

Bericht: AISOOP

Analyse der Informationsarchitektur, Schnittstellen und Organisation der Online-Plattformen

Rudolf Mumenthaler
Karsten Schuldt
Bruno Wenk

Eine Studie der HTW Chur
Gefördert im Rahmen des Programms
„Wissenschaftliche Information“ (SUK 2013-2016 P-2)
von swissuniversities

Chur, Februar 2017

Inhalt

Inhalt.....	2
Summary.....	3
1. Einleitung	4
2. Übersicht der untersuchten Online-Plattformen.....	5
3. Berichte zum Status Quo	5
3.1 Bericht zu technischen Fragen der Plattformen.....	6
3.1.1 Funktionsumfang der Plattformen.....	6
3.1.2 Schnittstellen.....	6
3.1.3 Metadaten	7
3.1.4 Open Data.....	8
3.1.5 Archivierung	9
3.2 Bericht zur Governance der Plattformen.....	9
3.3 Nutzung der Plattformen	10
3.3.1 Ungeklärt: Wie sollte die Nutzung erhoben werden?.....	10
3.3.2 Beispiele für Probleme bei der Erhebung der Nutzungszahlen.....	11
3.3.3 Nutzungsgruppen.....	11
3.4 Bericht zur Sicht der Personas auf die Plattformen	12
3.4.1 Beschreibung der Personas.....	12
3.4.2 Beschreibung der Plattformen aus Sicht der Personas	14
4. Effektiver Regelbetrieb bei den vorhandenen Plattformen, Potentiale bei der Dokumentation.....	16
4.1 Potentiale im Status Quo	17
4.2 Potentiale für vermutete Entwicklungen.....	18
4.3 Europeana.....	20
5. Fazit.....	21
Literatur	23
Anhang A: Umfrage.....	25

Summary

Im Projekt AISOOP (Analyse der Informationsarchitektur, Schnittstellen und Organisation der Online-Plattformen) wurden die in der Schweiz vorhandenen Plattformen zur Vermittlung von digitalen Medien mit nationaler Ausrichtung analysiert, um eine Empfehlung hinsichtlich ihrer Weiterentwicklung abzugeben. Auf der Grundlage von Interviews und einer Umfrage bei den Betreibern der Plattformen – hauptsächlich wissenschaftliche Bibliotheken – wurde der IST-Zustand erhoben und mit möglichen Anforderungen von Nutzerinnen und Nutzern (in Form von Personas) verglichen. Dabei wurde eine gewisse Diskrepanz zwischen den Ankündigungen bezüglich der Nutzung durch Forschende und den real an die Plattformen gestellten Anforderungen ihrer Nutzerinnen und Nutzer festgestellt. Eine Literaturanalyse zeigte, dass die aktuell diskutierten Potentiale von datengetriebener Wissenschaft und insbesondere der Digital Humanities in der Realität deutlich bescheidener und vor allem heterogener aussehen. Bisher wurden die Betreiber der Online-Plattformen in der Schweiz kaum von den angekündigten Forschenden der Digital Humanities kontaktiert. Deren Bedürfnisse sind gemäss bisheriger Untersuchung äusserst heterogen. Zudem besteht ein Teil ihrer Forschungsarbeit darin, eigene Anwendungen oder Plattformen zu entwickeln.

Die IST-Analyse der Plattformen ergab, dass diese in der Regel nicht nur für die Online-Publikation der Medien und Dokumente sorgen, sondern dass sie auch den gesamten Workflow von der Digitalisierung über die Erschliessung und Vermittlung bis zur Archivierung abdecken. Diese Workflows und auch die User-Interfaces sind auf den jeweiligen Dokumenttyp zugeschnitten. Die Plattformen sind primär auf den direkten Zugriff durch einzelne User ausgerichtet. In jüngster Zeit findet eine Entwicklung zu Open Data statt, indem die Inhalte wenn möglich unter offenen Lizenzen oder gemeinfrei angeboten werden. Meistens ist eine OAI-PMH-Schnittstelle eingerichtet, welche den Bezug der Metadaten (im Dublin Core-Format) ermöglicht. Mit weitergehenden Anforderungen wurden die Plattformen bisher nicht oder nur sehr selten konfrontiert. Dies hat sich bislang auch mit den im Programm „Wissenschaftliche Information“¹ neu geförderten Projekt NIE-INE und mit dem Projekt DaSCH (Data and Service Center for the Humanities) der SAGW nicht geändert.² Die Bereitschaft zu einer Weiterentwicklung ist grundsätzlich vorhanden, wie auch die Entwicklungen seit Einreichung der ersten Fassung dieses Berichts (Oktober 2016) bis zur Publikation der definitiven Fassung (Februar 2017) verdeutlichen. Ein grosser Vorteil der aktuellen Situation besteht darin, dass hinter den Plattformen Bibliotheken stehen, die diese aus eigenem Interesse zur Vermittlung und Archivierung ihrer eigenen Bestände nachhaltig gesichert betreiben.

1 Das Programm „Wissenschaftliche Information: Zugang, Verarbeitung und Speicherung“ wird von swissuniversities durchgeführt. Nach einer ersten Förderperiode (SUK 2013-2016 P-2) läuft es 2017-2020 unter gleichem Titel und der Nummer „P-5“ weiter. Das Programm vergibt projektgebundene Beiträge des Bundes (<http://www.swissuniversities.ch/isci>).

2 Es ist nach den Erfahrungen aus Anfragen von Forschenden, die in den Interviews im Rahmen dieser Studie geführt wurden, auch nicht zu erwarten, dass dies passiert. Das DaSCH geht mit der Archivierung der Forschungsdaten ein anderes Problem als die hier untersuchten Plattformen an. Obgleich es z.B. in der „Strategie Digitalisierung“ des Programms „Wissenschaftliche Information“ den Wunsch zu geben scheint, beides in flexiblen Plattformen zu verbinden, ist dies nichts, was bislang aus der Forschung als reale Anforderung formuliert wird. Während der Interview-Phase dieser Studie beschloss das Programm die Förderung des Projektes NIE-INE (Nationale Infrastruktur für Editionen – Infrastructure nationale pour les éditions), und der SNF bewilligte die Förderung von 23 Editionsprojekten

([http://www.snf.ch/SiteCollectionDocuments/Editions_Approved_projects_of_the_first_call_\(September_2016\).pdf](http://www.snf.ch/SiteCollectionDocuments/Editions_Approved_projects_of_the_first_call_(September_2016).pdf)). Alle diese Vorhaben können in Zukunft einen Einfluss auf die Arbeit der in dieser Studie untersuchten Plattformen haben. Im Rahmen der Interviews wurde aber bislang nie auf diese eingegangen. Wir haben die möglichen Effekte dieser neuen Angebote bei den Personas weiter unten berücksichtigt.

Als Empfehlung sieht die Projektgruppe deshalb vor, die Weiterentwicklung der Plattformen in Richtung offener Schnittstellen und Standardisierung zu fördern (z.B. Implementierung von IIIF). Denkbar wäre es, dass die Plattformen das Europeana Data Model übernehmen, wodurch die entsprechenden APIs für verschiedene Nutzergruppen für neue Formen der Nutzung eingesetzt werden könnten. Dies würde auch zu erhöhter Sichtbarkeit der Plattformen und der Bestände führen. Der Anstoss für die Weiterentwicklung sollte von den jeweiligen Plattformen kommen, die sich um Fördermittel bewerben können. Für audiovisuelle Medien gibt es noch keine vergleichbare national ausgerichtete Online-Plattform. Gerade für die Zugänglichkeit und Nutzung von Inhalten aus kleineren GLAM-Institutionen wäre eine offene Plattform sinnvoll. Ein entsprechendes Projekt könnte durch das Programm ausgeschrieben werden.

1. Einleitung

Das Projekt AISOOP (Analyse der Informationsarchitektur, Schnittstellen und Organisation der Online-Plattformen) hatte zum Ziel, die vorhandenen Plattformen, welche hauptsächlich von Wissenschaftlichen Bibliotheken in der Schweiz kollaborativ betrieben werden, daraufhin zu analysieren, ob sie für zukünftige Nutzungen gut aufgestellt sind. Ziel war es, herauszuarbeiten, wie diese Plattformen in die zukünftige Forschungsdateninfrastruktur der Schweiz integriert werden können. Dabei sollte benannt werden, was beim heutigen Status Quo schon möglich ist und was das Potential hat, weiterentwickelt zu werden. Ebenfalls sollten Lücken und Hindernisse für die Nutzung der Plattformen und deren Inhalte durch Forschende aufgezeigt werden.

Dabei galt es in diesem Projekt, zwei widersprüchliche Vermutungen zu überprüfen. Sowohl das White Paper als auch die „Strategie Digitalisierung“ des Programmes“ Wissenschaftliche Information: Zugang, Verarbeitung und Speicherung“ (SUK 2013-2016 P-2) beschreiben die Plattformen als „Datensilos“ und vermuten, dass sie es in ihrer jetzigen Form nicht ermöglichen, die in ihnen enthaltenen Daten und Dokumente einfach und in grossen Mengen zur Verfügung zu stellen sowie ausreichend recherchierbar und nutzbar zu machen. Diese Vermutung geht also davon aus, dass die Plattformen in der jetzigen Situation nicht für die zukünftige wissenschaftliche Nutzung – die als datengetrieben beschrieben wird – aufgestellt seien. Hingegen gibt es aus dem Feld der Bibliotheken der Schweiz die Vermutung, dass die Plattformen grundsätzlich gut aufgestellt seien, dass sie die in ihnen enthaltenen Daten bestmöglich verwalten und präsentieren sowie auf alle vorhandenen Anforderungen eingehen würden.

