

Mitgliederversammlung 2024

05.06.2024

Tagesordnung

1. Begrüßung und Feststellung der Beschlussfähigkeit
2. Tätigkeits- und Kassenbericht des Vereinsvorstands für den Zeitraum Mai 2023 bis Juni 2024
3. Bericht der Kassenprüfer:innen
4. Bericht des Releasemanagements
5. Bericht des Dokumentationsmanagements
6. Aussprache zu den Berichten
7. Entlastung des Vorstands
8. Wahl des Vorstands
9. Wahl des Releasemanagements und des Dokumentationsmanagements
10. Wahl der Kassenprüfer:innen
11. Aktivitäten des Community Boards
12. Mögliche Integration von OCR-D in Kitodo e. V.
 - a) Verstetigung der kooperativen Pflege und Weiterentwicklung für die zentralen Werkzeuge in OCR-D (Projektleitung OCR-D)
 - b) Bericht Review der zentralen Softwarekomponenten
 - c) Diskussion
13. Verschiedenes

TOP 1: Begrüßung und Feststellung der Beschlussfähigkeit

TOP 2:
Tätigkeits- und Kassenbericht des Vereinsvorstands für
den Zeitraum Mai 2023 bis Juni 2024

Robert Strötgen (TU Braunschweig) / Jürgen Christof (TU Berlin)

Mitgliederentwicklung seit Mai 2023

Austritte

- keine

Eintritte

- keine

→ aktuelle Mitgliederzahl: 44

Aus der Arbeit des Vorstands

Geschäftsordnung vom 12.10.2012

Fünf Vorstandssitzungen (20.09.2023, 07.03.2024 und 24.05.2023 virtuell, 11.07.2023 und 12.12.2023 in Berlin)

Mitgliederversammlung 23.05.2023 (Hannover)

Schwerpunkte

- Entwicklungsfonds
- Kitodo und OCR-D
- Praxistreffen

Entwicklungsfonds – Rückblick auf die ersten drei Jahre

Ausgaben (136.914,26 €)

- 2021: 32.620,28 €
- 2022: 49.006,58 €
- 2023: 55.287,40 € (noch nicht alle Rechnungen eingegangen)

Aufteilung

- Kitodo.Production: 6 Projekte (74.398,80 €)
- Kitodo.Presentation: 5 Projekte (62.515,46 €)

Auftragnehmer

- effective WEBWORK, knopflogik, MIK-Center, Zeutschel

Schwerpunkte

- Versionspflege (Java, TYPO3, Hibernate Search, ...)
- Test Coverage
- IT-Sicherheit
- Usability

Entwicklungsfonds – Projekte 2023

Kitodo.Presentation

- Update and improve Solr compatibility
> <https://github.com/kitodo/kitodo-presentation/issues/594>
- Fix security issues and warnings (and improve code styling)
> <https://github.com/kitodo/kitodo-presentation/issues/893>



Kitodo.Production

- Test coverage
> <https://github.com/kitodo/kitodo-production/issues/5539>
- OpenSearch/HibernateSearch
<https://github.com/kitodo/kitodo-production/issues/5546>
<https://github.com/kitodo/kitodo-production/issues/4896>



Entwicklungsfonds – vierte Runde 2024

■ Entwicklungsprojekte priorisiert

1. (25 Punkte) - high range
Kitodo.Production 5935 : Metadata editor improvement – Bundle
<https://github.com/kitodo/kitodo-production/issues/5935>
2. (21 Punkte) – medium
Kitodo.Presentation 900 : Support for TYPO3 12 LTS
<https://github.com/kitodo/kitodo-presentation/issues/900>
3. (18 Punkte) - low to medium
Import von Archiv-Metadaten aus EAD-DDB-XML um hierarchische Vorgänge automatisch zu erstellen
<https://github.com/kitodo/kitodo-production/issues/5984>

- ## ■ Ausschreibungen veröffentlicht, laufen bis zum 16.06.2024
- <https://www.kitodo.org/software/entwicklung/ausschreibungen>

Entwicklungsfonds – Sonderfall Tomcat-Update

- Update Tomcat für Kitodo.Production, damit verbunden auch Update Java PrimeFaces
- Entscheidung im Vorstand
- Gründe: dringend und in absehbarer Zeit notwendig
- Separate Ausschreibung im Januar 2024
- Unerwartet
 - Hohe Komplexität (und hohe Kosten)
 - Nicht abwärtskompatibel (mit Community abgestimmt)
- Aufteilung in zwei Lose (Tomcat und PrimeFaces getrennt)
- Vergabe an zwei Anbieter mit Gesamtauftragssumme 11.971,40 €

Herzlichen Dank allen, die den Entwicklungsfonds unterstützen!

