

„Das ist nicht in unserer Verantwortung“

- Strategien zur nachhaltigen Bereitstellung lebender Systeme

Patrick Helling | Data Center for the Humanities (DCH), Universität zu Köln

5. SaxFDM-Tagung | Freiberg | 17. September 2024

1

Problemstellung



A Catalog of Digital Scholarly Editions (Sahle)

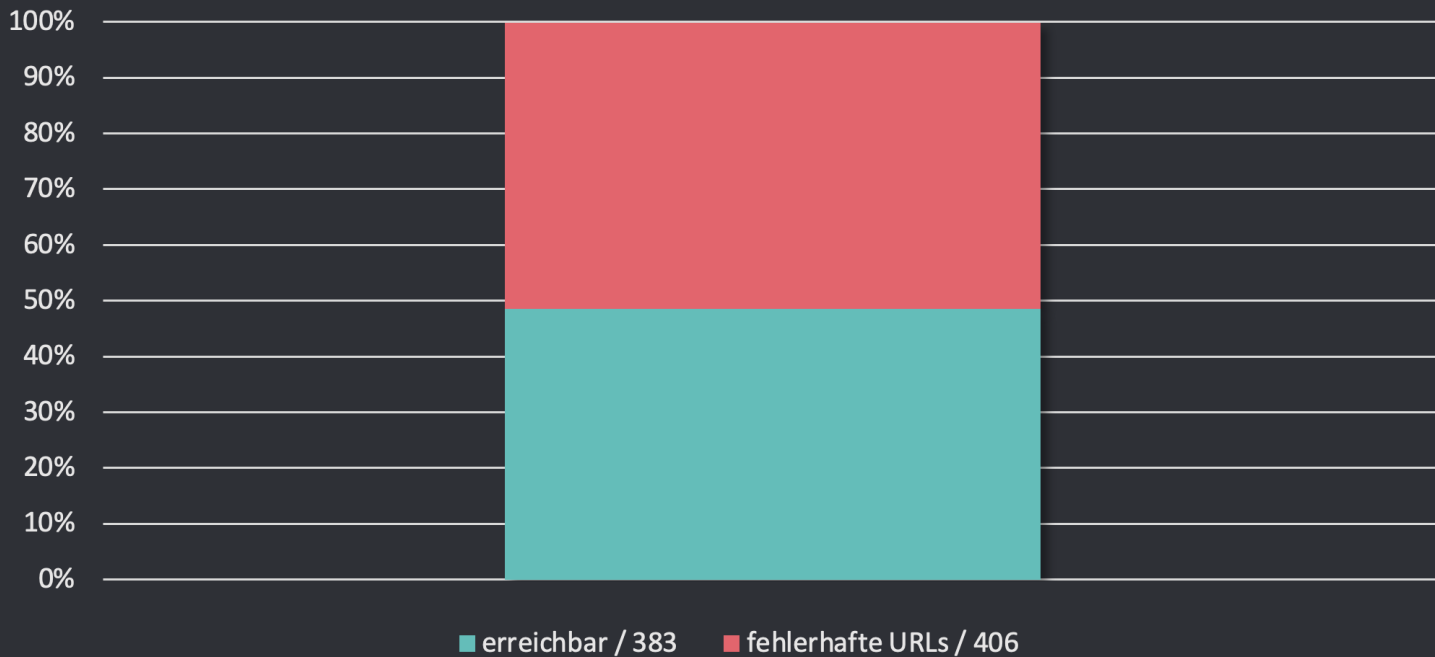
unvollständige (aber sehr umfangreiche) Liste Digitaler Editionen

- insgesamt 905 Digitale Editionen (Stand 28. August 2024)
- durchsuchbar nach verschiedenen Kategorien, bspw. edendum, subject, period, language

