

*[Dieser Fortsetzungsantrag wurde am 22. Oktober 2021 bei der DFG eingereicht. Für alle im Text genannten Personen mit Positionen gibt der Antrag den Planungsstand bei der Einreichung wieder.]*

## Handschriftenportal (HSP)

Fortsetzungsantrag

### **Beschreibung des Vorhabens**

**Fortsetzungsantrag zum DFG-Projekt Handschriftenportal (HSP)**

<b>1</b>	<b>AUSGANGSLAGE</b> .....	<b>1</b>
1.1	Ausgangslage und eigene Vorarbeiten.....	1
1.2	Projektbezogenes Publikationsverzeichnis.....	4
1.2.1	Veröffentlichte Arbeiten aus Publikationsorganen mit wissenschaftlicher Qualitätssicherung, Buchveröffentlichungen sowie bereits zur Veröffentlichung angenommene, aber noch nicht veröffentlichte Arbeiten.....	4
1.2.2	Andere Veröffentlichungen mit und ohne wissenschaftliche Qualitätssicherung .....	4
1.2.3	Patente.....	4
<b>2</b>	<b>ZIELE UND ARBEITSPROGRAMM</b> .....	<b>5</b>
2.1	Voraussichtliche Gesamtdauer des Projekts .....	5
2.2	Ziele.....	5
2.3	Arbeitsprogramm und Umsetzung.....	6
2.3.1	Querschnittsaufgaben.....	6
	AP 1 – Projektmanagement.....	6
	AP 2 – Kommunikations- und Öffentlichkeitsarbeit .....	6
	AP 3 – Usability und User Experience.....	7
	AP 4 – Schema-/Skriptpflege.....	8
	AP 5 – Nutzer:innenverwaltung .....	8
	AP 6 – Infrastruktur .....	9
2.3.2	Einzelmodule .....	10
	AP 7 – Normdatenmodul .....	10
	AP 8 – Nachweis, Erfassung, Veröffentlichung .....	11
	AP 9 – Import .....	14
	AP 10 – Nachweis externer Daten.....	14
	AP 11 – Export.....	15
	AP 12 – Kollaborativer Arbeitsplatz .....	16
	AP 13 – Rechercheerweiterung und -optimierung.....	19
	AP 14 – Redaktion .....	20
	AP 15 – KI-Begleitung.....	23
2.3.3	Auswirkungen.....	23
	AP 16 – IIF-Hosting.....	23
	AP 17 – Richtlinien Handschriftenerschließung .....	24
<b>3</b>	<b>LITERATURVERZEICHNIS</b> .....	<b>25</b>
<b>4</b>	<b>BEGLEITINFORMATIONEN ZUM PROJEKTKONTEXT</b> .....	<b>25</b>
4.1	Allgemeine ethische Aspekte .....	25
4.2	Maßnahmen zur Erfüllung der Förderbedingungen und Umgang mit den Projektergebnissen .....	25
4.3	Erklärungen zur Erfüllung der Förderbedingungen.....	25

## **Beschreibung des Vorhabens – Projektanträge im Bereich „Wissenschaftliche Literaturver- sorgungs- und Informationssysteme“ (LIS)**

### **LIS-Förderprogramm oder Ausschreibung: Digitalisierung und Erschließung**

#### **Fortsetzungsantrag zum DFG-Projekt Handschriftenportal (HSP)**

Dr. Achim Bonte (Generaldirektor Staatsbibliothek zu Berlin - Preußischer Kulturbesitz)

Prof. Dr. Ulrich Johannes Schneider (Direktor Universitätsbibliothek Leipzig)

Dr. Klaus Ceynowa (Generaldirektor Bayerische Staatsbibliothek München)

Prof. Dr. Peter Burschel (Direktor Herzog August Bibliothek Wolfenbüttel)

Prof. Dr. Elke Greifeneder (Stellvertretende Direktorin Institut für Bibliotheks- und Informationswissenschaft [IBI], Humboldt-Universität zu Berlin)

---

## **1 Ausgangslage**

### **1.1 Ausgangslage und eigene Vorarbeiten**

Seit Oktober 2018 entsteht im Rahmen eines DFG-Projekts das deutsche [„Handschriftenportal“](#) (HSP) als zeitgemäßer und technisch zukunftsweisender Zentralnachweis für Text- und Bildinformationen zu Buchhandschriften, dessen innovatives Potential bereits jetzt international Beachtung findet. Das HSP ist als materialspezifisches und epochenübergreifendes Nachweissystem für alle Objekte konzeptioniert, die dem Überlieferungstyp ‚Europäische Buchhandschrift‘ zugehören, und ersetzt den bisherigen Zentralkatalog [Manuscripta Mediaevalia](#) (MM). Perspektivisch ist eine Zusammenführung mit dem Nachweis für orientalische Handschriften (Qalamos: connecting manuscript traditions [vormals [Orient-Digital](#)]) vorgesehen.

Die **zentralen Charakteristika** des HSP-Projekts sind:

- die verteilte Entwicklung durch vier wichtige Handschriftenbibliotheken mit leistungsfähigen IT-Abteilungen: [SBB-PK](#) als Projektleitung und langfristige Betreiberin, in der Entwicklung zuständig für Datenhaltung und -erfassung; [UBL](#): Entwicklung von Recherche, Präsentation und digitalem Arbeitsplatz für Nutzer:innen; [HAB](#): Skript- und Schemapflege, Datenkonversion; SBB, UBL, HAB und [BSB](#): Verbesserung und Vereinheitlichung der Datenbasis durch eine auf die vier Häuser verteilte Datenredaktion
- die Projektdurchführung nach dem Prinzip der agilen Softwareentwicklung mit einem häuserübergreifenden Projektteam aus inhaltlicher Fachseite für die Anforderungserstellung und starkem IT-Entwicklungsanteil für die angemessene Umsetzung innovativer Technologien
- die Nutzung von Open Source-Software und Datenmodellierung nach Linked Open Data-Prinzipien (LOD); Verwendung der im Handschriftenbereich international etablierten Standards TEI-XML für textuelle Inhalte und IIIF für Bilddaten
- eine größtmögliche Transparenz nach außen, konsequente projektbegleitende Einbeziehung der internationalen wissenschaftlichen und infrastrukturellen Zielcommunities und Ausrichtung an deren Bedarfen
- die Begleitung durch die Handschriftenzentren und deren Wissenschaftlichen Beirat, enge Kooperation mit anderen handschriften-spezifischen Portalanbietern des deutschsprachigen Raums und deren Einbeziehung in den international besetzten Wissenschaftlichen Beirat des HSP

- die kontinuierliche Berücksichtigung von Usability-Kriterien längs der gesamten Entwicklungsarbeit, Zusammenarbeit mit professionellen Partner:innen für User Experience

Der **Projektstand mit Stichtag 30. September 2021** wird im beigegebenen Bericht (s. Anlage 1) detailliert dargelegt und lässt sich wie folgt zusammenfassen:

- Die Infrastruktur für Projektentwicklung und Datenverwaltung an der SBB-PK ist eingerichtet.
- Die zentrale Repräsentanz der Handschrift im System, das ‚Kulturobjektdokument‘ (KOD), ist technisch etabliert und wird mit den normierten objekterschließenden ‚Kerndaten‘ aus der Datenredaktion gefüllt.
- Im ‚Nachweis‘, der zentralen Datenablage und -verwaltung, können Handschriftenbeschreibungen strukturiert im TEI-Format abgelegt und von dort an die Präsentation ausgeliefert werden.
- Nutzer:innen können sich mit unterschiedlichen Rollen (Erschließer:in, Redakteur:in) am Nachweis anmelden.
- Die Umwandlung von aus MM stammenden Handschriftenbeschreibungen in TEI-XML und deren Einspielung ins HSP ist implementiert und wird iterativ mit immer neuen Datenbeständen verbessert.
- Der Normdatenservice zur Verwaltung von Personen, Körperschaften, Orten, Sprachen und Beziehungen steht zur Verfügung und ist durch Nachweis und Präsentation über eine HTTP REST-Schnittstelle abfragbar.
- Das Erfassungssystem für die manuelle Erstellung neuer TEI-konformer Handschriftenbeschreibungen liegt in einer ersten Version vor.
- Im Zuge der Datenredaktion wurden und werden OCR-bearbeitete Handschriftenbeschreibungen aus 257 gedruckten Handschriftenkatalogen als Ergänzung des bisherigen Datenbestands von MM so aufgearbeitet, dass ihre Darstellung als Katalogimages und die Recherche in ihren Volltexten möglich ist. Gleichzeitig sind bislang 80% des Gesamtbestands an Handschriftenbeschreibungen in MM auf ihre Kerndaten hin für die Recherche ausgezeichnet und vereinheitlicht.
- Seit dem 7. Dezember 2020 ist die ‚[Testversion](#)‘ der Präsentation mit Recherche und Anzeige online (inkl. englischer Nutzer:innenführung), seitdem profitiert das Projekt von den Rückmeldungen der Nutzer:innen, wird iterativ optimiert und mit neuen Features versehen. Derzeit sind Daten zu 4.342 Handschriften online. Weitere 1.148 Handschriftendokumente stehen in der Stage-Umgebung für die Veröffentlichung im nächsten Sprint zur Verfügung.
- Der IIIF-Viewer [Mirador](#) ist für die Darstellung von Texten angepasst, so dass Nutzer:innen im ‚Arbeitsplatz‘ Digitalisate und Beschreibungen in einer einheitlichen Präsentationsumgebung nutzen können.
- Die Integration des Content Management Systems (CMS) für die redaktionellen Seiten der künftigen Portal-Website ist in Entwicklung.
- Projektarbeiten und -stand werden durch eine eigene Kommunikations- und Öffentlichkeitsarbeit gezielt und kontinuierlich in die Nutzungszielgruppen kommuniziert und deren Rückmeldungen koordiniert ins Projektteam rückgespiegelt. Die [HSP-Projektwebsite](#) ist seit 18. Februar 2020 online und informiert aktuell mit [News-Seite](#) und eigenem [Blog](#). Der Twitter-Account ([@hsprt1](#)) ist seit August 2020 eingerichtet und hat mit Stand 30.09.21 1.388 Follower. Interessierte können außerdem einen Newsletter abonnieren.

Aufgrund des aktuellen Projektstands ist davon auszugehen, dass **bis zum Ende der 1. Förderphase**, die angesichts unterschiedlicher Einzelaufzeiten in den Teilprojekten und aufgrund von

Corona-Verlängerungen (vgl. Bericht) in Abstimmung mit der DFG-Geschäftsstelle für den 31. März 2022 angesetzt ist, folgende Funktionalitäten bereitstehen:

- HSP-Präsentation mit Recherche, ergänzt um statische Informationsseiten und Social Media-Einbindung
- Erfassungssystem in einer Basisversion
- Bereitstellung TEI- und MXML-Importschnittstelle, TEI-Exportschnittstelle
- Bereitstellung aller aus MM migrierten Datenbestände mit voller Recherchemöglichkeit
- Recherche in Volltext und Kerndaten der 257 während des Projekts digitalisierten Handschriftenkataloge mit Anzeige der Katalogimages

Bereits im **Erstantrag** war die **Perspektive** eines Fortsetzungsantrags vorgesehen. Nach damaliger Planung sollten in einer 2. Förderphase besonders im Fokus stehen:

- der Nachweis relevanter Handschriftenbestände aus anderen Systemen
- die auf das Digital Age und die HSP-Möglichkeiten abgestimmte Aktualisierung der [DFG-Richtlinien Handschriftenkatalogisierung](#) (DFG-Richtlinien)
- die Weiterentwicklung des Arbeitsplatzes zu einem interaktiv nutzbaren und für die Publikation von User-generiertem Content (Annotationen) offenen Instrument
- die Ausdehnung der Datenbearbeitung auf weitere besonders rechercherelevante Informationen jenseits der Objekt-Kerndaten inkl. Erstellung von Fach-Thesauri
- die Entwicklung eines Tools, um aus dem HSP-Datenbestand heraus Katalogpublikationen zu generieren

Den Gegebenheiten der agilen Entwicklung entsprechend, haben sich innerhalb der Projektlaufzeit Aufgabenfelder konkretisiert, wurden Umgewichtungen notwendig und haben sich neue Erfordernisse für die fachgerechte Optimierung des HSP-Angebots in einer 2. Projektphase ergeben. Einzelne im Erstantrag geplante Teilarbeitspakete ließen sich bislang nicht in vollem Umfang umsetzen, was aber durch kostenneutrale Laufzeitverlängerungen abgefangen wird (vgl. Bericht). Das umfangreiche Feedback aus den Zielcommunities hat ebenfalls wichtige Anregungen gegeben, die in diesen Antrag eingeflossen sind.

