Der Pre-Ingest – Digitale archivische Vorfeldarbeit

Veranstaltungsreihe:
Willkommen in der digitalen Welt

27.08.2025

Lukas Heck

Supportstelle DAN-Magazinpartnerschaften
Brandenburgisches Landeshauptarchiv

Agenda

Kapitel 1

Digitale Archivierung? Muss das sein?

Kapitel 2

Vorarbeiten zur Übernahme von Archivgut

Kapitel 3

Inhaltliche Bearbeitung von Archivgut

Neues aus dem DAN-Verbund

- Die Aufgabe der Supportstelle ist eine Landesaufgabe und die dazugehörige Stelle wird durch das Land Brandenburg finanziert.
 - Für ihr Kommunalarchive als Magazinpartner fallen in Zukunft nur Kosten für den eigentlichen Betrieb der DIMAG-Software an.
- Attraktive Preisgestaltung für Sie!

Kapitel 1

Digitale Archivierung?

Muss das sein?

Warum dieser Vortrag?

Elektronisches Archivgut gibt es in jeder Verwaltung: Seit Jahrzehnten bereits schrittweise Umstellung auf digitales Arbeiten in der Verwaltung – archivwürdige Unterlagen entstehen nicht nur in einem Dokumentenmanagementsystem:

→ Digitale Ablagen in Ordnersystemen ohne DMS (Behörden, private Nachlässe), Digitale Fotos, Digitale Karten und Pläne, Fachverfahren (Datenbanken), Webseiten, Digitale Publikationen (graue Literatur), Pressestelle

→ Digitale Archivierung ist keine Datensicherung, sondern archivarische Pflicht- und Fachaufgabe:

Daten müssen in einem **menschenverstehbaren Kontext** dauerhaft gesichert und nutzbar gehalten werden, um sie auch **in 50 oder 100 Jahren** noch **lesen** und **verstehen** zu können.

Anforderung an die Archive! Analog wie Digital

Archive sichern die...

Integrität: Archivalien sind unverändert.

Vollständigkeit: Keine Archivalien wurden nachträglich entfernt.

Authentizität: Archivalien stammen vom angegebenen Urheber.

Lesbarkeit: Archivalien können angezeigt und interpretiert werden.

Was bedeutet: Pre-Ingest

Pre-Ingest

Pre = Vor, „Vorarbeit“

Ingest = „Aufnehmen“, Vorgang der Archivierung

Was bedeutet: Pre-Ingest

Der *Pre-Ingest* umfasst **alle Aktivitäten vor der eigentlichen Archivierung**, die notwendig sind, um digitale Objekte, ihre Metadaten und ihre Struktur so vorzubereiten, dass sie technisch, organisatorisch und rechtlich archivfähig sind:

- Bewertung
- Datentransfer
- Formaterkennung, Formatvalidierung, Formatprotokollierung, Formatkonvertierung
- Vorbereitung für den Upload nach DIMAG (Formierung, Erschließung)

Kapitel 2

Vorarbeiten zur Übernahme von Archivgut

Vorüberlegungen

Um was für Daten handelt es sich?

Ist es ein unstrukturierter Nachlass, mit relativ wenig Daten?

Oder ist die Abgabe des Bauamtes mit riesigen Karten und Plänen, welche viel Speicherplatz in Anspruch nehmen?

Oder möchte der Bürgermeister Ihnen Fotos vom Besuch bei den Stadtwerken übergeben?

Ausgangssituation: Analoge vs. Digitale Übernahme

Unterlagen wurden bereits als archiwürdig bewertet und können nun übergeben werden. Best Case Szenario:

Analoge Unterlagen

Anlieferung/Abholung in jeglicher Form

Digitale Unterlagen

Wie kann eine Übergabe erfolgen?

Maßnahmen für den sicheren Datentransfer

- Pakete bilden (große Datenmengen aufteilen)
- Verschlüsselung bei sensiblen Daten
- Transportsicherung: Abgabeliste und Hashwerte erzeugen
 - Per E-Mail
 - Per Datenträger
 - Per Post
 - Persönliche Übergabe
 - Per Upload auf ein Cloud System

- Löschen bei der abgebenden Stelle erst nach erfolgreicher Sicherung und Qualitätsprüfung im Archiv.

