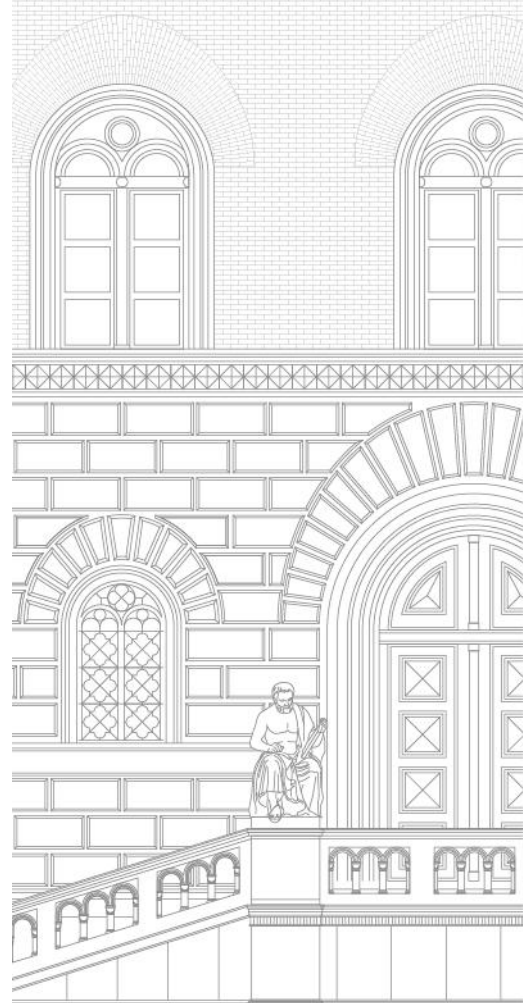
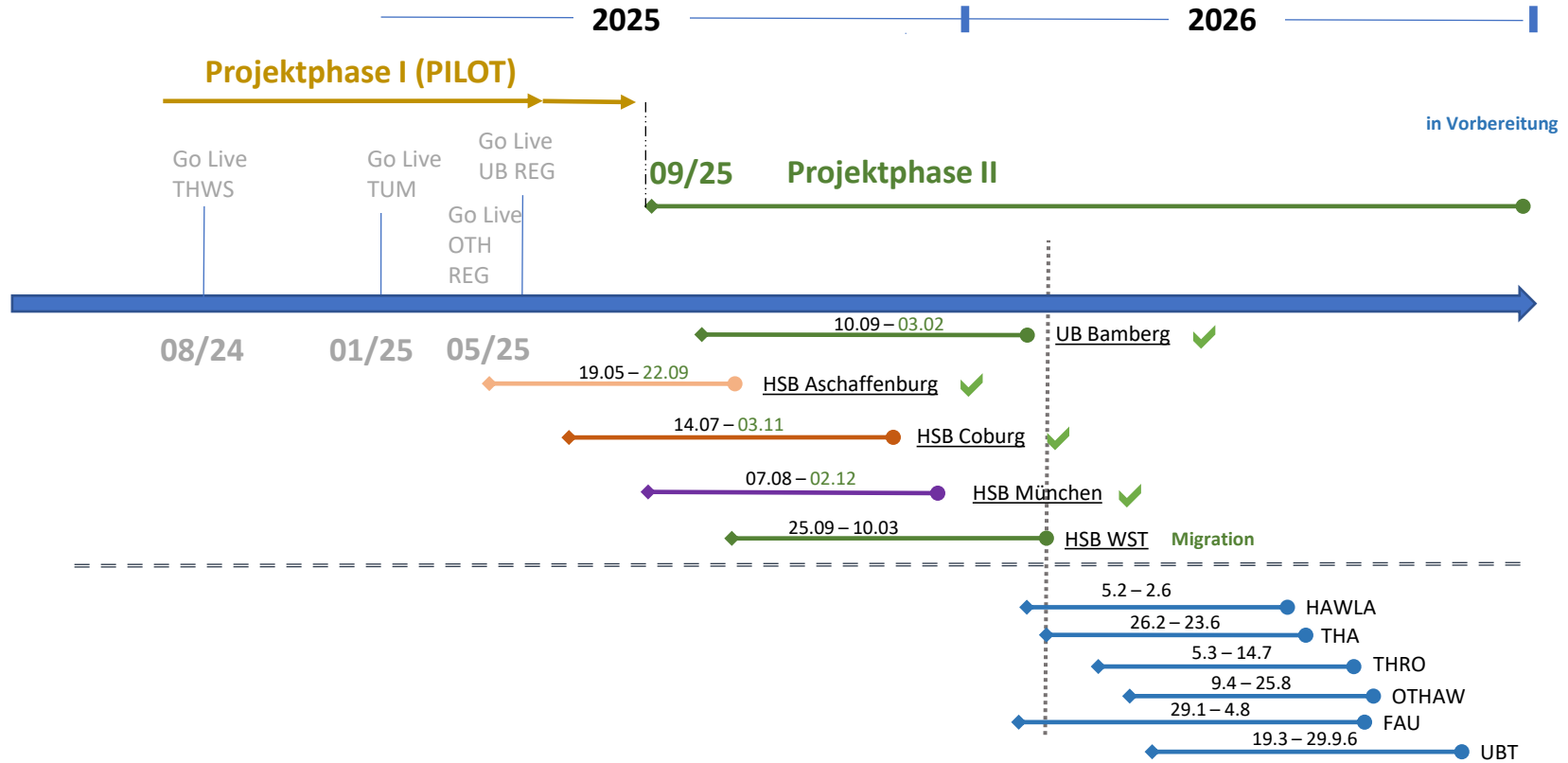