Im Projekt AISOOP wurde dieses Ziel mit einer kombinierten Methodik angegangen. Es wurden Personas formuliert, die Einrichtungen oder Personen repräsentieren, welche die Plattformen in unterschiedlicher Form ansprechen oder nutzen könnten. Aus der Sichtweise dieser Personas wurden Anforderungen an die Plattformen formuliert. Die Personas basierten auf der vorhandenen Literatur sowie Erfahrungen der Forschenden. In einer Reihe von Interviews und einer Umfrage unter Verantwortlichen für die Plattformen wurde die technische Infrastruktur, die Governance der Plattformen sowie deren Wahrnehmung der tatsächlichen Herausforderungen erfragt. Gleichzeitig wurde in der Literatur zur tatsächlichen Praxis datengetriebener Wissenschaft, insbesondere den Digital Humanities, nach tatsächlichen Anforderungen an Datensammlungen und der tatsächlichen Forschungspraxis recherchiert.

Dieser Bericht stellt die Ergebnisse der Studie dar und ist wie folgt aufgebaut. Das anschliessende Kapitel (2) stellt die untersuchten Online-Plattformen dar, darauf folgt ein Kapitel (3), welches deren Status Quo differenzierter aufzeigt. Hier findet sich auch die Beschreibung der Plattformen aus Sicht der Personas (3.4). Anschliessend (4) wird beschrieben, was die Plattformen heute leisten und welche Potentiale für eine zukünftige Entwicklung realistisch vorhanden sind. Im letzten Kapitel (5) wird ein Fazit gezogen und Empfehlungen für die weitere Arbeit der Plattformen gemacht.

2. Übersicht der untersuchten Online-Plattformen

Auswahlkriterium für die untersuchten Plattformen war, dass sie für die Wissenschaft und Öffentlichkeit der gesamten Schweiz digitale Sammlungen von Medien anbieten und auch in der Schweiz betrieben werden. Die im Projekt AISOOP untersuchten Online-Plattformen entstanden vorrangig im Rahmen des Projektes "e-lib.ch: Elektronische Bibliothek Schweiz", welches von 2008 bis 2011, finanziert von der damaligen Schweizerische Universitätskonferenz (SUK), dem ETH-Rat und dem Bundesamt für Berufsbildung und Technologie (BBT), die Zusammenarbeit der Wissenschaftlichen Bibliotheken in der Schweiz fördern sollte. Alle Plattformen bieten online die Sammlung spezifischer digitalisierter Dokumente und dazugehöriger Metadaten mit der Perspektive der Langzeitarchivierung an. Die Dokumente stammen aus den Sammlungen schweizerischer Bibliotheken und anderer Einrichtungen der Schweiz. Ebenso werden alle Plattformen kollaborativ betrieben, wobei es immer führende Einrichtungen gibt, bei denen zumeist die konkrete Infrastruktur (Server, Personal etc.) der Plattformen angesiedelt ist und welche den inhaltlichen Lead übernehmen, sowie beigeordnete Einrichtungen, welche den führenden Einrichtungen zuarbeiten, zum Beispiel indem sie Sammlungen digitalisieren, zur Verfügung stellen und finanzielle Beiträge leisten.

Die folgende Tabelle gibt eine Übersicht zu diesen Plattformen.

Name der Plattform	Fokus der Plattform	Leitende Einrichtungen
e-codices.ch	Vorrangig mittelalterliche Handschriften aus schweizerischen Bibliotheken, Archiven, Museen und privaten Sammlungen	Universität Freiburg
e-manuscripta.ch	Handschriften aus schweizerischen Bibliotheken und Archiven	Zentralbibliothek Zürich, beteiligte Institutionen: ETH-Bibliothek Zürich, Universitätsbibliothek Basel
e-rara.ch	Drucke des 15. bis 20. Jahrhunderts	ETH-Bibliothek Zürich, beteiligte Institutionen: ZB Zürich, UB Basel, UB Bern, Bibliothèque de Genève u.a.
E-periodica.ch	Historische und laufende schweizerische Zeitschriften	ETH-Bibliothek Zürich, beteiligt: Schweizerische Nationalbibliothek
Memobase	Portal für audiovisuelles schweizerisches Kulturgut und Helvetica	Memoriav
Scriptorium	Zeitungsplattform für den Erhalt des Kulturgutes des Kantons Waadt	BCU Lausanne
Schweizer Presse Online	Historische schweizerische Zeitungen	Schweizerische Nationalbibliothek, beteiligt: Kantonsbibliotheken

3. Berichte zum Status Quo

Die untersuchten Plattformen haben alle einen gut funktionierenden Betrieb aufgebaut, welcher die im Normalbetrieb anfallenden Aufgaben erfüllt und jeweils auf einen langfristigen Betrieb hin ausgerichtet

ist. Bis auf e-codices.ch sind alle Plattformen, durch ihre Anbindung an Bibliotheken oder Förderung durch den Bund oder Kantone, langfristig finanziell und infrastrukturell abgesichert. Sie sind der Projektphase entwachsen, reagieren aber auf weitere Anforderungen. In diesem Kapitel werden die technischen Prozeduren der Plattformen (3.1) sowie deren Governance (3.2) beschrieben. Auffällig wurde im Laufe des Projektes, dass die technischen Fragen nicht zu vernachlässigen sind, aber dass die meiste Arbeit in der Governance geleistet werden muss. Anschliessend werden die Plattformen aus Sicht verschiedener Personas – welche Einrichtungen und Einzelpersonen aus der Schweiz repräsentieren – beschrieben. (3.3)

3.1 Bericht zu technischen Fragen der Plattformen

3.1.1 Funktionsumfang der Plattformen

Die Plattformen dienen in der Regel nicht nur dem Zugang zu den digitalisierten Inhalten für Nutzerinnen und Nutzer, sondern bieten auch eine Unterstützung des gesamten Workflows von der Digitalisierung über die Strukturierung bzw. Beschreibung der Metadaten bis hin zur Präsentation im Web und der Langzeitarchivierung. Die Plattformen sind aus dem konkreten Bedarf der beteiligten Institutionen in Bezug auf die jeweils bearbeiteten Dokumenttypen hin entwickelt worden und dementsprechend mit ihren Funktionen auf diese spezifischen Anforderungen zugeschnitten. So unterscheiden sich zum Beispiel der Grad der Strukturierung wie auch die verwendeten Metadatenschemata je nachdem ob Bücher (z.B. in e-rara), Zeitungen (Scriptorium), Zeitschriften (wie in E-periodica), Manuskripte (e-manuscripta, e-codices) oder audiovisuelle Medien (Memobase) erschlossen werden.

Die meisten Plattformen werden von kommerziellen Anbietern entwickelt (Olive, Semantics/Visual Library, Satzrechenzentrum/Agora, Fedora/Memobase, Geneza/MediaInfo). Die Ausnahme bildet e-codices mit einer Eigenentwicklung. E-codices hatte ursprünglich die Software-Lösung kleio (Universität Köln) übernommen, dann aber 2008 eine eigene Anwendung auf der Basis von Open Source Technologien entwickelt. Bei der Weiterentwicklung sind die Einrichtungen also abhängig davon, dass die Anbieter die Anforderungen aufnehmen und umsetzen. Es gäbe auch offene Systeme auf dem Markt, doch haben sich die Plattformen im Rahmen von Systemevaluationen aus jeweils für die Institution überzeugenden Gründen für die genannten Systeme entschieden.

Name der Plattform	Technische Plattform
e-codices.ch	Eigenentwicklung
e-manuscripta.ch	Visual Library (Semantics)
e-rara.ch	Visual Library (Semantics)
E-periodica.ch	Agora (Satz-Rechenzentrum)
Memobase	Fedora
Scriptorium	MediaInfo (Geneza)
Schweizer Presse Online	Olive (Olive Software)

3.1.2 Schnittstellen

Grundsätzlich verfügen alle Plattformen über die Möglichkeit, Schnittstellen – sowohl offene als auch spezifisch solche zu Bibliothekssystemen – anzubieten. Nicht bei allen ist dies bislang umgesetzt; dies aber vor allem aus mangelnder Nachfrage, nicht wegen technischen Schwierigkeiten.

Nach Möglichkeit werden die bibliographischen Metadaten über eine entsprechende Schnittstelle aus dem Bibliothekssystem übernommen. Bei e-rara und e-manuscripta erfolgt die Übernahme mittels einer SRU-Schnittstelle (Nachfolger von Z39.50). Diese Schnittstellen werden momentan vorwiegend dazu verwendet, in einem externen System (z.B. dem Bibliothekssystem) vorhandene Metadaten zu übernehmen, um doppelte Aufnahmen zu verhindern und um den Aufwand bei der Erfassung zu reduzieren. Bei e-manuscripta gibt es auch Schnittstellen zu den Archivsoftwares CMiStar und Scope.

Die Schnittstellen sind nach aussen hin nur zum Teil offen. E-rara, e-manuscripta, E-periodica und e-codices bieten zum Beispiel eine OAI-PMH-Schnittstelle zum Bezug der Metadaten und Links zu den Digitalisaten an. Die Inhalte von e-rara sind z.B. auch in Gallica nachgewiesen und recherchierbar. Die Inhalte von e-codices sind auf Gallica, Europeana, Manuscripta Mediaevalia, recherche Isidore, Manuscriptorium u.a. zu finden. Zusätzlich werden Schnittstellen für spezifische Anwendungen spezifiziert. So werden die Inhalte von E-periodica (also die einzelnen, in XML ausgezeichneten Zeitschriftenartikel) via OAI-PMH geharvestet, über eine sogenannte Pipe in den Index des Discovery-Tools Primo übernommen und sind so im NEBIS oder im Wissensportal der ETH-Bibliothek suchbar. Ein Interesse von Anwenderinnen und Anwendern an der Nutzung dieser offenen Schnittstellen wurde bisher nur sehr selten geäussert.

