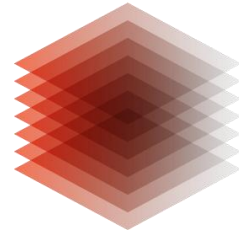

LEIBNIZ-INFORMATIONSZENTRUM
TECHNIK UND NATURWISSENSCHAFTEN
UNIVERSITÄTSBIBLIOTHEK



TIB

Preservation-as-a-Service LZA Dienstleistung der TIB

Thomas Bähr
Leiter Bestandserhaltung & LZA
7. Bibliothekskongress, Leipzig, 19. März 2019

Agenda



3 Perspektiven einer Dienstleistung

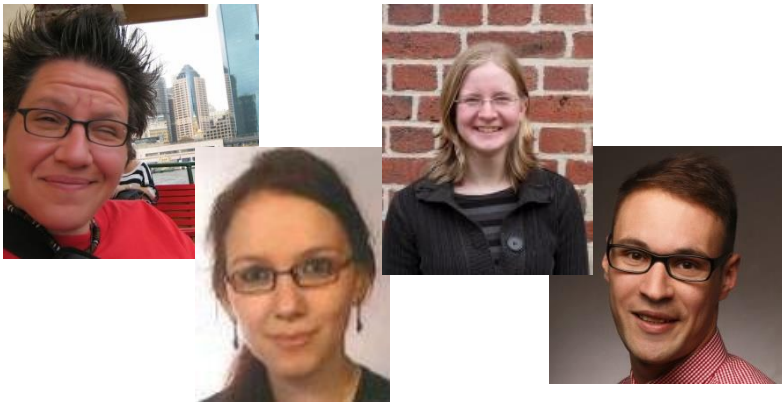
1. Eigener Betrieb
2. Konsortialangebot
3. Preservation-as-a-Service

