

# Themen und Ergebnisse vom Barcamp 2024

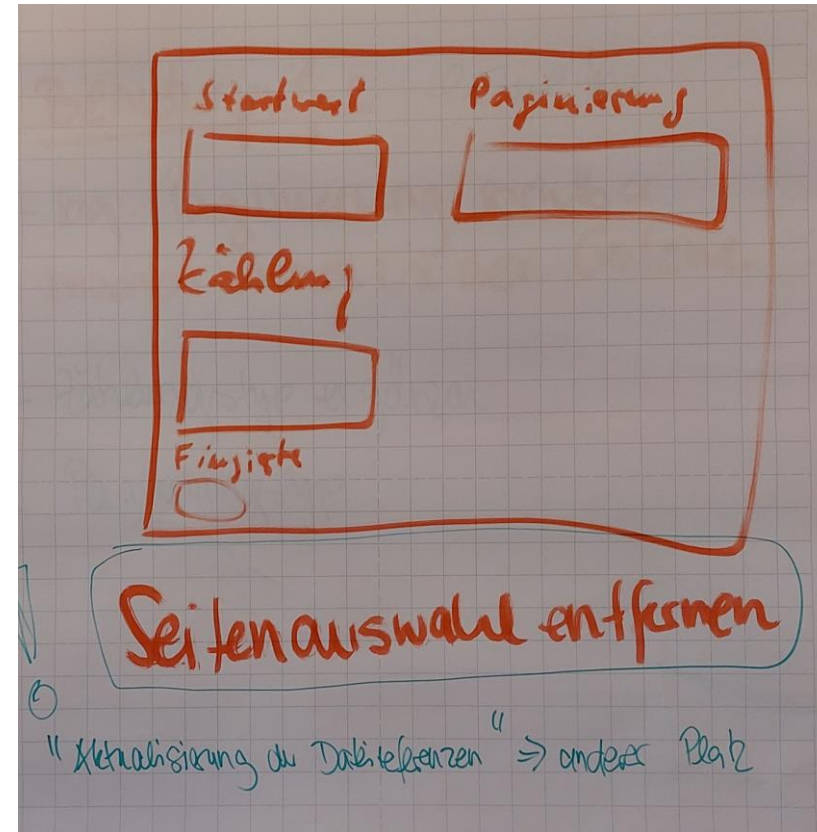
---

# Paginierungseditor – was fehlt? Was kann man verbessern?



# Diskurs & Ergebnisse

- Dokumentation in einer [Github-Diskussion](#)
- Community-Board hat Ideen mit Releasemanager vorbesprochen
- Formulierung eines [Issues](#) für Bewerbung beim Entwicklungsfonds
- Auswahl beim Entwicklungsfonds für Beauftragung
- Umsetzung fast abgeschlossen – Teil von Version Kitodo.Production 4.0



# Schutzfristeditor (wie Paginierung)



## FÄLLE

- URHEBERRECHT / PERSONENSCHUTZRECHT / GEHEIMHALTUNG
- GAR KEINE VERÖFFENTLICHUNG / EINGESCHRÄNKTE V.

## ZIELE K.PROD / METADATENEDITOR

### 1. ÜBERNAHME METADATEN (REGELSATZ)

→ ANWISUNG IM EDITOR

→ MAPPING AUF SEITENELEMENT

↓ VERSCHIEDENE  
TYPEN,  
BASISBESTIMMUNG

- ABGLEICH / RÜCKGABE METADATEN AUS / OPA  
→ WIDERSPRÜCHE?

### 2. "MUSTER"-WERKZEUG IM EDITOR

- MÖGLICHE MEHRFACHBELEGUNG

- NUR  
BILDBEREICH

### 3. VISUELLE DARSTELLUNG AN BILDANZEIGE

### 4. AUSGABE / ANWISUNG IN METADATEN FÜR WEITERE SCHRITTE / SYSTEME

## ÜBE (ÜBERNAHME) METADATEN

### 1. AUZUNEHMENDE METADATEN

- SCHUTZFRIST ENDATUM

- (GGF. BERECHNET) (TYP)

- ENTSTEHUNGSJAHR (?) → BERECHNUNG IN KINO

- FRIST TYP (URHEBERRECHT,...)