SCHIMMEL

Virenscan vs. Quarantäne/Schimmelprüfung

Unterlagen wurden für Sie ins Archiv angeliefert.
Was ist zu tun?

Analog

- Akten werden in einer Quarantäne zwischengelagert.
- Akten werden bei Bedarf auf Schimmelsporen überprüft.
- Liegt Schimmelbefall vor, werden diese entweder gereinigt oder die Archivierung wird abgebrochen.

Digital

- Bei der Übernahme externer Datenträger ist ein Viren Scan notwendig.
- Überprüfung mit hauseigenem Virens Scanner. (ggf. Hilfe durch IT-Abteilung).
- Werden Viren festgestellt, ist die Archivierung abubrechen.

VIREN SCAN

Prüfsummen bzw. Hash Werte erheben

- Abgebende Stelle hat Daten ohne Prüfsummen übergeben. Prüfsummen sollten für das Ablieferungsverzeichnis mit **checksum** oder **Quickhash GUI** erstellt werden.
- Um Kopierfehler auszuschließen, müssen regelmäßig die Prüfsummen aller Dateien zwischen dem Originaldatenträger und der erstellten Kopie miteinander verglichen werden.
- **Analog:** Überprüfung des Bestandes mit den vorliegenden Akzessionslisten.

Beispiel mit MD5 Hash. Es können auch beliebige andere Hashfunktionen genutzt werden z. B. SHA 256/512.

Containerformate entpacken

- Was sind Containerformate? Containerformate (z.B. ZIP, 7z, gz, tar) umhüllen eine oder mehrere Dateien.
- Vor einer Bearbeitung müssen Containerformate gefunden und entpackt werden.
- Kompressionsformate sind nach dem Entpacken zu löschen.

Anlegen von Arbeitsdateien

- Kopieren des Ausgangsverzeichnis („orig“, d. h. Original-Dateien) in ein Arbeitsverzeichnis („work“).
- Das „orig“ Verzeichnis stellt bei der Archivierung das Backup da und kann bei fehlerhaften Hash Werten, wiederum kopiert werden und für die Archivierung verwendet werden.
- Achtung: Ausreichend Speicherplatz auf Servern bzw. Festplatten vorhalten, um ein „orig“ und „work“ Verzeichnis ablegen zu können.
Empfehlung von 1 TB Speicherplatz für den Anfang
- **Empfehlung:** Die „orig“ & „work“ sollten auch wenn die Archivierung komplett abgeschlossen ist, noch einige Monate vorgehalten werden und können dann gelöscht werden.

Beispielhafter Übergabeprozess

Kapitel 3

Inhaltliche Bearbeitung von Archivgut

Kreislauf des Pre-Ingest

Hinweis: Vorgänge beim Pre-Ingest können sich auch wiederholen, z. B. wenn eine technische Bearbeitung an Dateien vorgenommen wird und damit eine weitere Formatvalidierung durchgeführt werden muss, um die Eigenschaften der Datei zu validieren.

Analyse der Abgabe

„Was habe ich da eigentlich vor mir?“

In dieser Phase verschaffen sich die Mitarbeitenden einen Überblick über die Dateisammlung und versuchen Erkenntnisse über die Struktur und die Dateien zu gewinnen.

Digitale Archivierung ist eine archivische Fachaufgabe, tauschen Sie sich mit den Kollegen dazu aus.

Analyse mit Archifiltre

Datenanalyse

- Im Analogen ist es von hoher Bedeutung sich mit der Abgabe auseinanderzusetzen und eine Vorstrukturierung durchzuführen.
 - Strukturierung eines Nachlasses
 - Kellerraum voller Akten

- Mit Software wie z. B. **Archifiltre** oder **Treesize** lassen sich Datenstrukturen visualisieren.
 - Beinhaltet die Abgabe ungewöhnliche Strukturen (Verschachtelungen), Dateiformate, etc.
 - Muss die vorliegende persönliche Fileablage vom Bürgermeister neu geordnet werden?

