



Masse als Herausforderung

Ein Überblick zur digitalen Archivierung im Bundesarchiv



Akten



Bilder



Filme



Töne



Karten

Dr. Sebastian Gleixner

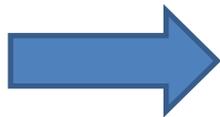
Geschichtlicher Überblick

- Erfahrung mit digitalen Unterlagen bereits seit der Wiedervereinigung (Übernahme von Daten aus DDR-Stellen)
- 2008-2018: Wirkbetrieb des OAIS-konformen Digitalen Archivs für elektronische Akten (Problem: Verzögerte Einführung von DMS in der Bundesverwaltung)
- Seit 2014: Einführung des Digitalen Magazins
 - Zunächst nur zur Präsentation von Digitalisaten über invenio
 - Einführung von zusätzlichen Funktionen, u.a. für die Sicherung der Digitalisate im Bandspeicher
- Seit 2016: Projekt BASYS 3-Magazin zur Etablierung einer umfassenden digitalen Archivierung im Bundesarchiv
- Seit 2018: Aufbau eines neuen Speichersystems (ausbaufähig bis zu 120 PB)

Wege von potentielltem Archivgut ins Bundesarchiv

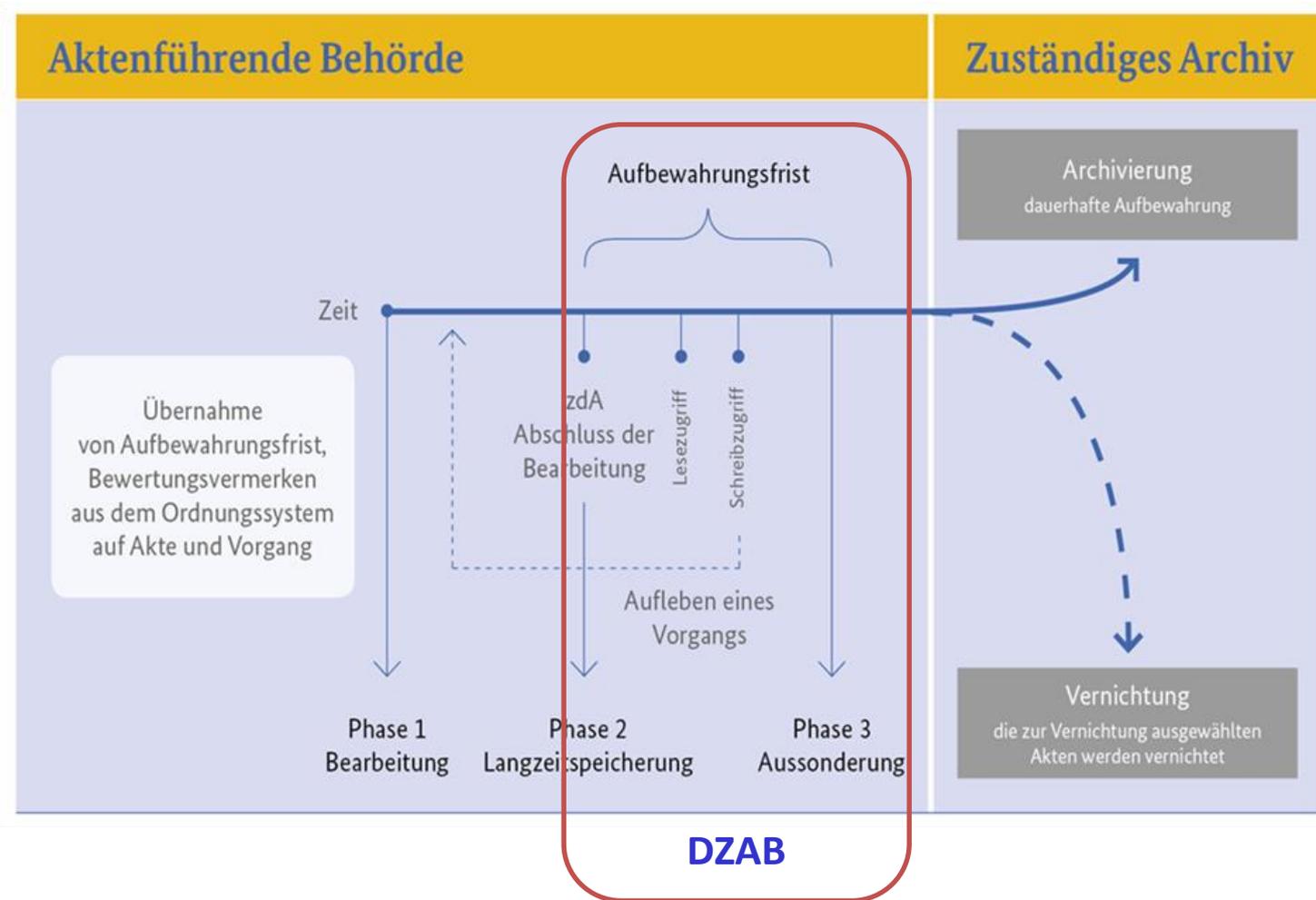
Wege von Archivgut ins Bundesarchiv

- Digitales Zwischenarchiv des Bundes (DZAB)
- Behördliche Abgaben ohne DZAB
- Nichtbehördliche Unterlagen mit Bezug zur Bundesrepublik Deutschland (Nachlässe, Verbände)
- Digitalisierung intern/extern



unterschiedliche Formen von SIPs

Das Digitale Zwischenarchiv des Bundes (DZAB)