Nutzung der Datendienste in VuFind-OPACs (GND-Verweisformen) und FOLIO-Systemen mittels OAI-Harvesting (ZDB-Holdings)

Christian M. Strohmaier



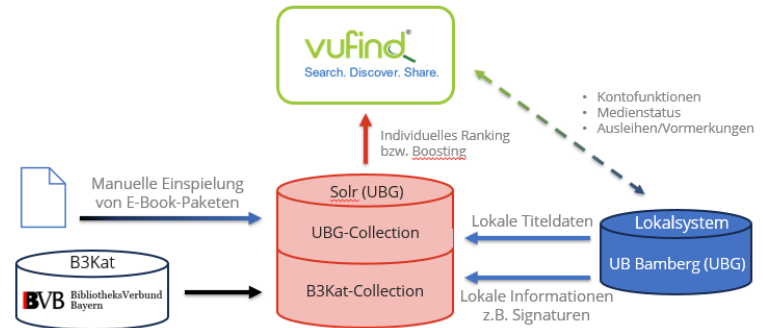
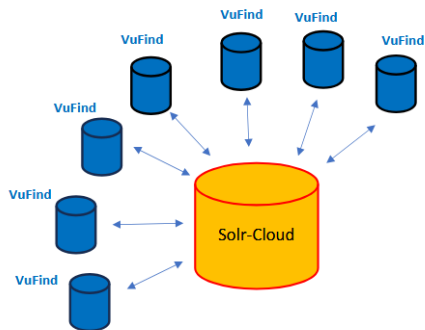
Großprojekt im BVB: Ablösung von SISIS/TouchPoint/Infoguide durch FOLIO/Vufind



Das neue Betriebskonzept für den Index

Eine SOLR-Collection für alle OPACs, primär auf Basis des B3Kat

- zusätzliche Anreicherungen
 - lokale Informationen (z.B. Signaturen)
 - Inhaltsverzeichnisse, Abstracts u.ä.
 - **GND-Verweisformen**
- Bei Bedarf pro Bibliothek eine zusätzliche Collection



Wir nutzen Solr-as-a-Service

Technik unseres Dienstleisters (LRZ)

Hardware

- 3 Kubernetes Cluster – Prod, Test, Infrastruktur
- 54 Server/VMs: 1580 CPUs, 4 TB Ram, 78 TB Disk
 - Nur Produktionsumgebung auf aktueller Hardware