- CCS GmbH
- effective WEBWORK GmbH
- Gottfried Wilhelm Leibniz Bibliothek – Niedersächsische Landesbibliothek
- Staats- und Universitätsbibliothek Hamburg
- Staatsbibliothek zu Berlin
- Sächsische Landesbibliothek – Staats- und Universitätsbibliothek Dresden
- Universitätsbibliothek Braunschweig
- Universitätsbibliothek der Technischen Universität Berlin
- Universitätsbibliothek Erlangen-Nürnberg
- Universitätsbibliothek Mannheim
- Württembergische Landesbibliothek Stuttgart

Temporärer Umzug der Geschäftsstelle

- Katja Selmikeit (UB TU Berlin) setzt ab Herbst für 1 Jahr aus
- Magdalena Eberle (SUB Hamburg) übernimmt in dieser Zeit die Geschäftsstelle

- Vielen Dank an alle Beteiligten!

Kitodo-Praxistreffen 2023

- Praxistreffen an der Staatsbibliothek zu Berlin
- Viele Beiträge und Barcamp



Praxistreffen – Barcamp

Rechtmanagement

Kitodo Production

- Zugriffrechte auf Projektleitung auf ausserhalb ✓
- * Erfassung von Rechten u. Beschränkungen (z.B. KRM)
- Export direkt neben dem Ablauf von Endprodukten
- Aktualisierung + Rechte = Aktualisierung im Metadaten (KRM)

Kitodo Präsentation

- * Zugriffrechte auf Ebene von Einzelarten / Gruppen / Metadaten
- Nutzung von Metadaten und Rechte, um besser zu arbeiten (KRM)
- Problem: UML beachtet nicht auf KRM, z.B. KRM
- Lösung: KRM als Proxy → KRM über KRM (z.B. KRM)

⇒ Kitodo Reporting? / Kitodo Print?

⇒ Was ist Standard der WSC

DINI AG KRM Gruppe "Köln"

Kitodo Community Board

⇒

- * Idee für Modernisierung = DINI-Projekt

Strukturbausteine

Strukturtyp + Metadaten
 nicht in einem Schritt
 (Metadaten in der Datenbank als konfigurierbare Objekte)

Prozess

- noch nicht ganz abseits / klarer
- in Zukunft mit KRM
- Integration von KRM auf Datenrechner

Metadaten

- z.B. in Gruppen / KRM
- Support / Erweiterung / KRM
- Support / Erweiterung / KRM

Glossar

- Semantische / KRM
- Integration von KRM in KRM

Datensicht

- Komplexer / KRM
- einfaches / KRM
- Struktur / KRM
- KRM / KRM

parallele Arbeit in KRM → KRM

Usability

Workflows

Wo fängt Kitodo an?

- nach dem Scannen
- mit dem Einlesen der O-Aufnahmen
- Versand an den Dienstleister
- nach Einlesen mit Anlagen des Vorgangs
- nach Prüfung Antrag Digitalisierung (ES)

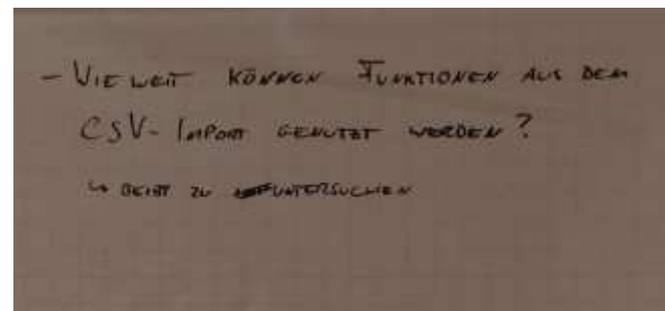
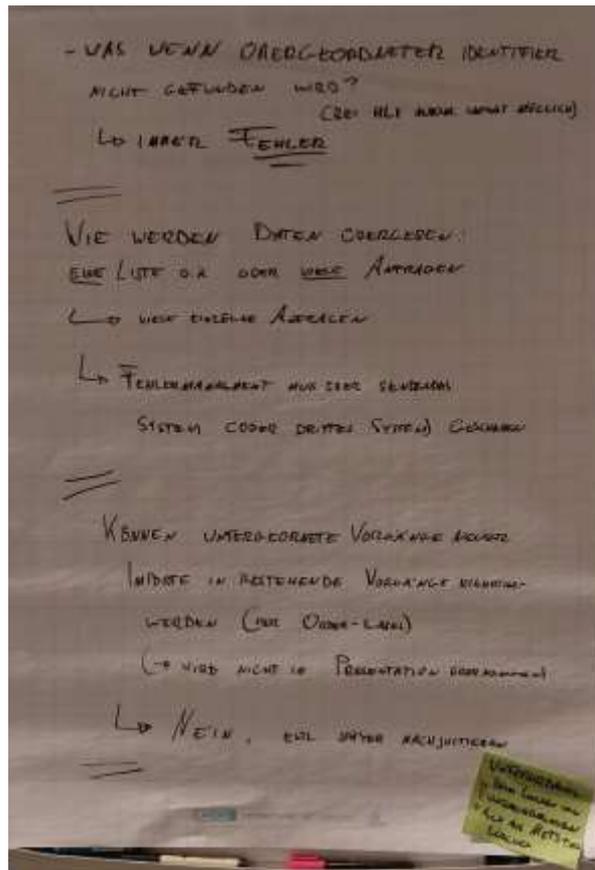
Welche Schritte?