Eigener Betrieb – Organisationsstruktur

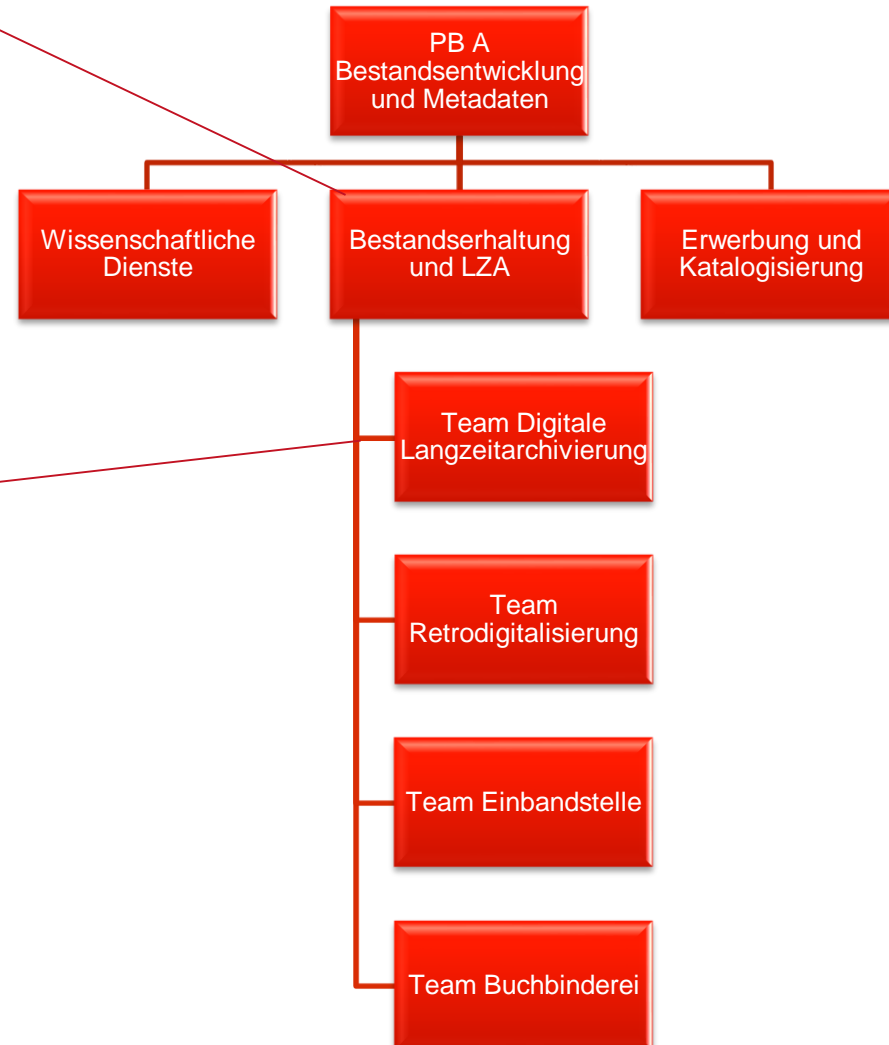
Bereichsleitung



Team LZA



EDV / Administrator



Eigener Betrieb – Technische Ressourcen

3 Umgebungen – DEV, TEST, PROD

- Jede Umgebung ist redundant ausgelegt (2, 4, 4 Application Server + eigene DB + eigener AppServer)
- Solaris (virtualisierte Zonen), Oracle, Debian
- ZFS Storage via NFS angebunden gespiegelt (neues Stagesystem mit Einbindung Uni-RRZ aktuell in Planung)

Integrationen (Auswahl)

- Katalog: hbz Aleph / GBV PICA
- Repositories: DSpace, GMS Harvester; Submission Application / SIP Packer, easyDB Anbindung, goobi2Rosetta Plugin
- Plugins: DROID, JHOVE, itext Migrationstool, Sophos Virensan
- Datenlieferung: Deposit API, OAI-PMH, SFTP, Datenträger

Eigener Betrieb – Ressourcen / Know-how

Welche Ressourcen sind erforderlich?

- 1 MA Dokumentation, Policies, Zertifizierungen, Community Watch
- 1 MA Technical Analyst – Formatspezialist, Business-Analyst, Technology Watch
- 1 MA Systemadministration
- 1 MA Entwicklung, Anwendungsadministration
- 1 MA Leitung (anteilig), Organisation, Verhandlungen, Verwaltung, Recht, Beratungen, Organisation Watch

Welches Know-how ist erforderlich?

- Kenntnisse der Prozesse, Bestände und Systeme im Haus
- Konzeption und Anwendung von Datenstrukturen
- Datei-Formate (Aufbau, Fehler, Analysetools, Risiken)
- LZA-Standards und Prozesse (OAIS, Migration, Emulation)
- Erwartungen der Kunden (Designated Community, Significant Properties)
- Metadatenformate und –standards (PREMIS, DC, METS, XML, RDF)
- Grundlegende Scriptkenntnisse (und fleißige studentische Hilfskräfte)

Eigener Betrieb – Organisatorische Faktoren

Bestände

- Dissertationen, graue Literatur, Deutsche Forschungsberichte, AV, 3D Objekte, CD/USB Images, ...

Verständnis LZA als organisatorischer Prozess

- Qualitätskontrolle, auch außerhalb der LZA
- Archivierung als Bestandteil von Lizenzvereinbarungen
- Beratung von Teams und Datenlieferanten

Preservation Policy - öffentliche Beschreibung der Prozesse:

<https://wiki.tib.eu/confluence/display/lza/Digitale+Langzeitarchivierung+an+der+TIB>

Zertifizierung

- Data Seal of Approval (2015)
- nestor Seal (2017)
- Core Trust Seal (geplant 2019)