Folio: Microservice-basiert, mandantenfähig

- 75+ Container, API Gateway, Datenbankcluster, Kafka-Cluster, Elasticsearch Cluster
- 9 Folio Umgebungen, 72 Mandanten

VuFind:

14 Mandanten, 2 Solr-Clouds, 9 Datenbank Cluster

- Automatisierte Installation / Management über Helm



Kompass für unsere Entwicklungen

arbeiten mit Standard-Software/Board-Mitteln

konfigurieren besser als programmieren

Verwendung von Standard-Schnittstellen

keine Vorverarbeitungsschritte

Uncle Bob (DE-588)114440964:

"Every line of code you write is code you have to maintain."



Was bedeutet das übertragen auf unsere Arbeiten am Index?

Vufind-Harvester holt B3Kat-Daten aus unserem OAI-Repository

Vufind-Indexer

- sorlmarc (minimale Erweiterungen)
- marc_local.properties
- Java-Plugins

Folio-Exporte als Atomic-Updates in Dynamic-Fields

Anpassungen schema.xml



Wie finden die Leute auch Albrecht den Großmüthigen?

<small>GND</small>	
Link zu diesem Datensatz	https://d-nb.info/gnd/118647571
Person	Albrecht V., Bayern, Herzog
Adelstitel	Herzog
Geschlecht	männlich
Andere Namen	Albrecht, der Großmütige Albert, der Großmütige Albert V., Bayern, Herzog Albert, Bayern, Herzog Albert, von Baiern Albrecht, der Grossmütige Albrecht, Rhein, Pfalzgraf Albertus V., Bavaria, Dux Albrecht V., Herzog von Bayern ((VD-16)) Albert V, Duc de Bavarie (Französisch, Code: fre) Albert V, Duke of Bavaria (Englisch, Code: eng) Albert, der Grossmütige, Herzog von Bayern Albert, der Grossmütige, Duc de Bavarie Albert, der Grossmütige, Duke of Bavaria Albrecht V, Duc de Bavarie Albrecht V, Herzog von Bayern Albrecht V, Duke of Bavaria Albert V, Herzog von Bayern



Ansatz zur Anreicherung (in Produktion)

Erweiterung der Indexierungs-Pipeline im Schema: `SynonymGraphFilter`

Recycling aus dem derzeitigen Gateway Bayern

GND-Dump der DNB -> `SynonymGraphFilter`

Veröffentlichung des GND-Dump halbjährlich

=> schleichende Aktualisierung der Daten

`SynonymGraphFilter` im Hauptspeicher der Solr-Nodes

⇒ Extrem schneller Zugriff bei der Indexierung

⇒ Treibt HW-Kosten in die Höhe, ungünstig für Kubernetes



neuer Ansatz (in Vorbereitung)

Lokaler Spiegel in Authority Collection

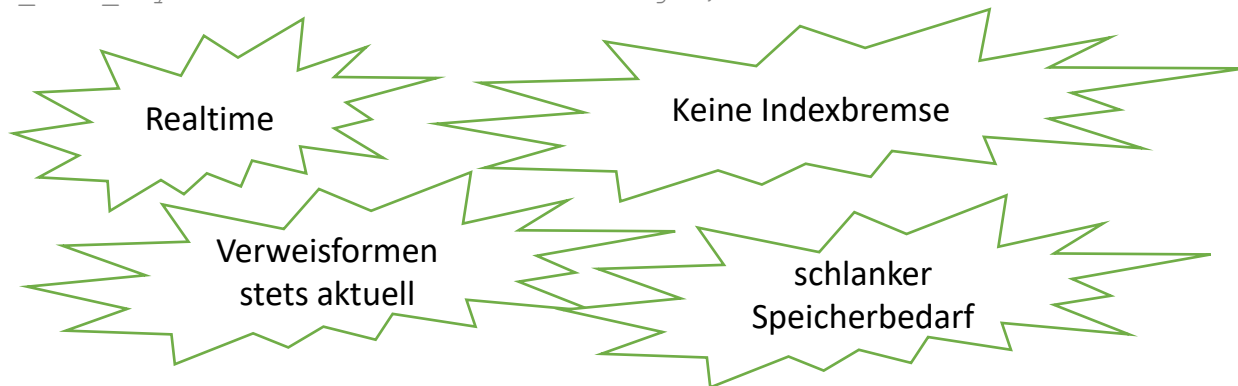
- Vufind **Standard-Collection**, aber **Schattendasein**
 - **Harvesting von der DNB**
=> Daten tagesaktuell
 - **Mapping gewünschter Marc-Felder**
 - Eigenes Java-Plugin beim Import (mit Cache) –
für jeden Record wird in Authority Collection nachrecherchiert
- ⇒ **Indexierung des B3Kat etwa halb so schnell**
- ⇒ **Vernachlässigbarer Platzbedarf im Kubernetes-Cluster**



