

# Absichtserklärung

im Rahmen der Ausschreibung „Implementierung der OCR-D-Software zur Volltextdigitalisierung“

## Antragsteller

Antragsteller sind die

- SLUB Dresden, vertreten durch den Generaldirektor Dr. Achim Bonte
- UB Mannheim, vertreten durch die Ltd. Bibliotheksdirektorin Dr. Sabine Gehrlein
- UB Braunschweig, vertreten durch die Ltd. Bibliotheksdirektorin Katrin Stump.

Verantwortliche Projektleiter sind

- Sebastian Meyer, IT-Stabsstelle Digitale Bibliothek, SLUB Dresden
- Stefan Weil, kommissarischer Leiter der Abteilung Digitale Bibliotheksdienste, UB Mannheim
- Robert Strötgen, stellv. Direktor der UB Braunschweig und Leiter der Abteilung IT und forschungsnahe Services

## Projektziele

Kitodo ist mit dem Workflowmodul Kitodo.Production und dem Präsentationsmodul Kitodo.Presentation eine weit verbreitete und offene Lösung für die Massendigitalisierung in kulturbewahrenden Einrichtungen, die für große und kleine Institutionen passende Betriebsmodelle erlaubt. Ein auf den Werkzeugen und Workflows von OCR-D basierendes Verfahren zur Texterkennung muss deshalb der Flexibilität der verschiedenen Betriebsmodelle, der komplexen Workflows sowie der bedarfsgerechten Skalierbarkeit für kleine bis sehr große Digitalisierungsprojekte gerecht werden. Das geplante Vorhaben der Antragsteller verfolgt drei aufeinander aufbauende und sich komplementär ergänzende Projektziele, die im Ergebnis den Einsatz von OCR-D in der Massendigitalisierung mit Kitodo ermöglichen sollen.

### 1. Aufbau und Dokumentation einer web-basierten und skalierbaren OCR Appliance

Auf Basis der Werkzeuge und Verfahren von OCR-D sowie insbesondere der auf GitHub bereitgestellten Integration *ocrd\_all*<sup>1</sup> soll eine einfach zu installierende Serverlösung zur automatisierten Durchführung einer hoch-qualitativen Texterkennung im Kontext der Massendigitalisierung aufgebaut werden. Die Bereitstellung der Lösung soll in Form einer leicht nachnutzbaren und gut dokumentierten Appliance (z. B. Docker) erfolgen. Zusätzlich soll eine ausführliche Dokumentation Expert\*innen ermöglichen, die Komponenten alternativ manuell zu installieren und so erweiterte Konfigurationsmöglichkeiten und optimierte Betriebsmodi (etwa den Einsatz von GPUs) zu nutzen.

Um einen verteilten Betrieb sowie die Bereitstellung des OCR-Dienstes durch einen dedizierten Dienstleister zu ermöglichen, soll die OCR Appliance ausschließlich über Web-Schnittstellen mit Workflow-Systemen integriert werden können. In Abstimmung mit dem Koordinierungsprojekt sowie aufbauend auf den bestehenden OCR-D-Spezifikationen<sup>2</sup> werden die Schnittstellen dabei möglichst generisch spezifiziert, um auch von anderen Software-Lösungen neben Kitodo implementiert werden zu können.

Die Implementierung der Web-Schnittstellen auf Seiten des OCR-Dienstes soll in einem eigenen Software-Modul als OCR Controller erfolgen, um im Bedarfsfall einen skalierbaren Parallelbetrieb mehrerer, auch verteilter OCR Appliances zu ermöglichen. Der Controller soll einerseits die zentrale Schnittstelle zur Integration in Workflow-Systeme bereitstellen und gleichzeitig bedarfsgerecht bis zu einer konfigurierbaren Höchstgrenze mehrere OCR Appliances zur Verteilung der Prozesslast aktivieren und so ein skalierbares OCR-Cluster steuern können.

### 2. Implementierung eines OCR-Moduls für Kitodo.Production

Um die oben beschriebene OCR Appliance flexibel in Digitalisierungsworkflows integrieren zu können,

---

<sup>1</sup> vgl. [https://github.com/OCR-D/ocrd\\_all](https://github.com/OCR-D/ocrd_all)

<sup>2</sup> vgl. <https://ocr-d.de/de/spec/>

soll ein OCR-Modul für *Kitodo.Production* entwickelt werden. Das Modul soll analog zu bestehenden Modulen wie dem Meta- und Strukturdateneditor als unabhängiger Arbeitsschritt in beliebigen Workflows aufgerufen werden können.

