

Langfristige Datenqualität als Landesaufgabe - LZA Landesinitiative Niedersachsen

LaVaH Workshop

Goethe-Universität Frankfurt

Thomas Bähr

Frankfurt a. M., 25. September 2025

Wann fängt „Langzeit“ an?

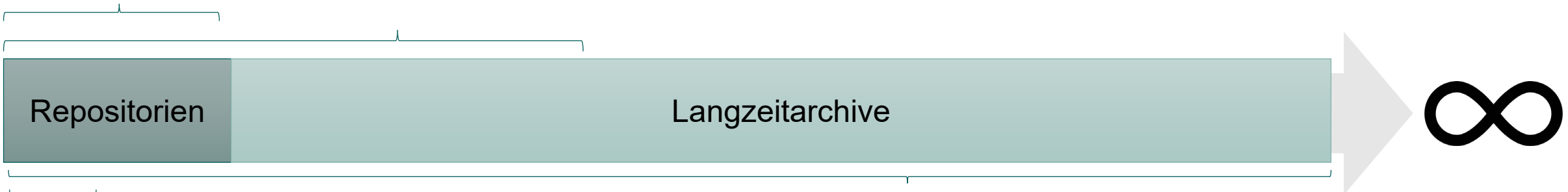
Repositorien sind keine Langzeitarchive!

10 Jahre

z. B. Forschungsdaten
aus Drittmittelprojekten;
elektronische Rechnungen

25 Jahre

z. B. Daten aus medizinischen Studien;
Fertigungs- und Konstruktionsdaten im Zusammenhang
mit der Einhaltung von Sicherheits- und Umweltnormen