Wer kann uns bei noch besserem Ansatz helfen?

Solr beherrscht sog. Join-Queries

```
{!join method=crossCollection from=gnd to=topic fromIndex=authority  
v=see_also_keywords:"Albrecht der Großmüthige"}
```



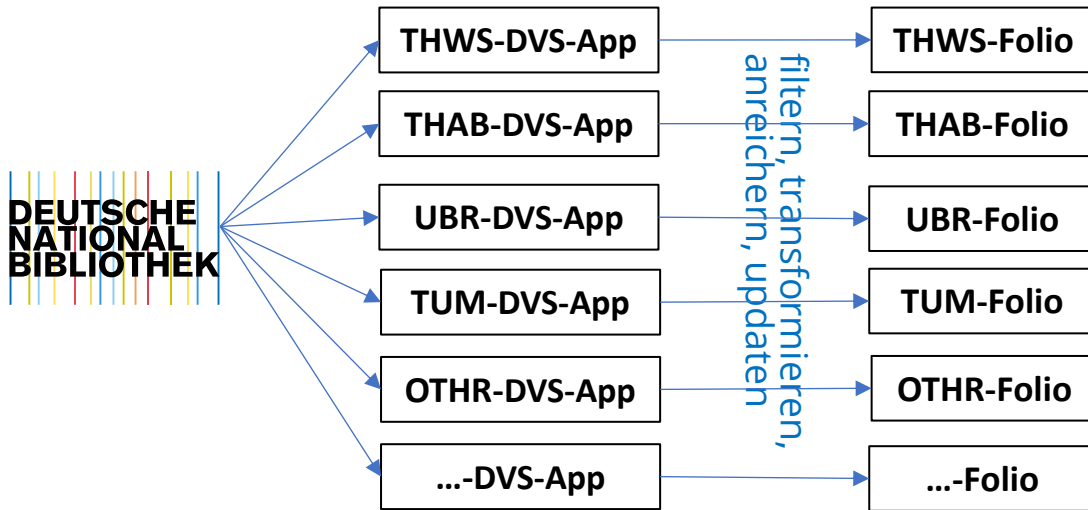
Aber leider:

- Einbau per Konfiguration in Vufind nicht gelungen
- Damian Katz konnte nicht für einen Einbau in PHP erwärmt werden



Nutzung der ZDB-OAI-Schnittstelle

Bsp.: <https://services.dnb.de/oai/repository?set=zdb:holdings&verb=ListRecords&metadataPrefix=MARC21-xml&from=2026-03-04T10:00:01Z&until=2026-03-04T10:30:01Z>



DSV-App: <https://gitlab.lrz.de/folio-bvb-public/katalog-datenversorgung>



Können ZDB-Löschungen länger aufgehoben werden?

Hintergrund:

Während der Migration kann eine Lücke von ein bis drei Monaten entstehen.

Laut

https://www.dnb.de/EN/Professionell/Metadatendienste/Datenbezug/OAI/content/faqOai9_akk.html sind Löschungen jedoch nur eine Woche lang abrufbar

Wäre es möglich den Zeitraum von einer Woche auf drei Monate auszudehnen? Permanent?



Vielen Dank für Ihre Aufmerksamkeit!

Christian.Strohmaier@bsb-muenchen.de