- UB Leipzig: Daten sind schon da, KRM ist schon im Projekt integriert
- KRM: KRM, KRM, KRM

Wo endet Kitodo?

- Bestandhaltung (am Ende UML am Anfang)
- Was wird in Kitodo?
- Kitodo ist Modernisierung (Einleitung)

Praxistreffen – Barcamp



Kassenbericht

Einnahmen / Ausgaben des Vereins 2023

Übertrag aus 2022 (= Kontostand 31.12.2022)	61.509,42 €
Beiträge	84.200,00 €
Releasemanagement	20.000,00 €
Entwicklungsfonds	42.542,50 €
Website	2.261,00 €
Gerichtskosten	66,67 €
Bankgebühren	184,44 €
Übertrag nach 2024 (= Kontostand 31.12.2023)	80.654,81 €
Nachrichtlich: Kontostand am 25.04.2024	54.676,33 €

Danke für ein weiteres erfolgreiches Jahr!

- Allen Mitgliedern
- Katja Selmikeit für die Arbeit in der Vereinsgeschäftsstelle und dem Releasemanagement unter Leitung von Sebastian Meyer und Arved Solth sowie Matthias Kissler für das Dokumentationsmanagement
- Allen, die aktiv an der Erweiterung und Verbesserung des gemeinsamen Kitodo-Quellcodes mitgearbeitet haben
- An die Kolleginnen und Kollegen im Vorstand der jetzt endenden Amtszeit!
- Unseren Kassenprüfern

Kontakt

Kitodo. Key to digital objects e. V.

Robert Strötgen – Vorsitzender

Stefan von der Heide – Stv. Vorsitzender

Stefan Kwasnitza – Stv. Vorsitzender

Jürgen Christof – Schatzmeister

Kerstin Wendt – Schriftführerin

Katja Selmikeit - Geschäftsstelle

Vereinsregisternummer: 5803 (Amtsgericht Dresden)

c/o Technische Universität Berlin

Universitätsbibliothek

Fasanenstraße 88

10623 Berlin



TOP 3: Bericht der Kassenprüfer

Gerrit Gragert (Staatsbibliothek zu Berlin) /
Hans-Christian Pust (Württembergische Landesbibliothek)

TOP 4:

Bericht des Releasemanagements

Rechenschaftsbericht für den Zeitraum 05/2023 bis 06/2024

Arved Solth (eww) / Sebastian Meyer (OCC)

Kitodo Releasemanagement

Kitodo.Production - Aktuell

Kitodo.Production 3.5.1 (04.09.2023)

Bugfix-Release für Kitodo.Production 3.5.x, der u.a. die Kompatibilität zu MySQL 5.7 wiederherstellt.

Die Weiterentwicklung der Version 3.5 ist mit dem Release von Kitodo.Production 3.6.0 eingestellt, so dass 3.5.1 der letzte Release der Version 3.5.x ist.

Kitodo.Production 3.6.0 (22.09.2023)

- Metadateneditor (Multimedia-Support etc.)
- Kalender-Tool (Import & Export)
- Überarbeitetes Filter-Menü

Seit dem Release von 3.6.0 sind **82 weitere Commits** mit Fehlerbehebungen und Fixes für potenzielle Sicherheitslücken im Branch 3.6.x gemacht worden. Diese Commits resultierten in drei weiteren Patch-Releases für Kitodo.Production 3.6.x.

Der letzte Release 3.6.3 ist letzte Woche erschienen.

Kitodo Releasemanagement

Kitodo.Production - Geplant

Kitodo.Production 3.7.0

- Verzögert wegen ein paar notwendiger, ausstehender Bugfixes
- Erweiterte Funktionen für Digitalisierung von Multimedia-Dateien
- Verbesserte Test-Abdeckung (Entwicklungsfonds 2023)

Geplante Weiterentwicklungen:

- HibernateSearch (Entwicklungsfonds 2023)
- Update auf Tomcat 10 und Java 17
- Erweiterungen des Metadaten-Editors (Entwicklungsfonds 2024)
- Massenimport aus EAD-Dateien (Entwicklungsfonds 2024)

Großer Dank gebührt allen Entwickler:innen und Tester:innen, die sich an der Weiterentwicklung und dem Review von Pull Requests beteiligen!