- LIZENZ WERK / METADATEN

- FRIST LÄNGEN

VORH.: - MANDATEN EINSTELLUNG?

- PROJEKT EINSTELLUNG?

- REGELSATZ? !

(PER INCLUDE)?

- SCHALTER PRIORITÄT

- METADATENREPRÄSENTATION IN K

- SUCHE: "HASENBARGO" / ZEITBEREICHE

↳ SCHLÜSSERSPÄLLE IN ÜBERSICHT  
(KENNZEICHNUNG BESCHRÄNKUNGEN)

- KEINE (!) AUTOMATISCHE ÜBERNAHME?  
AUF BILDEREINE → ANWISUNG KEIN KONTAKT  
VORBEREITUNG / ANZEIGE

- DIFFERENZIERUNG AUF STRUKTURRAUM  
(z.B. EINGETRAGENE FELD)

(- WIDERSPRUCH BEHANDLUNG  
- HINWEIS VS. BLOCHIEREN  
- NACHGELEGTER FOLGESCHRITT ANDERE ROLLE  
AUTOMAT.

### 2. AUS OBERER EBENE FOLGEN?

- BEREICHSAUSWAHL  
MIT ÜBERSCHREIBSCHUTZ

- AUS ANGABEN BILDEREIB  
BERECHNEN  
MIT ÜBERSCHREIBSCHUTZ

- LEEREN

- KITO DO SCRIPT?

### 3. VISUELLE DARSTELLUNG

- Roter Rahmen um  
GESPERRETE BILDER

- RÜCKMELDUNG ÜBER KONTAKT

EVTL. PER AUTOMATISCH AUSGE  
LÖSTEN KORREKTURSCHRITT

### 4. AUSGABE

- EVTL. UNTERE EBENE  
AUFFÜLLEN

- METS / MODS → FEL  
ANT

- INFORMATIONEN Z  
ABGLEICH / ÜBERNAHME

# Schnittstellen jenseits von ActiveMQ und KitodoScript



# Schnittstellen

Welches sind unsere Objekte?

Aufgaben, Benutzer, Vorgänge, Meta-/Strukturdaten, Projekte

Welches sind unsere Interaktionen?

- ~~Aufgaben~~ abschließen, zurückholen, Status abfragen  
Aufgabe
- Statistiken abfragen, Listen erzeugen
- Workflows aktualisieren
- Daten reinholen, Daten rausholen (muss z.B.)  
z.B. aufbereitete OCR, Images, andere KI Fälle oder DMS
- Vorgänge anlegen, abfragen, abschließen
- Metadatenkorrektur / -ergänzung (Batch?)
- Import aus Fremdsystemen (z.B. Import Prozess)
- Transaction-Logs
- **Authentifizierung?**  
**Absicherung der API**

Waram brauchen wir Sie?

- ↳ Abbau von Komplexität ↳ direktzugriff auf DB vermeiden?
- ↳ Automatisierung (auch in Zusammenarbeit mit Fremdsystemen)  
z.B. Ogi On Demand ↳ man sollte nicht gezwungen sein sich "Lösungen" zurechtzubasteln
- ↳ Know-how-Re-Use ~~hängen~~
- ↳ Auf neue Anforderungen reagieren ohne händische Bearbeitung
- ↳ Alle Funktionen sind da, aber verstreut, "inaccessible"  
(GUI only, Wiki, ~~script~~ Soennecken Admin UI)

## Schnittstellen

(technische Persp.)