Aufgabe des OCR-Moduls ist die automatisierte Übermittlung der METS-Datei des zu prozessierenden Digitalisats an den OCR Controller (siehe oben) sowie die Verfolgung und Visualisierung des Prozessfortschritts bis zum Abschluss der Texterkennung. Anschließend sollen die OCR-Ergebnisse in die Vorgangsdaten übernommen und standardkonform referenziert werden sowie der Workflow-Schritt automatisch abgeschlossen werden. Das OCR-Modul muss zudem sicherstellen, dass die eingangs übermittelte METS-Datei alle für die Texterkennung notwendigen Parameter und Metadaten enthält bzw. diese ggf. ergänzen und auch die OCR-Ergebnisse gegen entsprechende XML-Schemata (i. d. R. ALTO) validieren. Im Fehlerfall muss der Workflow unterbrochen und der Fehler auf geeignete Weise an den Anwender kommuniziert werden.

Das OCR-Modul soll zudem in einer Weise implementiert werden, die auch die wiederholte Durchführung einer OCR erlaubt. D. h. auch wenn für einen Vorgang bereits OCR-Daten vorliegen, soll dennoch eine erneute Texterkennung durchgeführt werden können. Dabei soll konfigurierbar sein, ob die bestehenden Ergebnisse ersetzt oder ergänzt werden sollen. Dadurch können Verbesserungen im OCR-Verfahren auch bereits prozessierten Retrodigitalisaten zu Gute kommen.

### **3. Implementierung eines OCR-Moduls für Kitodo.Presentation und den DFG-Viewer**

Auch mit der Verfügbarkeit einer quelloffenen, freien OCR-Lösung für die Massendigitalisierung wird die weitgehend vollständige Texterkennung aller bereits digitalisierten Bestände noch viel Zeit in Anspruch nehmen. Um Nutzenden in der Zwischenzeit dennoch im Bedarfsfall zeitnah den Volltext eines Werks zur Verfügung stellen zu können, soll auch in der Präsentationsoberfläche *Kitodo.Presentation* die Möglichkeit geschaffen werden, über die oben beschriebene OCR Appliance eine Texterkennung durchzuführen. Diese dynamisch von Nutzenden ausgelöste Texterkennung ermöglicht analog zu bestehenden Angeboten wie „Digitisation on Demand“ eine bedarfsgerechte „OCR on Demand“.

Dieser Dienst soll auch im DFG-Viewer, dem von der SLUB Dresden betriebenen und auf *Kitodo.Presentation* basierenden freien Webdienst zur Anzeige von Digitalisaten aus verteilten Repositorien, zur Verfügung gestellt werden. Die SLUB Dresden wird die dazu notwendige Infrastruktur aufbauen und die Anbindung an ein freies, von Mitgliedern der Kitodo-Community bereitgestelltes OCR-Cluster vornehmen, dessen erste Instanzen initial die Projektpartner zur Verfügung stellen werden.

Da der Texterkennung dasselbe Verfahren und dieselbe Parametrisierung wie im regulären Digitalisierungsworkflow zugrunde liegen, können die Ergebnisse ebenso in der Datenhaltung persistiert und etwa langzeitarchiviert werden. Letztlich stellt die nutzergesteuerte Texterkennung also lediglich eine alternative Priorisierung dar, jedoch kein grundlegend verschiedenes Verfahren, dessen Ergebnisse weniger hochwertig oder valide sind. Ein besonderes Augenmerk bei der Implementierung muss daher auf dem Zwischenspeichern, der Bewertung der OCR-Qualität durch die Nutzenden und dem Rückfluss der Daten in die regulären Digitalisierungsworkflows liegen – insbesondere bei einem verteilten Betrieb der Kitodo-Komponenten oder mandantenfähigen Installationen wie dem DFG-Viewer, bei dem die OCR-Ergebnisse abhängig vom Datenanbieter an unterschiedliche Systeme übermittelt werden müssen.