Einige Plattformen planen die Öffnung hin zur Europeana – e-codices kooperiert seit 2010 –, was auch die Übernahme des Europeana Data Model bedingen würde (Scriptorium). (Siehe 4.3)

Bei Zeitschriften, AV-Medien und Zeitungen (Ausgaben ab dem 20. Jahrhundert) schränken die Urheberrechte die Nutzung ein. Entsprechend sind offene Schnittstellen nicht unbedingt erwünscht. Bestimmte, urheberrechtlich geschützte Werke können aus diesem Grund in der Memobase+ nur an dedizierten PCs in bestimmten Institutionen genutzt werden. Beim Projekt letempsarchives.ch der EPFL wurden die Daten zwar als Linked Open Data aufbereitet, aber für die Nutzung muss auch hier die Zustimmung des Verlags eingeholt werden.

Relativ neu ist die Thematik offener Schnittstellen für den Zugriff auf Digitalisate. Hierfür bietet IIIF (International Image Interoperability Framework) einen viel versprechenden Lösungsansatz, momentan für Bilddateien. E-Codices ist bei dieser Initiative federführend dabei. Bei e-rara und e-manuscripta zieht man die Implementierung im Rahmen der Weiterentwicklung in Betracht.

Bei den Abklärungen mit den Betreibern der Plattformen wurde auch klar, dass die Bibliotheken bislang (oder bis vor kurzem) die digitalen Objekte nicht als Forschungsdaten im eigentlichen Sinn betrachtet haben.³ Aus Sicht der Bibliotheken hat man eigene Bestände digitalisiert, online publiziert und archiviert, um sie langfristig einer breiten Nutzung verfügbar zu machen. Forschende als Nutzende wurden natürlich als wichtige Zielgruppe betrachtet. Jedoch gelten (oder galten) die Digitalisate als eine Form von Medien und Sekundärliteratur, wie sie von den Bibliotheken zur Verfügung gestellt werden. Es ist eine neue Entwicklung, die vor allem aus dem Bereich der Digital Humanities getrieben wird, wonach diese Inhalte als primäre Quellen und Rohdaten für die Forschung genutzt werden können. Wir kommen bei der Diskussion um Open Data auf diesen Aspekt zurück.

Plattform	Schnittstellen
E-rara	OAI-PMH mit Metadaten in Dublin Core, METS, MODS oder MARCXML
E-manuscripta	OAI-PMH mit Metadaten in Dublin Core, METS, MODS oder MARCXML
E-codices	IIIF, TEI-P5, OAI-PMH-Schnittstelle mit Metadaten in Dublin Core
E-periodica	OAI-PMH mit Metadaten in Dublin Core, OpenURL

3.1.3 Metadaten

Bibliothekarische Standards auf der Ebene der Beschreibung der einzelnen Dokumente werden in der Regel eingehalten. Bei der Strukturierung der Inhalte (z.B. von Zeitungen und Zeitschriften) orientierte man sich weniger an Standards, sondern eher an den Strukturen und wissenschaftlichen Fragestellungen an die jeweiligen Medien. So wurden auch für die spezifischen Nutzerinnen und Nutzer zugeschnittene Methoden eingesetzt. In E-periodica werden die Inhalte in einer an METS angelehnten XML-Struktur erfasst, ohne dass der Standard explizit unterstützt wird. In e-rara und e-manuscripta werden MODS, MARC-XML und METS-XML unterstützt. E-codices wendet IIIF an, die Beschreibungen werden in TEI kodiert. Bei Schweizer Presse Online werden die Metadaten in XML strukturiert.

³ Vergleichbar werden Forschungsdaten durch das Projekt r3data definiert: „Research data means information objects generated by scholarly projects for example through experiments, measurements, surveys or interviews.“ (<http://www.re3data.org>)

Eine mögliche Entwicklung geht in Richtung METS/ALTO, damit die Metadaten aus der Europeana aufgenommen werden könnten. Scriptorium unterstützt bereits XML METS/ALTO.

Der Ausbau und die Öffnung in Richtung Linked Open Data steht zumeist – ausser bei der Memobase – aktuell nicht zur Diskussion, da hier eine entsprechende Nachfrage durch die Nutzerinnen und Nutzer fehlt. Dieses Thema wird im Bibliotheksbereich der Schweiz vom Projekt linked.swissbib angesprochen.

3.1.4 Open Data

Für die Nutzung durch die Wissenschaft ist die Form der Bereitstellung der Inhalte ein entscheidender Faktor. Klare Lizenzbezeichnungen, möglichst als offene Lizenzen (Creative Commons oder Public Domain) bieten den Nutzerinnen und Nutzern die Gewissheit, in welcher Form die Daten genutzt werden können. Open Data wird auch von den Betreibern von Plattformen als wichtige Voraussetzung bewertet. Aktuell halten sich jedoch noch immer einschränkende Nutzungsbedingungen auf den Plattformen, welche für Forschende ein Hemmnis darstellen.

Bezüglich der Umsetzung von Open Data gestaltet sich ein heterogenes Bild. Nicht alle Plattformen stellen die Digitalisate von gemeinfreien Werken unter freien Lizenzen zur Verfügung. Derzeit findet aber offenbar ein Sinneswandel statt und die Zeichen stehen auf einer weiteren Öffnung der Plattformen. Auch wenn noch kaum Open Data Policies bestehen, so ist ein Bewusstsein für die Thematik vorhanden und man arbeitet an entsprechenden Strategien. Hindernisse in der Umsetzung werden vor allem in fehlenden Ressourcen und mangelndem technischen Know-how gesehen, wie eine Studie im Rahmen einer Bachelorarbeit an der HTW Chur gezeigt hat (Humbel 2016).

ETH-Bibliothek: Die Umstellung weitgehend vollzogen hat die ETH-Bibliothek, die sowohl ihr Bildmaterial in E-Pics wie auch andere Digitalisate "möglichst offen" anbietet und entsprechend kennzeichnet. So werden alle älteren Werke, bei denen das Copyright abgelaufen ist, als Public Domain gekennzeichnet. Inhalte, bei denen die Bibliothek das Urheberrecht besitzt, werden als CC-BY veröffentlicht.

E-rara: Die Digitalisate wurden bis zum Zeitpunkt der Untersuchung nicht mit CC-Lizenzen gekennzeichnet. Gemäss den Nutzungsbedingungen standen die Dokumente „für nicht-kommerzielle Zwecke in Lehre und Forschung sowie für die private Nutzung frei zur Verfügung. Einzelne Dateien oder Ausdrucke aus diesem Angebot können zusammen mit diesen Nutzungsbedingungen und den korrekten Herkunftsbezeichnungen weitergegeben werden“. Zum Download stehen die PDF-Dokumente zur Verfügung, höhere Auflösungen im Format TIFF können bei der jeweiligen Bibliothek bezogen werden, wobei hier Gebühren anfallen.

Nachtrag: seit anfangs 2017 wird in e-rara die angekündigte Open Data-Strategie umgesetzt: Neu ist bei allen Dokumenten die Lizenzierungsart angegeben. Wo möglich werden die Lizenzen Public Domain Mark 1.0 oder CC BY-SA 4.0 vergeben. Weitere Informationen finden sich in den Nutzungsbedingungen (<http://www.e-rara.ch/wiki/termsOfUse>).

E-Manuscripta: Hier sind die Bedingungen vergleichbar mit E-Rara: *Seit anfangs 2017 wird eine Open Data-Strategie angewandt (<http://www.e-manuscripta.ch/wiki/termsOfUse>).*

E-Codices: Die digitalisierten Handschriften sind entweder unter Public Domain (z.B. Uni Basel) frei verfügbar oder unter der Lizenz CC-BY-NC zur nichtkommerziellen Nutzung lizenziert. Im zweiten Fall muss bei einer Verwendung für kommerzielle Zwecke die Zustimmung der jeweiligen Bibliothek eingeholt werden. Die Entscheidung zur Lizenz treffen die Einrichtungen selber, nicht e-codices. Aus Sicht der Open Data Community sollten diese urheberrechtsfreien Werke als Public Domain veröffentlicht werden.

Aus der Perspektive von Open Data kann konstatiert werden, dass auf den drei Plattformen, die vorwiegend urheberrechtsfreie Inhalte anbieten, die Nutzung durch Einschränkungen zum Untersuchungszeitpunkt (Mitte 2016) noch erschwert war. Zwar war die Nutzung für Lehre und Forschung explizit erlaubt, doch fehlte ein klares Statement zu Open Data. Nur bei e-codices wurden Metadaten explizit unter der Lizenz CC-BY (und zum Teil als Public Domain) publiziert. Mittlerweile haben e-rara und e-manuscripta die Open Data-Strategie nach den Wünschen der Forschung umgesetzt.

Der Zugang zu Digitalisaten mit höheren Auflösungen ist beschränkt – dies auch aus technischen Gründen. Auf den Plattformen werden in der Regel die komprimierten Nutzungsformate gespeichert und publiziert, während die unkomprimierten und hochauflösenden Archivkopien zum Teil auf anderen Speichermedien lagern und nicht online abgerufen werden können.

Bei den anderen Plattformen (E-Periodica, Zeitungsplattformen, AV-Medien) bestehen in der Regel Einschränkungen durch das Urheberrecht, so dass hier keine Publikation als Open Data möglich ist.

3.1.5 Archivierung

Bei der Digitalisierung werden gemäss den internationalen Standards für die Digitale Langzeitarchivierung unkomprimierte TIFF-Files in einer druckfähigen Auflösung (300 dpi, bei Handschriften mindestens 600 dpi) hergestellt, welche die Grundlage für die Archivierung bilden. In e-rara, E-periodica und e-manuscripta werden diese TIFF-Files mit den Metadaten in METS-XML als Archivkapsel (AIP), wenn vorhanden inklusive der OCR-Daten, gespeichert. Für die Online-Nutzung werden verschiedene JPEG-Formate generiert, zum Teil auch ein PDF. Zudem werden für jeden Artikel oder jedes Buch DOIs als Persistent Identifier vergeben.