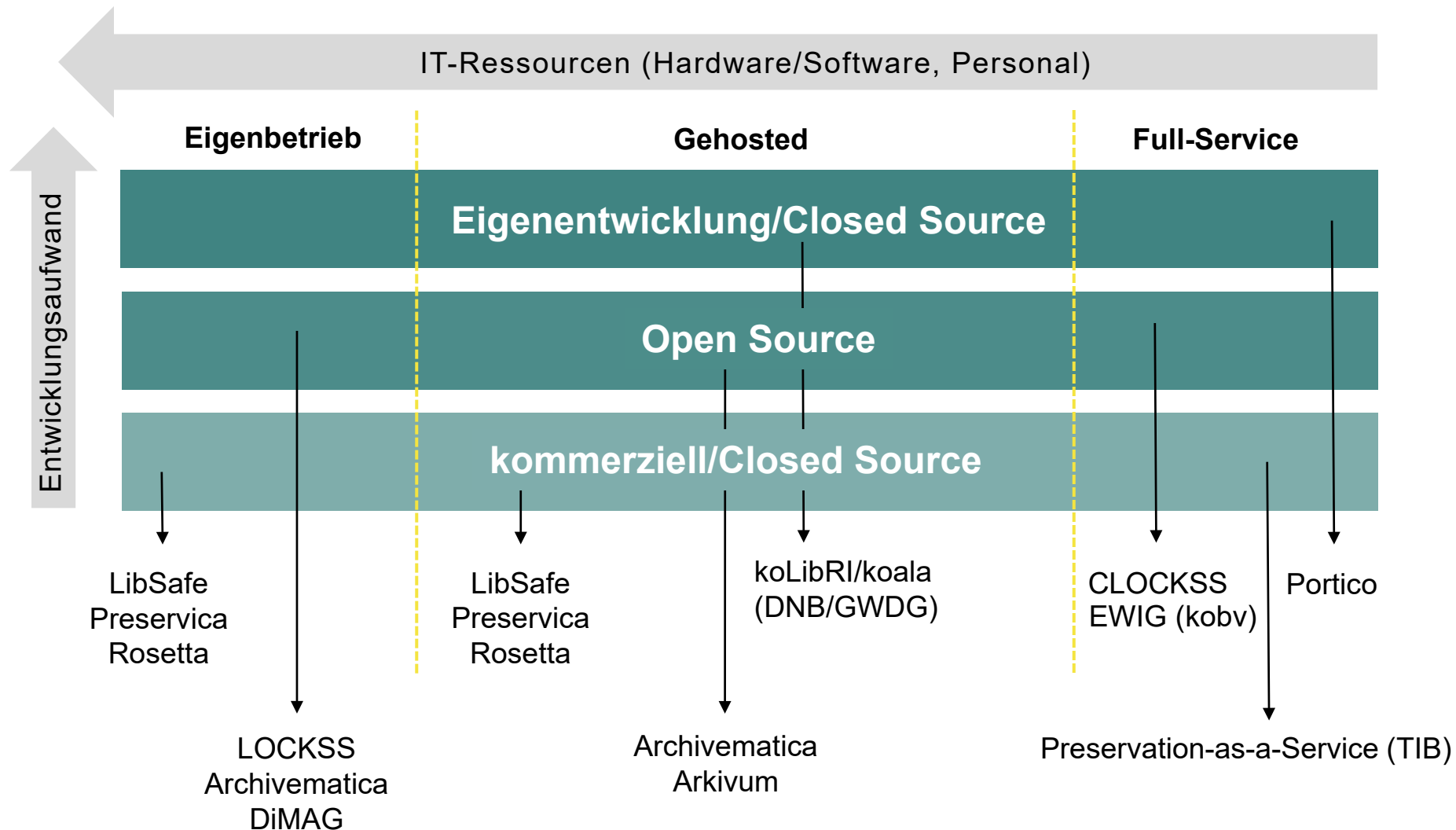
0 bis 5 Jahre

z. B. institutionelle Forschungsdaten (je nach Policy);
Internet Service Provider & Telekommunikationsdaten
(je nach Land)

Unbegrenzt

z. B. digitales Kulturerbe; wissenschaftliche
Arbeiten; ausgewählte Daten der öffentlichen
Hand; Informationen zur Lagerung radioaktiver
Abfälle

Allgemeine Landschaft von LZA-Lösungen



Lösungen unterscheiden sich weiter durch unterschiedliche Unterstützung von:

- LZA-Prozessen
- Content-Typen/Dateiformaten
- Offenen Schnittstellen
- Standards
- Erweiterbarkeit

Software unterstützt die LZA, löst sie aber nicht.

Digitale Langzeitarchivierung an der TIB

3.4 Mio.
Archivpakete

31.7 Mio.
Dateien

296
Dateiformate

- **Aktiv seit 2009**, Produktivbetrieb seit 2015
- Regelmäßige Zertifizierungen als **vertrauenswürdiges digitales Langzeitarchiv** (2015, 2017, 2020, 2022, 2025)
- **Personelle Ressourcen:**
14 Personen Team LZA + 6 Personen anteilig in Erwerbungs-teams und IT
- **Infrastrukturelle Ressourcen:**
3 Systemumgebungen; Produktivsystem mit 4 Application Servern + db-Server;
Softwarekern: Rosetta von Ex Libris
- **Netzwerke:**
107 Kooperationspartner aus über **25 Ländern**