- Framework verwenden mit Standards <sup>Richtig großer Schritt!</sup>
  - ↳ automatische Dokumentation > curl Beispiele
  - ↳ Webentferne
  - ↳ Wie sieht die "Response" aus? • "Selbstverwaltung"
- Wie commiten wir uns?
  - ↳ Prioritäten setzen und direkt umsetzen
  - ↳ für neue Features sollte die API direkt mitberücksichtigt werden
  - ↳ sukzessive Core-Features in die API bringen
  - ↳ Attraktivität + Selling Point API

Wie kommen wir dahin?

- ↳ Github-Diskussion?
- ↳ Wiki um Bedarfe zu sammeln und zu konzipieren
- ↳ Community-Bond um Bedarf zu sammeln?
- ↳ Entwicklungsland oder Drittwelt?
- ↳ nach umfangreichen Pflichtenheft die Hauptbedarfe zuerst umsetzen
- ↳ Welche Software/Welches Framework passt und ist zukunftsfähig?  
(z.B. Verbreitung, Integration etc.)

OpenAPI?

---

Verschiedene Fundstellen in den Protokollen des Community Boards:

- *Die UB Braunschweig und die USB Köln sondieren mit der Firma Imageware die automatische Anlieferung vom Scanner an Kitodo. Ziel ist, existierende Kitodo Vorgänge automatisiert mit den Bilddaten zu beliefern. Das Verfahren befindet sich noch in der Evaluation.*
- *Prozessautomatisierung wie diese würden insbesondere von einer Kitodo-API profitieren.*
- *Anders als ActiveMQ könnte eine API die relevante ID/Vorgangskennung, ID eines Schrittes und andere Parameter wie Name des Dateiverzeichnisses für tif-Bilder für einen Prozess aus der Datenbank ermitteln.*

- 
- *Nur lesend in der ersten Stufe?*
  - *Ideensammlung aus Marburg für nächste Sitzung*
  - *Armin: Benötigt werden vor allem aktuelle Statusabfragen und entsprechendes abarbeiten*

# Einsatz KI in Kitodo



Prod.

Strukturierung

Priorisierung anhand von Nutzerverhalten

Bild-, Video- & Audioanalyse

• Paginierung

• Qualitätsicherung

• Metadaten

• Aufwandschätzung

• Automatisierung von Dingen, die man wiederholt  
machen

• Dokumentation

Extern

- OCR-Verbesserung

Verdichtung

Presentation

• Sensitivity-Hinweis

• Anonymisierung Metadaten

• ChatGPT-Zugänglichkeit

• Volltext-Abstracts / Podcast

• Empfehlungen

• Beschreibungstexte aus den Metadaten

• Übersetzungen

• Schreiberegler in Presentation

# Performance Optimierung / Skalierung

---



# Performance

48GB RAM 6 Kerne

Viele Rdn pro Person  $\rightarrow$  schlägt auf Performance

$X_{ms} = X_{mk} \cdot \text{seconds?}$

Elasticsearch auf gleicher Maschine?

Kitodo-Problem oder JAVA/Tomcat-Problem?

oder Workflow-Problem: - WF kleiner machen!  
- Script-Takt optimieren } weniger  
Wartezeiten

zu lange Laufzeit!

(Festplatte (jährlich ändern & organisieren) / Manuelle Nacharbeiten  $\rightarrow$  schlecht)



Reiter Daten  $\rightarrow$  Aktion

" - Batch - " / Stapelverarbeitung



Batchverarbeitung

# DB:

MySQL/MariaDB (partitionierte evtl. auch mit PostGIS)

MySQL  $\rightarrow$  InnoDB

Indizes erstellen : slow-query-log

SQL-Schema? HQL-Queries ausbauen

Rollentrennung? ALLOCATION/Trennung

evtl. ELS absetzen / RAM limit begrenzen!

evtl. DB Backup

WF kürzen! Active MQ! Batchverarbeitung

Hardware

4-8 Kerne / 32GB RAM

GitHub  $\rightarrow$  Wiki  $\rightarrow$  Performance Best