Kitodo Releasemanagement

Kitodo.Presentation

Releases

Kitodo.Presentation 4.1.0 (16.12.2023)

- prototypische Unterstützung für 3D-Objekte und AV-Medien (SLUB)

Kitodo.Presentation 5.0.0 (23.05.2024)

- Verbesserte Test-Abdeckung (Entwicklungsfonds 2022, EWW)
- Verbesserte Code-Qualität (Entwicklungsfonds 2023, Zeuschel)
- Unterstützung für TYPO3 v10 und v11 (Entwicklungsfonds 2023, EWW)

Kitodo.Presentation 5.0.1 (31.05.2024)

- Bugfixes

Kitodo Releasemanagement

Kitodo.Presentation

Statistik

Seit der letzten Mitgliederversammlung sind **229 Pull Requests** mit Code von **12 Entwickler:innen** eingeflossen, davon allein 159 Pull Requests von Beatrycze Volk (SLUB).

Es wurden **315 Dateien** geändert, wobei **33.585 Zeilen** Quellcode hinzugefügt und **8.733 Zeilen** entfernt wurden. Das entspricht dem Umfang eines Romans mit über **1.000 Normseiten**.

Herzlichen Dank allen Entwickler:innen, die im vergangenen Jahr zu Kitodo.Presentation beigetragen haben!

Kitodo Releasemanagement

Kitodo.Presentation

Ausblick

- Entwicklungsfonds
 - 2 Pull Requests für Entwicklungsfonds 2022 stehen noch aus (Test-Abdeckung)
 - 3 Pull Requests für Entwicklungsfonds 2023 liegen vor (Solr-Verbesserungen)
 - Ausschreibung für Entwicklungsfonds 2024 wurde veröffentlicht (TYPO3 v12)
- Laufende und geplante Projekte
 - Verbesserung der 3D-Unterstützung (Markus Weigelt, SLUB)
 - Verbesserung der AV-Unterstützung (Florian Schölzel, SLUB)
 - Unterstützung für rechtebewehrte Objekte (SUB Hamburg, UB Mannheim, SLUB)
- Umstellung des Release-Zyklus auf halbjährliche Feature Releases
 - Feature Freeze jeweils gegen Ende Q1 bzw. Q3
 - Veröffentlichung jeweils gegen Ende Q2 bzw. Q4
 - Bugfix-Releases wie gehabt nach Bedarf

Kitodo Releasemanagement

Kitodo.Presentation

Kitodo Coding Guidelines

- Mit Verbesserung der Test-Abdeckung in Kitodo.Production und Kitodo.Presentation wurden automatisierte Tests fester Bestandteil des Entwicklungsworkflows auf GitHub.
- Ergänzt werden sie um statische Code-Analysen, die Coding Standards überprüfen, Schwachstellen entdecken sowie einheitliche Formatierungen und durchgehende Dokumentation erzwingen können.
- **Ein Code Review durch das Releasemanagement erfolgt erst, wenn dieser Workflow erfolgreich durchlaufen wurde!**
- Kitodo Coding Guidelines wurden zuletzt 11/2021 aktualisiert und spiegeln diese Änderungen noch nicht adäquat wider. Eine Überarbeitung erfolgt in den nächsten Monaten.
- **Automatisierte Tests ersetzen keine manuellen Funktionstests!**