<http://www.digitale-edition.de/>

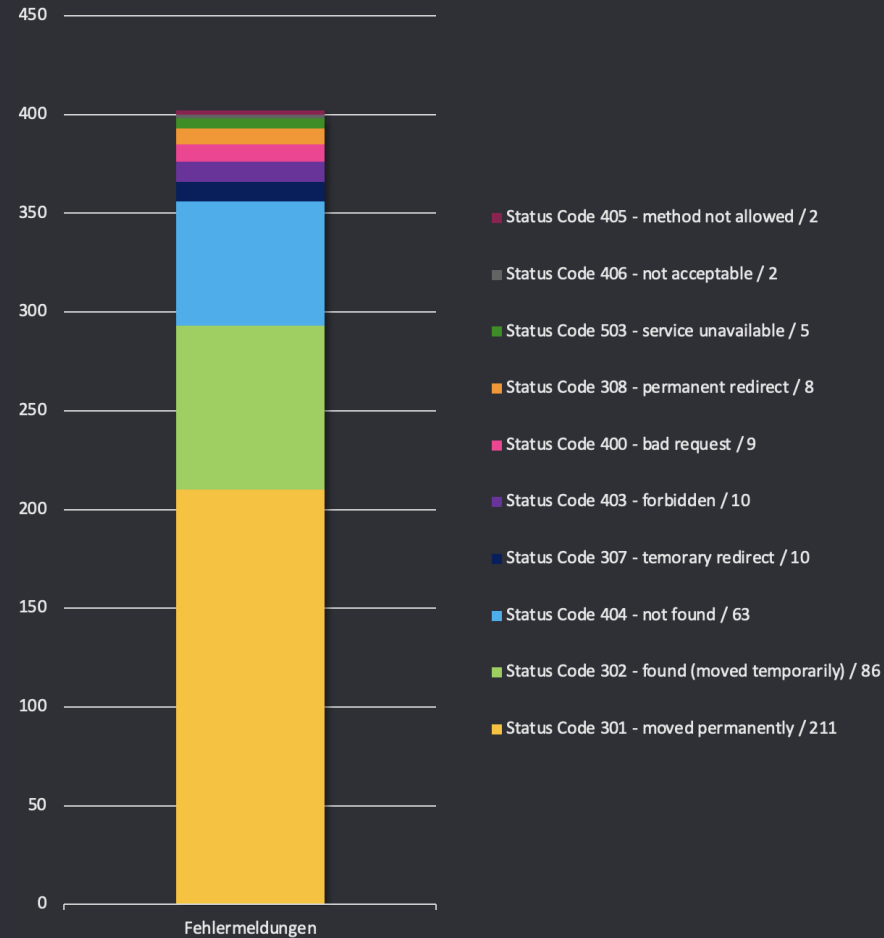
Zugänglichkeit digitaler Editionen

Gesamtzahl der gelisteten digitalen Editionen





Fehlermeldungen



N=406

- Quantitative Umfrage zum Umgang mit lebenden Systemen¹
in den Digital Humanities

○ Zielgruppe: 16 geisteswissenschaftliche Datenzentren im deutschsprachigen Raum

- Hosting lebender Systeme?
- Technische und organisatorische Rahmenbedingungen?
- Gründe für (nicht) Hosting lebender Systeme?

...tatsächlich haben nur vier Datenzentren an der Umfrage teilgenommen

¹Dynamische Umgebungen, Zugangspunkte zu/Repräsentationsformen für Forschungsdaten, die permanent kuratiert und betreut werden müssen. Hauptfunktion von lebenden Systemen ist die Sicherstellung der Zugänglichkeit zu Forschungsdaten in einer spezifischen Form.

● Rahmenbedingungen für das Hosting lebender Systeme

Projekt-
beteiligung

Spezifische Technologiestacks

Lauffähigkeit auf
spezifischer Infrastruktur

Solange es keine
Mehrarbeit erzeugt

dynamisch

statisch

Datenzentrum A



Datenzentrum B



Datenzentrum C



Datenzentrum D



nicht
leistbar

2

Existierende Strategien für das Management von lebenden Systemen

● Existierende Strategien für das Management lebender Systeme

1

Investition ausreichender finanzieller/personeller Ressourcen

- Skaliert nicht mit wachsender Anzahl lebender Systeme
- Wer bezahlt so etwas?