Die Projektpartner haben aus den damit konturierten Bedarfen eine Anforderungsanalyse erstellt, die die Grundlage des vorliegenden Fortsetzungsantrags bildet. Die **Schwerpunkte der geplanten Projektarbeiten in einer 2. Förderphase** sind:

- Weiterentwicklung des Normdatenmoduls
- Nachweis von Daten aus externen Systemen
- IIF-Beratungs- und Hosting-Services
- Neufassung Richtlinien Handschriftenerschließung
- Personalisierter und kollaborativer Arbeitsplatz
- Ausbau des Nachweismoduls mit Datenerfassung und Nutzerverwaltung
- Rechercheerweiterung und -optimierung
- Datenverbesserung und Thesaurus-Erstellung

Die gesamte Entwicklung soll dabei weiterhin durch eine aktive Öffentlichkeitsarbeit an die Zielgruppen transparent vermittelt und durch eine professionelle Usability- und User-Experience-Begleitung nutzungsgerecht gestaltet werden. Schließlich sollen die Erfahrungen aus dem BMBF-Projekt [QU-RATOR](#) genutzt werden, um ein Konzept für einen produktiven Einsatz von Künstlicher Intelligenz (KI) im HSP zu erarbeiten.

Um sicherzustellen, dass das HSP auch nach Ende der Drittmittelförderung ein laufend gepflegtes und für Innovationen offenes Produkt bleibt, haben die Direktionen der vier beteiligten Einrichtungen sich am 4. Oktober 2021 auf folgenden Aufgabenzuschnitt verständigt:

Der technische Betrieb des Systems erfolgt durch die SBB-PK Berlin. Dies schließt Hosting, Wartung, Aktualisierungen, Feature-Anpassung und Bug-Fixing ein. Die BSB steht für eine finanzielle Beteiligung an diesen technischen Aufgaben bereit. Die fachlich-inhaltliche Betreuung (Beratung, Anforderungserhebung, Innovationsmanagement, Austausch mit Partnern, Schema- und Skriptpflege etc.) erfolgt über eine Redaktion, die maßgeblich mit Personal der SBB und der BSB sowie zu einem kleineren Teil der UBL und der HAB besetzt ist. Als Pilotanbieter der IIF-Hosting- und Beratungsservices sind die UBL und die BSB dauerhaft in die HSP-Infrastruktur integriert. Im Laufe der 2. Phase werden diese Zusagen aktualisiert und formalisiert.

Es besteht außerdem Einigkeit, dass das HSP weiterhin von einem Wissenschaftlichen Beirat aus Forschung, Infrastruktureinrichtungen, anderen Portalanbietern und den Handschriftenzentren begleitet werden soll, um so die dauerhafte Ausrichtung an den Bedarfen der Zielcommunities zu gewährleisten. Insbesondere die Ausweitung der internationalen Reichweite des HSP soll dieser Beirat unterstützen und begleiten.

## **1.2 Projektbezogenes Publikationsverzeichnis**

### **1.2.1 Veröffentlichte Arbeiten aus Publikationsorganen mit wissenschaftlicher Qualitätssicherung, Buchveröffentlichungen sowie bereits zur Veröffentlichung angenommene, aber noch nicht veröffentlichte Arbeiten.**

-

### **1.2.2 Andere Veröffentlichungen mit und ohne wissenschaftliche Qualitätssicherung**

20 Blogbeiträge auf <https://handschriftenportal.de/blog/>

3 Blogbeiträge auf <https://mittelalter.hypotheses.org/author/carolinhahnhsb>

1 Tagungspräsentation auf <https://doi.org/10.5281/zenodo.4912650>

1 Bericht zum Card Sorting-Test: <https://edoc.hu-berlin.de/handle/18452/23848>

### **1.2.3 Patente**

-

## 2 Ziele und Arbeitsprogramm

### 2.1 Voraussichtliche Gesamtdauer des Projekts

Das Projekt läuft seit 1. Oktober 2018. Die hier beantragte Projektfortsetzung ist wieder auf drei Jahre angelegt.

Geplanter Projektbeginn: 1. April 2022

### 2.2 Ziele

Nachdem in der 1. Förderphase das System in seinen grundständigen Komponenten und Funktionalitäten (Datenhaltung, -erzeugung und -format; Präsentation und Recherche; Vervollständigung des Datenbestands) aufgebaut wurde, soll in der 2. Phase das innovative Potential des Portals entfaltet werden. Auf der Basis der aktualisierten Anforderungsanalyse sind die zentralen Ziele hierbei Vernetzung, Öffnung für Interaktionsmöglichkeiten und Ausdifferenzierung im Bereich von Datenmodell, -generierung, -recherche und -pflege. Diese drei Schwerpunkte des Fortsetzungsantrags beinhalten folgende Arbeitspakete:

Vernetzung:

- Weiterentwicklung des Normdatenmoduls und Aufbau eines Knowledge Graphen als zentraler logischer Knoten für Datenbeziehungen (AP 7)
- Schaffung von Exportschnittstellen und Erweiterung des Imports (AP 9, AP 11)
- Nachweis von Daten externer Anbieter (AP 10)
- IIF-Dienstleistungen für Einrichtungen ohne eigene oder ohne IIF-fähige Präsentation (AP 16)
- breit abgestimmte Aktualisierung der Richtlinien Handschriftenerschließung im Digitalen Zeitalter (AP 17)

Öffnung für Interaktionsmöglichkeiten:

- Entwicklung eines kollaborativen Arbeitsplatzes (AP 12)
- Öffnung für user-generated Content (Annotation/Publikation) (AP 12)

Ausdifferenzierung:

- Ausbau des Erfassungsmoduls (AP 8)
- Versionierung und Publikation von Beschreibungen (AP 8)
- Ergänzung des Datenmodells (Kuturobjektkomponente, virtuelle Kulturobjekte, weitere Entitäten des Nachweises) (AP 8)
- Aufbau der Nutzer:innenverwaltung (AP 5)
- neue Rechercheeinstiege in der Präsentation (Normdatensuche, Katalogsuche, Visualisierung, Personalisierung) (AP 13)
- umfassende Normdatenreferenzierung im nochmals ergänzten Gesamtdatenbestand inkl. Thesaurus-Erstellung (AP 14).

Eine Übersicht über die künftigen technischen Module des HSP und ihre Beziehungen zueinander gibt Anlage 2. Aufgrund der komplexen Beziehungen zwischen den Modulen wird im vorliegenden Antrag auf eine pauschale Gruppierung in Back- und Frontoffice bewusst verzichtet.

Der vorliegende Antrag wurde mit den Wissenschaftlichen [Beiräten des Handschriftenportals](#) und der [Handschriftenzentren](#) diskutiert und wird von ihnen unterstützt.

Dank der verbesserten Erfassung und der innovativen Recherche kann in der 2. Förderphase die vollständige Abschaltung der Altsysteme von Manuscripta Mediaevalia (MM) erfolgen: Mit Online-schaltung des Handschriftenportals bei Abschluss der 1. Phase wird 2022 MM als Sucheinstieg ersetzt; die verbesserte Erfassung löst im Anschluss die bisherige Dateneingabe per MXML ab.

Schon jetzt zeichnet sich ab, dass sich im Verlauf der 2. Förderphase ergänzende Aufgaben im Bereich des Nachweises externer Daten, in Zusammenhang mit dem User Interface für den Knowledge Graph, in Hinblick auf Verwendungsszenarien von KI-Verfahren bei der Datenredaktion und bezüglich der Erzeugung von digitalen Katalogpublikationen aus dem Datenbestand ergeben werden, die gegebenenfalls über einen Ergänzungsantrag noch in das Projekt aufgenommen werden, vgl. unten bei den einzelnen APs.

## **2.3 Arbeitsprogramm und Umsetzung**

### **2.3.1 Querschnittsaufgaben**

#### **AP 1 – Projektmanagement**

Ziel des AP ist die Fortführung des agilen Projektmanagements aus der 1. Förderphase, welches sich als effizient erwiesen hat.

Die Teilprojekte der beteiligten Institutionen sind durch regelmäßige Meetings in unterschiedlichen Konstellationen miteinander verzahnt und werden von der Gesamtprojektleitung (Dr. Robert Giel SBB / Dr. Christoph Mackert UBL) koordiniert, für technische Entscheidungen steht das Technical Board (Malgorzata Asch SBB / Leander Seige UBL) zur Verfügung. In den beiden Hauptentwicklungsinstitutionen SBB und UBL ist die Rolle des Product-Owners zentral, um die Vermittlung zwischen Fachseite und Entwicklerteams, ein detailliertes Anforderungsmanagement mit Entwicklung von User Stories, die Releaseplanung und den Abgleich mit der übergreifenden Projektvision zu leisten. In der 1. Projektphase hat sich gezeigt, dass die tatsächlichen Aufwände hierfür zu gering kalkuliert waren, was durch massive Eigenleistungen abgedeckt werden musste. Für die 2. Projektphase ist daher geplant, dass an der SBB weitere Personen modulbezogen in die Rolle des Product-Owners aufgenommen werden: Die Kommunikation mit fachlichen Ansprechpartner:innen innerhalb des Projektteams wird weiterhin in Eigenleistung wahrgenommen, künftig von Dr. Robert Giel (SBB) und Dr. Carolin Schreiber (BSB). Auf diese Weise wird die BSB als künftige Mitbetreiberin des Portals aktiv in das Gesamtmanagement in der Entwicklungsphase eingebunden. Für Anforderungsspezifikation und -management (Unterstützung Product-Owner und Scrum-Master) ist an der SBB Frau Magdalena Luniak vorgesehen, die bereits im Projekt tätig ist. An der UBL wird die Aufgabe der Product Ownerin weiterhin von Annika Schröer übernommen werden.

#### **AP 2 – Kommunikations- und Öffentlichkeitsarbeit**

Die projektbegleitende Kommunikations- und Öffentlichkeitsarbeit (ÖA) wird fortgesetzt, um transparent und kontinuierlich nach außen über das Projekt zu informieren und Nutzer:innen-Feedback anzustoßen. Darüber hinaus wird in der 2. Projektphase ein Fokus auf der Organisation von Umfragen und Workshops sowie der Erarbeitung von digitalem Informationsmaterial für Nutzer:innen liegen.



Die begleitende Kommunikation über das Projekt erfolgt via Twitter-Account, die HSP-eigene News-Seite, den Newsletter und bei Bedarf über Blog-Artikel, auch die einschlägigen Mailinglisten werden bespielt. Die ÖA beobachtet zudem das Angebot einschlägiger Fachtagungen, auf denen das HSP vorgestellt werden kann, und unterstützt Bewerbung und Durchführung. Die sozialen Medien YouTube, Facebook und Instagram werden im Bedarfsfall über die jeweiligen Accounts der SBB bespielt. Zudem wird weiterhin die Plattform [Zenodo](#) für die Veröffentlichung von Präsentationen genutzt.