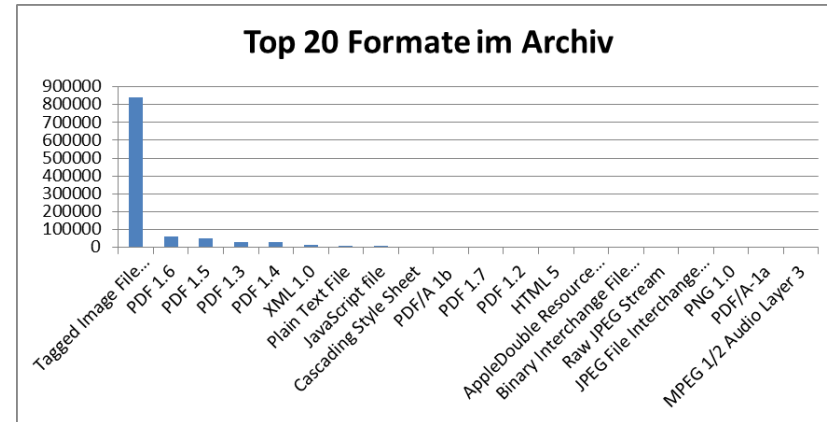


Netzwerkarbeit

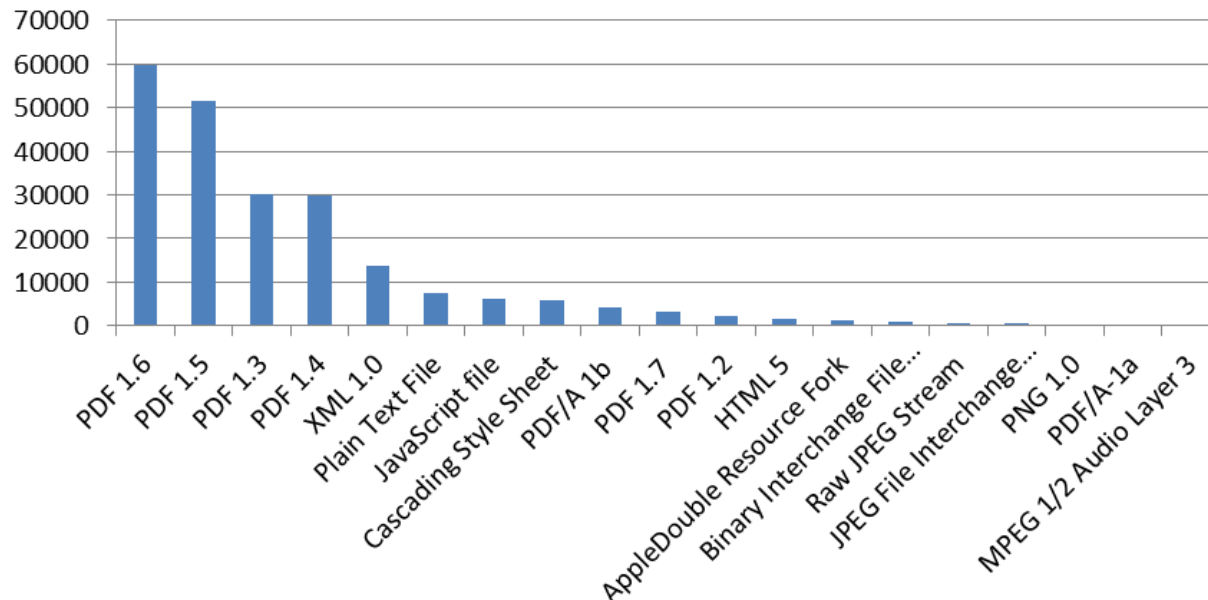
- Nestor, Open Preservation Foundation, Deutschsprachige Rosetta Anwender Gruppe, Rosetta Steering Committee, ...

Eigener Betrieb – Datenstand

- Dateien: 1.064.264
- Intellektuelle Einheiten: 91.209
- Formate: 98
- Zuwachs 2019: > 430 TB (Digitalisate)



Top 20 Formate im Archiv (ohne TIFF)



Konsortialangebot

TIB Hosting von Rosetta im Verbund der 3 Zentralen Fachbibliotheken

Bereitstellung und Wartung der HW/SW, Kommunikation mit dem SW-Dienstleister, Support, zentrale Konfigurationen, ...

MANDANTEN konfigurieren eigene Workflows und definieren Archivierungsprozesse im System. ZBW und ZB MED arbeiten als MANDANTEN im System der TIB.



Konsortialnehmer – Ressourcen / Know-how

Welche Ressourcen sind erforderlich?