2

Restriktion von Technologiestacks

- Zu wenige Community-Standards
- Könnte im Widerspruch zur Wissenschaftsfreiheit stehen

3

Virtualisierung durch Technologien, bspw. Docker

- Bedarf spezifischer Infrastruktur
- Bedarf qualifiziertes Personal

4

Hosting lebender Systeme in statischer Form

- Nicht immer eine geeignete Strategie
- Verlust bereits entwickelter (notwendiger) Funktionalitäten

3

Ein reales Beispiel



Digitale Datenbank

Aufnahmen von Sprecher:innen einer spezifischen Sprache

700 Interviews je 2 Stunden

300 GB Daten, hauptsächlich .wav und .mp3

Zugänglichkeit über Website

+ zusätzliche Software, u. a. Lame, MP3splt, FFmpeg

● ...genauer gesagt

- Datenbank und Website wurden ursprünglich auf einem Windows Server an einem Institut einer anderen Universität gehostet
- Nachdem der Professor die Universität verlassen hatte, gab es niemanden, der/die sich um das lebende System hätte kümmern können
- Der Professor dachte, er stellt das Überleben der Ressource sicher, indem er sie an die Bibliothek seiner Universität übergibt

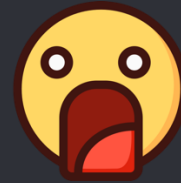
Die Jahre zogen ins Land...

- Die Bibliothek wusste nicht mehr, wie sie die Ressource weiterhin betreuen sollte (und genau genommen gab es auch nie eine Finanzierung dafür)

- ...genauer gesagt

- Die Bibliothek informierte den Professor Emeritus, dass sie die Ressource innerhalb eines Monats abschalten würde

... und Panik brach aus



● Lösung!?



Lobbyarbeit

Überzeugung verschiedener Stakeholder, u. a. der Universität, des lokalen Rechenzentrums, der Leitung des Datenzentrums



