

Ein Digitales Archiv für alles? Verbundlösung Berlin-Brandenburg mit

Rolf Däßler

Fachbereich Informationswissenschaften

daessler@fh-potsdam.de

KBD | Koordinierungsstelle

BRANDENBURG-DIGITAL

Hintergrund

- Der **Arbeitskreis Brandenburg.digital** arbeitet seit 2007 spartenübergreifend zu den Fragen der digitalen **Präsentation von Kulturgut** im Land Brandenburg.
- Ihm gehören verschiedene Landeseinrichtungen und Verbände der Sparten Archiv, Bibliothek, Museum, Denkmalpflege und Gedenkstätten sowie wissenschaftliche Einrichtungen an.

Webadresse: brandenburg-digital.org

Kooperationsnetzwerk

Brandenburg.digital

- Brandenburgisches Landeshauptarchiv
- Brandenburgisches Landesamt für Denkmalpflege und Archäologisches Landesmuseum
- Brandenburgisches Literaturbüro
- Fachhochschule Potsdam
- Filmmuseum Potsdam
- Filmuniversität Babelsberg Konrad Wolf
- Deutsches GeoForschungsZentrum (GFZ)
- Deutsche Gesellschaft für Geowissenschaften
- Haus der Brandenburgisch-Preußischen Geschichte
- Koordinierungsstelle Brandenburg-digital
- Landesfachstelle für Archive und öffentliche Bibliotheken im Brandenburgischen Landeshauptarchiv
- Museumsverband des Landes Brandenburg
- Sorbisches Institut, Zweigstelle Cottbus
- Stadt- und Landesbibliothek Potsdam
- Stiftung Brandenburg
- Stiftung Brandenburgische Gedenkstätten
- Stiftung Preussische Schlösser und Gärten Berlin-Brandenburg
- Universitätsbibliothek der Europa-Universität Viadrina Frankfurt (Oder)
- Universitätsbibliothek der Universität Potsdam
- Zentrum für Zeithistorische Forschung Potsdam



KOOPERATIONSNETZWERK BRANDENBUR.DIGITAL

DIE DIGITALE PRÄSENTATION VON KULTURGUT IM LAND BRANDENBURG

Historisches Erbe ist in besonderem Maße geeignet, kulturelle Identität zu stiften. Es trägt entscheidend dazu bei, Brandenburg in Deutschland und der Welt sichtbar zu machen. (Kulturpolitische Strategie 2012)

Wer nicht digital präsent ist, wird nicht mehr wahrgenommen. (Johanna Rachinger, Generaldirektorin der Österreichischen Nationalbibliothek 2006)

(2014) Becher, Jürgen; Michalowsky, Ulrike; Paschke, Ralph; Preuß, Ulf; Scholz, Michael: **Die Digitale Präsentation von Kulturgut im Land Brandenburg**. Potsdam, Fachhochschule Potsdam. (Bericht)
URN: [urn:nbn:de:kobv:525-5433](https://nbn-resolving.org/urn:nbn:de:kobv:525-5433)

Kooperative digitale Langzeitverfügbarkeit/-
nutzbarkeit des kulturellen Erbes im Land
Brandenburg – Konzept für eine Verbundlö-
sung in Kooperation mit dem ZIB Berlin

Der Aufbau und die Beständigkeit digitaler Infrastruk-
turen und Abläufe sind für viele Einrichtungen im Be-
reich Kulturerbe nur durch kooperative Lösungswege
realisierbar.

(Die digitale Präsentation von Kulturgut im Land Brandenburg 2014)

Zielstellung (1)