Bewertung

Für die Bewertung von digitalen Unterlagen ist wie im Analogen, entsprechende Vorarbeit nötig:

- In welchem Format übernehme ich?
 - Welche Merkmale / Funktionen will ich erhalten?
 - Auswertbarkeit / Nutzbarkeit?
 - Gibt es Dokumentationen/Handbücher?
 - Archivfähigkeit?
 - Sind die Daten noch lesbar?
 - Sind es ausführbare Dateien?
 - Speicherplatz?
- Auf diese Fragen wird in den folgenden Folien eingegangen.

Datei- und Ordnernamen normalisieren

- Machen sie sich mit den Ordnerstrukturen und den Dateinamen vertraut und normalisieren Sie diese für ihre Arbeit, um die weiteren Schritte der Archivierung der Erschließung zu erleichtern.
- Im Analogen ist dies ähnlich zu einer Strukturierung der Abgabe.
- Protokollieren Sie die Umbenennungen:
Alt: Alter Datei- und Ordnername

Besuch Senftenberg

Neu: Neuer Datei- und Ordnername

Besichtigung Schloß Senftenberg. Betriebsausflug Bauamt des Kreis OSL

Formaterkennung am Beispiel von Papier

Was habe ich vor mir? Ist
Papier gleich Papier?

Entspricht das Papier einer
Standardgröße oder wird aufgrund
von DIN A2 oder A1 andere
Verpackungen benötigt?

Ist das Papier
archivfähig? Besteht die
Gefahr, dass Papier bei
Benutzung zerstört
wird?

Formaterkennung mit Softwareunterstützung

Der Vorgang der Formaterkennung ist unerlässlich!

	Beispiel_4	19.05.2025 14:41	Adobe Acrobat-Dokument
	Beispiel_5	19.05.2025 14:41	Adobe Acrobat-Dokument
	bild_jp2.jp2	19.05.2025 14:42	JP2-Datei
	bild_jpg	19.05.2025 14:43	JPG-Datei
	bild_png	19.05.2025 14:42	PNG-Datei
	Dokument_3	19.05.2025 14:40	Microsoft Word-Dokument
	Dokument_4	19.05.2025 14:40	Microsoft Word-Dokument
	Dokument_5	19.05.2025 14:40	Microsoft Word-Dokument
	Dokument_7	19.05.2025 14:40	Microsoft Word-Dokument
	Dokument_9	19.05.2025 14:40	Microsoft Word-Dokument
	Praesentatio...	19.05.2025 14:40	Microsoft PowerPoint-Präsentation

	Beispiel_1.pdf	pdf	Acrobat PDF 1.3 - Portable Do...	1.3
	Beispiel_2.pdf	pdf	Acrobat PDF 1.3 - Portable Do...	1.3
	Beispiel_3.pdf	pdf	Acrobat PDF 1.3 - Portable Do...	1.3
	Beispiel_4.pdf	pdf	Acrobat PDF 1.3 - Portable Do...	1.3
	Beispiel_5.pdf	pdf	Acrobat PDF 1.3 - Portable Do...	1.3
	bild_jp2.jp2	jp2	JP2 (JPEG 2000 part 1)	
	bild_jpg.jpg	jpg	JPEG File Interchange Format	1.01
	bild_png.png	png	Portable Network Graphics	1.0
	Dokument_3.docx	docx	Microsoft Word for Windows	2007 onward
	Dokument_4.docx	docx	Microsoft Word for Windows	2007 onward
	Dokument_5.docx	docx	Microsoft Word for Windows	2007 onward
	Dokument_7.docx	docx	Microsoft Word for Windows	2007 onward
	Dokument_9.docx	docx	Microsoft Word for Windows	2007 onward
	Praesentation_1.pptx	pptx	Microsoft Powerpoint for Win...	2007 onward
	Praesentation_10.pptx	pptx	Microsoft Powerpoint for Win...	2007 onward
	Praesentation_6.pptx	pptx	Microsoft Powerpoint for Win...	2007 onward
	Praesentation_8.pptx	pptx	Microsoft Powerpoint for Win...	2007 onward

PDF	Nach PDF/A-2b konvertieren
PDF/A-1a	Nicht konvertieren
PDF/A-1b	Nicht konvertieren
PDF/A-2a	Nicht konvertieren

Es genügt nicht, die Dateiendung bzw. die Eigenschaften der archivwürdigen Dateien anzusehen.