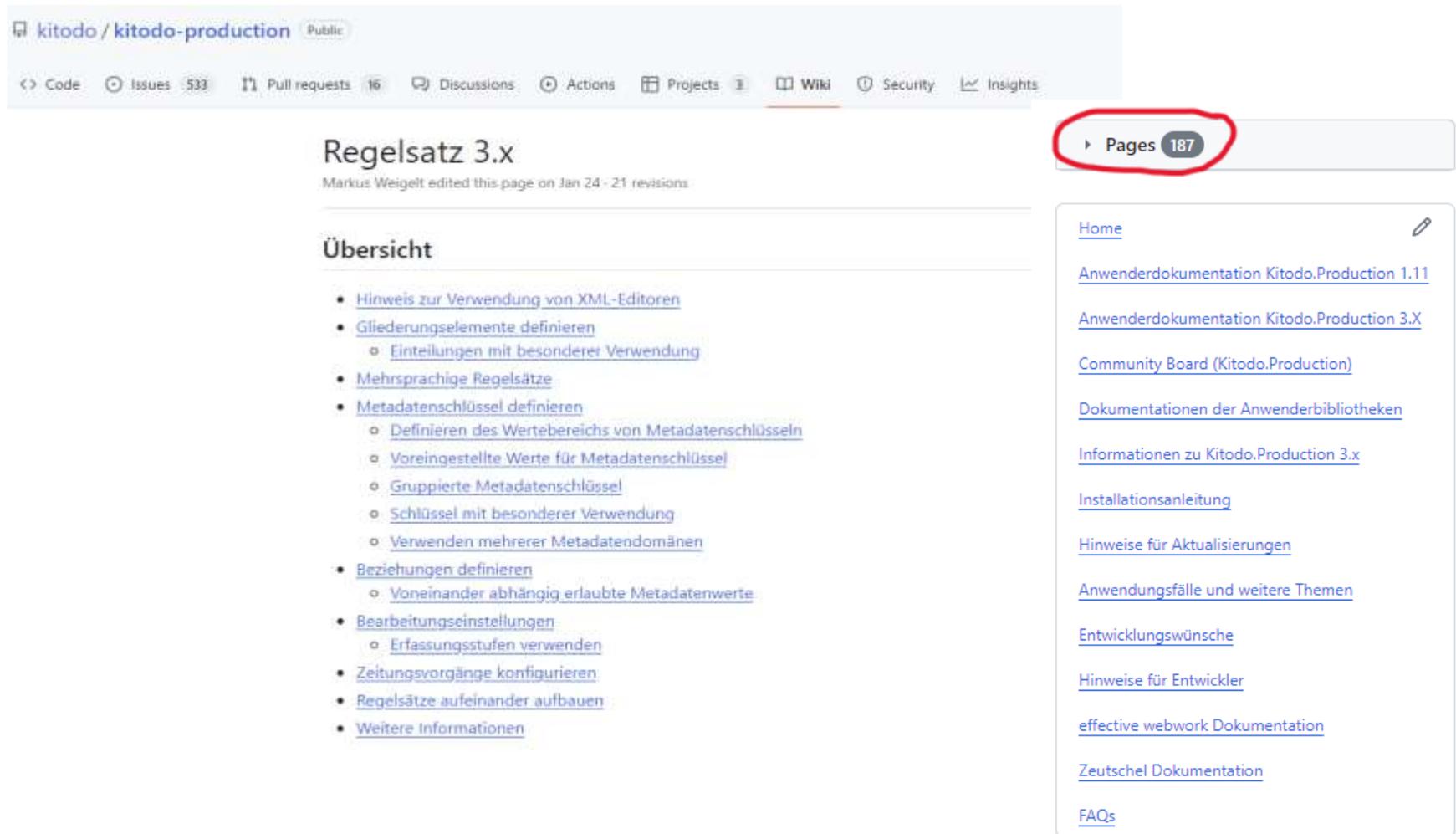
TOP 5: Bericht des Dokumentationsmanagements

Matthias Kissler (TU Braunschweig)

Dokumentationsmanagement – Stand der Dinge

- Stand: Kitodo 3.x in BS im produktiven Einsatz (aktuell in der Version 3.6.2)
- Stand: Umzug der alten Dokumentation von readthedocs endgültig abgeschlossen
- Stand: Community Board hat neue Impulse für eine zielgerichtete Dokumentations-Verbesserung gebracht
- Ziel: Bessere Update-Dokumentation bei Release einer neuen Version (z.B. 3.5.x → 3.6.0)
- Ziel: Weiterer kontinuierlicher Ausbau und bessere Strukturierung der Dokumentation auch im Zusammenspiel mit dem Community Board
- GitHub-Wiki: <https://github.com/kitodo/kitodo-production/wiki>

Dokumentation auf GitHub



kitodo / kitodo-production Public

<> Code Issues 533 Pull requests 16 Discussions Actions Projects 3 Wiki Security Insights

Regelsatz 3.x

Markus Weigelt edited this page on Jan 24 - 21 revisions

Übersicht

- [Hinweis zur Verwendung von XML-Editoren](#)
- [Gliederungselemente definieren](#)
 - [Einteilungen mit besonderer Verwendung](#)
- [Mehrsprachige Regelsätze](#)
- [Metadatenschlüssel definieren](#)
 - [Definieren des Wertebereichs von Metadatenschlüsseln](#)
 - [Voreingestellte Werte für Metadatenschlüssel](#)
 - [Gruppierte Metadatenschlüssel](#)
 - [Schlüssel mit besonderer Verwendung](#)
 - [Verwenden mehrerer Metadatendomänen](#)
- [Beziehungen definieren](#)
 - [Voneinander abhängig erlaubte Metadatenwerte](#)
- [Bearbeitungseinstellungen](#)
 - [Erfassungstufen verwenden](#)
- [Zeitvorgänge konfigurieren](#)
- [Regelsätze aufeinander aufbauen](#)
- [Weitere Informationen](#)

Pages 187

[Home](#) 