An der ETH-Bibliothek ist vorgesehen, die AIPs im System Rosetta für die Digitale Langzeitarchivierung zu speichern. Zurzeit werden die Daten aus e-rara aufgenommen, die aktuell 70 Terabyte umfassen, E-periodica soll im Laufe des Jahres 2017 aufgenommen werden. Anzumerken ist, dass die ETH-Bibliothek nur die eigenen Archivkapseln aus e-rara und e-manuscripta in Rosetta archiviert. Die Kapseln der Partner werden diesen zur individuellen Speicherung jeweils zur Verfügung gestellt. Auch im Projekt Schweizer Presse Online sind die Kantonsbibliotheken selbst für die Langzeitarchivierung der von ihnen gelieferten Inhalte zuständig.

3.2 Bericht zur Governance der Plattformen

Die Plattformen sind organisatorisch sehr unterschiedlich aufgebaut. E-rara ist zum Beispiel ein gemeinsames Projekt der Partner ETH-Bibliothek, Zentralbibliothek Zürich, UB Basel, UB Bern und Bibliothèque de Genève. Gehostet wird die Plattform an der ETH-Bibliothek. Über diese Partner nutzen weitere Bibliotheken e-rara: von der BCU Lausanne über die BPU Neuchâtel bis zur Biblioteca Salita dei Frati (<http://www.e-rara.ch/wiki/libraries>). Andere Bibliotheken können zu bestimmten Konditionen eigene Bestände in e-rara einbringen (<http://www.e-rara.ch/wiki/aboutERara>). Seit Ende 2013 ist das e-lib.ch-Projekt e-rara praktisch abgeschlossen. In 2015 wurde es formell in den Routinebetrieb überführt. Basis dafür waren die von den verantwortlichen Gremien beschlossenen Vorgehensweisen, die ausgearbeitet, verschriftlicht und angewendet wurden (u.a. Vereinbarungen, Kostenmodell, Betriebskonzept). Auf dieser formellen Basis kann die Integration der Bestände aus «kleineren» Institutionen aufbauen. Wie bereits erwähnt, wird e-rara.ch als Produkt betrieben, wohingegen die aktuelle Struktur noch sehr projektorientiert aufgebaut ist. Der aktuelle Steuerungsausschuss e-rara.ch wird die Entscheidung treffen, wie diese Organisationsstruktur weiterentwickelt wird.

Bei e-manuscripta sind es weniger Partner (ZB Zürich, ETH-Bibliothek, UB Basel). Diese aktiven Partner betreuen weitere passive Partner/Anwender wie z.B. die ZHB Luzern, die über die UB Basel indirekt teilnimmt. Bei E-periodica ist neben der ETH-Bibliothek selber die Schweizerische Nationalbibliothek der hauptsächliche Kooperationspartner. In der Regel handelt die ETH-Bibliothek die Konditionen und Kosten mit den Anbietern (Verlagen) der jeweiligen Zeitschrift aus. Auf der Homepage von e-rara und e-manuscripta wird jeweils eine Kontaktadresse (E-Mail) angegeben, bei der sich interessierte Institutionen melden können. Bei e-rara werden auch die Kosten für eine Teilnahme angeführt.

E-codices ist zurzeit immer noch ein Projekt der Universität Freiburg, soll aber in den nächsten Jahren in die Trägerschaft der wichtigsten Handschriftenbibliotheken übertragen werden. Aktuell sind über 70 Sammlungen im In- und Ausland beteiligt, u.a. UB Basel, Stiftsbibliothek St. Gallen, ZB Zürich, Bibliothèque de Genève.

Bei Schweizer Presse Online wird die Anwendung Olive vom Rero gehostet. Die Partner (v.a. Kantonsbibliotheken) bezahlen einen Beitrag für das Hosting (Fr. 2500/TByte/Jahr). Interessierte Partner

wenden sich an die Nationalbibliothek, welche die Vorgaben erläutert und eine Vereinbarung mit dem Partner, der SNB und dem Rero abschliesst.

Die Plattformen haben alle neben der eigenen personellen und technischen Infrastruktur Arbeitsabläufe etabliert, die es ermöglichen, die Dienstleistungen der Plattform auch anderen Einrichtungen anzubieten. Nicht immer ist von aussen einfach zu klären, welche Stelle der Plattform dafür angesprochen werden muss; bei Schweizer Presse Online scheinen die teilnehmenden Einrichtungen bislang auf die Nationalbibliothek und Kantonsbibliotheken eingeschränkt zu sein.

Eine wichtige Aufgabe, welche die Plattformen wahrnehmen, ist die Beratung anderer, meist kleinerer Einrichtungen, die Einbindung deren Dokumente und Metadaten in die jeweilige Plattform und die Organisation von Digitalisierungsprojekten.

3.3 Nutzung der Plattformen

In diesem Projekt interessierte auch die Frage nach der Nutzungsintensität der Plattformen und der Erhebung der Nutzung bzw. des Mehrwerts der Plattformen für die Forschung. Die IST-Erhebung hat gezeigt, dass die Plattformen je nach Inhalt und Format unterschiedliche Kennzahlen erheben. Vorgaben seitens der Schweizerischen Bibliotheksstatistik⁴ (oder anderer Institutionen) gibt es diesbezüglich nicht. Der Indikator "Digitale Medien der Universitätsbibliotheken" wurde zwar vor kurzem neu definiert, beinhaltet aber nach wie vor keine Standardisierung bezüglich der Messung der Zugriffszahlen.

3.3.1 Ungeklärt: Wie sollte die Nutzung erhoben werden?

Es wird in der Bibliotheksstatistik unterschieden zwischen Elektronischen Zeitschriften (ohne Eigendigitalisate), E-Books (inklusive Eigendigitalisaten, also z.B. aus e-rara), digitalen AV-Medien und digitalen Einzeldokumenten (Manuskripte, Bilder, Karten), deren Nutzung jeweils nach unterschiedlichen Kriterien erhoben wird. Die Nutzung der E-Books bezieht sich auf den Download einer Datei und/oder die Betrachtung am Bildschirm.⁵ Die Nutzung der digitalen Einzeldokumente wird dagegen nicht spezifiziert.

Standardisiert ist die Nutzung elektronischer Zeitschriften und von E-Books im COUNTER-Standard. Dieser bezieht sich jedoch auf Produkte kommerzieller Verlage und beschränkt sich auf die primäre Nutzung in Form des Downloads der PDF-Versionen von Zeitschriftenartikeln oder E-Books-Kapiteln (<http://www.projectcounter.org/about/counter-for-libraries/>). Eventuell wäre er aber auch für die hier betrachteten Plattformen sinnvoll. Grundsätzlich ist die Frage der zu erhebenden Nutzungsdaten bislang nicht geklärt, wird aber in der Praxis auch nicht gefordert. Zu beachten wäre bei einer Diskussion um Nutzungsdaten die Aufgabe der Plattformen, vor allem das historische Erbe der Schweiz digital sichtbar zu machen, die sich nicht direkt in Nutzungszahlen übersetzen lässt. Die Plattformen versuchen die Nutzung im Rahmen der Nutzungsbedingungen so zu regeln, dass ein Feedback an die Plattform erfolgen soll. Bei e-manuscripta heisst es:

„Erwünscht ist das Setzen des persistenten Links DOI aus dem Angebot von e-manuscripta.ch [...]. Bei einer Vervielfältigung oder Verwertung ist für jedes Dokument oder Teile daraus der korrekte Zitiernachweis anzugeben. Der Zitiernachweis findet sich in den Titelseiten. Die Zustellung eines Belegexemplares an die besitzende Institution ist erwünscht.“

Es ist zu vermerken, dass die Öffnung der Plattformen in Richtung Open Data es noch schwieriger macht, die effektive Nutzung der Inhalte zu verfolgen. Die Lizenz Public Domain erlaubt die freie Nutzung und Reproduktion, wodurch die Rückmeldung an die Bibliothek höchstens auf freiwilliger Basis erfolgt.

4 <https://www.bfs.admin.ch/bfs/de/home/statistiken/kultur-medien-informationsgesellschaft-sport/erhebungen/chbs.html>

5 Vgl. zur Definition der Variablen <http://www.bfs.admin.ch/bfs/portal/de/index/themen/16/02/02.html>

Sobald die Dokumente auch über andere Zugänge und Plattformen verfügbar gemacht werden können, verliert die besitzende Bibliothek die Kontrolle über die Nutzung. Der Standard IIF soll hier die Rückkoppelung von Drittanwendungen zur ursprünglichen Besitzerin der Inhalte ermöglichen, ist aber bislang nur für Bilddateien standardisiert.

3.3.2 Beispiele für Probleme bei der Erhebung der Nutzungszahlen

Die Erhebung der Kennzahlen erfolgt entsprechend auch in den Online-Plattformen abhängig vom Medienformat und den Funktionen der Plattformen unterschiedlich. So werden bei E-periodica die Views ausgewiesen, da die Downloads der PDF-Versionen nur eine sekundäre Nutzungsform darstellen und diese zudem sehr häufig durch verdächtige Quellen (vermutlich Bots) erfolgen. Die Standardnutzung der Plattform besteht darin, dass online in den Seiten geblättert wird und diese am Bildschirm gelesen werden. Einzelne Strukturelemente wie Artikel, Werbung etc. können dann (zusätzlich) als PDF-Dokument heruntergeladen werden. Ähnlich sieht es bei den Plattformen e-rara und e-manuscripta aus: hier werden ebenfalls die Views ausgewertet, der Download von PDF-Versionen wird als sekundäre Nutzungsform angesehen.