Prüfen mit der Software **DROID** ob das vorliegende Format für die Archivierung geeignet ist.

Risikoabschätzung treffen und ggf. Erhaltungsmaßnahmen einleiten zu können

Formaterkennung durchführen

Learning:

Eine rein analoge Betrachtung Dateiendung bzw. der Eigenschaften der archivwürdigen Dateienanzusehen genügt nicht!
Es ist **immer** eine Formaterkennung ist notwendig:

- ✓ Überprüfung der Datei-Formate mit Software z. B. mit DROID (Digital Record Object Identification)
- ✓ Um zu prüfen, ob das vorliegende Format für die Archivierung geeignet ist.
- ✓ Risikoabschätzung: Ist das Format langzeitstabil? Wenn nicht müssen Erhaltungsmaßnahmen eingeleitet werden.

Formate vereinheitlichen

- Welche Formate muss ich migrieren?
- Welche Formate kann ich direkt ins digitale Archiv übernehmen?

Format	Empfehlung	Begründung
JPEG	Nicht konvertieren	Verlust schon vorhanden, JPEG ist ein bekanntes Format
PDF	Nach PDF/A-2b konvertieren	Nicht archivfähig
PDF/A-2u	Nicht konvertieren	Archivformat
Word	Nach PDF/A-2b konvertieren	Nicht archivfähig
MSG	Nach PDF/A-2b konvertieren	Nicht archivfähig

Empfehlungen zu
Formatfragen

KOST Koordinationsstelle für die dauerhafte Archivierung elektronischer Unterlagen DE FR IT Suchen

Aktuell
Standards
Dateiformate (KaD)
Projekte
Publikationen
Tools & Services

Einleitung
Empfehlung
Text
Bild
Audio
Video
Tabellenkalkulation
Strukturierte Daten
Datenbanken
Hypertext
GIS
CAD/CAM

Dateiformate (KaD) > Empfehlung

Empfehlung der KOST

Der Katalog archivischer Dateiformate (KaD) kann von Archiven und verwandten Institutionen als Grundlage für die Formulierung eigener Formatempfehlungen benutzt werden. Aufgrund der im KaD detailliert erläuterten Analysen, Bewertungen und Schlussfolgerungen empfiehlt die KOST für die verschiedenen Formatkategorien die folgenden Dateiformate als gegenwärtig archivtauglichste Lösung:

Text

Text soll möglichst bereits in der semi-aktiven Phase in ein PDF-Format (idealerweise PDF/A-2u) umgewandelt und so archiviert werden. Eine Konvertierung von PDF-Dateien in ein PDF/A-Format ist zu einem späteren Zeitpunkt meist problemlos möglich.

- Details zum Format **PDF-A-2**
- Details zum Format **PDF-A-1**
- Details zum Format **PDF**
- Analyse der **Textformate**

Katalog archivischer Dateiformate der KOST (Koordinationsstelle für die dauerhafte Archivierung elektronischer Unterlagen KOST)
Der Katalog archivischer Dateiformate beschreibt wichtige und verbreitete Dateiformate und analysiert ihre Eignung für die digitale Archivierung.

Formatkonvertierung von PDF-Dokumenten

- Nur ein PDF/A stellt sicher, dass
 - PDF-Dateien nicht nachträglich bearbeitet werden können.
 - Keine Verschlüsselung mit Passwörtern vorliegt, damit die Inhalte jederzeit zugänglich sind.
 - Keine Einbettung von Video- und Audiodateien, da sonst ggf. externe Software zum Darstellen oder Abspielen benötigt wird.