Vom Content zur Dienstleistung



TIB

Archivierung des
TIB-Contents



Kund:innen

Preservation-as-a-
Service



Fachcommunitys

Entwicklung für TIB-
Communitys



LZA-Community

Entwicklung
internationaler
Lösungen

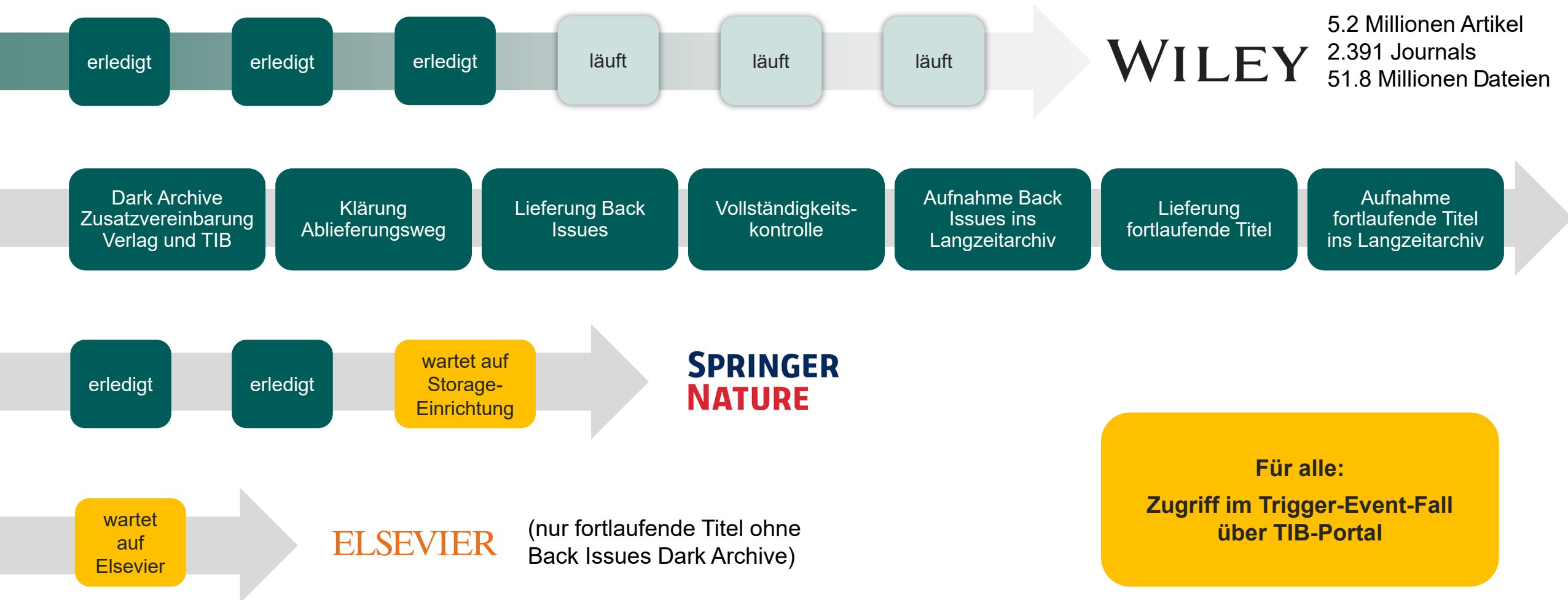


Deutschland

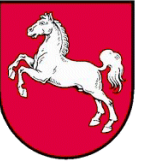
nationale
Archivierungs-
aufgaben

LZA als kooperative Informations-Infrastruktur

IST-Stand DEAL Dark Archives an der TIB



LZA Landesinitiative Niedersachsen



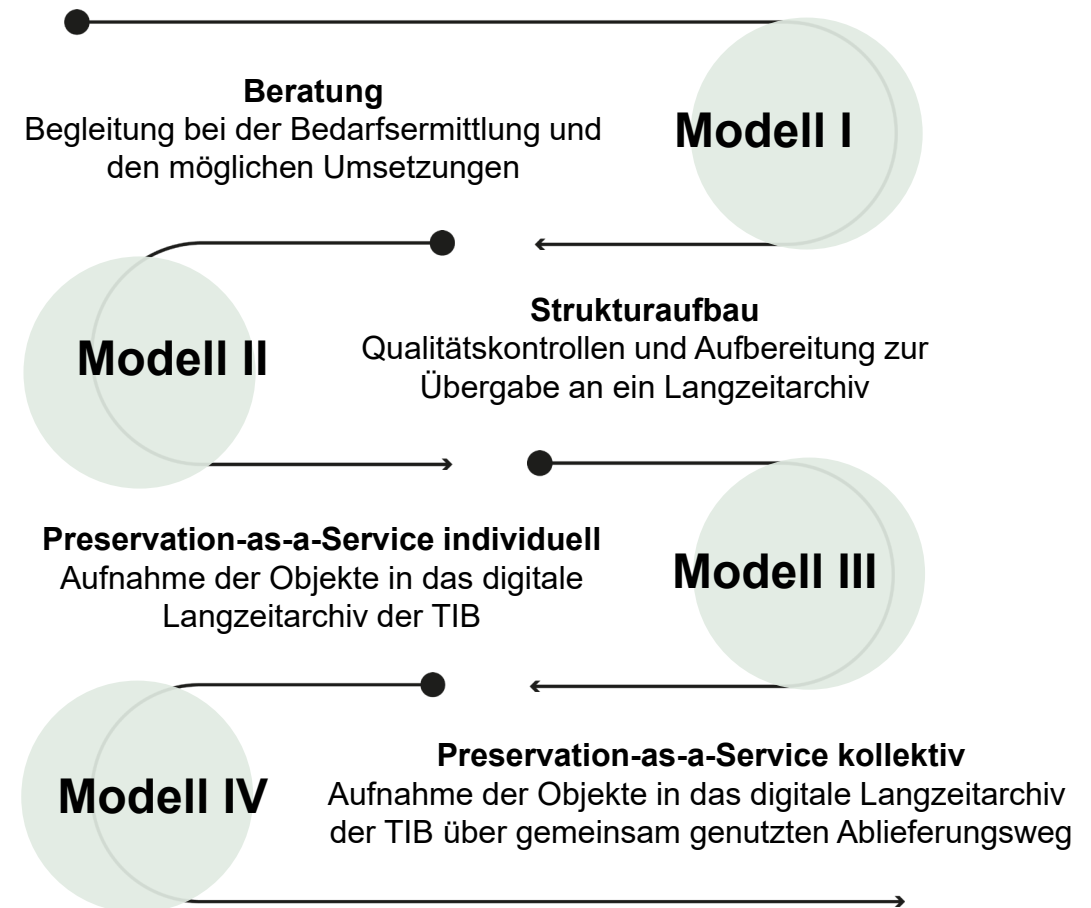
Ziel: Bedarfsgerechte Unterstützung in der digitalen Langzeitarchivierung für 30 niedersächsische wissenschaftliche Bibliotheken, Landesbibliotheken und Landesmuseen.