Der Service des DZAB

- Beweiswerterhalt
- Reduktion des Speicherbedarfs im DMS der Behörden → Kostenreduktion
- Unterstützung bei der Einlagerung durch das Bundesarchiv
- Recherche u. Abruf unabhängig vom DMS
- Erfüllung der Anbietungspflicht
- Frühzeitige Bewertung

Vor allem ein Problem der Masse

- **Digitales Zwischenarchiv des Bundes (DZAB) (bis zu 40 GB pro Behörde/Jahr)**
- **Behördliche Abgaben ohne DZAB (bis zu 40 GB pro Behörde/Jahr)**
- **Nichtbehördliche Unterlagen mit Bezug zur Bundesrepublik Deutschland (Nachlässe, Verbände) -> nicht planbar**
- **Digitalisierung intern/extern (bis zu 10 PB im Jahr, v.a. durch die Filmdigitalisierung)**



Masse an Daten als entscheidendes Kriterium

Planung Archivierung

- Weiterhin Aufbau eines Bandspeichers auf Grundlage von Oracle mit Ziel eines Aufnahmevermögens von ca. 120 PB (skalierbar auf den zehnfachen Wert (1.2 ExaByte) bei Nutzung von LTO-8); Datendurchsatz im Moment bei 600 MB/s (bei Tests bis zu 1 GB/s; bei Ausbau auf 640 Bandlaufwerke skalierbar auf 790 TB/h).
- Weiterhin Ausbau des Digitalen Magazins; derzeit Einlagerung der Daten (v.a. Digitalisate; Bilder gesondert) im Format BagIt für alle Archivaliengattungen
- In Vorbereitung: Ausschreibung einer Ingest-Software zur OAIS-konformen Paketbildung (SIP, AIP, DIP) und Archivierung für alle Archivaliengattungen (digitalisiert/genuin digital)

Planung Bereitstellung

- Weiterer Ausbau der Bereitstellung über die Internetplattform invenio; derzeit Lagerung der DIPs in einem Fedora Repository
 - invenio greift zur Anzeige des DIPs auf den dazugehörigen Viewer zu.
 - Momentan v.a. für den Bereich Akte; soll für alle Archivaliengattungen weiter ausgebaut werden
 - Für Bilder im Moment noch eigenes System
- Bereitstellung von Kopien der Originaldateien über Datenträger und BArchBox (one way) – weiterer Ausbau von cloudbasierten Übermittlungsmöglichkeiten ist in Prüfung
- Grundsätzlich ist in diesem Bereich zu überlegen, inwieweit Big Data und Künstliche Intelligenz (KI) der Recherche und Bereitstellung neue Möglichkeiten eröffnen.

Planung Bestandserhaltung

- Zwei Bereiche: speichertechnische und fachtechnische Bestandserhaltung
- Speichertechnische Bestandserhaltung (z.B. Überprüfung der Bänder)
 - Wird von der Speicherinfrastruktur „selbstheilend“ umgesetzt
 - Wird im Zusammenhang mit dem Ausbau der Speicherinfrastruktur geregelt
- Fachtechnische Bestandserhaltung (z.B. Schutz vor Formatobsoleszenz)
 - Derzeit v.a. Formatmonitoring
 - Allgemeines Digital Preservation Management Schritt für Schritt im Aufbau

Zusammenfassung und Ausblick

- Die Masse an digitalen Daten und die dazu notwendige Performanz der Systeme ist ein wesentlicher Faktor für Ingest, Archivierung, Bestandserhaltung und Bereitstellung.
- Digitale Magazinierung ist genauso wie Formate, Speichertechnologien oder Anwendungsmöglichkeiten von Big Data und KI im dauernden Wandel.
- Um mit diesem Wandel mitzuhalten und technisch adäquate Lösungen zu bieten, sind je nach Zielrichtung sowohl Flexibilität und kleinere Anpassungsschritte als auch langfristige Entwicklungsrichtlinien wichtig.
- Die sich durch diesen Wandel stets ändernden Rahmenbedingungen („Kundenwünsche“) müssen immer wieder evaluiert werden, um diese Entwicklungsrichtlinien anzupassen und Fehlentwicklungen zu vermeiden.
- Standards sind die Grundlagen („Leitplanken“) für die Weiterentwicklung.

Danke fürs Zuhören!

Präsentation:

Titel: Masse als Herausforderung. Ein Überblick zur digitalen Archivierung im Bundesarchiv

vorgetragen von: Dr. Sebastian Gleixner

vorgetragen am: 25.04.2019

Kontaktdaten:

Ansprechpartner/-in: Dr. Sebastian Gleixner

Telefon: 0261/505-206

Email: s.gleixner@bundesarchiv.de

Anschrift: Bundesarchiv
Potsdamer Str. 1
56075 Koblenz
Referat IT 4