### **4. Weitere Projektaktivitäten**

Zur Erreichung der oben beschriebenen Ziele des Projekts verfolgt das Vorhaben noch eine Reihe Aktivitäten, etwa die Durchführung von Community-Workshops zur Anforderungserhebung und Dissemination, eine ausführliche Dokumentation der Projektergebnisse sowie die Evaluierung eines allgemein verfügbaren OCR-Dienstangebots. Hier werden im Projekt die enge Anbindung an die vitale Community und den Verein "Kitodo. Key to digital objects" e. V. genutzt. Sämtliche Projektergebnisse werden entsprechend der Vereinssatzung quelloffen und zur freien Nachnutzung auf GitHub zur Verfügung gestellt.

## **Anwendungsszenarios**

Die OCR Appliance soll alle Anwendungsszenarios und Betriebsmodelle unterstützen, die auch mit Kitodo möglich sind und in der Praxis eingesetzt werden. So sollen grundsätzlich alle Software-Komponenten (*Kitodo.Production*, *Kitodo.Presentation*, OCR Controller und OCR Appliance) ausschließlich webbasiert miteinander kommunizieren, so dass sowohl ein lokaler Betrieb als auch eine verteilte Infrastruktur möglich ist.

Darüber hinaus sollen alle Komponenten mandantenfähig konzipiert werden, um mit einer einzigen Instanz mehrere Anwender bedienen zu können. Dies erfordert unter anderem die Möglichkeit zur Authentifizierung einzelner Anwender und Mechanismen zur Volumenkontrolle.

Der verteilte und mandantenfähige Betrieb ermöglicht unter anderem Service-Architekturen wie sie etwa regionale Verbünde bereitstellen ebenso wie kommerzielle Dienstleistermodelle oder den verteilten Betrieb eines OCR-Clusters innerhalb einer Community.

Grundsätzlich soll beim Aufbau der OCR Appliance sowie aller Betriebsmodelle sichergestellt werden, dass keine exklusiven Abhängigkeiten zu Kitodo entstehen, so dass sie auch mit anderen Workflow-Managementsystemen oder im wissenschaftlichen Kontext eingesetzt werden können. Der Kitodo e. V. prüft zudem, ob der Aufbau eines freien, community-gestützten OCR-Dienstes möglich ist.

## Vorarbeiten

Alle drei Projektpartner verfügen über umfangreiche Vorerfahrungen im Einsatz und der Entwicklung von Kitodo sowie OCR-D.

Die SLUB Dresden führte das Konsortialprojekt zur Weiterentwicklung von Kitodo.Production, auf dessen Ergebnissen das hier skizzierte Projekt aufbauen soll, während die UBs in Mannheim und Braunschweig als Piloteinrichtungen die neue Kitodo-Version einsetzen.

In der zurückliegenden OCR-D-Förderphase war die UB Mannheim als Projektnehmer mit einem erfolgreichen Entwicklungsprojekt beteiligt und hat maßgeblich zum Aufbau der *ocrd\_all*-Integration beigetragen, die im nun geplanten Projekt zum Einsatz kommen soll. Die SLUB Dresden und die UB Braunschweig sind im wissenschaftlichen Beirat vertreten und so ebenfalls in die Projektaktivitäten involviert. SLUB Dresden und UB Mannheim haben zudem an der Pilotevaluierung der OCR-D-Projektergebnisse teilgenommen.

Durch die Vertretung der UB Braunschweig im Vorstand des Kitodo e. V. besteht nicht zuletzt auch eine enge Verbindung zur Community und eine Beteiligung an der zwischen Kitodo e. V. und den Koordinatoren von OCR-D geschlossenen Kooperationsvereinbarung<sup>3</sup>, die im Rahmen des hier skizzierten Projekts in konkrete Maßnahmen umgesetzt werden soll.

## Grobe Abschätzung des Mittelbedarfs

Die Projektpartner rechnen mit einem Gesamtaufwand von 60 Personenmonaten in der Personalkostenkategorie Doktorandin/Doktorand und Vergleichbare über 24 Monate Laufzeit zuzüglich Sachmitteln in Höhe von ca. 15.000 EUR für Reise- und Veranstaltungskosten sowie Infrastrukturmaßnahmen im Kontext des DFG-Viewers.

05. Mai 2020



---

Dr. Achim Bonte  
SLUB Dresden



---

Dr. Sabine Gehrlein  
UB Mannheim



---

Katrin Stump  
UB Braunschweig

---

<sup>3</sup> vgl. <https://ocr-d.de/de/2019/11/20/kooperation-mit-kitodo.html>