Die Zahlen aus dem Jahresbericht der ETH-Bibliothek belegen eine hohe Nutzungsintensität, auch im Vergleich zu den lizenzierten Inhalten (E-Journals und E-Books). (ETH-Bibliothek 2016, eigene Darstellung)

Zugriffe auf	2013	2014	2015
Bibliotheks-Website (Visits)	976 221	1 568 214	1 526 110
Lizenzierte E-Journals (PDF-Downloads)	3 814 850	5 246 336	5 072 384
E-Books (PDF-Downloads)	2 836 437	3 484 070	5 215 962
ETH E-Collection (Visits)	2 056 808	3 213 537	3 285 286
E-Periodica (Visits)	374 561	3 971 222	6 108 052
e-rara.ch (Visits)	647 720	844 933	1 026 664
Datenbanken (Sessions, Searches)	417 702	367 008	367 392

Dabei ist noch zu relativieren, dass bei der E-Books-Nutzung gemäss COUNTER der Download eines einzelnen Kapitels registriert wird.

3.3.3 Nutzungsgruppen

Da man sich für die offen nutzbaren Inhalte nicht registrieren muss, gibt es keinen Nachweis der Nutzerinnen und Nutzer. Der Trend zur Publikation der Digitalisate als Open Data hat zur Folge, dass die Herkunft der Nutzung nicht mehr festgestellt werden kann. Klar ist, dass die Inhalte weltweit auf Interesse stossen (bei e-rara erfolgt die höchste Nutzung aus den USA) und dass ein breites Publikum mit unterschiedlichen Interessen die Plattformen nutzt. Eine genaue Aussage über die Nutzung kann man bei den Anfragen für Digitalisate in Druckqualität machen, die oft von Forschenden (oder Journalistinnen/Journalisten) für Publikationen verwendet werden. Die Nationalbibliothek stellt fest, dass die digitalisierten Zeitschriften häufig von Historikerinnen und Historikern genutzt werden. Auch bei e-manuscripta und e-codices ist die Nachfrage durch Forschende in Form von direkten Anfragen erkennbar.

Bisher gibt es kaum Anfragen von Forschenden, welche die Inhalte anders nutzen möchten, als über die klassische Recherche und den Download einzelner Dokumente hinaus. Bei e-manuscripta kommt das vor, wobei die vorhandenen Schnittstellen dafür genutzt werden können. Jüngstes Beispiel ist das HTR-Projekt READ in Innsbruck (<http://read.transkribus.eu>).

Auf die Frage, ob der Bedarf von Forschenden im Voraus erhoben wurde, meint die BCU bezüglich der Plattform Scriptorium: "No. They are very happy with the functionalities and just use it like hell (10'000 visitors a day)." (Antwort in der Umfrage im Projekt)

Andere Projekte, die im Rahmen des Programms "Wissenschaftliche Information" von swissuniversities gefördert werden (z.B. DLCM, <https://www.dlcm.ch>), oder das DaSCH

(<http://dhlab.unibas.ch/dasch/>) zielen darauf ab, direkt Forschungsdaten zu managen oder darüber hinaus Forschungsdaten zu archivieren und sind deshalb auf eine direkte Zusammenarbeit mit Forschenden ausgerichtet. Die Praxis, welche sich bislang bei den in dieser Studie untersuchten Plattformen durch die tatsächliche Arbeit in Forschungseinrichtungen, mit Forschenden und der Öffentlichkeit etabliert hat, ist allerdings eine, die sich auf andere Fragen – Langzeitarchivierung, Aufbereitung von Metadaten, Unterstützung anderer Einrichtungen etc. – fokussiert.

3.4 Bericht zur Sicht der Personas auf die Plattformen

Um die tatsächlichen Anforderungen zu bündeln, die an die Plattformen gestellt werden, wurden Personas gebildet. Dieses Vorgehen bündelt die Voraussetzungen und Fragen aus Zielgruppen in einzelnen Kunstfiguren, welche immer eine Ziel- oder Anspruchsgruppe repräsentieren. Aufgabe ist es, diese Kunstfigur möglichst nahe an der tatsächlichen Gruppe zu halten, die sie repräsentieren soll (und nicht z.B. die eigene Vorstellung davon, was diese Gruppe fordern sollte, einfließen zu lassen, sondern nur die tatsächlich festzustellenden Forderungen). Sind diese Personas ausreichend gesättigt konstruiert worden, ist es möglich, Angebote etc. aus deren Sicht zu analysieren. Dies wurde im vorliegenden Fall getan.

In einem ersten Schritt wurde bestimmt, welche Personas zu bilden sind. Dabei wurden nicht alle möglichen Personas gebildet (beispielsweise keine für "normale Bürgerinnen und Bürger"), sondern nur solche, die im Hinblick auf Forschung und Forschungsinfrastruktur sinnvoll sind. Dies ergibt sich aus dem Blickwinkel des Programmes "Wissenschaftliche Information", in dessen Rahmen die vorliegende Studie erstellt wurde. Im Alltag haben die Plattformen auch mit anderen Einrichtungen, beispielsweise Verlagen, zu tun und müssen ebenso auf deren Anforderungen reagieren.

Im nächsten Schritt wurden die Personas angereichert, d.h. durch die vorhandene Literatur wurde versucht zu beschreiben, welche Interessen diese an den Plattformen selber haben werden. Dabei wurde darauf geachtet, die konstruierten Personas an der realen Praxis der Gruppen, die sie repräsentieren, zu orientieren. Dies ist z.B. bei der Persona des Digital Humanists wichtig geworden, bei der nicht die vor einigen Jahren in der Literatur gemachten Vorhersagen, was die Digital Humanities in Zukunft sein würden, zur Basis genommen wurde, sondern die tatsächlich in der Forschungsliteratur sichtbar werdende Praxis. Diese unterscheidet sich relevant von den Vorhersagen, traf am Ende aber z.B. auch eher die Erfahrung der untersuchten Plattformen.

Nachdem die Personas angereichert waren, wurden in einem weiteren Schritt alle Plattformen analysiert. Diese Analyse geschah einmal vom Schreibtisch aus, d.h. aus der Situation heraus, aus der die meisten Einrichtungen und Personen die Onlineplattformen wahrnehmen werden. Anschliessend wurden im Rahmen der Interviews und Umfrage eruiert, wie die Plattformen auf bestimmte Anforderungen aus Sicht der Personas reagieren würden. Festgehalten werden kann, dass die Plattformen bei einem einfachen Zugriff auf die jeweilige Webpräsenz als verschlossener erscheinen als bei einer direkten Nachfrage. Letztere führte eigentlich immer zu schnellen Lösungen.

Im Folgenden werden zuerst die verwendeten Personas dargestellt.

3.4.1 Beschreibung der Personas

Big Data Scientist. Beispiel Digital Humanities

- Will Text and Data Mining machen, arbeitet in einem internationalen Projekt (mit dem MIT)
- Will grosse Textmengen analysieren und neue Erkenntnisse gewinnen
- Braucht dafür Zeitschriftenplattformen, um historische Zeitschriften zu durchforsten
- Benötigt OCR-erkannte Texte, die frei genutzt und ausgewertet werden können
- Benötigt auch lizenzierte Inhalte, rechtliche Fragen müssen geklärt werden

Wissenschaftshistorikerin

- Forscht zu Johannes Bernoulli (Basler Mathematiker), benötigt Material aus der UB Basel (Bernoulli-Archiv), aber auch aus anderen Quellen (z.B. Leibniz-Briefe)
- Benötigt Briefe und Manuskripte in digitaler Form, will sie editieren, transkribieren, mit Varianten aus verschiedenen Quellen vergleichen können
- Benötigt auch digitalisierte Bücher (Werke von Bernoulli), die mit OCR bearbeitet sind
- Will mit Forschenden in Dresden und Sankt Petersburg zusammenarbeiten
- Hat noch keinen Zugang zu einer Plattform für die Edition, überlegt sich, ob sie eine eigene aufbauen oder eine bestehende nutzen soll
- Fragt sich, was mit den Ergebnissen geschieht: Wo kann sie diese ablegen und sichern? Können diese mit der Plattform verlinkt werden, aus der die Originaldokumente stammen?

AV-Medien: Ethnologe (Kulturforschung, europäische Ethnologie)

- Benötigt Tonaufnahmen und Filmmaterial/Videos zur Darstellung von Urlaub (Ferien) in der Schweiz
- will diese Aufnahmen annotieren
- möchte auf einzelne Ausschnitte verlinken
- möchte Filme als Forschungsdaten verwalten

Kleinere Kantonsbibliothek. Thurgauer Kantonsbibliothek

- Digitalisiert Nachlässe (Fotos, Handschriften, Zeitungsausschnitte, usw.), Zeitungen, Zeitschriften und Bücher
- Benötigt Repository für die Datenspeicherung und Vermittlung der Medien an ihre Nutzerinnen und Nutzer
- Möchte Digitalisate im OPAC/Discovery nachweisen (eine Suche für alle Inhalte)
- Fragt sich, ob sie eine eigene Plattform benötigt oder bei einer anderen mitmachen kann (soll? muss? will?)
- Will Metadaten für einen zentralen Katalog zur Verfügung stellen (welche Anforderungen werden an einen solchen zentralen Katalog/Plattform/Suchoberfläche gestellt?)
- In der „Strategie Digitalisierung des Programms „Wissenschaftliche Information“ SUK 2013-2016 P-2 wird die Integration von Digitalisierungsprojekten kleinerer Einrichtungen (und dieser Einrichtungen im Allgemeinen) in die vorhandenen Plattformen als zentrales Ziel benannt

Kleines Museum. Beispiel: Ortsmuseum Bergün

- Digitalisiert Objekte (3D) und Filme (Videos) und regionale Schriften
- Kann das nicht selbst ausführen
- Benötigt Repository für die Datenspeicherung und Vermittlung
- Kann (soll? muss? will?) Metadaten für einen zentralen Katalog zur Verfügung stellen
- In der Digitalisierungsstrategie des Programms SUK P-2 wird die Integration von Digitalisierungsprojekten kleinerer Einrichtungen (und dieser Einrichtungen im Allgemeinen) in die vorhandenen Plattformen als zentrales Ziel benannt