Implementation

eine simplen Suchfunktionalität



Migration

der Ressource auf einen Linux Server



Upload und Testing

Ablage des lebenden Systems und der Daten auf verschiedenen Infrastrukturen



Reduktion

von dynamischen Funktionalitäten

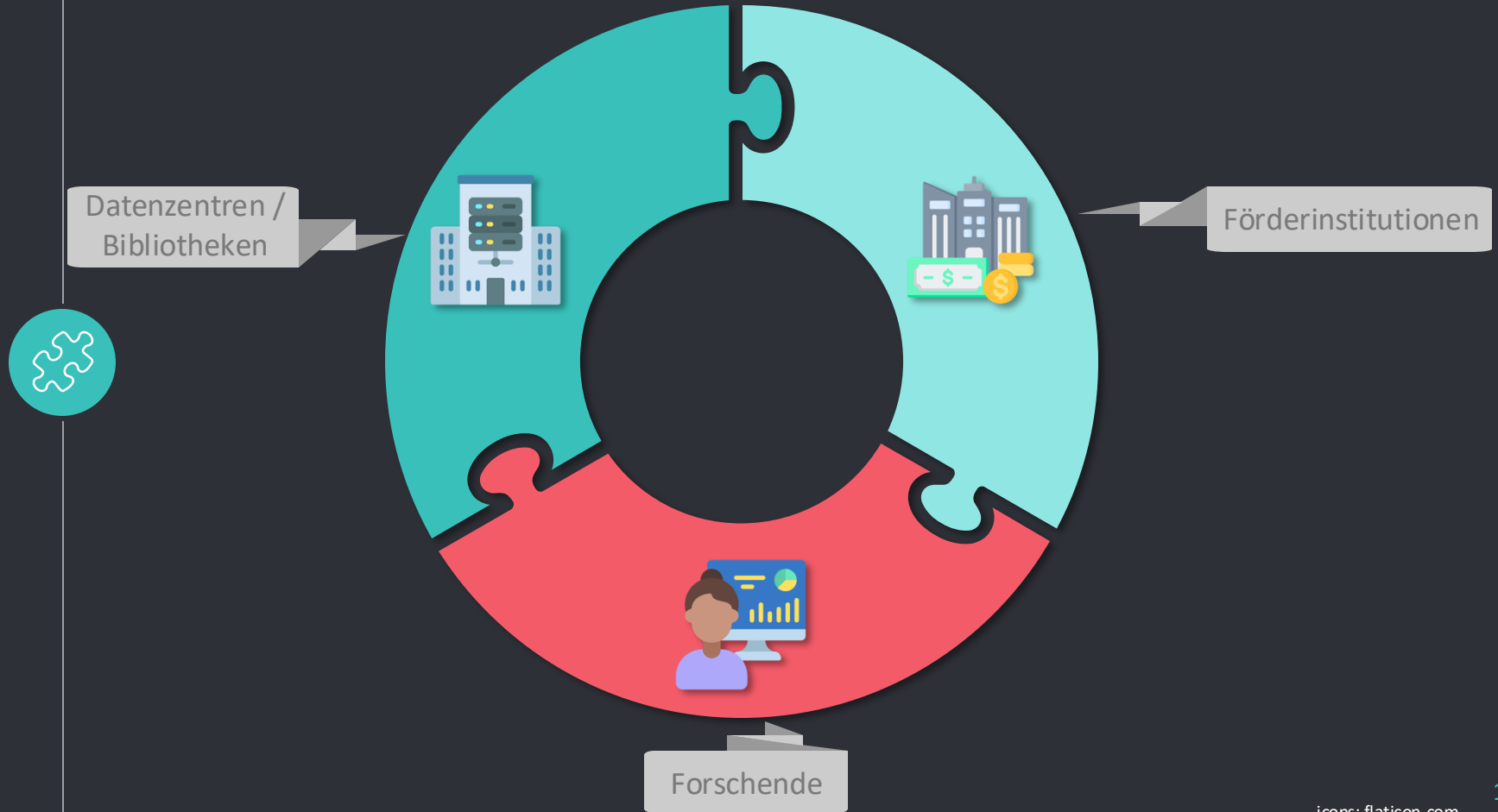


Anpassungen

...work in progress

Kernfrage

Wer ist verantwortlich für die nachhaltige Planung, Entwicklung und das langfristige Hosting sowie die Archivierung von lebenden Systemen?



Verantwortungen, Rollen und Motivation

Forschende

- Expert:innen in ihrem Fachbereich
- verantwortlich für die nachhaltige Entwicklung lebender Systeme
- erhalten i.d.R. öffentliche Gelder
- intrinsische Motivation
 - persönliche Investition in Forschung
 - Erhaltung der eigenen Forschung



- Expertise in Softwareplanung und -entwicklung
- Unterstützung durch Expert:innen und Institutionen
- ausreichend Geld für die Entwicklung nachhaltiger lebender Systeme (ohne Kürzung des Etats für Forschung)

Verantwortungen, Rollen und Motivation

Förderinstitutionen

- fördern Forschung durch die Vergabe öffentliche Gelder
- koordinieren den Review-Prozess für Drittmittelanträge



- Fähigkeit zur Investition in nachhaltige Drittmittelanträge und Lösungen
- Nachhaltige Konzepte sollten im Review-Prozess von Drittmittelanträgen berücksichtigt werden
- Gelder sollten für die Entwicklung und das langfristige Hosting / die Archivierung von lebenden Systemen vergeben werden (ohne Kürzung des Etats für Forschung)