Die ÖA unterstützt außerdem sowohl die Redaktion des HSP bei einlaufenden Anfragen vor allem während der vollständigen Ablösung von MM als auch das AP Usability und User Experience (s. AP 3) bei der Bewerbung von Usability-Tests und Interviews sowie der Verbreitung der Ergebnisse.

Ein wichtiger neuer Aspekt der ÖA in der 2. Projektphase wird die Kommunikation mit einzelnen Fachcommunities im Rahmen der Thesaurus-Entwicklung und Facettenbildung darstellen. Die ÖA organisiert in diesem Zusammenhang Umfragen und Workshops (s. AP 14).

Rechtzeitig vor Ende des Projekts wird außerdem eine zentrale, von der ÖA organisierte Auswertungsveranstaltung die Nutzungsmöglichkeiten des HSP vor Vertreter:innen der verschiedenen Zielgruppen vorstellen und nochmals Gelegenheit für intensives Feedback bieten.

### **AP 3 – Usability und User Experience**

Ziel des AP ist die systematische Einbeziehung von Aspekten der Usability und User Experience (UX) längs der gesamten Entwicklung, um die intuitive Bedienbarkeit und ein positives Nutzungserlebnis auch für die neuen Features zu gewährleisten und gleichzeitig die Entwicklung besonders effizient zu gestalten. Unter Usability wird die effektive und effiziente Nutzung durch die Zielgruppen verstanden, unter UX die Qualität der Erfahrung, die Nutzende mit einem System machen. Zu dieser Qualität der Erfahrung gehören die Elemente der guten Benutzbarkeit, aber auch die Erwartungshaltung vor und Erfahrung mit dem Produkt nach Nutzung; auch die Barrierefreiheit wird wie in der 1. Phase geprüft. Das AP arbeitet dabei eng mit dem AP 2 zusammen, um die kontinuierliche Einbeziehung der Zielcommunities durch Befragungen und Workshops zu begleiten und zu ergänzen.

Die Ziele des Arbeitspakets werden gemäß dem in Absprache mit der DFG entwickelten Usability-Konzept aus der 1. Projektphase umgesetzt und die Zusammenarbeit mit entsprechend spezialisierten Partner:innen verstärkt. Dies betrifft zum einen das Webdesign für Präsentation/Arbeitsplatz und Nachweis/Erfassung, das wieder zusammen mit einem professionellen Anbieter weiterentwickelt werden soll. Die kontinuierliche UX-Begleitung wird durch die neu hinzugekommene Mittragsstellerin [Prof. Dr. Elke Greifeneder](#) umgesetzt. Die [Humboldt-Universität zu Berlin](#) (HU) stellt für das Arbeitspaket Soft- und Hardware sowie das [iLab](#) als Forschungslabor für UX-Tests.

Das AP baut auf den UX-Tests der 1. Förderphase auf (s. Anlage 1), verschiebt den Fokus aber weg von der Gesamtnutzung des Portals hin auf die neu zu entwickelnden Features. Dies betrifft insbesondere die Datenerfassung (AP 8, geplante Methode: Eyetracking und Think-Aloud), die komplexe Nutzungssituation des personalisierten kollaborativen Arbeitsplatzes (AP 12, geplant: Customer Journey Mapping), die Normdatennutzung in Erfassung und Recherche (AP 8, 13, geplant: First Click Testing) und die erweiterten Rechercheoptionen (AP 13, geplant: Tree Jack Testing).

#### **AP 4 – Schema-/Skriptpflege**

Ziel des AP ist die Weiterentwicklung und Pflege aller Schemata und Skripte, die im Rahmen des Portals benötigt werden, sowie die Dokumentation der Datenstrukturen und ihres Einsatzes bei der Erfassung.

Grundlage der Erfassung sowie des Datenimports ist eine HSP-spezifische TEI-Customisation, die in einer sogenannten ODD-Datei (= One Document Does it all) dokumentiert ist. Das ODD repräsentiert damit als Regelwerk die Erfassungs-, Darstellungs- und Transformationsmöglichkeiten des HSP. Eine daraus generierte Schema-Datei im Format RelaxNG ermöglicht die Validierung aller Daten im HSP.

Das ODD muss in der 2. Phase kontinuierlich fortgeschrieben und angepasst werden. Insgesamt wird die Erfassung auf Beschreibungsvolltexte ausgerichtet, die eine relativ flache Strukturierung aufweisen und in denen Entitäten und Begriffe ausgezeichnet werden, die für die Recherche und Facettierung genutzt werden.

Da die HSP-interne TEI-Struktur über den bisherigen TEI-Stand an einigen Stellen programmatisch hinausgeht (z. B. Einbände als eigenständige msPart-Komponenten, Parallelisierung von Inhalts- und Formbeschreibung bei illuminierten Handschriften), muss ein entsprechender Anpassungsprozess in der TEI-Community initiiert werden.

Im AP werden alle Transformationsskripte entwickelt und gepflegt. Dazu gehören Skripte für die Schnittstellen des Normdatenmoduls (AP 7), für den Datenimport aus Bibliothekssystemen (AP 9) und von externen Anbietern (AP 10) sowie für den Datenexport (AP 11). Neben die bereits entwickelte Konversion aus MXML und beliebigen externen TEI-Formaten treten in der 2. Phase Transformationen aus und zu MARC21.

Der Aufwand für die Schema- und Skriptpflege hat sich im Lauf der ersten Projektphase gegenüber den ursprünglichen Planungen als deutlich höher erwiesen. Dem wird mit einer Aufstockung der beantragten Personalmittel und einem entsprechend erhöhten Eigenanteil Rechnung getragen.

#### **AP 5 – Nutzer:innenverwaltung**

Ziel des AP ist das sichere und komfortable Management der persönlichen Daten von Nutzer:innen mittels einer graphischen Oberfläche. Zu diesem Zweck wird die Verwaltung dieser Daten auf zwei Systeme verteilt. Die Registrierung und Authentifizierung der Nutzer:innen wird weiterhin über das Identitymanagement auf Datenbasis eines LDAP-Verzeichnisses durchgeführt. Die Verwaltung der digitalen Identitäten und Anwendungsberechtigungen bzw. Rollen (Role Based Access) wird davon unabhängig in einem separaten System implementiert und betrieben. Dieses erlaubt insbesondere die Zuweisung von Nutzer:innen und -gruppen zu bestimmten Rollen, die Änderung der persönlichen Daten sowie die Auswahl eines Authentifizierungsverfahrens (OAUTH2, Shibboleth [SAML2]).

Die eigenständige Nutzer:innenverwaltung ermöglicht die unabhängige Auswertung der persönlichen Daten durch jede Softwarekomponente des HSP. Darüber hinaus genügt diese Konzeption aktuellen Sicherheits- und Datenschutzerfordernungen, da durch sie lediglich das Identitymanagement im Internet verfügbar sein muss, nicht aber die Nutzer:innenverwaltung selbst. Außerdem wird das Modul nur durch entsprechend autorisierte Mitarbeiter:innen verwendet. Damit erhöht diese Trennung in zwei Teilsysteme die Sicherheit und Flexibilität des Gesamtsystems.

## **AP 6 – Infrastruktur**

### Betrieb der Infrastruktur und Administration

Für das gesamte System des neuen Handschriftenportals muss an der SBB-PK weiterhin eine Infrastruktur aus VM-Servern, Datenbanken, Netzwerken, Applikations- und Indexservern sowie einer Message-Queue-Lösung betreut und erweitert werden. Diese Infrastruktur umfasst alle drei Umgebungen: Test, Stage und Produktion. Sie ist in der SBB-PK bereits im produktiven Einsatz und wird für die 2. Projektphase angepasst.

Die benötigte IT-Infrastruktur wird als Nutzung virtueller Server im größeren Rechnerverbund der SBB-PK realisiert. Dieser Aufbau ist kostengünstiger als eine separate Hardware-Infrastruktur, da er auf die optimale Nutzung vorhandener Rechnerkapazitäten hin ausgelegt ist. Die beantragten Mittel decken die der SBB hierfür tatsächlich entstehenden Kosten und liegen deutlich unter denjenigen für die Beschaffung von separater Hardware oder für die Anmietung der notwendigen Rechenleistung bei privaten Anbietern (s. Anlage 5).

Für den Betrieb des Gesamtsystems werden außerdem Plattform-Services genutzt und betreut. Dazu gehören: zentrales Logging (ELK-Stack), zentrales Monitoring (Nagios) und Service-Registry (Netflix-Eureka) mit Routing.

### Entwicklungsunterstützende Systeme (Erweiterung und Betreuung)

Als Voraussetzung für eine effiziente Softwareentwicklung in verteilten Teams wurde zentral bei der SBB-PK eine Entwicklungsinfrastruktur aufgebaut. Diese wird während der gesamten Dauer der 2. Phase des Projektes erweitert und betreut. Zu den entwicklungsunterstützenden Systemen gehören: Continuous Integration und Continuous Deployment mit Jenkins, SonarQube und Nexus, Versionsverwaltung mit Git/GitLab sowie das Ticketsystem Redmine.

## 2.3.2 Einzelmodule

### AP 7 – Normdatenmodul

Ziel des AP ist es, das Potential der im HSP bereitgestellten fachspezifischen Daten möglichst umfangreich nutzbar zu machen, und zwar sowohl innerhalb des Portals als auch für die Vernetzung mit externen Datenquellen. Dies wird erreicht durch ihre Einbindung in eine eigene Normdatenumgebung, in der sowohl schwerpunktmäßig GND-Daten als auch die in großem Umfang vorliegenden HSP-eigenen Normdaten zusammengefasst und gemeinsam verwaltet werden. Das bestehende Normdatenmodul soll in der 2. Phase daher zu einem Knowledge Graph ausgebaut werden. Es bildet damit die Grundlage eigenständiger Rechercheeinstiege in Nachweis und Präsentation sowie eine Plattform für die Integration von LOD-Szenarien und Analysetools. Auf diese Weise erhält das Normdatenmodul eine zentrale Bedeutung für die Unterstützung der Erfassung und Recherche und die Vernetzung mit anderen wissenschaftlichen Informationsangeboten. Um ein möglichst GND-konformes Datenangebot zu gewährleisten, besteht das Normdatenmodul des HSP künftig aus zwei miteinander kommunizierenden Services: dem fachlichen Knowledge Graph-Service und dem GND-Service, der die zentrale Aufgabe hat, normierte HSP-Daten in hoher Qualität über die GND zu führen (vgl. Anlage 6).

#### Knowledge Graph

Die Normdatenobjekte des HSP beziehen sich auf die fachlich relevanten Entitäten und Begriffe in den Beschreibungen und KODs. Sie werden in einer Graphdatenbank verwaltet, um Objekte und Beziehungen fachspezifisch modellieren und hinterlegen zu können. Aus dieser heraus werden sie Nachweis/Erfassung und Präsentation zur Verfügung gestellt.