- [Anwenderdokumentation Kitodo.Production 1.11](#)
- [Anwenderdokumentation Kitodo.Production 3.X](#)
- [Community Board \(Kitodo.Production\)](#)
- [Dokumentationen der Anwenderbibliotheken](#)
- [Informationen zu Kitodo.Production 3.x](#)
- [Installationsanleitung](#)
- [Hinweise für Aktualisierungen](#)
- [Anwendungsfälle und weitere Themen](#)
- [Entwicklungswünsche](#)
- [Hinweise für Entwickler](#)
- [effective webwork Dokumentation](#)
- [Zeuschel Dokumentation](#)
- [FAQs](#)

TOP 6: Aussprache zu den Berichten

TOP 7: Entlastung des Vorstands

TOP 8: Wahl des Vorstands

TOP 9:
Wahl des Releasemanagements und des
Dokumentationsmanagements

TOP 10: Wahl der Kassenprüfer:innen

TOP 11: Aktivitäten des Community Boards

Aktivitäten des Community Boards

Unser Vorsatz beim Start in der MV 2023

„Ein **Community Board** soll den kontinuierlichen Austausch der Anwender:innen stärken und eine bessere Abstimmung zwischen Softwareentwicklung und der konkreten Nutzung von Kitodo im Alltag unterstützen.“

Bisherige Aktivitäten:

- Das Community Board tagt monatlich – z. T. zusätzliche Treffen in Untergruppen zu Detailfragen
- Auftrag, aktuelle Zusammensetzung und alle Protokolle → [Dokumentation im Wiki](#) von Kitodo.Production

Aktivitäten des Community Boards

- Beratung aktueller Themen aus der laufenden Softwareentwicklung (GitHub-Diskurse usw.) → direkte Adressierung der Mitglieder des Boards aus GitHub heraus
- Planung und Vorbereitung von bisher 2 virtuellen **Community-Treffen** für alle interessierten Anwender:innen
- *Bisherige Arbeitsergebnisse:*
 - Erarbeitung einer detaillierten Leistungsbeschreibung zum Reimport von Katalogdaten
 - Aufarbeitung des Diskurses aus der Barcamp-Runde zum Metadateneditor beim Praxistreffen 2023 und erfolgreiches Einbringen in den Entwicklungsfonds

Aktivitäten des Community Boards

- *Was machen wir gerade – was haben wir uns vorgenommen:*
 - Aufarbeiten des zweiten Community-Treffens (Thema Metadateneditor) – gemeinsames Ranking ?
 - Diskussion zum Paginierungseditor und zur Visualisierung der Hierarchien in der Galerie beim Praxistreffen

TOP 12: Mögliche Integration von OCR-D in Kitodo e. V.

Robert Strötgen
Johannes Mangei
Reinhard Alenhöner
Arved Solth



Kitodo und OCR-D

- OCR für viele Nutzer:innen von Kitodo von großer Bedeutung
- Von Anfang an vielfältige Kooperationsbeziehungen (z. B. über Personen im OCR-D-Beirat)
- Befassung mit dem Thema u.a. bei
 - Vorstandssitzungen
 - Praxistreffen 2022 und 2023
 - Mitgliederversammlungen
 - Austausch zu Software-Review
 - OCR-D-Veranstaltungen



OCR-D

Koordinierte Förderinitiative zur Weiterentwicklung
von Verfahren der Optical Character Recognition (OCR)



Kooperationsvereinbarung

- Kooperationsvereinbarung seit 2019, erweitert 2020
- Ziel der Kooperation
 - Integration von OCR-D in Kitodo.Production
 - Austausch über Schnittstellen und Entwicklungen
 - Unterstützung und Beratung zur Verstetigung von OCR-D
→ „In diesem Rahmen wird auch geprüft, inwieweit eine Einbeziehung von Kitodo in das OCR-D-Verstetigungskonzept für beide Seiten vorteilhaft sein könnte.“



LETTER OF INTENT

Kooperation bei der koordinierten und nachhaltigen Entwicklung und Bereitstellung von OCR- Softwarelösungen zur Massenvolltextdigitalisierung

Der Verein Kitodo. Key to digital objects e.V. mit seiner Software Kitodo.Production und das DFG-Koordinierungsprojekt zur Weiterentwicklung von Verfahren der Optical Character Recognition (OCR-D), koordiniert an der Herzog August Bibliothek Wolfenbüttel, streben eine vertiefte Kooperation bei der koordinierten und nachhaltigen Entwicklung und Bereitstellung von OCR-Softwarelösungen zur Massenvolltextdigitalisierung an.