Förderung: 1.853.345 € durch Niedersächsisches Ministerium für Wissenschaft und Kultur

Laufzeit: 1. April 2025 bis 31. März 2030

Umsetzung:

- 6 Institutionen pro Jahr
- von Beratung bis Preservation-as-a-Service
- Steuerungsgruppe und verschiedene Arbeitsgruppen
- flankiert durch Andockprojekte



Zielsetzungen

- Klares Bekenntnis zu einem kooperativen Modell in der Langzeitarchivierung auf Landesebene
- Abgestimmte, effiziente Ressourcennutzung und Senkung der Kosten(steigerungen)
- Sicherstellung eines qualitativ hochwertigen, nachhaltig abgesicherten Angebots von skalierbaren Infrastrukturen und kooperativen Services
- Möglichkeit für die einzelnen Institutionen, aus verlässlichen und überzeugenden Angeboten einen passenden Service auszuwählen
- Schaffung klarer und praxistauglicher (und schlanker!) Strukturen und Organisationsformen mit rechtlich abgesicherten Kooperations- und Verrechnungsmodellen
- Grundlage für die Aushandlung eines strukturellen Finanzierungsplanes innerhalb der Institutionen und mit dem Land Niedersachsen

Häufige Ausgangssituation in den Institutionen

- Kaum personelle Ressourcen und wenig technisches Know-how vorhanden
- Unterschiedliche Anforderungen aus der Universität – Bibliotheksbestände, Forschungsdaten, Archivbestände
- Datenstruktur selten für die LZA nachnutzbar
- direkte Zuordnung der Dateien zu den Metadaten in der Regel nicht vorhanden
- Digitalisierungsprojekte: Nutzungskopien gut aufbereitet und mit Metadaten verknüpft - Master-Dateien schlecht strukturiert, ohne direkten Verweis auf Metadaten
- häufig mehr als eine Repräsentation der Dateien
- kontrollierten Redundanz im Sinne von Mehrfachkopien oder ungewollten Redundanz aufgrund nicht dokumentierter Datenablage
- Verzeichnissysteme sind sehr heterogen
- Export der Metadaten meist möglich aber Metadatenstandards, -formate und Vollständigkeit sehr unterschiedlich
- Bestände meist nicht oder nur unzureichend dokumentiert, Vollständigkeit und Qualität oft unklar

LZA Optionen und notwendige Ressourcen

Eigener Betrieb

- Eigene Hardware und Software
- Min. 5 MA
- Administration HW, SW, Systemumgebung
- Kenntnisse über: Prozesse, Bestände, 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

Kooperativer Betrieb

- Min. 3 MA (beim Partner)
- Teilweise Administration SW
- Kenntnisse über: Prozesse, Bestände, 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

Kuratierte Dienstleistung

- Min 1 MA (beim Kunden)
- Kenntnisse über: Prozesse, Systeme im Haus, evtl. historische Bestandskenntnis, Erschließungsstandards
- Mitarbeit bei: Konzeption und Anwendung von Datenstrukturen (Objekte, Metadaten)
- Hilfreich: Grundlegende Scriptkenntnisse oder fleißige studentische Hilfskräfte für die Aufbereitung der Datenpakete Schnittstellen

Qualitätssicherung durch kuratierte Langzeitarchivierung

Kooperative LZA

- Klare Teilung von Verantwortung und Ressourcen, Netzwerkbildung
- Wissensaufbau bei Kooperationspartnern zwingend erforderlich
- Kostensenkung durch zentrale Hard- und Software

Kuratierte LZA

Wie kooperative LZA aber zusätzlich:

- Verantwortung für Inhalte liegt bei Partnerinstitution
- Unterstützung bei Planung, Analyse, Objektauswahl und –bewertung, Strukturierung je nach Möglichkeiten und Ressourcen der Partner – bei Bedarf Übernahme der Aufgaben durch Dienstleister
- Erweiterte Dienstleistungsangebote wie Schnittstellenentwicklung, Datenträgermigration, usw.
- Deutliche Entlastung der Partnerinstitutionen - hohe Anforderungen an Know-how und Personalressourcen beim Dienstleister

Die LZA Landesinitiative basiert auf einem kooperativen Ansatz der durch eine kuratierte LZA erweitert wird.