Gemeindearchiv Schwarzenberg/BE

- Digitalisiert und bereitet historischen Bestand auf (Zimmerwald-Konferenz)
- Benötigt Repository für die Datenspeicherung und Vermittlung
- In der Digitalisierungsstrategie des Programms SUK P-2 wird die Integration von Digitalisierungsprojekten kleinerer Einrichtungen (und dieser Einrichtungen im Allgemeinen) in die vorhandenen Plattformen als zentrales Ziel benannt

Openresearchdata.ch

- Will einen Metadatenkatalog für Forschungsdaten aus Schweizer Forschungsprojekten zur Förderung der Sekundärnutzung dieser Daten aufbauen
- Benötigt dazu die Metadaten im CKAN-Format, müsste aber sinnvollerweise das DCAT Application Profile for Switzerland implementieren, um die Kompatibilität mit der Open Government Data Plattform opendata.swiss zu gewährleisten
- Stellt kein Repository für die Forschungsdaten zur Verfügung; diese müssen nach wie vor von den einzelnen beteiligten Institutionen dezentral gespeichert werden

3.4.2 Beschreibung der Plattformen aus Sicht der Personas

Big Data Scientist. Beispiel Digital Humanities

Der Forschende hat, als Digital Humanist, Erfahrungen mit Metadaten, mit dem Programmieren und auch der Arbeit mit Schnittstellen. Seine Recherche auf den schweizerischen Plattformen zeigt ihm keine klar definierten Schnittstellen auf, deshalb wechselt er direkt auf die Europeana, welche diese Schnittstellen bietet und auf einer eigenen Seite beschreibt (<http://labs.europeana.eu/api>). Für seine Arbeit scheint dies ausreichend.⁶

Bei der Auswertung der Daten bemerkt er allerdings, dass ihm schweizerische Zeitschriften gänzlich fehlen. Da sich seine Forschungsarbeiten auf Zeitschriften aus ganz Europa beziehen, will er diese Lücke schliessen. Deshalb wendet er sich per Mail direkt an die Plattformen und erhält anschliessend durch diesen persönlichen Kontakt Zugriff auf die Schnittstellen. Probleme lassen sich, auch da der Forschende selber Erfahrungen und technisches Wissen mitbringt, relativ schnell lösen. Auf Nachfrage wird ihm erklärt, dass er seit Jahren der erste Forschende sei, der solche Fragen stelle. An sich sind die Plattformen dazu in der Lage, die gefragten Metadaten über standardisierte Schnittstellen auszuliefern. Bei grösserer Nachfrage würden sie dies auch transparenter machen.

Nachtrag: die ETH-Bibliothek klärt momentan die Anforderungen ab, die eine Öffnung der Plattform E-Periodica für Text and Data Mining bedeuten würde.

Wissenschaftshistorikerin

Die Historikerin recherchiert in mehreren Plattformen und versteht dies auch als Teil ihrer Forschungsarbeit. Die einzelnen Dokumente kann sie händisch herunterladen und bearbeiten. Keine der Plattformen bietet eine Integration aller Dokumente und auch keine Möglichkeit, diese kollaborativ zu bearbeiten.

Sie recherchiert weiter nach vorhandenen Möglichkeiten, ihre Arbeit digital zu organisieren. Sie entscheidet sich für eines der vorhandenen Software-Pakete für diese Aufgabe (TextGrid), organisiert sich einen Account für das von der Software genutzte Repository (DARIAH), lädt die bei den schweizerischen (und anderen) Plattformen heruntergeladenen Dokumente in dieses und arbeitet anschliessend damit. Ihr Kontakt mit den Plattformen beschränkt sich auf das Recherchieren und Downloaden. Sie findet sich selber zu Recht und fordert keine Hilfe von Seiten der Plattformen ein.

Im Falle der Bernoulli-Briefe aus der UB Basel müssen ohnehin noch grosse Bestände digitalisiert werden. Am meisten würde es der Forscherin dienen, wenn sie zusammen mit der Bibliothek für die Digitalisierung dieser Bestände aus einem Fonds finanzielle Mittel beantragen könnte. Auf ihre Anfrage verweist die Bibliothek nämlich auf andere Digitalisierungsprojekte, die eine höhere Priorität geniessen.

6 Grundsätzlich wird in der Literatur zur Digital Humanities (z.B. Burdick et al. 2012) davon ausgegangen, dass Forschende, die bislang nicht mit Daten gearbeitet hätten, im Zuge der Digital Humanities Fragen formulieren würden, die nur mit Daten – insbesondere grossen Datenmengen – zu beantworten sind. Die tatsächlich publizierte DH-Literatur und Studien zur Praxis der DH (Bender 2016, Antonijević 2015) zeigen aber, dass die Forschenden, die Forschungsfragen formulieren, die grosse Datenmengen zu Beantwortung benötigen, technisch in der Lage sind, mit diesen umzugehen. Andere Forschenden scheinen weiterhin andere Fragen zu stellen.

Nachtrag: im Rahmen des Projekts NIE-INE wird eine nationale Infrastruktur für digitale Editionen aufgebaut, die von der Bernoulli-Edition genutzt werden kann. Ein automatisierter Zugriff auf Digitalisate ist bei diesen Editionsprojekten nicht vorgesehen.

AV-Medien: Ethnologe (Kulturforschung, europäische Ethnologie)

Dem Forschenden wird schnell klar, dass er auf schweizerischen Plattformen vor allem textuelle Medien findet. Einzig die Memobase und die Schweizerische Nationalphonothek bieten Film- und Tonaufnahmen.

Die Recherche ist energieverbrauchend aufwändig. Die Dokumente in der Memobase sind jeweils nur über mehrere Links zu erreichen und unterliegen, da sie verschiedenen Institutionen gehören, unterschiedlichen Nutzungsbedingungen. Die Tondokumente sind praktisch nur an den AV-Arbeitsstationen in verschiedenen Bibliotheken der Schweiz oder der Phonothek selber zu hören.

Auch für Bildbestände, die den Ansprüchen des Ethnologen entsprechen (möglichst hochaufgelöste Fotos), fehlt bislang eine übergreifende, schweizweite Plattform. Die meisten Institutionen betreiben hier eigene Systeme oder versuchen bestehende Plattformen für digitale Bilder zu verwenden. Am ehesten könnte das Bildinformationssystem e-pics (<http://www.e-pics.ethz.ch>) an der ETH, das von der ETH-Bibliothek betrieben wird, den Anforderungen eines Forschenden entsprechen. Allerdings ist auch diese Plattform für die eigenen Bestände der ETH-Bibliothek konzipiert – wobei sie auch von der Schweizerischen Fotostiftung genutzt wird. Eine übergreifende Suche ist jedoch momentan nicht möglich. Immerhin ist der Forschende hier auf die Postkartensammlung Feller gestossen und hat alte Ansichtskarten aus der Schweiz entdeckt, die er nun in verschiedenen Formaten downloaden und unter einer Creative Commons-Lizenz (CC-BY oder Public Domain) nutzen kann. Die Recherche nach Postkarten führt ihn auch zum Projekt Swiss Postcards der Schweizerischen Nationalbibliothek (<https://www.nb.admin.ch/aktuelles/01952/04772/index.html?lang=de>). Mehrere Kantonsbibliotheken stellen hier historische Postkarten zur Verfügung. Für den Forscher ist jedoch der Zugriff ausschliesslich über Apps untauglich.

Der Ethnologe ist einen solchen schwierigen Zugang gewöhnt, kann damit aber auch nur qualitativ arbeiten; quantitative Arbeiten sind so nicht möglich, da er gar nicht genügend unterschiedliche Dokumente für quantitative Fragestellungen findet. Grundsätzlich ist dies das Vorgehen der Ethnologie, der Forschende fragt sich gleichwohl, ob die Versprechen der Digital Humanities und des Digitalen Zeitalters auf den schnellen Zugriff auf viele unterschiedliche Medien nicht auch einmal für die Ethnologie gelten müssten. Es gibt in seiner Wissenschaft einige Texte mit Beispielen für solche Arbeit mit Daten, diese sind aber immer noch eine Ausnahme. Vielleicht bleiben sie es auch weiterhin.

Kleinere Kantonsbibliothek. Thurgauer Kantonsbibliothek

Bei den Zeitungen arbeitet die Kantonsbibliothek bereits mit der SNB im Projekt Schweizer Presse Online zusammen. Hier werden einige lokale Zeitungen, die im Auftrag der Kantonsbibliothek Thurgau bei einem externen Dienstleister digitalisiert wurden, gegen eine jährliche Gebühr aufgenommen. Eine Integration der Volltextsuche in den eigenen Katalog ist nicht möglich. Einige Zeitschriften lokaler Verlage wurden von der ETH-Bibliothek für die Plattform E-periodica digitalisiert und in diese Plattform integriert. Eine Suche ist über NEBIS und somit auch dem Swissbib möglich. Für einen Bestand an Alten Drucken könnte die Bibliothek mit der Zentralbibliothek (ZB) Zürich zusammenarbeiten: die ZB würde die Bücher im Digitalisierungszentrum digitalisieren und als eigenständige Kollektion in e-rara veröffentlichen – analog zu einem Bestand der Kantonsbibliothek Graubünden. Im eigenen Katalog werden die Werke mit dem Link auf das Digitalisat ergänzt, wodurch sie auch über die lokale Recherche direkt suchbar werden.

Bei allen Projekten bleibt die Verantwortung für die digitale Langzeitarchivierung bei der Kantonsbibliothek selber. Sie erhält die Archivfiles samt Metadaten als Archivkapseln und müsste diese nun in ein Langzeitarchiv geben können. Diese Möglichkeit fehlt jedoch.