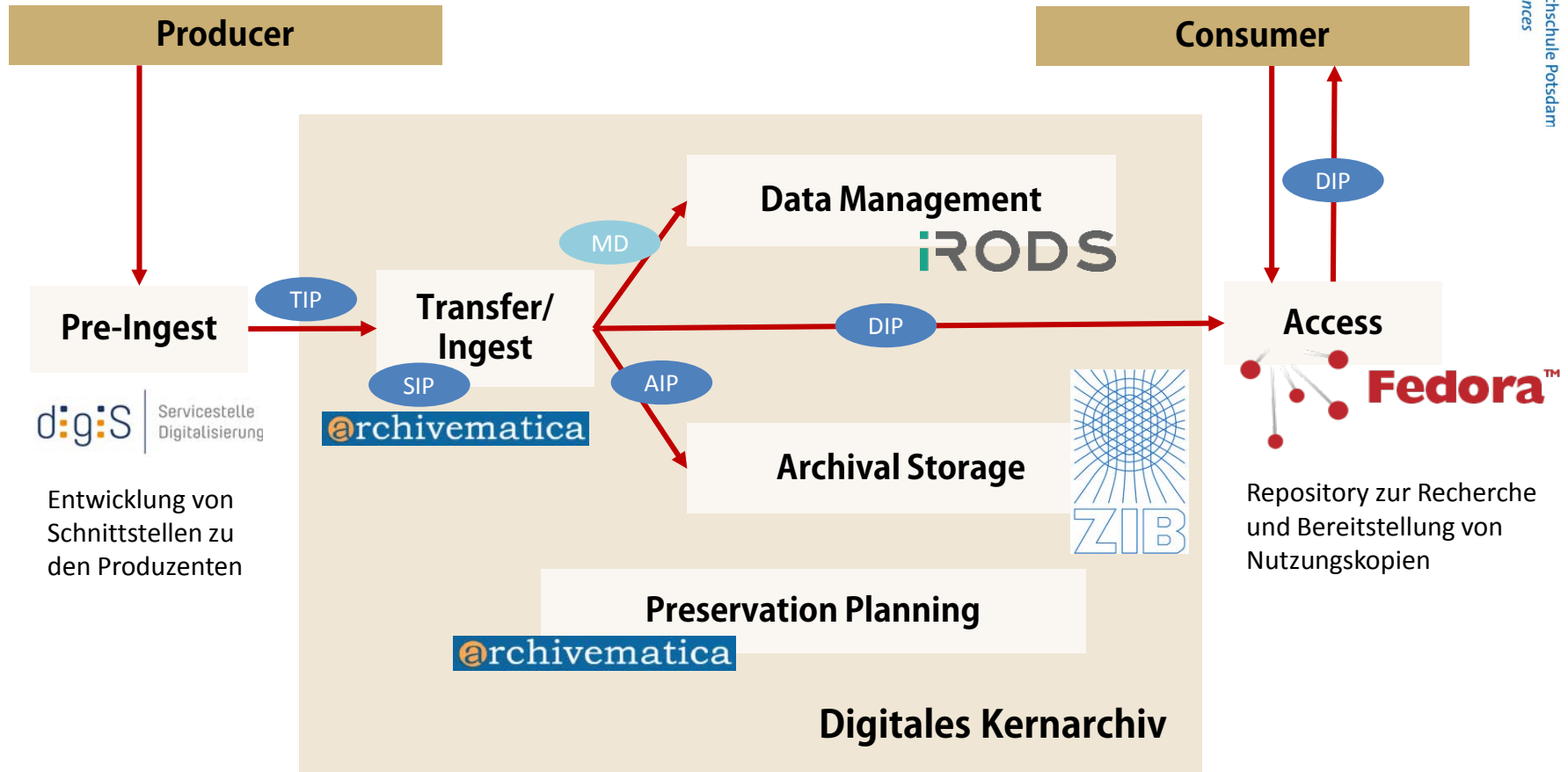
- Aufbau einer Infrastruktur zur langfristigen und nachhaltigen Sicherung von **Digitalisaten** insbesondere aus **kooperativen Projekten**.
- Archivlösung für **kleinere und mittlere Einrichtungen** ohne eigene Infrastruktur um digitale Objekte langfristig und nachhaltig aufzubewahren.
- Archivische **Verbundlösung** in Kooperation der Koordinierungsstelle **brandenburg.digital (KBD)** und der **Servicestelle Digitalisierung Berlin (digiS)** am **Zuse Institut Berlin (ZIB)**.

Zielstellung (2)

- Bereitstellung von Metadaten und Digitalisaten für die **DBB**.
- Bereitstellung von **persistenten Nutzungskopien** digitaler Objekte kleinerer und mittlerer Einrichtungen ohne eigene Infrastruktur um digitale Objekte langfristig und nachhaltig **zu präsentieren**.
- Unterstützung der **Eigenverantwortlich** der im Rahmen der DBB abliefernden Einrichtungen für die **digitale Langzeitarchivierung** digitaler Objekte.