Die Konvertierung von **doc, xls, ppt, docx, pptx, odt, ods oder odp** ins **PDF/A-Format** ist leichter als gedacht und funktioniert z. B. mit **Adobe Acrobat** oder kostenlosen Alternativen wie **PDF24**.

Konvertierung von AV-Medien

- Die Lebensdauer von AV-Medien wie CD & DVD ist begrenzt.
- Konvertierung der DVD mit Software möglich.
 - Software mit grafischer Oberfläche: Handbrake (nur für Videos)
 - Software auf Kommandozeile: FFmpeg

Da die anfallende Datenmenge bei Videos verhältnismäßig groß ist, werden im Consumerbereich, zum Beispiel bei DVDs oder Videoportalen, Videos heutzutage verlustbehaftet komprimiert.

Handbrake bietet verlustfreie Videocodecs wie **FFV1**. Dadurch entstehen Daten von ca. dem 5-10 fachen des Ausgangswert.

Formatvalidierung

- Am Ende der Konvertierung sollte auch eine Validierung stattfinden.
- Bei der Validierung wird die Formatspezifikation mit dem vorliegenden Objekt abgeglichen.
- Um das zu überprüfen, existieren verschiedene Validierungstools. Hierbei muss aber wieder je nach Format nach Tools unterschieden werden:

PDF/A: VeraPDF oder AdobePreflight

AV-Medien Audio, Foto, Video: JHOVE

z. B. Spezialisiert auf Matroska video/FFV1: MediaConch

Formatprotokollierung/ Aufbereitungsbericht

Aufbereitungsbericht

Ablieferungsverzeichnis: Q:\2025-04-03_Archivierung_Amtsblatt\orig\Amtsblätter 2024

GeschZ: Geschäftszeichen des Vorgangs.

Betreff: Bestandsbildner: Abgabe von Amtsblätter der Pressestelle des der fiktiven Stadt Fakenstedt. Die Amtsblätter haben eine lokalhistorische Bedeutung und müssen zur Dokumentation des Verwaltungshandeln archiviert werden.

Datum der Übernahme in das Archiv: xx.xx.xxxx

Endarchivbestand: Archivbestand (Bestandskürzel).

Beginn der Eingangsbearbeitung: xx.xx.xxxx

Abschluss der Eingangsbearbeitung: Datum, Abschluss der Bearbeitung.

Bearbeiter*in: Herr Mustermann

Bestandsverwalter*in: Frau Musterfrau

Zugang	Prozessschritt: Zugang in Ablieferungsverzeichnis sichern
Verzeichnis: Dateipfad	Wann: 03.04.2025 Wer: xxx
Bemerkungen: Klicken oder tippen Sie hier, um Text einzugeben.	

Zugang	Prozessschritt: Virenprüfung
Verzeichnis: Dateipfad	Wann: Datum Wer: xxx
Bemerkungen: Virenprüfung mit Software zur Virenprüfung durchgeführt, keine Malware gefunden.	

Zugang	Prozessschritt: Eingangsrevision
Verzeichnis: Dateipfad	Wann: Datum Wer: xxx
Bemerkungen: Durchsicht Dateiformate (z. B. .txt, .pdf); Kein Übergabeverzeichnis vorliegend.	

Zugang	Prozessschritt: Prüfsummen erheben
Verzeichnis: Dateipfad	Wann: Datum Wer: xxx
Bemerkungen: Abgebende Stelle hat Daten ohne Prüfsummen übergeben. Prüfsummen wurden erstellt für das Ablieferungsverzeichnis mit checksum/Quickhash GUI: Dateipfad zur Prüfsummen Software mit gehashten Daten	

Zugang	Prozessschritt: Zugang in Übernahmeregister registrieren
Verzeichnis: Dateipfad	Wann: Datum Wer: xxx
Bemerkungen:	

Zugang	Prozessschritt: Prüfsummen validieren
Verzeichnis: Dateipfad	Wann: Datum Wer: xxx
Bemerkungen: Dateipfad zur Datei mit Hashwert angeben	

Alle Bearbeitungen beim Pre-Ingest sollten dokumentiert werden! Nur so lässt sich die Integrität, Vollständigkeit, Authentizität und Lesbarkeit der Daten dauerhaft nachweisen.