Es gibt für alle betreffenden Bestände eine Plattform, die von der Kantonsbibliothek Thurgau genutzt werden könnte – ausser für die Bildbestände. Einzelne Bilder könnten in e-manuscripta integriert werden, aber für grössere Fotosammlungen ist die Plattform nicht geeignet. Eine gemeinsame Darstellung des kulturellen Erbes des Kantons ist so nicht möglich. Dafür müsste eine eigenständige

Plattform aufgesetzt werden, welche die Bestände aus den verschiedenen Plattformen wieder aggregiert, z.B. mit Hilfe solcher Software wie Omeka. Möglich wäre zudem die Publikation der gemeinfreien Bilder auf Wikimedia und die Erschliessung als Online-Ressource im Bibliothekssystem.

Kleines Museum. Beispiel: Ortsmuseum Bergün

Memobase liefert nur einen Metakatalog. Das heisst, dass die AV-Medien in einer anderen Plattform erfasst und gespeichert werden müssten. Ein entsprechendes, im Hinblick auf Langzeitdatenarchivierung vertrauenswürdige Repository ist nicht verfügbar. Bisherige Lösungen sind nur für einzelne Institutionen gedacht oder für spezifische Fachinhalte (v.a. im Bereich Kunst). Möglich ist auch die Nutzung von Drittanbietern (z.B. kommerzielle Videoportale), die aber für eine solche Aufgabe nicht vorgesehen sind und z.B. gerade keine langfristige Sicherung der Daten garantieren. Das Museum würde dem vertrauenswürdige Lösungen, als Vorbild wäre z.B. das Hosting von Videos beim Bundesarchiv direkt zu nennen, vorziehen.

Das Museum kann für textuelle Dokumente für Beratungen bei den Einrichtungen anfragen, welche die anderen Plattformen betreiben, z.B. bei E-Periodica für die Digitalisierung und Online-Publikation einer lokalen Zeitschrift. Für andere Objekte oder Medien fehlen bislang Plattformen.

Gemeindearchiv Schwarzenberg/BE

Ein gemeinsam betriebenes oder von kleineren Institutionen/Archiven nutzbares Repository gibt es nicht. Die Gemeinde muss versuchen, mit dem Staatsarchiv zusammenzuarbeiten, wobei sich das Staatsarchiv zumeist auf den gesetzlichen Auftrag konzentriert und nur kantonsrelevante Inhalte bearbeitet. Für andere Inhalte, auch solche, die lokalgeschichtlich eine grosse Bedeutung haben, gibt es bislang im Archivbereich keine Lösung.

Openresearchdata.ch

Die Online-Plattformen können ihre Metadaten momentan weder im CKAN-Format noch im Format des DCAT Application Profile for Switzerland zur Verfügung stellen. Die Plattformbetreiber müssten selbst auf ihre jeweils eigenen Gegebenheiten zugeschnittene Konversionsfunktionen entwickeln. Um den dafür nötigen Aufwand abschätzen zu können, müssten zunächst die unterschiedlichen technischen Details abgeklärt werden; der zu erwartende Aufwand darf jedenfalls nicht unterschätzt werden.

Es ist fraglich, ob sich der Aufwand lohnen würde. Wenn die Plattformbetreiber den Zugang zu ihren Metadaten vereinheitlichen möchten, wäre ein zukunftsgerichtetes internationales Format besser geeignet als ein nationales. Eine sinnvolle Weiterentwicklung von openresearchdata.ch scheint deshalb eher fraglich.

4. Effektiver Regelbetrieb bei den vorhandenen Plattformen, Potentiale bei der Dokumentation

Die untersuchten Plattformen haben alle, für ihren jeweiligen Bereich, einen funktionierenden Regelbetrieb aufgebaut, der den jetzigen Anforderungen an sie entspricht. Grundsätzlich ist eine weitere Entwicklung möglich und wird zum Teil auch betrieben, der Grossteil der Arbeit wird allerdings in der alltäglichen Arbeit – insbesondere der Einarbeitung von Materialien – aufgebracht. Entsprechend sind auch die Funktionalitäten zur Unterstützung des Workflows (von der Auswahl über die Digitalisierung zur Erschliessung, Strukturierung bis hin zur Archivierung) von grosser Bedeutung.

Die Literatur zur Entwicklung von digitalen Sammlungen und zur Entwicklung digitaler Wissenschaft ist seit langem voll von regelmässig wiederholten Voraussagen über die möglichen Entwicklungen in diesen Bereichen. Aus diesen Vorhersagen – insbesondere die, dass in Zukunft mit grossen Daten-

mengen gearbeitet würde – werden Anforderungen abgeleitet, beispielsweise die nach möglichst guten und möglichst einheitlichen Metadatenstandards und nach offenen Schnittstellen. Es ist eine wissenschaftspolitische, keine technische Fragestellung, ob diesen Voraussagen gefolgt wird und damit Infrastrukturen gefördert werden sollten, die z.B. zur Vereinheitlichung beitragen. Aus der jetzigen Situation, d.h. den realen Anforderungen der Forschung an die Plattformen sowie den tatsächlich unter dem Schlagwort Digital Humanities publizierten Forschungen, oder den Personas, die in dieser Studie konstruiert wurden, ergibt sich dies zurzeit nur bedingt.⁷

In ihrer jetzigen Arbeit nehmen die Plattformen ein solches Interesse von Seiten der Forschung nicht wahr. Sie sind grundsätzlich darauf vorbereitet, auf entsprechende Anforderungen zu reagieren. Dies belegen die aktuell laufenden Projekte zur Weiterentwicklung der Plattformen: bei e-rara werden OCR sowie Volltextrecherche eingeführt, bei e-manuscripta die Annotationsfunktionalität. Bei e-periodica prüft man die Nutzung durch Text and Data Mining, und e-codices baut die Plattform nach Kriterien der Offenheit um.

Die Plattformen sind zudem mit der Bearbeitung konkreter Fragen, insbesondere von anderen Einrichtungen, die Dokumente digitalisieren und in die Plattformen integrieren wollen, beschäftigt. Hiermit erfüllen sie schon ein Ziel, welches in der „Strategie Digitalisierung“ des Programms „Wissenschaftliche Information“ genannt wird. Soweit ersichtlich, sind die Nutzenden bislang mit den Angeboten zufrieden. Die Nutzungszahlen haben jedenfalls in den letzten Jahren stark zugenommen.

Zu erwähnen ist auch, dass das Personal hinter den Plattformen mit den relevanten Diskussionen vertraut ist und beispielsweise die wichtigsten Argumente für einheitliche Metadaten – als nicht offen besprochenes, aber immer im Hintergrund bekanntes Vorbild ist das Europeana-Datenmodell zu nennen – schon mehrfach präsentiert bekommen haben. Gleichwohl haben sie sich für unterschiedliche Metadaten entschieden, da sie sich in ihren realen Arbeitsumgebungen (Bibliotheken, Archive, Museen auf der institutionellen sowie Forschende und andere Nutzende auf der anderen Seite) als sinnvoll erwiesen.

4.1 Potentiale im Status Quo

Die Plattformen sind grundsätzlich gut aufgestellt, auch für potentielle Veränderungen. Sie sind in den meisten Fällen z.B. auch in der Lage, grössere Datenmengen auszuliefern, wenn es entsprechende Anfragen gäbe.

Allerdings sind diese Potentiale bislang wenig sichtbar für Forschende, die sich den Plattformen von aussen nähern. Es wäre sinnvoll, dies zu ändern und klarer anzugeben, was an Metadaten und Daten in den Plattformen vorhanden ist, wie auf sie zugegriffen werden könnte, welche Schnittstellen frei nutzbar sind und welche eingerichtet werden könnten. Grundsätzlich geht es eher um die Dokumentation des Vorhandenen, weniger um neue Angebote der Plattformen. Dies könnte auch zu einer anderen Nutzung führen, indem das Angebot die Nachfrage erst erzeugt.

Die Plattformen könnten zudem überlegen, ob sie selber konkreter auf Forschende zugehen und ihre Daten anbieten sollten. Bislang reagieren sie offenbar auf Anfragen, die an sie gerichtet werden, und dies zur Zufriedenheit der fragenden Forschenden. Es wäre aber auch möglich, jeweils Strategien zu entwickeln, die Daten mehr bei anderen möglicherweise interessierten Forschenden sichtbar zu machen (Publikationen in Fachzeitschriften, Information auf Kongressen, Workshops etc.), wobei gerade letzteres in den letzten beiden Jahren von einigen Bibliotheken (und Museen) bereits angegangen wurde, indem man sich aktiv an der Organisation von Hackathons beteiligte.

Auch die Teilnahme interessierter Institutionen ist heute bei allen Plattformen möglich. Verschiedene Kantonsbibliotheken sind Partner bei e-rara, e-manuscripta, e-codices, E-periodica und Schweizer Presse Online. Die Möglichkeiten scheinen bekannt. Ein Hindernis sind dabei die für kleinere Institutionen recht hohen Kosten für die Digitalisierung und Aufbereitung der Bestände. Dies gilt auch für

⁷ Richtig ist, dass in andere Wissenschaften als den Humanities grosse Datenmengen für Forschungsfragen benutzt werden. Diese Dokumente, die auf den untersuchten Plattformen angeboten werden, sind aber für diese Wissenschaften – z.B. die Physik – kaum interessant.

Forschende (siehe Persona Wissenschaftshistorikerin), die vor allem Finanzierungsmöglichkeiten für die Digitalisierung interessanter Bestände benötigen.

4.2 Potentiale für vermutete Entwicklungen

Ein wichtiges Potential ist die Einbindung von nicht-textuellen Dokumenten. Eventuell ist dies über weitere Plattformen, die an anderen Einrichtungen als den Bibliotheken angesiedelt sein könnten, umzusetzen. Grundsätzlich ist dafür die Memobase vorgesehen, allerdings ist die Memobase (vergleichbar mit Europeana oder der Deutschen Digitalen Bibliothek) ein Metakatalog und bietet kleineren Institutionen keine Plattform zur Publikation und Archivierung audiovisueller Medien.