#### Diese Graphdatenbank

- ermöglicht erstens eine HSP-spezifisch fachliche und potentiell mehrsprachige **Begrifflichkeit**, die von den in der GND vorhandenen Ansatzformen in begründeten Fällen abweichen und variante Bezeichnungen als Verweisungsformen hinterlegen kann. Auch die Referenzierung anderer Normdatenrepositorien wie aktuell schon GeoNames ist möglich.
- Die Normdatenobjekte können im Graphen zweitens durch individuell herstell- und qualifizierbare **Beziehungen** entitätenübergreifend in einen semantischen Zusammenhang gebracht werden. Auf diesem Weg können einfache Listen sowie mono- und polyhierarchische Konzeptstrukturen (Thesauri/Ontologien) modelliert werden. Derzeit finden hier die GND-Ontologie und schema.org Verwendung; bei Bedarf können weitere Ontologien wie etwa CIDOC-CRM hinzugezogen werden.
- Durch die Vernetzung dieser Normdaten mit den Kulturobjekten innerhalb der Graphdatenbank können drittens Beziehungen zwischen Personen, Orten, Körperschaften, Sprachen, Werken oder Konzepten (Sacherschlussbegriffen) und den Handschriften als eigentlichem Gegenstand des HSP sichtbar und zum Ausgangspunkt weiterer Recherchefunktionen gemacht werden (vgl. AP 13).

Die **Anlage und Bearbeitung** der Normdatenobjekte und ihrer Beziehungen im Knowledge Graphen können auf mehreren Wegen erfolgen. Bei einer manuellen Erfassung von Normdaten im Nachweis werden Nutzer:innen durch eine Visualisierung der Beziehungen in der Verortung des gesuchten Begriffs in einer Ontologie unterstützt (s. u. AP 8). Um größere Mengen von Normbegriffen ins System zu integrieren, kann eine automatische Anlage dieser Daten über einen Import initiiert werden.

Wenn im Zuge der Einzelerfassung einer Beschreibung ein Normbegriff nicht im Graphen gefunden wird, erfolgt zunächst eine Anfrage beim GND-Service und ggf. eine Übernahme aus der GND in den Graphen. Fällt auch die GND-Anfrage negativ aus, kann manuell ein neues Objekt angelegt und über den GND-Service an die GND übermittelt werden. In analoger Weise werden beim Import von Beschreibungen – speziell bei der Normdatenreferenzierung im Bereich der Datenredaktion – Analyseprozesse auf Normdaten durchgeführt und ggf. neue GND-Objekte in den Graphen integriert (Normdatenerkennungsalgorithmus). Auf diesem Weg wird der Datenbestand des Graphen sukzessive den konkreten Bedürfnissen der Nutzer:innen des HSP angepasst.

Die Normdaten des HSP werden an einer öffentlichen **Schnittstelle** der Graphen in gängigen Formaten auch als LOD zur Verfügung gestellt (z. B. HTML, RDF, JSON). Dort werden Nutzer:innen über SPARQL individuell definierte Datensets exportieren können.

Alle Objekte der Graphdatenbank sind permanent im Web adressierbar.

### GND-Service

Die Kommunikation des Graphen mit der GND wird durch den GND-Service sichergestellt. Er stellt den gesamten GND-Bestand in Form eines Dumps zur Verfügung, der automatisch und regelmäßig über die DNB-GND-API (via OAI-PMH) synchronisiert wird. Die Objekte des GND-Services werden bei Bedarf über die Kafka-Message-Queue des HSP in die Graphdatenbank vermittelt. Bei diesem Vorgang werden die GND-Attribute der Normdatenobjekte aktualisiert, mögliche HSP-eigene Attribute bleiben hingegen unverändert (Merging). Aus dem Graphen nimmt der GND-Service auf dem gleichen Weg neue und veränderte Objekte aus dem HSP entgegen. Die im HSP angelegten Normdaten können auf diese Weise der Aktualisierung und Ergänzung von GND-Daten dienen, was durch eine GND-Redaktion zu betreuen ist. Für diese Prozesse wird eine graphische Benutzeroberfläche im GND-Service konzipiert und implementiert.

### AP 8 – Nachweis, Erfassung, Veröffentlichung

Ziel des AP ist es, den Nachweis als die zentrale Verwaltung aller fachlichen Entitäten des HSP gemäß den neuen fachlichen Anforderungen der 2. Projektphase weiterzuentwickeln. Sowohl für bestehende Entitäten (Kulturobjekte, Beschreibungen, Kataloge) als auch für neu zu integrierende (Kulturobjektkomponenten, virtuelle Kulturobjekte, Annotationen, Projekte, Sammlungen und Lizenzen) werden die Funktionalitäten des Nachweises erweitert. Im Nachweis werden Personalisierungsoptionen in Form eines Dashboards geschaffen. Graphische Interfaces sind auch für die Eingabe- und Anzeigemöglichkeiten des Normdatenmoduls erforderlich (s. AP 7). Für die Erschließung von Handschriften zentral ist der voll funktionale Ausbau des als Basisversion vorliegenden Erfassungseditors, bei dessen Design flexible und differenzierte Erschließungsoptionen umgesetzt werden. Schließlich muss der Nachweis es ermöglichen, Änderungen bereits publizierter Beschreibungen in verlässlich adressierbaren ‚Auflagen‘ zu publizieren.

## Nachweis

Angemeldeten Erschließer:innen oder Redakteur:innen wird auf der Einstiegsseite des Nachweises ein personalisiertes **Dashboard** zur Verfügung gestellt, das den Zugriff auf relevante Informationen im HSP bietet. Zum Ende der 1. Phase bereits vorhandene Informationen („Mein Profil“, „Meine Beschreibungen“) werden rollenspezifisch um weitere Optionen ergänzt: zuletzt publizierte Beschreibungen, in der Präsentation mit Annotationen versehene Beschreibungen, zur Bearbeitung freigegebene Beschreibungen, statistische Angaben zu den eigenen Beschreibungen bzw. Projekten. Redakteur:innen haben zusätzlich Zugriff auf ausgewählte Informationen verschiedener Institutionen und auf die letzten Importe sowie offene Importvorgänge. Lokale Redakteur:innen erhalten einen Überblick über die Beschreibungen oder Projekte der eigenen Einrichtung. Den Leiter:innen der Handschriftenzentren und anderen Projektverantwortlichen steht so ein Werkzeug der Projektsteuerung zur Verfügung. Anhand von markierten Kulturobjekten werden Nutzer:innen eigene ToDo-Listen noch zu bearbeitender oder zu annotierender Kulturobjekte zusammenstellen können.

Für alle Entitäten des Nachweises und des Normdatenmoduls werden im Nachweis sortier- und filterbare Listenanzeigen sowie Detaildarstellungen implementiert. Das Design dieser Oberflächen wird im Sinne eines möglichst geschlossenen Erscheinungsbildes des HSP an das der Präsentation angeglichen und für die Anforderungen von PCs und Tablets optimiert (vgl. AP 3). Die Nachweis-Suche wird um neue Filter bzw. Facetten bedarfsgerecht erweitert.

Über das Kulturobjekt und die Beschreibung hinaus werden die folgenden neuen fachlichen Entitäten und die zugehörigen Funktionalitäten im Nachweis implementiert.

Mit der **Kulturobjektkomponente (KOK)** als Teil des Kulturobjekts wird es möglich, auch selbständige Teile einer Handschrift (Einbände, Faszikel, Fragmente) als eigene Entitäten anzulegen und miteinander in Beziehung zu setzen. Die Zugehörigkeit einer Kulturobjektkomponente zu einem Kulturobjekt muss zugewiesen, angezeigt und gelöscht werden können. Die Kerndaten der Kulturobjektkomponenten können künftig aus den Beschreibungskomponenten über Attributsreferenzen vermittelt werden.

Für Codices discissi und andere verteilt erhaltene Kulturobjekte ermöglicht das HSP künftig die Identifizierung und Adressierung dieser ehemaligen Objekte in Form von **virtuellen Kulturobjekten**, deren Existenz durch die Referenzierung mindestens zweier Kulturobjekte oder Kulturobjektkomponenten begründet ist.

Mit der neuen Möglichkeit für Nutzer:innen, in ihrem personalisierten kollaborativen Arbeitsplatz eigene Inhalte als **Annotationen** im Portal zu veröffentlichen, ergibt sich die Notwendigkeit, diese Entität mit ihren Referenzen im Nachweis zu speichern (vgl. AP 12). Publizierte Annotationen sind für alle Erschließer:innen sichtbar. Redakteur:innen können darüber hinaus gemeldete Missbrauchsfälle löschen (s. u. AP 12, Qualitätssicherung).

Für das effiziente Datenmanagement werden die drei neuen Entitäten Projekt, Sammlung und Lizenzen eingeführt, mit denen sich auch verbesserte Recherchemöglichkeiten in der Präsentation ergeben.

Um den organisatorischen Rahmen, in dem sich Handschriftenerschließung häufig bewegt, auch im HSP abbilden zu können, wird die Entität **Projekt** eingeführt. Ein solches Projekt kann zu einer oder mehreren beteiligten Institutionen und Personen referenzieren. Neue Beschreibungen können ei-

nem Projekt zugewiesen oder wieder aus diesem entfernt werden. Auch sind projektbezogene Statistiken möglich. Entsteht als Projektergebnis eine Katalogpublikation, so kann aus dem Projekt bzw. den Projekten auf das Katalogobjekt referenziert werden.

Da eine Handschrift jeweils unterschiedlichen Gruppen zugewiesen werden kann (z. B. Signaturensegment, Provenienzen, thematische Sachgruppe), wird die **Sammlung** im HSP zu einer Entität, mit der die betreffenden Kulturobjekte verknüpft werden können. Ein Kulturobjekt kann zu mehreren Sammlungen gehören. Einer Sammlung können ihrerseits Normdaten (Bsp.: Orte, Körperschaften, Personen) zugewiesen werden, diese können untereinander in qualifizierten Beziehungen stehen.

Die im HSP angebotenen Erschließungsinformationen unterliegen potentiell unterschiedlichen Lizenzierungsvorgaben. Um Aktionen auf Beschreibungen und Kataloge lizenzabhängig gestalten zu können, werden im Portal **Lizenzen** als eigene Entitäten verwaltet.

Bei der Anzeige der **Kulturobjekte** müssen die Beziehungen zu diesen neuen Entitäten berücksichtigt werden. Außerdem werden Kulturobjekte mit externen Angeboten verlinkbar, inklusive einer Typisierung dieser Links. Die Anzeige des Kulturobjekts im Nachweis kann als PDF-Dokument oder TEI-Datei ausgegeben werden.

Die Anzeige der Metadaten der **Beschreibungen** im Nachweis wird um die Angabe ihrer Version sowie ihrer publizierten Auflage erweitert. Außerdem werden ggf. das Projekt, in dem eine Beschreibung erstellt wurde, sowie Annotationen auf die Beschreibung angezeigt. Den Beschreibungen können typisierte Links zu externen Systemen beigegeben werden. Um die Bearbeitung einer Beschreibung durch Dritte zu ermöglichen (z. B. Normdatenverknüpfung), wird die Freigabe einer Beschreibung durch ‚Teilen‘ implementiert. Bei importierten Beschreibungen kann der Beschreibungstyp im Nachweis von ‚extern‘ zu ‚intern‘ verändert werden, um eine Bearbeitung im System zu ermöglichen. Beschreibungen können als PDF-Dokument oder TEI-Datei ausgegeben werden.

## Erfassung

Der vollfunktionale Ausbau der bisherigen Basisversion der Erfassung ermöglicht die schnelle Ablösung der MXML-Erfassung von MM.

Die Erfassung von Beschreibungen erfolgt über einen **Beschreibungseditor**. Grundlage der Weiterentwicklung ist dabei das HSP-ODD. Fachliche Anforderungen, HSP-ODD und technische Realisierung befinden sich dabei in kontinuierlicher agiler Entwicklung, wobei die Konsistenz der Daten jederzeit gewährleistet sein muss. Auf diese Dynamik muss der Beschreibungseditor flexibel reagieren und in einem permanenten Abstimmungsprozess jeweils angepasst werden.