Verantwortungen, Rollen und Motivation

Datenzentren / Bibliotheken

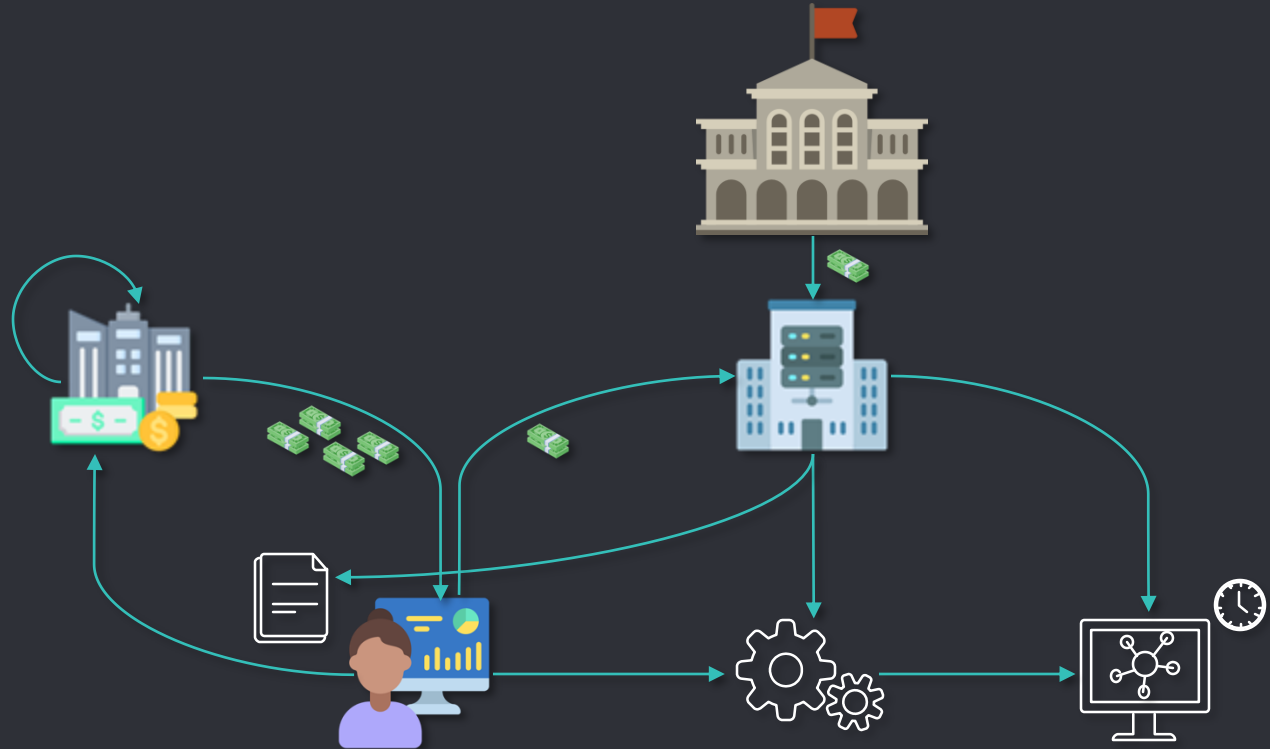
(und ihnen übergeordneten Institutionen)

- Auftrag zur Unterstützung von Forschenden
- übernehmen Forschungsergebnisse (u. a. Forschungsdaten)
- sind von institutionellen Interessen geprägt



- sollten spezifische Unterstützungsangebote aufbauen, bspw. Beratungsservices, Planungs- und Entwicklungsunterstützung, Software Management Pläne
- sollten Trainingsangebote für nachhaltiges Forschungsdatenmanagement anbieten
- sollten nachhaltige Infrastrukturen für das Hosting lebender Systeme bereitstellen
- brauchen Dauerstellen und Dauerfinanzierung

- Orchestrierung aller Stakeholder



● Adressierung bestehender Strategien (und ihrer Restriktionen)

1

Investition ausreichender finanzieller/personeller Ressourcen

- Skaliert nicht mit wachsender Anzahl lebender Systeme
- Wer bezahlt so etwas?

→ zielorientiertere Förderung in nachhaltige Konzepte und Infrastrukturen

Restriktion von Technologiestacks

- Zu wenige Community-Standards
- Könnte im Widerspruch zur Wissenschaftsfreiheit stehen

2

→ Entwicklung von Standards durch die Forschungscommunity gemeinsam mit Datenzentren und Infrastrukturen

Virtualisierung durch Technologien, bspw. Docker

- Bedarf spezifischer Infrastruktur
- Bedarf qualifiziertes Personal

3

→ Etablierung nachhaltiger Datenzentren, Infrastrukturen und Services zur Unterstützung von Forschenden/Forschung

Hosting lebender Systeme in statischer Form

- Nicht immer eine geeignete Strategie
- Verlust bereits entwickelter (notwendiger) Funktionalitäten

4

→ Vermeidung der Entwicklung zu komplexer Funktionalitäten von Beginn an



Vielen Dank!
Fragen?
Oder Lösungen?

Patrick Helling

Data Center for the Humanities (DCH)

University zu Köln

Tel.: +49 221 / 470 1307

Mail: patrick.helling@uni-koeln.de