Vorteile kuratierter Langzeitarchivierung

Unterstützung bei der Umsetzung von Richtlinien und Standards

- Aufbau hausinterner Richtlinien auf der Basis allgemeiner Standards

Prozesse und Verfahren

- Orientiert sich ausschließlich an den Bedarfen und Möglichkeiten der Partnerinstitutionen

Qualitätssicherung

- Kontinuierliche Qualitätssicherung beim Dienstleister durch wiederholte Zertifizierungen
- Übernahme von Qualitätsstandards (hauptsächlich strukturell) durch Partner

Monitoring und Bewertung

- Preservation Management (Technology Watch, Community Watch, Preservation Planning) als zentrale Dienstleistung
- Dokumentationen interner Prozesse, Entwicklung von Policies
- Reporting aus der LZA

Die kuratierte LZA stellt bedarfsgerechte individuelle Lösungen zur Verfügung.

LZA ist Qualitätssicherung

3 verbreitete Irrtümer:

- (Verlags) Daten sind vollständig
- (Internet) Datenquelle ist verlässlich
- Rechte sind geklärt



Prüfung Vollständigkeit

- Fehler in den Inventories
- Abgleich (Meta-)Daten untereinander
- Kontrolle Ordnerstruktur
- Kontrolle Inhalt (XML/PDF)
- Kontrolle leere/sehr kleine Dateien
- Abgleich Anzahl der Dateien, Größe des Journals, Eindeutigkeit ISSN
- uvm...

~ 10 % Fehler unterschiedlicher Schwere

Prüfung Quelle

- Abgleich zwischen Publikations-Daten (OAI PMH) und Publikations-Daten (Website)
- temporäre oder dauerhafte Fehler in Website
- Vollständigkeit bei digitalisiertem Altbestand unklar
- nicht verlinkte Anhänge

Was gehört zur Publikation (Artikel, Supplements, ...)?

Prüfung Rechte

- Klärung Datenschutzstufe
- Liegen die Rechte für Archivierung vor (Download, Speichern, Migration)?
- Sind die Lizenzbedingungen für jedes einzelne Objekt bekannt?
- Vertrag für die gesamte Lieferung oder Rechteauszeichnung pro Publikation in den Metadaten?
- 12 verschiedene Lizenzen hinterlegt

Übernahmepolicies zwingend erforderlich!

Finanzierung der Digitalen Langzeitarchivierung

Rechtlichen Rahmen schaffen

- Gesetzliche Verpflichtung zur Archivierung für bestimmte Datenarten - Finanzierungspflicht klar regeln
- Zuständigkeiten für die Archivierung und Finanzierung müssen klar definiert werden, um Doppelungen zu vermeiden und die Effizienz zu steigern.

Finanzielle Instrumente entwickeln

- Finanzierung muss in die langfristige Haushaltsplanung von Bund und Ländern integriert werden
- vorausschauende Bedarfsplanung erforderlich

Öffentliche Sensibilisierung

- Bewusstsein für die Bedeutung der LZA bei politischen Entscheidungsträgern schärfen
- Ausbildung von Fachkräften im Bereich der digitalen Langzeitarchivierung ist essentiell

Politischen Konsens und eine **langfristige strategische Planung** sind für eine **nachhaltige Finanzierung** der digitalen Langzeitarchivierung notwendig.

Digitale Langzeitarchivierung ist eine **Daueraufgabe**, die personell und finanziell aufwändig ist und nicht nebenbei geleistet werden kann und leistet einen wesentlichen Beitrag zur **Wissenssicherheit!**

Vielen Dank!

Fragen?



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

Kontakt

Thomas Bähr

thomas.baehr@tib.eu

Weitere Infos unter: <https://www.tib.eu/de/publizieren-archivieren/digitale-langzeitarchivierung>



Creative Commons Namensnennung 3.0 Deutschland
<http://creativecommons.org/licenses/by/3.0/de>