Eine zentrale Aufgabe ist die **Einbindung des Normdatenmoduls** (Vorschlagslisten, Verknüpfung, Anzeige), um Textinformationen entsprechend auszeichnen zu können. Weiterhin ist vorgesehen, im Editor eine **semantische Formatierung** zu implementieren (z. B. die Kursivierung zur Auszeichnung von Textzitatens oder Petit-Schrift). Für verschiedene Beschreibungstypen (Texthandschrift, illuminierte Handschrift) werden entsprechende Templates angeboten. Für die Darstellung von **Sonderzeichen**, wie sie für die Transkription handschriftlicher Texte erforderlich sind, wird ein unicodefähiger Zeichensatz angeboten, aus dem eine spezifische Auswahl in einer virtuellen Tastatur oder als Short-Cuts möglich sein soll. Validierungsfehler werden in verständliche **Fehlermeldungen** übersetzt, die eine sinnvolle Korrektur ermöglichen. Für das effiziente Erstellen von Beschreibungen wird ein:e Nutzer:in je eigene **Beschreibungstemplate**s anlegen und verwalten können. Diese sollen als Liste angezeigt und bei Neuanlage einer Beschreibung aufgerufen werden können.

Zu den bereits definierten Beschreibungstypen für die Texterschließung und die Erschließung illuminierten Handschriften mit ihren zugeordneten Beschreibungskomponenten tritt nun neu die Möglichkeit, auch in sich unstrukturierte Beschreibungsinformationen (Kurzbeschreibungen, ältere Inventarisierung o. ä.), in Form einer **globalen Komponente** zu erfassen. Eine spätere Strukturierung dieser Texte gemäß dem Regelwerk des HSP bleibt möglich.

Im Sinne einer möglichst schnellen Nutzung der HSP-Erfassung durch datengenerierende Institutionen soll im ersten und dritten Projektjahr je ein Workshop zur speziellen Anwendungsschulung in Präsenz angeboten werden. Darüber hinaus wird ein kontinuierliches virtuelles Beratungs- und Anleitungsangebot realisiert.

### Veröffentlichung

Beschreibungen können in der Erfassung so lange verändert werden, bis sie in der HSP-Präsentation veröffentlicht werden. Bis dahin werden einzelne Arbeitsstände als ‚Versionen‘ gespeichert. Mit Veröffentlichung einer Beschreibung ist diese dann als fixe ‚Auflage‘ persistiert und zitierbar. Veröffentlichte Beschreibungen können im Nachweis nicht mehr inhaltlich bearbeitet oder gelöscht werden. Ergibt sich Änderungs- oder Ergänzungsbedarf, kann die Beschreibung als neue Auflage publiziert werden. Um die Entwicklung einer Beschreibung transparent zu machen und alle Referenzen auf unterschiedliche Auflagen einer Beschreibung dauerhaft funktional zu halten, werden in Nachweis und Präsentation jeweils auch alle älteren Auflagen einer Beschreibung bei Bedarf zur Anzeige gebracht.

Die Veröffentlichung von Erst- und Folgeauflagen wird also als bewusster Schritt mit einer eigenen Logik verstanden und umgesetzt. Hierfür kann ein:e Erschließender:in die Beschreibungen sammeln und bei einem/r lokale:n Redakteur:in anmelden. Der Status des aktuellen Publikationsprozesses wird für Erschließender:innen in ihrem Dashboard nachvollziehbar sein.

### AP 9 – Import

In der 1. Projektphase war der Import mit dem vordringlichen Ziel entwickelt worden, die Beschreibungsdaten aus MM ins HSP zu migrieren. Daher wurden prioritär die Formate MXML und TEI-XML umgesetzt. Um künftig auch den Import von Daten aus bibliothekarischen Systemen zu ermöglichen, wird in der 2. Phase der Import im Format MARC21 ermöglicht.

Um den automatisierten iterativen Import externer Daten realisieren zu können, wird außerdem ein Harvesting von OAI-PMH-Schnittstellen implementiert. Dabei wird der Import für den Nachweis der Handschriftendaten der ÖNB um deren Mehrstufigkeit anzupassen sein (s. AP 10).

### AP 10 – Nachweis externer Daten

Das Handschriftenportal versteht sich als Dienstleister für den Nachweis von Buchhandschriften in deutschen Sammlungen. Daher soll es auch Handschriften aus diesen Sammlungen nachweisen, deren Erschließungsinformationen nicht primär im Portal selbst gepflegt werden. Außerdem werden weitere wichtige nationale Webangebote des deutschsprachigen Raums auch über das HSP sichtbar gemacht.

Auf einer ersten Stufe ist das Ziel der **Nachweis der andernorts erschlossenen Kulturobjekte** und ihrer Erschließungsadressen im HSP. Dazu werden die im Partnersystem erschlossenen Hand-



schriften mit denjenigen des HSP abgeglichen und – sofern noch nicht vorhanden – mit ihren Minimaldaten als Kulturobjekte registriert. Vorhandene und neue Kulturobjekte werden um den Link ins betreffende Ursprungssystem ergänzt. Schließlich werden die referenzierten Digitalisate auf Dubletten im HSP hin geprüft, importiert und verknüpft. Als zweite Stufe über den reinen Nachweis externer **Erschließungsinformationen** hinaus ist die Anzeige dieser Informationen selbst im HSP vorgesehen.

Mit folgenden Betreibern bestehen folgende Absprachen:

Die **Österreichische Nationalbibliothek (ÖNB)** erschließt ihren bedeutenden Bestand an abendländischen Buchhandschriften über ihr Bibliothekssystem ALMA. Die ÖNB ist an einer möglichst umfassenden Präsentation dieser Daten auch über das HSP interessiert. Es geht um ca. 20.000 Datensätze mit etwa 2.000 Referenzen auf Digitalisate, die ab 2022 IIF-fähig werden. Über eine OAI-PMH-Schnittstelle stehen die Daten in MARC21 für eine regelmäßige Übernahme zur Verfügung. Für den Import ist eine Spezifizierung erforderlich, da etwa 5.500 dieser Dokumente mehrteilig sind und weitere 45.000 Einzeltexte umfassen. Dieses hierarchische Verhältnis muss nach dem Import im HSP in dessen Komponentenstruktur abgebildet werden.

Das Schweizer Portal **e-codices** befindet sich zum Zeitpunkt der Antragstellung hinsichtlich seines künftigen Betriebs in einer Umbruchsituation. Die e-codices nutzenden Bibliotheken, die den Fortbestand des Systems in einer von ihnen getragenen Infrastruktur wünschen, beabsichtigen, in Eigenleistung die Formate der beiden Systeme in Kooperation mit dem HSP zu mappen (e-codices-TEI – HSP-TEI) sowie mit Unterstützung des HSP probeweise Kulturobjekte zu registrieren und Beschreibungsimporte ins HSP anzustoßen. Damit soll ein regelmäßiger Datenimport ins HSP vorbereitet werden. In Austausch mit den Schweizer Partnern soll außerdem eine französischsprachige Nutzer:innenführung des HSP implementiert werden.

Der **Handschriftencensus (HSC)** bietet Informationen zu etwa 26.000 mittelalterlichen deutschsprachigen Handschriften und Fragmenten weltweit. Für das HSP werden schnittstellenbasiert die Informationen zu Kulturobjekten in deutschen, österreichischen und Schweizer Sammlungen berücksichtigt und dabei entweder neue KODs angelegt oder HSC-Informationen an bestehenden KODs als externe Beschreibungen verlinkt. Abhängig von der Schnittstelle des HSC wird auch eine Übernahme von Erschließungsinformationen angestrebt.

Ein analoges Vorgehen ist auch für das **Wasserzeichen-Informationssystem (WZIS)** geplant, für das derzeit ein DFG-Antrag zur Weiterentwicklung in Vorbereitung ist, der mit dem HSP abgestimmt wird.

Ein entsprechender Austausch ist ebenfalls mit **manuscripta.at**, dem zentralen Portal für mittelalterliche Handschriften in Österreich (z. Z. ca. 45.000 Kulturobjekte), geplant und in Gesprächen vorbereitet. Voraussetzung ist, dass manuscripta.at eine Schnittstelle einrichtet.

Mit weiteren Anbietern (**Archivportal-D, KdiH, Einbanddatenbank, RISM, Qalamos**) bestehen bereits Kontakte. Konkrete Planungen zum Datenaustausch werden sich aber erst im Lauf der 2. Projektphase ergeben. Wenn sich Bedarfe hierzu im Verlauf der 2. Projektphase konkretisieren, werden diese in einem Ergänzungsantrag aufgenommen.

## AP 11 – Export

Für Kulturobjekte und Beschreibungen des HSP wird den Nutzer:innen in Präsentation und Nachweis ein manuell angestoßener PDF-Download zur Verfügung gestellt (s. AP 8). Darüber hinaus

werden die Daten des Systems auch für einen maschinellen Export angeboten. An der OAI-Schnittstelle des Nachweises werden die Erschließungsdaten außer in TEI-XML in den Formaten MARC21 und Dublin Core zur Verfügung gestellt. Um den Bedarf nach gezielten maschinellen Datenabfragen (z. B. durch HSC oder WZIS) in diesen Formaten zu bedienen, wird außerdem eine SRU-Schnittstelle implementiert. An dieser werden unterschiedliche Abfrageparameter zur Verfügung gestellt. Konkrete Nutzeranforderungen werden zu Beginn der 2. Phase erhoben; weitere Lieferformate können im Projekt definiert werden. Die Objekte des Normdatenmoduls werden als JSON-LD und XML über eine SPARQL-API angeboten (s. AP 7).

## **AP 12 – Kollaborativer Arbeitsplatz**

In der 2. Phase soll der bestehende ‚Arbeitsplatz‘ der Präsentation von einem Anzeigetool für Rechercheergebnisse zu einer personalisierbaren kollaborativen Arbeitsumgebung erweitert werden, in dem Nutzer:innen Inhalte produzieren, zusammenstellen, austauschen und publizieren können. Waren bislang die angebotenen Inhalte fast ausschließlich von den Bibliotheken und Handschriftenzentren generiert, öffnet sich das HSP nun dem vielfach artikulierten Bedürfnis nach interaktiven Nutzungsmöglichkeiten (vgl. schon Müller/Haye 2011 und Magen/Schreiber 2015): Es wird ein Instrument des wissenschaftlichen Austausches und ermöglicht der Handschriftencommunity, das im HSP dokumentierte Wissen zu erweitern und zu aktualisieren.

Grundlage zur Nutzung dieser erweiterten Funktionalitäten ist ein persönlicher Account im HSP. Angemeldeten Nutzer:innen werden die erweiterten Funktionalitäten als zusätzliche Menüs bzw. Auswahlmöglichkeiten angeboten. Nutzer:innen können Inhalte in Form von Annotationen und Inhaltszusammenstellungen zunächst für sich selbst anlegen, sodann portal-intern mit anderen angemeldeten Nutzer:innen teilen und schließlich stabil referenzierbar veröffentlichen.

**Account:** Zur Nutzung der kollaborativen und personalisierten Funktionen (s. AP 13) von Präsentation und Suche sowie zur Erstellung von Mikropublikationen ist ein HSP-Account notwendig. Aus der Präsentation und Suche heraus sind eine An-/Abmelde-Funktion verfügbar und der Anmeldestatus ersichtlich. Benutzer:inneninformationen werden aus der Nutzer:innenverwaltung übernommen (s. AP 5).

**Erstellung von Inhalten:** Annotationen sollen zunächst der Fachdiskussion von Portalinhalten dienen, um beispielsweise eine bisherige Datierung begründet zu korrigieren. Dazu werden sie explizit auf einen bestimmten Abschnitt eines Beschreibungstextes bzw. einen Bildausschnitt eines IIIF-fähigen Digitalisats (Canvas-Bereich, auch externer Anbieter) bezogen, der im Mirador-Arbeitsplatz ausgewählt und dort kommentiert wird. Von der KOD-Übersichtsseite aus können außerdem Annotationen zu einer gesamten Handschrift angelegt werden. Um einen wissenschaftlichen Diskurs zu ermöglichen, können publizierte Annotationen wiederum von Nutzer:innen annotiert werden. Annotationen werden auf einem Arbeitsplatz-Server gespeichert und sind durch das dort implementierte Rechtemanagement zunächst nur für die erstellenden Nutzer:innen selbst sichtbar.