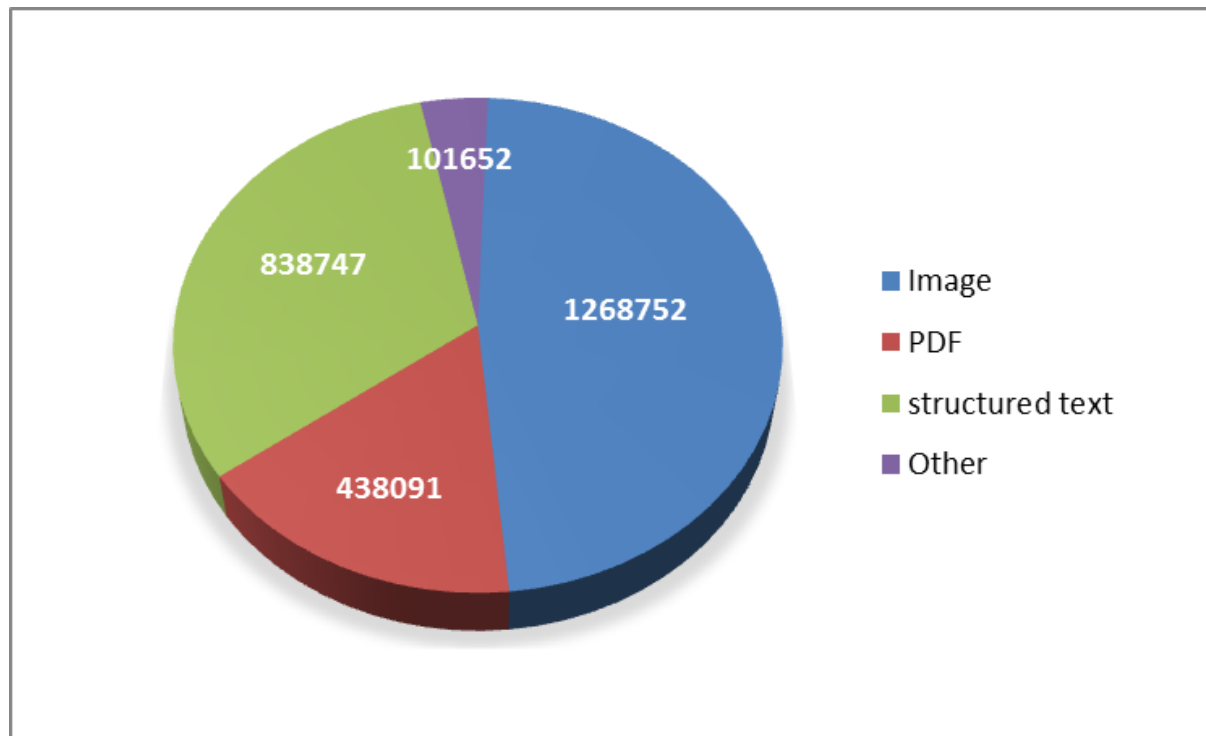
- 1 MA Dokumentation, Policies, Zertifizierungen, Community Watch, Technical Analyst, Technology Watch
- 1 MA Entwicklung (anteilig) für Schnittstellen zu eigenen Systemen)
- 1 MA Leitung (anteilig), Organisation, Verwaltung, Recht, Organisation Watch

Welches Know-how ist erforderlich?

- Kenntnisse der Prozesse, Bestände und Systeme im Haus, bei älteren Beständen auch historische Bestandskenntnis
- Konzeption und Anwendung von Datenstrukturen
- Datei-Formate (Aufbau, Fehler, Analysetools, Risiken)
- LZA-Standards und Prozesse (OAIS, Migration, Emulation)
- Erwartungen der Kunden (Designated Community, Significant Properties)
- Metadatenformate und –standards (PREMIS, DC, METS, XML, RDF)
- Grundlegende Scriptkenntnisse und fleißige studentische Hilfskräfte für die Aufbereitung der Datenpakete

Konsortialangebot – Datenstand

- Dateien: 2.647.242
- Intellektuelle Einheiten: 575.669
- Formate: 127
 - Davon ca. 50% mit weniger als 10 Instanzen im Archiv
- Volumen (nur PRESERVATION MASTER Dateien): ~ 3.6 TB



Preservation-as-a-Service

Full-SERVICE Angebot

TIB übernimmt Mandantenrolle für **DIENSTLEISTUNGSNEHMER** und unterstützt in allen Prozessen. DIENSTLEISTUNGSNEHMER haben keinen direkten Zugriff auf das System.