Bearbeitetes Objekt	Prozessschritt
---------------------	----------------

Projekt Amtsblätter	Prozessschritt: Datei- und Ordernamen normalisieren
Verzeichnis: Dateipfad	Wann: Datum Wer: xxx
Bemerkungen: manuelle Umbenennung eines Dateinamens: Alt: Alter Datei- und Ordernamen Neu: Neuer Datei- und Ordernamen	

Projekt Amtsblätter	Prozessschritt: Dateiformaterkennung durchführen
Verzeichnis: Dateipfad	Wann: Datum Wer: xxx
Bemerkungen: Mit dem Tool „DROID v6.7.0 unter Verwendung der DROID_SignatureFile.xxx XXX MB / GB / TB entspricht XXX Dateien .pdf, fnt/19 (PDF 1.5), 100 Dateien .csv, x-fmt/18, 1 Datei	

Projekt Amtsblätter	Prozessschritt: Formatvalidierung durchführen
Verzeichnis: Dateipfad	Wann: Datum Wer: xxx
Bemerkungen: Formatmigration in PDF/A Software Adobe Acrobat Pro, für die PDF-Validierung veraPDF .	

Projekt Amtsblätter	Prozessschritt: Formatvalidierung durchführen
Verzeichnis: Dateipfad	Wann: Datum Wer: xxx
Bemerkungen: .csy-Formatvalidierung mit JHOVE Version X Betrifft 2 Dateien: - JHOVE 1.28: Zeichencodierung nicht erkannt → mit notepad++ Zeichencodierung auf UTF-8 migriert - JHOVE v1.28: Zeichencodierung UTF-8 erkannt und valide	

Projekt Amtsblätter	Prozessschritt: Formate auf Archivfähigkeit prüfen
Verzeichnis: Dateipfad	Wann: Datum Wer: xxx
Bemerkungen: Die Formate wurden im Rahmen der Übergabeverhandlungen mit der Landesvermessung abgestimmt und sind archivfähig.	

Projekt Amtsblätter	Prozessschritt: AIP bilden
Verzeichnis: Dateipfad	Wann: Datum Wer: xxx
Bemerkungen: Die vorgefundene Ordnerstruktur wurde aufgelöst und es wurden die Ordnerstrukturen für zwei Ingest-Tool -Projekte angelegt. Dabei wurden die Ordner mit den Primär- und Metadaten prüfsummengesichert verschoben (die .md5-Dateien wurden anschließend wieder gelöscht): - Erklärung wie Ordner neustrukturiert worden sind: Der Inhalt des Ordners „Archivierung Amtsblätter“ wurde verschoben in den Ordner „Amtsblaetter_Repräsentationen“ verschoben - Der Inhalt des Ordners „Amtsblätter Metadaten“ wurde verschoben in den Ordner „Amtsblaetter Verzeichnungsangaben“ Die neue Ordnerstruktur stellt sich folgendermaßen dar: Screenshot	

	Prozessschritt: Datentransfer nach DIMAG durchführen
--	---

Vorgezogene Erschließung

- Metadaten, welche für die Erschließung relevante Informationen beinhalten, können händisch, teilautomatisiert oder durch maschinelles Nachschlagen in einer Datenquelle ergänzt werden.
 - Inhaltliche oder technische Unzulänglichkeiten
 - Besondere Nutzungshinweise
 - Referenz auf etwaige Papierakten oder weitere digitale Bestandteile (z. B.: entnommen aus...) & Verweise auf eine zum Verständnis notwendige analoges Archivalie (z. B.: Handbücher etc.)

**Vielen Dank für Ihre
Aufmerksamkeit!**

Kontakt

Lukas Heck

Brandenburgisches Landeshauptarchiv
Supportstelle DAN-Magazinpartnerschaften
Referat 13 Digitale Archivierung

0331 5674-295

Lukas.Heck@BLHA.Brandenburg.de