Neben Annotationen können Nutzer:innen auch Zusammenstellungen von Inhalten anlegen und speichern. Der eigenen Sammlung aller Arten im Portal unterstützter Daten dient die ‚**Kollektion**‘: Sie kann aus Trefferlisten heraus sowie aus dem Arbeitsplatz erstellt werden und umfasst Handschriften (KODs), Beschreibungen, IIIF-Digitalisate (auch externer Anbieter) sowie Mikropublikationen, Normdaten und Kataloge. Zusätzlich zu diesen Kollektionen in Art von Lesezeichenlisten kann über die Speicherung von ‚**Workspaces**‘ der persönliche Arbeitsplatz selbst in einem bestimmten

Arbeitsstand gesichert und geräte- und zeitunabhängig nachgenutzt werden. Der Arbeitsplatz unterstützt mehrere Workspaces, womit die parallele Arbeit zu mehreren Themen ermöglicht wird. Bei der Speicherung dieser Workspaces wird der Status der geöffneten Inhalte beibehalten (Fenster-Anordnung, ausgewählte Seiten). Sowohl Kollektionen als auch Workspaces können individuell benannt und damit eindeutig referenziert werden.

Der **Verwaltung der eigenen gespeicherten Inhalte** dient eine spezielle **Oberfläche** innerhalb der Präsentationsumgebung. Analog zum Design verschiedener Cloud-Dienste sind die Inhalte dort aufgelistet, können umbenannt, zum Editieren geöffnet oder gelöscht werden. Von dieser Oberfläche aus werden außerdem das Teilen von Inhalten mit weiteren Nutzer:innen sowie eigene Publikationen ermöglicht.

**Freigabe zur kollaborativen Arbeit:** Gespeicherte private Annotationen, Kollektionen und Workspaces können über die Inhaltsverwaltung für ein oder mehrere explizit anzugebende Nutzer:innen mit Lese- oder Schreibrechten freigegeben werden, um beispielsweise in der Lehre eingebunden zu werden. Für mit Schreibrechten freigegebene Inhalte wird eine Lösung zur Vermeidung von Updatekonflikten implementiert. Nutzer:innen finden für sie freigegebene Inhalte, von den eigenen abgegrenzt, in ihrer Inhaltsverwaltung. Die Freigabe erfolgt über das im Arbeitsplatz-Server implementierte Rechtemanagement.

### Veröffentlichung

Annotationen können in einem weiteren Schritt als dezidierte Publikation öffentlich sichtbar und stabil zitierfähig gemacht und in die Suche des HSP einbezogen werden. Der aus der Inhaltsverwaltung heraus erfolgende Veröffentlichungsprozess lässt die Angabe von Mitautor:innen zu, um eine eindeutige Urheberschaft zu gewährleisten. Zur Veröffentlichung wird die jeweilige Annotation mit ihren Metadaten an das Nachweismodul übergeben, dort unter Referenz auf den Account des/r Nutzer:in persistiert und für die langfristig stabile Zitierfähigkeit mit einem Identifier versehen. Für publizierte Annotationen mit Bezug auf ggf. nicht dauerhaft stabile Bildausschnitte wird ein Lösungskonzept entwickelt. Die persistierte Veröffentlichung wird vom Nachweis KOD-bezogen an die Suche der Präsentation übermittelt, wo sie recherchierbar und auf der Übersichtsseite zur Handschrift aufgeführt wird. Über die erfolgte Publikation erhält der/die Nutzer:in eine Meldung, die in seinem News-Bereich aufgeführt wird (s. u. Qualitätssicherung).

## Transkriptionsunterstützung

Eine spezielle Form der Annotation, die nach dem bislang eingegangenen Feedback aus der Forschung eigens zu berücksichtigen ist, ist die Transkription. Angesichts der zahlreichen handschriftlich bezeugten Texte im HSP-Datenbestand, für die weder Abdrucke noch moderne Editionen vorliegen, kommt einer Transkriptionsmöglichkeit besondere Bedeutung zu. Im Rahmen des [Transkribathons ‚Faithful Transcriptions‘](#) wurden im Frühjahr 2021 die Annotationsfunktionen von Mirador 3 um TEI-XML-Exporte erweitert, um Transkriptionen direkt im Arbeitsbereich des HSP vornehmen zu können. Aufbauend auf den Ergebnissen dieses Projekts sollen die Annotationsfunktionen speziell für Transkriptionen erweitert werden.

Im Arbeitsplatz soll es möglich sein, zeilenweise Annotationen des Typs ‚Transkription‘ zu erstellen. Diese können nachfolgend als Text-Overlay direkt auf dem Digitalisat angezeigt sowie bei ihrer Publikation seitenweise zusammengefasst berücksichtigt werden. Eine Eingabeunterstützung für Sonderzeichen wie auch Hilfen zur Auszeichnung textueller Besonderheiten wie Abkürzungen sind weitere Anforderungen ebenso wie ein Gesamt-Export der Transkriptionen zu einem Digitalisat in einem üblichen Format. Der Erarbeitung des genauen Konzepts zur Transkriptionsunterstützung soll ein ausführlicher Anforderungsworkshop mit Vertreter:innen der Community vorausgehen. Ziel ist eine schlanke, gut nutzbare Anwendung, die allerdings nicht die Reife dezidierter Transkriptionssoftware erreichen kann. Ihre weitere Ausbaufähigkeit innerhalb des Community-Projekts von Mirador ist dabei ein wichtiger Aspekt, um für eine möglichst breite Zielgruppe eine Grundlage für andere IIF-basierte Transkriptionstools zu bieten.

## Qualitätssicherung

Wenn das HSP Inhalte veröffentlicht, die von Nutzer:innen generiert werden, wird die Frage der Qualitätssicherung zentral. Bei allen Vorstellungen der HSP-Pläne für eine 2. Phase vor wissenschaftlichem Publikum wurde dieser Punkt besonders intensiv diskutiert. Es ist klar, dass die HSP-Redaktion angesichts der fachlichen Spannweite und potentiellen Anzahl möglicher Annotationen keine Einzelprüfung von publizierten Inhalten leisten kann. Tatsächlich ist davon auszugehen, dass die annotierenden Nutzer:innen in der Regel die eigentlichen Spezialist:innen für die jeweilige Handschrift oder das jeweilige Fachthema sind. Auch ist handschriftenbezogenes Wissen nicht fix, sondern hermeneutisch hergeleitet und entwickelt sich im Forschungsdiskurs.

Die Qualitätssicherung wird daher nicht in Form von fachlich-inhaltlicher Prüfung erfolgen, vielmehr gewährleistet ein System von Elementen, dass die von Nutzer:innen publizierten Inhalte Teil des wissenschaftlichen Diskurses sind: Die Registrierung mit Klarnamen über ein Authentifizierungssystem möglichst mit ORCID-ID und/oder GND-Referenz wird den Kreis der publikationsfähigen Nutzer:innen wissenschaftsaffin eingrenzen. Mit der Anmeldung ist die Annahme von Nutzungsbedingungen verbunden, sodass Beiträge mit Verstößen (wie persönliche Beleidigung etc.) von der Redaktion gelöscht oder Nutzer:innen ausgeschlossen werden können. Zu prüfen sein wird, ob als Unterstützung der Redaktion der Einsatz eines maschinellen Screenings neuer Publikationen auf bestimmte Schlüsselwörter möglich ist. Ein Meldebutton wird auch Nutzer:innen die Anzeige von Verstößen ermöglichen. Qualitätssicherung soll schließlich durch eine besonders gute Sichtbarkeit von publizierten Nutzer:inneninhalten erfolgen: In der Recherche wird ein Filter für Handschriften ‚mit Annotationen‘ eingerichtet, die aktuellsten externen Inhalte werden auf der HSP-Webseite in einer Timeline angezeigt. Nutzer:innen werden ihrerseits bei der Publikation die Möglichkeit haben,

eine zielgruppenspezifische Verbreitung in ihren Social-Media-Kanälen anzustoßen, was die Attraktivität für das Publizieren im HSP erhöht und sie öffentlich wahrnehmbar macht. Auf diese Weise nähert sich das HSP einem selbstgesteuerten Open-Peer-Review-Verfahren an.

### **AP 13 – Rechercheerweiterung und -optimierung**

Dank der deutlichen Verbesserung der Normdatenreferenzierung im Zuge der Datenredaktionsinitiative der 2. Förderphase (s. AP 14) können auch die komplexe Suche und die Facettierung differenziert und ausgebaut werden, was einem breiten Bedürfnis antwortet, das auch im [Card Sorting-Test](#) feststellbar war. Neben der bisher angebotenen handschriftenzentrierten Suche existieren aber in der Fachcommunity noch weitere Recherchebedürfnisse, insbesondere in Bezug auf Handschriftenkataloge und Normdaten. Hierzu sollen weitere direkte Sucheinstiege angeboten werden. Die umfangreichen im Normdatenservice vorhandenen Informationen sollen außerdem aus den Beschreibungen und KODs heraus sicht- und ansteuerbar gemacht werden.

Im Zuge der Erweiterung des Arbeitsplatzes (s. AP 12) wandelt sich das HSP zu einer Arbeitsumgebung, in der sich Nutzer:innen längere Zeit aufhalten. Entsprechend mehr Raum soll Personalisierungsmöglichkeiten gegeben werden, um Konfigurations- und Wiederherstellungsaufwand der Umgebung für angemeldete Nutzer:innen zu minimieren.

**Normdatenübersichtsseiten:** Die Daten des Normdatenservices werden in Beschreibungsvolltexten sowie in den ‚Basisdaten‘ des KOD referenziert. Diese Referenzen werden künftig auf HSP-eigene Normdatenübersichtsseiten verlinken, wobei Vorschaufenster in der Beschreibungsansicht angeboten werden. Die Übersichtsseiten bieten eine Zusammenstellung der im Normdatenservice verfügbaren Informationen sowie die Möglichkeit, eine Suche nach dem jeweiligen Normdatum im HSP-Bestand auszulösen.

Da die Normdaten-Übersichtsseiten die HTML-Repräsentation der im Normdatenmodul verwendeten stabilen URIs (s. AP 7) sind, werden externe Referenzen auf HSP-Normdaten universell auflösbar.

**Normdatensuche:** Der Sucheinstieg für Normdaten basiert auf einem Index der durch das Normdatenmodul bereitgestellten Daten, wobei nur von HSP-Daten referenzierte Normdaten indexiert werden. Die Normdaten werden in einer separaten Trefferliste dargestellt und sind durchsuch- sowie nach spezifischen Kriterien filterbar. Die einzelnen Trefferkarten führen zu der jeweiligen Normdaten-Übersichtsseite. Der Knowledge Graph (s. AP 7) ermöglicht darüber hinaus perspektivisch eine visuell unterstützte Recherche in den Normdaten selbst sowie Abfragen zum Zweck der Datenanalyse, was in einem Ergänzungsantrag berücksichtigt werden soll.

**Katalogsuche:** Analog zur Normdatensuche wird ein weiterer Sucheinstieg für die Recherche nach und in Handschriftenkatalogen implementiert. Hierfür werden die Volltexte der Kataloge, inkl. Beschreibungen, Paratexte und Register indexiert. Die Trefferliste ist katalogspezifisch sortier- und filterbar. Die Trefferkarten führen an die entsprechende Stelle im Digitalisat des Katalogs, das im Arbeitsplatz angezeigt wird.