Grundlegendes Ziel der weitergehenden Kooperation ist es, inhaltliche und konzeptionelle Überschneidungen zwischen Kitodo und OCR-D zum Erzielen von Synergien in technischer und organisatorischer Hinsicht zu nutzen. OCR-D beabsichtigt in der kommenden dritten Projektphase, den entstandenen Prototyp zu einer stabil lauffähigen, breit eingesetzten sowie ausführlich dokumentierten Softwarelösung auszubauen, deren Support und Weiterentwicklung dauerhaft gesichert wird. Kitodo wird OCR-D bei der Konzepterstellung umfassend mit seinen eigenen Erfahrungen in diesem Bereich unterstützen. In diesem Rahmen wird auch geprüft, inwieweit eine Einbeziehung von Kitodo in das OCR-D-Verstetigungskonzept für beide Seiten vorteilhaft sein könnte. Im gegenseitigen Einverständnis tauschen sich die Kooperationspartner zudem über aktuelle sowie kommende Entwicklungen der OCR-Massenvolltextdigitalisierung aus.

Die Kooperationspartner dürfen die im Projekt entwickelten OCR-Softwarelösungen zur Massenvolltextdigitalisierung im Rahmen der für die OCR-D-Software verwendeten Lizenz Apache 2.0 nutzen, soweit keine anderen Nutzungs- und Patentrechte mit dieser Art der Verwertung damit verletzt werden.

Wolfenbüttel, den 25.09.2020

Johannes Mangel

OCR-D Projektkoordination
Herzog August Bibliothek Wolfenbüttel

Braunschweig, den 24.09.2020

Robert Strötgen

Vorsitzender des Vorstands von Kitodo e.V.



Kitodo-OCR-D Implementierungsprojekt

- OCR-D-Implementierungsprojekt in der dritten Förderphase "Integration von Kitodo und OCR-D zur produktiven Massendigitalisierung" (Production und Presentation) von SLUB Dresden, UB Mannheim und TU Braunschweig abgeschlossen
- Integration von OCR-D in Kitodo prototypisch gelöst und im Praxisbetrieb
- OCR-D-Prozess wird aus Kitodo.Production heraus über Skript gestartet, Ergebnisse zurückgeschrieben
- OCR-D-Prozess kann parametrisiert werden
- Uneingelöste Ziele des Projekts
 - Stabiler Zugriff über OCR-D-Web-API
 - Nutzerfreundliche und transparente Statusinformationen in Kitodo.Production
 - Einfache Integration über Service-Netzwerk
 - Geschäftsmodell über Verein (Steuerrecht)



DFG-Rundgespräch 2023

- Rundgespräch zur Zukunft der Produktion und Bereitstellung von Volltexten für die nationalbibliographischen Verzeichnisse VD 16, VD 17 und VD 18
 - Göttingen, 24./25. August 2023
 - Konzept zur Volltexttransformation der VD
 - Konzept zur Verstetigung von OCR-D
 - Erörterung der Möglichkeiten zur Kitodo-Integration
 - Aufnahme von Erfahrungen des Kitodo-Vereins



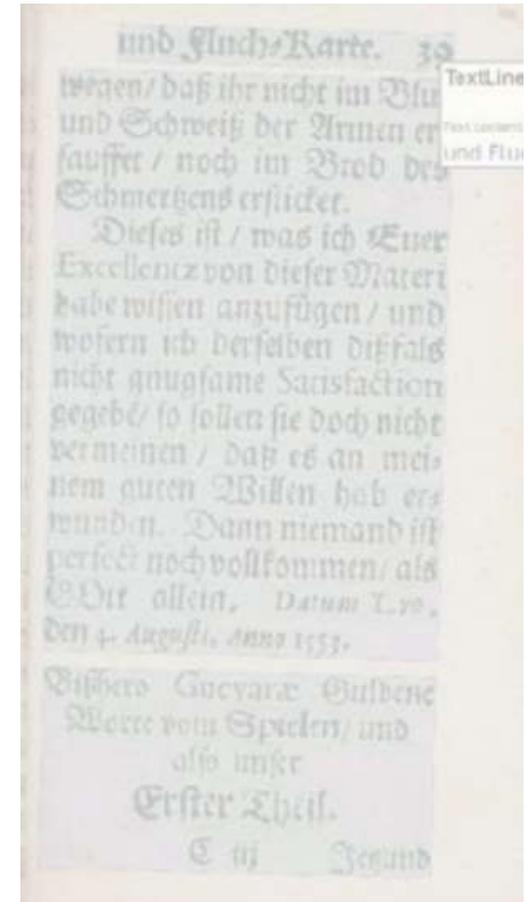
OCR-D-Konzepte und AWBI

- Rundgespräch
- Überarbeitung erster Entwürfe
- Einreichung der Entwürfe bei AWBI 10/2023
- Erörterung in der Herbstsitzung des AWBI
- Feedback des AWBI 01/2024
- Umsetzung von Vorarbeiten und erneute Überarbeitung, darunter: Vorbereitungen für Kitodo-Integration
- Erneute Einreichung der Konzepte 10/2024



Integration von OCR-D in Kitodo

- Teil des OCR-D-Verstetigungskonzepts
- Fahrplan bis zur heutigen Mitgliederversammlung
- Derzeitiger Stand:
 - Anwendung der Software
 - Release von V. 3.0: Weiterentwicklung der zentralen Werkzeuge
 - Finalisierung des Software-Reviews
 - Klären von organisatorischen Aspekten
 - Risiko-Abschätzung für Kitodo e.V.