BERATUNG



WORK-SHOPS



PRE-INGEST



ARCHIVIERUNG

PaaS – Art und Umfang

Dienstleistungsumfang

- Bereitstellung von Infrastruktur und Hardware inkl. Konfiguration, Monitoring und Backup
- Bereitstellung von Software inkl. kundenspezifischer Konfiguration und Qualitätskontrollen
- Beratung
- Pre-Ingest
- Ingest
- Archivierung
- Preservation Management
- Access
- DOI-Vergabe

Materialien in den entsprechenden Standardformaten

- Textuelle Materialien
- Nicht-textuelle Materialien (Bild, AV, 3-D)



<https://groups.google.com/forum/#!topic/digital-curation/SneLIDoqKdM>

PaaS – Service-Level

1. Level

- Beratung und Konzeptentwicklung zur Erstellung von SIP
- Pre-Ingest (Analyse und Strukturierung der Objektpakete und Metadaten)
- Ingest und Erstellung eines AIP gemäß Konzept
- Bitstream-Preservation
- Erstellung von Berichten und Statistiken
- Auslieferung von AIP an den Kunden (Trigger Event)

2. Level

Wie Level 1 plus:

- Technology Watch
- Preservation Management (ohne Erhaltungsmaßnahmen berechnet)

3. Level

Wie Level 2 plus:

- Access (permanenter Zugriff auf DIP)

<http://danbubanygolf.com/wp-content/uploads/2015/11/Bad-customer-service.jpg>





PaaS – Preservation Management

Technology / Community / Preservation Watch

- Existieren Risiken für Technologien / Format?
- Welche Formate erwarten meine Nutzer?
- Wird ein bestimmtes Format / eine Technologie nicht mehr genutzt?
→ Wissen über Technik / Zielgruppe sammeln



Preservation Management

- Welche Datenformate habe ich im System?
- Welche Risiken treffen auf die Objekte in meinem Bestand zu?
→ Wissen über das, was ich im Bestand habe



Preservation Planning & Action

- Muss / kann ich die Format anders zur Verfügung stellen?
- Welche Technologie / welches Format ist geeignet?
→ Eine Maßnahme planen und durchführen



PaaS – Ressourcen / Know-how

Welche Ressourcen sind erforderlich?

- 1 MA mit Bestandskenntnissen
- Zugangsmöglichkeit IT, Recht, Produzenten, ...
- Chefsache

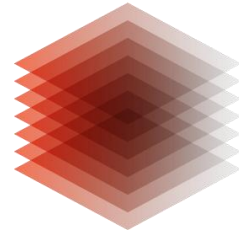
Welches Know-how ist erforderlich?

- Kenntnisse der Prozesse und Systeme im Haus, bei älteren Beständen auch historische Bestandskenntnis
 - Zwingend durch Kunden
- Konzeption und Anwendung von Datenstrukturen
 - Kunde oder DL durch TIB
- Metadatenformate und –standards im Bereich der digitalen LZA, aber auch der Erschließungsstandards der jeweiligen Domäne
 - Kunde oder DL durch TIB
- Grundlegende Scriptkenntnisse oder fleißige studentische Hilfskräfte für die Aufbereitung der Datenpakete Schnittstellen
 - Kunde oder DL durch TIB

PaaS – Benefits für den Kunden (Auswahl)

- Strukturierung der eigenen Datenablage
- Überprüfung der Zuordnung der Dateien zu den Metadaten
- bei älteren Digitalisierungsprojekten - strukturierten Datenablage für MASTER und Verweis auf Metadaten
- Klärung, wenn mehr als eine Repräsentation der Datei vorhanden
- häufig redundante Datenhaltung - kontrollierten Redundanz im Sinne von Mehrfachkopien oder ungewollten Redundanz aufgrund nicht dokumentierter Datenablage
- für Export der Metadaten wird Ergebnisqualität (Metadatenstandards und –format, Vollständigkeit der Metadaten).geprüft und bereinigt
- Bessere Dokumentation der Bestände
- Prüfung unklarer Rechte

LEIBNIZ-INFORMATIONSZENTRUM
TECHNIK UND NATURWISSENSCHAFTEN
UNIVERSITÄTSBIBLIOTHEK



TIB

Fragen?

thomas.baehr@tib.eu