**Normdatenbasierte Optimierung:** Die Recherche nach Orten und Sprachen wird durch die Nutzung des Normdatenservice optimiert. Informationen zu über- bzw. untergeordneten Konzepten in den Orts- sowie Sprachontologien werden für hierarchische Recherchen in den Handschriftendaten genutzt. Somit können bei der Suche nach ‚Deutschland‘ auch Handschriften mit dem hinterlegten Ort ‚Leipzig‘ gefunden werden, analog findet ‚Deutsch‘ auch Datensätze mit ‚Mittelhochdeutsch‘. Für

die Behandlung zeitlicher Veränderungen insbesondere im Bereich von Gebietskörperschaften müssen Lösungswege diskutiert werden.

### Visualisierungen

Durch visuelle Aufbereitung der HSP-Daten können direkt Zusammenhänge sichtbar gemacht und so neue Rechercheeinstiege geschaffen werden. Für die Daten zur Entstehungszeit sowie zu Entstehungs- und Besitzorten sollen Visualisierungen mittels Zeitstrahlen und Karten angeboten werden. Weiterhin sollen im Bereich der Filterung zum Einsatz kommen, indem etwa bei der Entstehungszeit die Verteilung der Handschriften als Diagramm eingeblendet wird und so direkt und intuitiv Rückschlüsse auf das Filterergebnis zulässt.

### Personalisierungen

Bei der wissenschaftlichen Arbeit an einem Thema ist es nötig, auf die Ergebnisse komplexer Suchanfragen zurückgreifen zu können. Innerhalb des persönlichen Arbeitsplatzes können angemeldete Nutzer:innen daher Suchanfragen in ihrem Account speichern, die beim erneuten Aufruf immer dynamisch auf den aktuellen Portaldateien ausgeführt werden und so auch Änderungen im Treffer-sample sichtbar machen. Sind statische Zusammenstellungen gewünscht, können gefundene Handschriften in Kollektionen gespeichert werden (s. AP 12). In Verbindung mit den neu eingeführten Speicher- und teilbaren Workspaces des kollaborativen Arbeitsplatzes stehen so umfangreiche Möglichkeiten zur Verfügung, Inhalte und deren Kombinationen schnell wieder für die Weiterarbeit bereitzustellen. Um auch die Konfiguration der Oberfläche über verschiedene Sitzungen und Geräte beizubehalten, sind von Nutzer:innen vorgenommene Einstellungen wie Sprache, Treffersortierung und kontrastoptimierte bzw. normale Darstellung ebenfalls im Account speicherbar.

### AP 14 – Redaktion

Ziel der Fortführung der Datenbearbeitung in der 2. Projektphase ist die Vereinheitlichung des Nutzungserlebnisses für den gesamten Datenbestand des HSP. Hierfür werden zum einen **Printkataloge** der jüngeren Vergangenheit ins System nachgeführt, die in der 1. Phase noch nicht berücksichtigt werden konnten. Zum anderen werden – als finaler Abschluss der Bearbeitung des bisherigen Datenbestands – die ins HSP migrierten Erschließungsdaten einem differenzierten Prozess der **Normdatenanalyse und -verknüpfung** unterzogen. Zugleich ermöglicht die Auswertung der vorhandenen Werte die Erstellung **kontrollierter Vokabulare** für die künftige Erfassung und Recherche. Die Redaktion beteiligt sich außerdem am AP 10 **Nachweis externer Daten** durch Abstimmung mit den Lieferanten, um die HSP-Konformität der eingehenden Daten zu gewährleisten.

Zu Mengengerüsten und Arbeitsschritten s. Anlage 7.

Die in der 1. Phase bewährte Organisation eines auf die vier Projekt-Bibliotheken verteilten Redaktionsteams soll beibehalten werden, so dass die bestehenden Kompetenzen und die etablierten Kommunikationsstrukturen weiterhin genutzt werden können. Da sich in der 1. Phase erwiesen hat, dass die Aufwände für Projektkoordinierung, Dokumentation, Fehleranalyse und Anforderungsanalyse für die IT-Unterstützung wesentlich aufwendiger waren als kalkuliert, wird dieser Teil des AP in der 2. Phase personell verstärkt.

### Ergänzende Katalogdigitalisierung und -bearbeitung

Das HSP soll den Nutzer:innen für alle Kulturobjekte, zu denen Handschriftenkataloge erschienen sind, ein möglichst homogenes Rechercheerlebnis bieten; aus diesem Grund sind ergänzende Digitalisierungen erforderlich (s. Anlage 8).

Im System fehlen bislang zum einen Digitalisate von **49 Handschriftenkatalogen**, die auf der Grundlage von genuin elektronischen MM-Beschreibungen erstellt wurden oder aktuell werden. Die Handschriftenkataloge enthalten zusätzlich die begleitenden Paratexte (Einleitung, individuell angepasste Register, zusammenfassende Gruppierung in kunsthistorischen Katalogen) sowie die Korrekturen aus der Drucklegung und haben daher forschungsrelevanten Mehrwert. Diese Kataloge werden digitalisiert, OCR-bearbeitet und für Recherche und Anzeige strukturiert. Recherchedaten liegen hier über die Primärerfassung in MM bereits im HSP vor.

Zusätzlich sollen **5 gedruckte Kataloge**, die seit Erstbeantragung des HSP nur im Druck erschienen sind und keine Datenentsprechung in MM haben bzw. nur mit Kürzestverweisen nachgewiesen sind, digitalisiert und OCR-bearbeitet werden. Bei diesen Bänden erfolgt zusätzlich zur Strukturierung der OCR-Volltexte die Auszeichnung und Homogenisierung der Kerndaten anhand der etablierten Workflows und Skripte aus der 1. Phase.

Ein Sonderfall sind **15 bisher nicht berücksichtigte Tafelbände** von Katalogen illuminiertes Handschriften. Hier sollen in den bereits vorliegenden Katalogimages und OCR-Volltexten die Abbildungen den jeweiligen KODs und den Beschreibungen des zugehörigen Katalogs zugeordnet werden.

Verbesserung der themenbezogenen Recherche durch Normdatenreferenzierung  
Mit der 2. Phase des HSP werden alle Handschriftenbeschreibungen in ihrem Volltext recherchierbar sein. Für eine differenzierte fachliche Recherche müssen – über die bereits in der 1. Projektphase homogenisierten Kerndaten des KOD hinaus – weitere besonders forschungsrelevante Werte in einem eigenen Arbeitsgang identifiziert und auf Normdaten referenziert werden. Diese Normierungsinitiative betrifft genannte Personen, Orte, Einrichtungen, Werke, Initien und Sachbegriffe, außerdem Handschriftenobjekte. Das HSP antwortet damit auf den Forschungsbedarf nach der Bereitstellung entsprechender Recherchedaten, wie er im Nutzungsfeedback immer wieder artikuliert wurde. So betrafen 21 von 59 Beiträge im Card Sorting-Test zu den HSP-Filtern fehlende Recherchemöglichkeiten zu Personen und Körperschaften.

Im AP muss zunächst eine einheitliche Ausgangsbasis für die Wertebearbeitung erreicht werden, bevor die zentrale Aufgabe der Normdatenreferenzierung durchgeführt wird.

**Nacherfassung von Katalogregistern:** In der 1. Projektphase wurden 78 Katalogbände mit knapp 20.200 Beschreibungen digitalisiert, die vor Projektbeginn noch nicht im Datenbestand von MM vertreten waren. Für diese Beschreibungen sind im HSP bislang zwar die Kerndaten normiert hinterlegt, nicht jedoch die in den strukturierten Initien- und Kreuzregistern enthaltenen Einträge ausgewertet, die recherherelevante Informationen in homogenisierter Form bieten. Gleiches wird für die 5 neu zu digitalisierenden Kataloge mit 1.877 Beschreibungen gelten. Aus früheren Retrokonversionsprojekten liegen noch 32 Kataloge mit insgesamt 24.168 Beschreibungen vor, deren Register ebenfalls nicht erfasst wurden. Insgesamt handelt es sich also um 115 Katalog-Register. In ihnen werden – unter Zuhilfenahme von Verfahren der Named-entity recognition (NER) – skriptbasiert Personen-, Orts- und Körperschaftsnamen, Initien sowie Werktitel identifiziert. Bei Sachbegriffen sind zusätzlich intellektuell-manuelle Markierungsverfahren nötig. Die Werte werden über einen Signaturenabgleich den zugehörigen Beschreibungen im HSP zugewiesen. Automatisierte Abläufe müssen intellektuell überprüft, korrigiert und ggf. ergänzt werden.



Weitere **6** OCR-bearbeitete **Katalogbände** der UBL werden über die HSP-Erfassung direkt ins System geführt und dort intellektuell auf Kern-, aber auch alle weiteren rechercherelevanten Daten hin ausgezeichnet.

Die Arbeiten für die zentrale Aufgabe der **Normdatenreferenzierung** erstrecken sich auf insgesamt 120.550 Beschreibungen (74.300 aus MXML, 24.150 aus Retrokonversionsprojekten, 20.200 aus 1. Phase, 1.900 aus 2. Phase). Sämtliche Referenzierungen erfolgen auf Entitäten im Graphen des Normdatenmoduls, wo neben den GND-Daten die eigenen fachlichen Konzepte des HSP verwaltet werden. Personen, Körperschaften, Geographika und literarische Werke werden teilweise über automatisierte, algorithmusgesteuerte Prozesse mit dem Normdatenmodul bzw. der GND abgeglichen (s. o. Knowledge Graph). Sind Werte mehrfach im System innerhalb von Beschreibungen präsent, erfolgt die Anlage eines neuen Normdatensatzes.

Bei deutschsprachigen Texten werden die Werktitel-Ansetzungen des Handschriftencensus für die Referenzierung genutzt. Für lateinische Werke werden die 21.417 in MM vorhandenen, repertori-ums-basierten Werknormdokumente algorithmusbasiert mit der GND abgeglichen.

Schließlich sollen Handschriften aus dem Datenbestand des HSP, die in Beschreibungen zitiert werden (,Handschriften, zitierte'), auf die IDs der entsprechenden KODs referenziert werden.

Für folgende fachlich einschlägige Sachbereiche werden Thesauri als Bezugspunkt der redaktionellen Referenzierung in breiter internationaler Abstimmung erstmals aufgebaut, wobei bestehende Lösungen (z. B. Thesauri zu Paläographie, Notation und Schreibsprachen der Projekte [Fragmentarium](#) und [Handschriftencensus](#)) nachgenutzt, erweitert und adaptiert werden. Eine Synchronisation mit der GND ist Teil der Arbeiten. Zugleich dienen die Thesauri der Normdatenverknüpfung in der künftigen Erfassung. Thesauri werden für die folgenden Themen entwickelt:

- europäische Schreibsprachen
- Schriften und Musiknotation
- Buchschmuck: Hier wurde bereits eine internationale Arbeitsgruppe gebildet.
- Buchkundliches, Material und Technik
- Textsorten/-gattungen, insbesondere für unfeste und nicht individualisierte Texte

Die Abstimmung mit den verschiedenen Fachcommunities soll auf insgesamt drei Workshops finalisiert werden: 1. Musikwissenschaft, 2. Kunstgeschichte, 3. Philologien.

#### Unterstützung des Nachweises externer Daten

Die Redaktion unterstützt Partner bei der Erstellung ihrer Kulturobjekte im HSP bei der Erfassung der hierfür erforderlichen Minimaldaten (Ort, Institution, Signatur). Hierzu gehört auch eine Kontrolle von Dubletten im Bereich der Kulturobjekte, die z. B. durch variante Ansetzungsregeln in verschiedenen Datenquellen entstehen können.