Ausblick

- Finalisierung der Konzepte
- Erörterung der Konzepte durch AWBI
- DFG-Förderung?
- Abstimmung mit Kitodo-Verein
- Kooperative Pflege der Software
- Kitodo-Integration

TOP 12 b: Review zentraler Software-Komponenten

Arved Solth (effective Webwork)

Kriterien

- Coding Guidelines
- Open Source-Software-Lizenzen
- Test-Abdeckung
- Dokumentation
- Code-Umfang
- Sicherheit
- Web-Interface
- Verwendete Software-Pakete

→ Vergleich mit bereitgestellten Angaben

Repositories

Untersuchte Repositories:

- „core“
- „ocrd_olena“:
- „ocrd_tesseract“
- „ocrd_fileformat“

Insgesamt 89 Repositories alleine in der „OCR-D“-GitHub-Organisation

Zustand

„core“:

- Version 3.0 geplant, verschoben (voraussichtlicher Release Sommer 24)
 - Bereinigung Code-Basis
 - Prozessierung auf Seitenebene
- Allgemein schnelle Bearbeitung von Issues und Pull Requests
- Ein paar alte PRs, die ggf. geschlossen werden könnten

„ocrd_olena“, „ocrd_tesseract“, „ocrd_fileformat“:

- Stabil; keine aktive Entwicklung bzgl. Funktionen
- Verbesserung der Testabdeckung noch in Arbeit

Coding Guidelines

- Programmiersprache: Hauptsächlich Python
- Coding Guidelines: angepasste Liste weit verbreiteter PEP8-Vorgaben
- Einhaltung auf hohem Niveau (81 Code Style-Verletzungen in über 10k Zeilen Code)
- Coding Guidelines sollten aber irgendwo explizit dokumentiert sein

Open Source-Lizenzen

- OCR-D: Apache 2-Lizenz
- Fast alle verwendeten Software-Abhängigkeiten stehen unter kompatiblen Lizenzen wie BSD, MIT oder auch Apache 2
- Nur 2 Abhängigkeiten haben Lizenzen, die mit Apache 2 **nicht kompatibel** sind:
 - „**frozendict**“: LGPL-3.0
 - „**paramiko**“: LGPL-2.1
- Ersatz-Bibliotheken finden?
- Tatsächliches Risiko durch Experten zu beurteilen

Test-Coverage & Security

Test-Coverage

- Hohe Testabdeckung in „core“ von 80%
- Unit- und Integration-Tests vorhanden
- Integration-Tests für andere Repositories aktuell in Arbeit
- Testausführung noch nicht für alle Betriebssystem problemfrei!

Security

- Isolation zwischen Docker- und Host-System gewährleistet hohe Abschottung
- Regelmäßige Aktualisierung der Docker-Images empfohlen

Dokumentation & Code-Umfang

Dokumentation:

- Ausführliche und gut geschriebene Dokumentation sowohl für Nutzer als auch für Entwickler
- Sowohl native Installation als auch Docker-Installation & -Nutzung dokumentiert
- Bis auf ein paar Ausnahmen die meisten Dokumentationsseiten auf Deutsch und Englisch verfügbar

Code-Umfang:

- Nur leichte Abweichungen zwischen zur Verfügung gestellten Angaben und tatsächlich ermittelten Code-Zeilen
- Erklärbar durch zwischenzeitliche Weiterentwicklungen

Web-Interface & Verwendete SW-Pakete

Web-Interface:

- Stabiler Zustand und für Produktiveinsatz bereit
- REST-Endpunkte für spezifische OCR-D-Workflows sowie einzelne Prozessoren
- Hohe Testabdeckung gewährleistet Funktionalität

Software-Pakete:

- Verwendete 3rd-Party Software-Komponenten in ausgereiftem Zustand
- Längere Release-Historie in meisten Fällen

Fazit

- **Betrachtete OCR-D-Repositories befinden sich in einem guten Zustand**
- Frage bzgl. inkompatibler Software-Lizenzen muss geklärt werden!
- Laufende Arbeiten: Release 3.0 abwarten

Kleinere Baustellen:

- Vollständig übersetzte Dokumentation
- Behandlung alter Pull Requests
- Testausführung für andere Betriebssysteme

TOP 13: Verschiedenes

Praxistreffen Herbst 2024

- Im Hessischen Landesarchiv Marburg
- Datum – Verschiebung möglich?: 18./19.11.2024
(bisher: 05./06.11.2024)
- Aufruf zur Beteiligung folgt!