OAIS-konformes digitales Archiv (DIGIS/ZIB)



Ingestwerkzeug

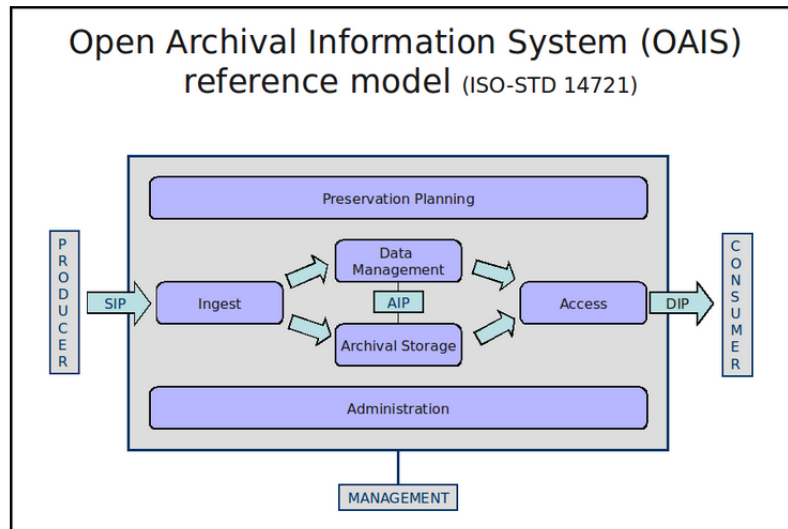


What is Archivematica?

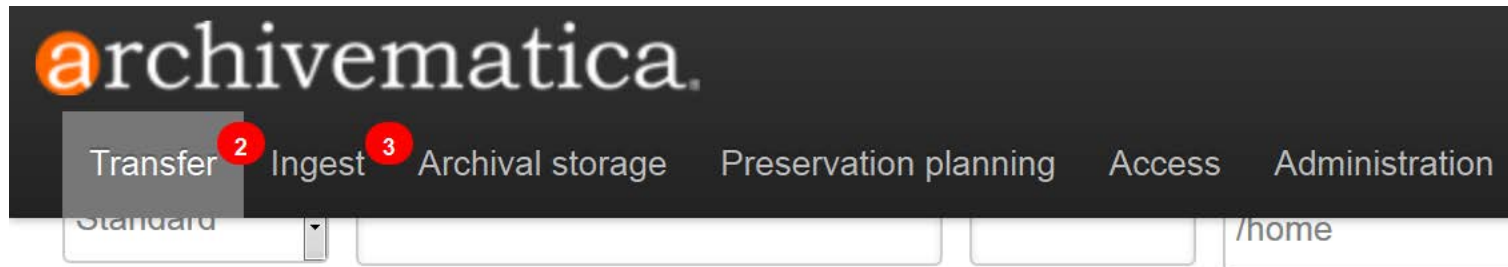
Archivematica is a free and open-source digital preservation system that is designed to maintain standards-based, long-term access to collections of digital objects. Archivematica is packaged with the web-based content management system AtoM for access to your digital objects.

Open source OAIS

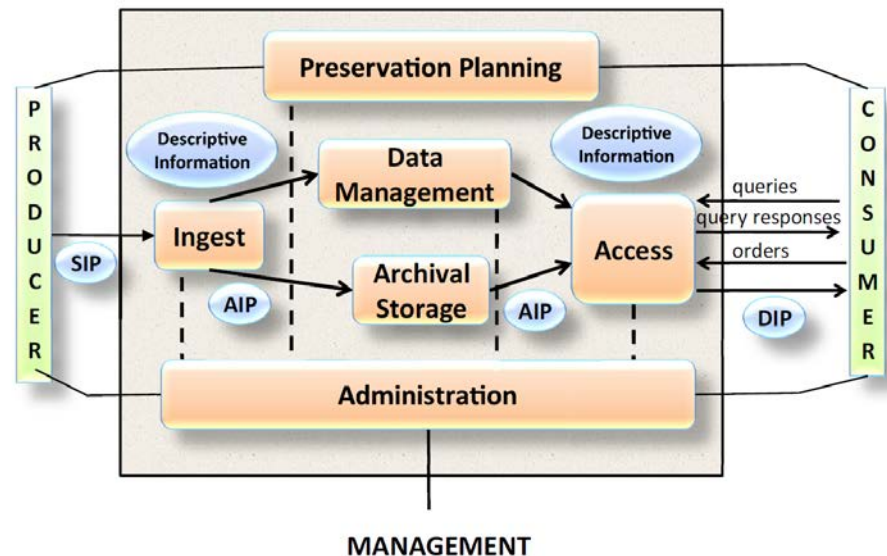
<https://www.archivematica.org>



Archivematica provides an integrated suite of free and open-source tools that allows users to process digital objects from ingest to archival storage and access in compliance with the [ISO-OAIS](#) functional model and other digital preservation standards and best practices. All of the Archivematica code is released under a [GNU Affero General Public License](#) and Archivematica documentation is released under a [Creative Commons Attribution-ShareAlike 4.0 International License](#).