Sollen über Minimaldaten hinaus Kerndaten an das HSP geliefert werden, wie dies aktuell für die ÖNB geplant ist, wird die Redaktion beratend tätig, um eine möglichst HSP-konforme Bereitstellung der Daten für die KODs zu erreichen und die Nutzung der Werte für die Recherche-Filter des HSP zu ermöglichen. Ein Regelwerk für die Kerndatenbereitstellung auf Basis der bereits vorhandenen internen Anleitungen aus der 1. Projektphase wird erarbeitet.



## **AP 15 – KI-Begleitung**

Eine annähernd vollständige manuelle Qualifizierung der aus MM migrierten bzw. in der ersten Projektphase neu generierten Erschließungsdaten ist aufgrund ihres Umfangs ausgeschlossen (s. AP 14 und Anlage 7). Zusätzliche Herausforderungen bezüglich der Datenqualität sind mit Blick auf künftige Erfassung sowie speziell Importe aus externen Systemen absehbar. Um langfristig die Qualität vorhandener und neu ins System zu integrierender Daten zu verbessern, sollen daher in der zweiten Phase Einsatzmöglichkeiten für Verfahren aus dem Arbeitsgebiet der KI im HSP geprüft werden. Konkreter Ansatzpunkt für diese Begleitung der HSP-Entwicklung sind die Erfahrungen und Ergebnisse des BMBF-geförderten Projekts QURATOR, das ebenfalls an der SBB-PK angesiedelt ist. Als Ergebnis wird ein Konzept für die möglichen Verwendungsszenarien von NER (Named Entity Recognition)- und EL (Entity Linking)-Verfahren in den Arbeitsschritten Erfassung, Import und Normdatenverknüpfung angestrebt. Eine produktive Umsetzung dieser Szenarien im Datenbestand ist nicht Gegenstand des AP. Um die Ergebnisse dieses AP in einen potentiellen Ergänzungsantrag aufnehmen zu können, sollte das AP in den ersten 18 Monaten der zweiten Phase bearbeitet werden.

### **2.3.3 Auswirkungen**

#### **AP 16 – IIIF-Hosting**

Um Digitalisate direkt im Portal anzeigen und mit ihnen interaktiv arbeiten zu können, ist deren IIIF-Fähigkeit Grundvoraussetzung. Für Einrichtungen, die selbst keine IIIF-Server betreiben oder – wie die meisten Kleinsammlungen – über keinerlei eigenes Präsentationssystem verfügen, wird das HSP daher IIIF-Services im Bereich Beratung und Hosting anbieten. In der 1. Förderphase wurde von der UBL ein entsprechendes Konzept entwickelt, das sich an den hauseigenen Workflows ausrichtet und technische Voraussetzungen, Lizenzierungsbedingungen und ein Kooperations-Geschäftsmodell beschreibt (s. Anlage 1). Dieses Konzept wird in der 2. Phase zusammen mit der BSB weiterentwickelt. Ziel ist dabei, eine verteilte nationale Struktur von IIIF-Hosting Services im Kontext der sich wandelnden Digitalisierungslandschaft anzustoßen.

Angestrebt wird die Definition einer dezentralen Infrastruktur, in der über die UBL und die BSB hinaus weitere auf Handschriftendigitalisierung spezialisierte Einrichtungen oder andere Infrastruktureinrichtungen IIIF-Beratung und/oder Hosting von IIIF-Digitalisaten anbieten. UBL und BSB werden hierfür gemeinsam Bedarfe, Akteure und mögliche Abläufe der praktischen Umsetzung ermitteln, im Kontext des HSP erproben und abstimmen. Das von der UBL entwickelte Konzept wird dabei bezüglich seiner überregionalen Einsetzbarkeit geprüft, ggf. modifiziert und an die Geschäftsgänge der verschiedenen Partner angepasst.

Zu den komplexen rechtlichen Voraussetzungen für Urheber- bzw. Bildrechte bei Digitalisaten, die zudem in jüngster Zeit größeren Veränderungen unterworfen waren, wird einmalig eine externe juristische Expertise eingeholt werden, die dem HSP gegenüber den kooperierenden Einrichtungen, die IIIF-Digitalisate anbieten möchten, geprüfte Aussagen in Urheberrechts- und Lizenzfragen ermöglicht.

Innerhalb des AP wird die UBL zusammen mit Kleinsammlungen aus vergangenen und laufenden Kooperationsprojekten des Leipziger Handschriftenzentrums das eigene IIIF-Hosting anhand von ca. 370 Handschriften testen und in den Regelbetrieb überführen; ein weiteres Kleinsammlungsprojekt ist während der 2. HSP-Phase geplant und wird ebenfalls berücksichtigt. Ebenso wird die BSB

zusammen mit Partnern aus verschiedenen Kooperationsprojekten das IIF-Hosting von ca. 250 Handschriften für andere bayerische Einrichtungen testen und einen entsprechenden Geschäftsgang für die Datenübernahme und Bereitstellung konzipieren. Geplant ist dabei auch, im Laufe des Projekts die Inhalte der Kaiser-Heinrich-Bibliothek der Staatsbibliothek Bamberg mit weiteren 165 Handschriften IIF-konform bereitzustellen. Zusätzlich wird ein Consulting-Konzept für die Implementierung von IIF an anderen bayerischen Institutionen entwickelt, die selbstständig die IIF-APIs auf Basis der technischen Entwicklungen der BSB einsetzen wollen.

### **AP 17 – Richtlinien Handschriftenerschließung**

Das internationale Renommee der Handschriftenerschließung in Deutschland verdankt sich maßgeblich den DFG-Richtlinien Handschriftenkatalogisierung, die einen hohen, an wissenschaftlichen Bedürfnissen ausgerichteten Erschließungsstandard gewährleisten. Ziel dieses APs ist es, die derzeit gültige Fassung der [DFG-Richtlinien aus dem Jahr 1992](#) sowohl an die veränderten Rahmenbedingungen des Digital Age als auch an die permanente Fortentwicklung der Erschließungspraxis anzupassen, die sich an den aktuellen Forschungsbedarfen orientiert und zu der neben Änderungen beim Umgang mit thematisch-inhaltlichen Aspekten und zunehmender Integration von naturwissenschaftlichen Methoden auch eine gestaffelte Erschließungstiefe gehört. Die Notwendigkeit der Aktualisierung wurde u. a. auch 2016 im Zuge der Abschlussbegutachtung der DFG-Pilotphase Handschriftendigitalisierung in Hinblick auf ein künftiges Handschriftenportal festgestellt. Mit dem erreichten Entwicklungsstand des HSP-Projekts können die Richtlinien nun mit den Anforderungen an Erfassung und Präsentation von Erschließungs- und Bilddaten im zentralen Nachweis abgestimmt werden. Zudem sind die Vernetzung mit anderen digitalen Angeboten (Portale für spezielle Handschriftentypen, grundwissenschaftliche Spezialdatenbanken) und Bezüge zur künftigen Nationalen Forschungsdateninfrastruktur ([NFDI](#)) zu klären.

In einem ersten Schritt wird die bestehende Praxis der Handschriftenerschließung anhand des publizierten Materials (Handschriftenkataloge, rein elektronische Einzelbeschreibungen) analysiert und nach den verschiedenen Erschließungstiefen strukturiert dokumentiert. Anschließend erfolgt eine Durchsicht der Dokumentation durch profilierte Erschließungsakteure in Deutschland und eine Abstimmung mit ihnen (Handschriftenzentren, Einzelbibliotheken wie UB Heidelberg, handschriftenerschließende Forschungsprojekte wie HSC und KdiH), um Ergänzungen und Differenzierungen zu integrieren. Die überarbeitete Fassung bildet dann die Grundlage für einen Abgleich mit den Strukturen und Erfassungsanforderungen des HSP.

Ziel ist u. a. ein möglichst homogenes Nutzungserlebnis im HSP, indem dieselben Kategorien in derselben Abfolge erscheinen, auch wenn abweichende Abfolgen derzeit für bestimmte Handschriftentypen im Bereich der Katalogpublikation noch üblich sind (mittelalterliche, neuzeitliche, griechische, illuminierte, Musikhandschriften). Eine Angleichung der Beschreibungsstruktur wurde bereits für mittelalterliche und neuzeitliche Buchhandschriften auf einem DFG-geförderten Workshop der SLUB Dresden zur Erschließung neuzeitlicher Handschriften befürwortet.

Außerdem erfolgt eine Abstimmung mit anderen überregionalen Nachweissystemen (z. B. Verbundkataloge, RISM, perspektivisch Qalamos), um Zuständigkeiten und Vernetzungsmodi zu vereinbaren. Auch auf internationaler Ebene werden die bestehenden Vernetzungen z. B. mit dem TEI-Konsortium genutzt, um die Ergebnisse an internationale Formate rückzubinden.

Die erarbeitete Fassung der Richtlinien stimmen Erschließungsakteure und HSP-Vertreter:innen schließlich auf einer Klausursitzung ab. Die daraus hervorgehende Vorschlagsversion wird über die

[Website der Handschriftenzentren](#) publiziert und die Forschung über einschlägige digitale Kanäle um Feedback gebeten. Auf der Basis dieser Rückmeldungen wird eine finale Fassung erstellt und der DFG zur Begutachtung vorgelegt. Nach Abschluss dieses Prozesses erfolgt die Publikation, ebenfalls über die Website der Handschriftenzentren.

Die Arbeiten in diesem AP werden an den Handschriftenzentren Leipzig und München koordiniert.

### **3 Literaturverzeichnis**

[Müller, Stephan; Hays, Thomas: Mittelalter-Philologie im Internet](#). 38. Beitrag: Digitalisierung mittelalterlicher Handschriften aus Sicht der Forschung. In: Zeitschrift für deutsches Altertum und deutsche Literatur 140 (2011), S. 416–420.

[Magen, Antonie; Schreiber, Carolin: Erste Tagung zur Pilotphase Handschriftendigitalisierung](#) vom 9. bis 10. Oktober 2014 an der Bayerischen Staatsbibliothek in München – Ein Tagungsbericht. In: ZfBB 62 (2015) 1, S.59 –62.

## **4 Begleitinformationen zum Projektkontext**

### **4.1 Allgemeine ethische Aspekte**

Datenschutz- und Persönlichkeitsrechte werden beachtet. Bei Erhebungen im Rahmen der User Experience (AP 3) wird die Ethik-Kommission der Philosophischen Fakultät an der Humboldt-Universität zu Berlin einbezogen. Bei ethischen Fragen steht zudem Prof. Greifeneder allen Projektbeteiligten als Ansprechpartnerin zur Verfügung.

### **4.2 Maßnahmen zur Erfüllung der Förderbedingungen und Umgang mit den Projektergebnissen**

Das HSP ist dem Open Access verpflichtet. Erschießende Textdaten sowie Bilddaten werden unter der jeweils freiest möglichen Lizenz bereitgestellt. Für eine möglichst breite Nachnutzung der Projektergebnisse wird der produzierte Quellcode auf GitHub als Open Source veröffentlicht und dokumentiert. Für die HSP-basierte Handschriftenererschließung werden die Richtlinien Handschriftenkatalogisierung in AP 17 gemeinsam mit der Community weiterentwickelt und abschließend auf der Website der Handschriftenzentren publiziert.

### **4.3 Erklärungen zur Erfüllung der Förderbedingungen**

Alle Antragsteller:innen verpflichten sich zur vollständigen Erfüllung sämtlicher Förderbedingungen. Dies schließt insbesondere das Einbringen der erforderlichen Eigenleistung (vgl. unten 5.9) sowie ein Commitment zur nachhaltigen Betreuung des HSP (vgl. oben 1.1) ein.