312 | ZAR - Verwaltungsstelle zur Abwicklung von Zahlungsansprüchen (für den Regierungsbezirk) | 6 | R |
| ZA_NURANG | 2 | R | 5313 | ZAK - Verwaltungsstelle zur Abwicklung von Zahlungsansprüchen (für den Landkreis) | 2 | K |
| ZA_NURANG | 2 | R | 5321 | ZLL - Verwaltungsstelle mit Leserecht für Zahlungsansprüchen (landesweit) | 8 | L |
| ZA_NURANG | 2 | R | 5322 | ZLR - Verwaltungsstelle mit Leserecht für Zahlungsansprüchen (für den Regierungsbezirk) | 6 | R |
| ZA_NURANG | 2 | R | 5323 | ZLK - Verwaltungsstelle mit Leserecht für Zahlungsansprüchen (für den Landkreis) | 2 | K |
| ZA_NURANG | 2 | R | 5900 | IZA - InVeKoS Zahlstelle des Landes | 8 | L |
| ZA_NURANG | 2 | R | 5901 | IZR - InVeKoS Zahlstelle (für den Regierungsbezirk) | 6 | R |
| ZA_NURANG | 2 | R | 5902 | IZL - InVeKoS Zahlstelle (für den Landkreis) | 2 | K |
| ZA_NURANG | 2 | R | 5910 | ILL - InVeKoS Zahlstelle mit Leserecht (landesweit) | 8 | L |
| ZA_NURANG | 2 | R | 5911 | ILR - InVeKoS Zahlstelle mit Leserecht (für den Regierungsbezirk) | 6 | R |
| ZA_NURANG | 2 | R | 5912 | ILK - InVeKoS Zahlstelle mit Leserecht (für den Landkreis) | 2 | K |
| ZA_NURANG | 3 | I | 1001 | IVK - InVeKoS-Betrieb | 1 | E |
| ZA_NURANG | 3 | I | 5003 | Z3 - ZID-Landeszentralstelle Zahlungsansprüche - ZA | 8 | L |
| ZA_NURANG | 3 | I | 5311 | ZAL - Verwaltungsstelle zur Abwicklung von Zahlungsansprüchen (landesweit) | 8 | L |
| ZA_NURANG | 3 | I | 5312 | ZAR - Verwaltungsstelle zur Abwicklung von Zahlungsansprüchen (für den Regierungsbezirk) | 6 | R |
| ZA_NURANG | 3 | I | 5313 | ZAK - Verwaltungsstelle zur Abwicklung von Zahlungsansprüchen (für den Landkreis) | 2 | K |
| ZA_NURANG | 4 | D | 10000 | DEL - Interner Typ, nur für TEST-System, kann nur von ZD zugeteilt werden kann, zum DELETE von Testdaten | 8 | L |
| ZA_NURANG | 5 | U | 5003 | Z3 - ZID-Landeszentralstelle Zahlungsansprüche - ZA | 8 | L |
| ZA_NURANG | 5 | U | 5311 | ZAL - Verwaltungsstelle zur Abwicklung von Zahlungsansprüchen (landesweit) | 8 | L |
| ZA_NURANG | 5 | U | 5312 | ZAR - Verwaltungsstelle zur Abwicklung von Zahlungsansprüchen (für den Regierungsbezirk) | 6 | R |
| ZA_NURANG | 5 | U | 5313 | ZAK - Verwaltungsstelle zur Abwicklung von Zahlungsansprüchen (für den Landkreis) | 2 | K |
| ZA_NURANG | 6 | S | 1001 | IVK - InVeKoS-Betrieb | 1 | E |
| ZA_NURANG | 6 | S | 5003 | Z3 - ZID-Landeszentralstelle Zahlungsansprüche - ZA | 8 | L |
| ZA_NURANG | 6 | S | 5311 | ZAL - Verwaltungsstelle zur Abwicklung von Zahlungsansprüchen (landesweit) | 8 | L |
| ZA_NURANG | 6 | S | 5312 | ZAR - Verwaltungsstelle zur Abwicklung von Zahlungsansprüchen (für den Regierungsbezirk) | 6 | R |
| ZA_NURANG | 6 | S | 5313 | ZAK - Verwaltungsstelle zur Abwicklung von Zahlungsansprüchen (für den Landkreis) | 2 | K |

Wie verändert die Format- und Schnittstellenfrage die Übernahme digitaler Daten?

Problemfelder in der Praxis:

4. Mehraufwand zur Identifizierung und Validierung des Abgabe- bzw. Archivformats
5. Einarbeitung in die Komplexität moderner Quellsysteme, um die Vollständigkeit der archivwürdigen Daten bei der Übernahme sicherzustellen

Lösungsansatz

Intensivierung der Beratungstätigkeit in den datenerzeugenden und -verwaltenden Stellen.
Diese muss so früh wie möglich einsetzen.

Was für Folgen hat die Format- und Schnittstellenfrage auf die Auswertbarkeit der elektronischen Unterlagen im Digitalen Archiv?

Bereits die Entscheidung über das Übernahmeformat bzw. die Beschaffenheit der Aussonderungsschnittstelle und gegebenenfalls die erste Migration in ein Archivformat im Rahmen des Ingest/Pre-Ingest bedeutet vielfach eine **Einbuße an Funktionalitäten und damit Verluste an Informationen und Auswertungsmöglichkeiten** gegenüber dem Quellsystem

Was für Folgen hat die Format- und Schnittstellenfrage auf die Auswertbarkeit der elektronischen Unterlagen im Digitalen Archiv?

Lösungsansätze

- **Eröffnung der Möglichkeit für den Archivbenützer im Rahmen des Access die archivierten Daten weiterzuverarbeiten**
- **Abstraktion von rein technischen Parametern zugunsten signifikanter Eigenschaften, die dauerhaft erhalten und damit den Archivbenützern auch in der Zukunft zugänglich gemacht werden sollen (vgl. nestor-Leitfaden zur digitalen Bestandserhaltung. Vorgehensmodell und Umsetzung, Version 2.0 von 2012)**



Fragen, Anregungen, Kritik?

**Dr. Michael Puchta M.A.
E-Mail: michael.puchta@gda.bayern.de
Tel.: 089/28638-2484**