Transfer
Ingest
Archival Storage
Preservation Planning
Access
Administration



Beispiele für Mikroprozesse

Transfer

Micro-service	Description
Generate METS.xml document [Expand]	Generates a basic METS file with a fileSec and structMap to record the presence of all objects in the <code>/objects/</code> directory and their locations in any subdirectories. Designed to capture the original order of the transfer in the event the user chooses subsequently to delete, rename or move files or break the transfer into multiple SIPs. A copy of the METS file is automatically added to any SIP generated from the transfer.
Extract packages [Expand]	Extracts objects from any zipped files or other packages. Extracts attachments from maildir transfers.
Scan for viruses [Expand]	Uses ClamAV to scan for viruses and other malware. If a virus is found, the transfer is automatically placed in <code>/sharedDirectoryStructure/failed/</code> and all processing on the transfer is stopped.
Clean up names [Expand]	Some file systems do not support unicode or other special characters in filenames. This micro-service removes prohibited characters and replaces them with dashes. Original filenames are preserved in the PREMIS metadata.
Characterize and extract metadata [Expand]	Identifies and validates formats and extracts object metadata using the File Information Tool Set (FITS). Also identifies file extensions, which are used for selecting normalization paths.

Ingest

Micro-service	Description
Prepare DIP [Expand]	Creates a DIP containing access copies of the objects, thumbnails and a copy of the METS file.
Upload DIP [Expand]	Allows the user to choose to upload the DIP to either ICA-AtOM or CONTENTdm.
Upload DIP to ICA-AtOM [Expand]	The user uploads the DIP to a selected description in ICA-AtOM.
Upload DIP to CONTENTdm [Expand]	The user uploads the DIP to a selected description in CONTENTdm.
Prepare AIP [Expand]	Creates an AIP in BagIt format. Indexes the AIP, then losslessly compresses it.
Store AIP [Expand]	Moves the AIP to <code>/sharedDirectoryStructure/www/AIPsStore/</code> or another specified directory. Once the AIP has been stored, a copy of it is extracted from storage to a local temp directory, where it is subjected to standard BagIt checks: <code>verifyvalid</code> , <code>checkpayloadoxum</code> , <code>verifycomplete</code> , <code>verifypayloadmanifests</code> , <code>verifytagmanifests</code> .

Beispiele für Mikroprozesse

Transfer	UUID	Transfer start time	
✓ test_0001	417730d8-206d-43b0-9641-130dfb6ba8ae	2016-03-01 16:52	 
▶ Micro-service: Create SIP from Transfer			
▶ Micro-service: Complete transfer			
▶ Micro-service: Examine contents			
▶ Micro-service: Validation			
▶ Micro-service: Characterize and extract metadata			
▶ Micro-service: Update METS.xml document			
▶ Micro-service: Extract packages			
▶ Micro-service: Identify file format			
Job: Identify file format		Completed successfully	

Task UUID: 5d9bf84e-ea64-4bcb-b9f9-37d1c57585ca
File UUID: d4312203-a958-4612-834a-93ef13369662
File name: MARBLES.TGA
Client: sandbox14_2
(exit code: 0)

Start time: Dienstag, 1. März 2016 16:57:05
End time: Dienstag, 1. März 2016 16:57:05
Created time: Dienstag, 1. März 2016 16:57:03
Duration: < 1 second(s)

[Show arguments](#)

```
identifyFileFormat_v0.0 "None" "/var/archivematica/sharedDirectory/watchedDirectories/workFlowDecisions/selectFormatIDToolTransfer/test_0001-417730d8-206d-43b0-9641-130dfb6ba8ae/objects/pictures/MARBLES.TGA" "d4312203-a958-4612-834a-93ef13369662" --disable-reidentify
```

STDOUT

```
IDCommand UUID: None  
File: (d4312203-a958-4612-834a-93ef13369662) /var/archivematica/sharedDirectory/watchedDirectories/workFlowDecisions  
/selectFormatIDToolTransfer/test_0001-417730d8-206d-43b0-9641-130dfb6ba8ae/objects/pictures/MARBLES.TGA  
Skipping file format identification
```

[Ingest](#) / [test_0001](#) / Micro-Services

Micro-Services

test_0001

Clean up names

- **Set file permissions:** Completed successfully

Prepare AIP

- **Copy submission documentation:** Completed successfully

Process metadata directory

- **Move metadata to objects directory:** Completed successfully
- **Assign checksums and file sizes to metadata :** Completed successfully

Verify transfer compliance

- **Verify mets_structmap.xml compliance:** Completed successfully

Clean up names

- **Load Dublin Core metadata from disk:** Completed successfully

Reject DIP

- **Move to the rejected directory:** Completed successfully

Store AIP

- **Store AIP location:** Completed successfully

Transcribe SIP contents

- **Transcribe:** Completed successfully

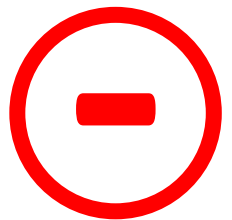
Erste Erfahrungen mit archivematica

- **INGEST-Werkzeug** zur Erzeugung von AIPs
- **OAIS-konform** mit Modifikationen
- **Workflow**-orientiert, **Archivobjekt**-orientiert
- hohe **Transparenz**, hohe **Flexibilität** bei Anpassung der Komponenten
- für **standardisierte** und **automatisierte** Übernahmeprozesse geeignet
- Bestandserhaltung - **Formatmigration** möglich (Bestandserhaltungs-Planung in der Entwicklung)
- Transferprozess erzeugt valide **SIPs** und kann vor Übernahme **Viren**-, **Format**- und **Validitäts-Überprüfungen** vornehmen

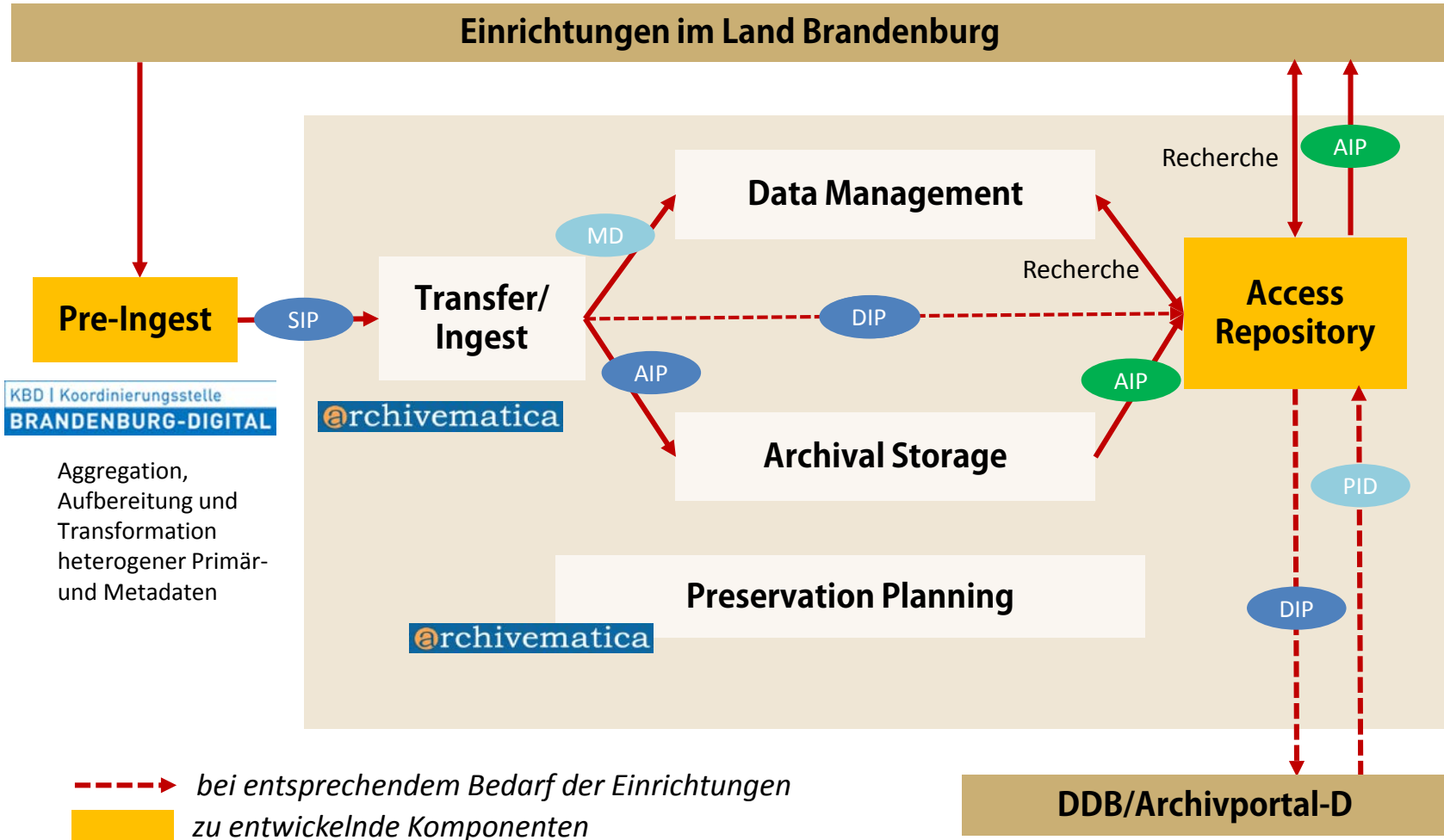


Erste Erfahrungen mit archivematica

- **Systemumgebung, Administration, Wartung** aufwändig
→ personeller Aufwand
- Unzureichende **Dokumentation**, wenig **Support** → personeller Aufwand
- **Automatisierte Übernahme von Metadaten** (METS oder DC/csv-Container) problematisch
- **Skalierbarkeit** bei Massenübernahme oder großen Objekten (audio/video) → leistungsfähige Technik
- **Einfaches Objektmodell** (Einzelobjekte und Sammlung von Einzelobjekten in einem Ordner)
- **Metadatenstruktur des AIPs** (viele Redundanzen, ungenügende Strukturierung der MD) → Nachbearbeitung erforderlich



Kooperative Verbundlösung - Architektur



Kooperative Verbundlösung Fazit (1)

- Spartenübergreifende **kooperative Archiv-Verbundlösungen** müssen einen **universellen Charakter** haben.
- Dazu sind bestimmte **OAIS-Kernkomponenten** erforderlich: ein digitales **Magazin mit einer Magazinverwaltung** zur Recherche und Bereitstellung der Archivpakete, ein weitgehend **automatisierter und standardisierter Ingest-Prozess** und ein **leistungsfähiger Archivspeicher**. Dazu die **digitale Bestands-erhaltungsplanung**.
- Für einen standardisierten Ingest-Prozess ist ein **Pre-Ingest-Prozess** erforderlich, über den eine **Aggregation, Aufbereitung und Transformation** heterogener Primärdaten und Metadaten erfolgt. Erzeugung eines **Mindest-Metadatensets**.
- Das digitale Archiv übernimmt dabei die **Verantwortung** für die langfristige **Formatmigration** sowie den **Erhalt der Authentizität und Integrität** der abgelieferten digitalen Objekte.
- Die **Verantwortung** für den Erhalt der **Information** bzw. der **Performance** bleibt dagegen beim Produzenten, der gleichzeitig auch der Konsument ist.

Kooperative Verbundlösung Fazit (2)

Eine **Archiv-Verbundlösung** ermöglicht es kleineren und mittleren Einrichtungen ohne eigene archivische Infrastruktur digitale Objekte **langfristig und nachhaltig** aufzubewahren.

Insbesondere sind kleinere Einrichtungen heute **nicht** mehr in der Lage:

- die **technische Infrastruktur** eines modernen Archivspeichers mit üblicher hierarchischer Speichertechnologie (z.B. RAID-Festplattensysteme und Bandroboter) vorzuhalten und
- die **erforderlichen personellen Ressourcen** für den langfristigen Betrieb einer digitalen Archivlösung bereitzustellen.

Eine Archivlösung für alles?



- durch Übernahme spartenabhängiger Metadaten **ohne manuelles Metadaten-Mapping**
- durch inhaltliche **Erschließung durch Produzenten**
- durch **Aggreagation und Normalisierung der Metadaten** im Pre-Ingest
- durch **Verzicht auf** Abbildung von **Archivtektonik** und **komplexen Archiv-Objektstrukturen** (Webseiten, Akten)
- durch eine **objektbezogene Übernahme** (im Gegensatz zur verzeichnisorientierten Übernahme)
- durch die **spezielle Organisationsform**: Archivobjekte bleiben in der Eigenverantwortlichkeit der abgebenden Institution