Politisch relevant ist, dass alle Plattformen auch den Sammlungsprofilen ihrer Trägereinrichtungen folgen, was zum Beispiel auf der Ebene lokaler Sammlungen in kleineren Bibliotheken, Archiven oder Museen oft heisst, dass für diese keine Plattformen vorgesehen sind, da sie z.B. zu wenig Bezug zum gesamten Kanton oder der gesamten Schweiz haben. Eine Lösung für die Digitalisierung und das langfristige digitale Angebot dieser Digitalisate (also ein Repository) wäre sinnvoll. Eine solche Plattform sollte die genannten offenen Standards und Schnittstellen unterstützen und von einem vertrauenswürdigen Anbieter langfristig betrieben werden.

Im Hintergrund solcher Debatten steht immer das Modell der Europeana, das sich ähnlich (als Metadatenkatalog von Digitalisaten, die anderswo angeboten werden) wohl auch für die Schweiz denken lässt. Ein Vorteil wäre die Einbindung verschiedener Medienformate über ein offenes Metadaten-schema. Für eine solche Lösung müsste allerdings, analog zur Europeana, eine Finanzierung und eine neue Trägerinstitution etabliert (vergleichbar zur Deutschen Digitalen Bibliothek) oder vorhandene (z.B. die Nationalbibliothek, das Bundesarchiv, das Landesmuseum oder die Memobase) ausgebaut werden.

Das White Paper des Programms „Wissenschaftliche Information“ (SUK 2013-2016 P-2) sowie die „Strategie Digitalisierung“, welche die Umsetzungsmassnahmen des White Paper im Bereich Digitalisierung ersetzt, gehen davon aus, dass die datengetriebene Wissenschaft – Forschung auf der Basis von grossen Mengen von Daten – in den nächsten Jahren an Fahrt aufnehmen wird. Die Plattformen spüren davon noch nichts, konzentrieren sich deshalb auf den Alltagsbetrieb. Für eine datengetriebene Wissenschaft wäre es wohl nötig, technische Lösungen für die Recherche in und die Auslieferung von grossen Datenmengen, im Idealfall als Open Data (Digitalisate) und Linked Open Data (Metadaten), bereitzustellen. Stimmen die Voraussagen zur kommenden datengetriebenen Wissenschaft, müssten die Plattformen eine Standardisierung von Metadaten und gemeinsame Beschäftigung mit eventuell neuen, aktuell vorgeschlagenen Formaten (z.B. IIF, TEI-CMI) anstreben. Grundsätzlich scheinen sie dafür gut aufgestellt zu sein.

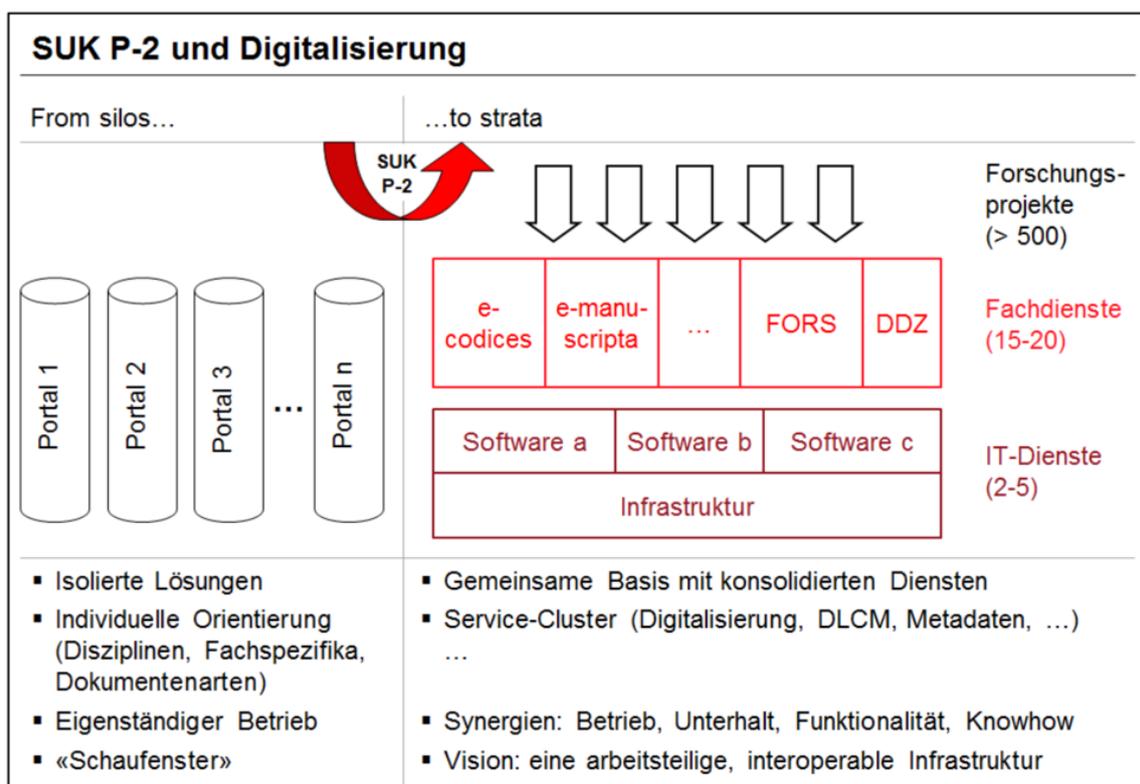
Die Plattform e-codices stellt in den letzten Jahre eine verstärkte Nutzung durch Wissenschaftlerinnen und Wissenschaftler in mehreren Projekten über ihre Schnittstellen fest.

Im Fall von Anfragen oder tatsächlichen Entwicklungen in diese Richtung scheinen die Plattformen zum grossen Teil in der Lage, schnell reagieren zu können. Gerade die von den Bibliotheken betriebenen Plattformen haben im Zuge ihres Aufbaus Formen der Kooperation etabliert, die gemeinsame Arbeiten in diese Richtung ermöglichen würden. Notwendig ist nicht so sehr ein Umbau der Plattformen, sondern wissenschaftspolitische Entscheidungen, den Vorhersagen zur datengetriebenen Wissenschaft zu folgen, die dann auch mit finanziellen Ressourcen untermauert werden müssen.

Allerdings, dies ist noch einmal zu betonen, zeigen sich weder bei den Anfragen an die Plattformen noch in der Literatur zu tatsächlich durchgeführten Forschungsprojekten, die sich als datengetrieben beschreiben lassen, eindeutige Hinweise darauf, dass Anforderungen an die Plattformen gestellt werden, die Daten anders aufzubereiten. Vielmehr scheinen diese Anforderungen aus Überlegungen hergeleitet, dass die datengetriebene Wissenschaft besser funktionieren könnte, wenn die Plattformen noch offener wären. Diese Vorhersagen sind aber nicht einfach mit der Realität der Plattformen oder der Realität in den Digital Humanities in Übereinstimmung zu bringen. Insoweit ist es nicht einfach, klar zu benennen, ob in dieser Richtung – bezüglich einer Infrastrukturentwicklung – Potentiale bestehen.

Es kann vermutet werden, dass eine einfache Integration der Inhalte der verschiedenen Plattformen in übergeordnete Suchzugänge aus Sicht der Nutzerinnen und Nutzer begrüsst würde. Aktuell benötigt man das Wissen darüber, welche Plattformen es für welche Inhalte gibt. Zudem kommt es durchaus vor, dass eine Forscherin (z.B. die Wissenschaftshistorikerin) für ihre Arbeit verschiedene Medientypen benötigt, beispielsweise Manuskripte aus e-manuscripta oder e-codices, alte Drucke aus e-rara und Zeitschriftenartikel aus E-periodica. Heute muss sie in verschiedenen Plattformen suchen, um die entsprechenden Dokumente zu finden. Aus technischer Sicht bieten sich hier ein Ausbau der Plattformen und eine Öffnung in Richtung übergeordneter Plattformen oder Portale auf der Basis von Linked Open Data an. Eine Weiterentwicklung in diese Richtung ist für die Betreiber der Plattformen heute eine Option, die aber nicht primär dem eigenen, sondern einem übergeordneten Interesse entspricht. Die Weiterentwicklung der Plattformen in diese Richtung müsste entsprechend von übergeordneter Stelle finanziell unterstützt werden.

Die in der „Strategie Digitalisierung“ des Programms „Wissenschaftliche Information“ skizzierte künftige Informationsinfrastruktur könnte pragmatisch so umgesetzt werden, dass sich die bestehenden, etablierten Plattformen als unabhängige, von mehreren Institutionen betriebene, fach- und medien-spezifische Anbieter für einen plattformunabhängigen Zugriff öffnen. Dadurch werden Synergien und bestehendes Know-how genutzt, und nur die oberste Schicht, bzw. die Schnittstellen zu einem übergreifenden Zugriff, müssten noch ausgebaut werden. Wobei dieser Zugriff wiederum über verschiedene Portale (oder auch direkt aus Anwendungen, die für Forschungsprojekte entwickelt wurden) erfolgen kann und künftig auch Plattformen einbinden kann, die heute noch nicht bestehen. Im Projekt linked.swissbib.ch wird für die Bibliothekskataloge (in denen zumindest teilweise auch die hier besprochenen Inhalte enthalten sind) eine entsprechende Lösung entwickelt.



„From silos to strata“: vereinfachte Darstellung der Förderziele

Abbildung „From silos to strata“ (Programm „Wissenschaftliche Information“, Strategie Digitalisierung)

Ausgehend vom in der Digitalisierungsstrategie skizzierten Schema lassen sich die bestehenden Portale als Fachdienste bezeichnen, die dezentral betrieben werden. Vor allem die offenen Schnitt-

stellen zu den Forschungsprojekten (bzw. zu Portalen) müssen noch standardisiert und implementiert werden. Rot dargestellt sind noch nicht oder noch unzureichend vorhandene Elemente.

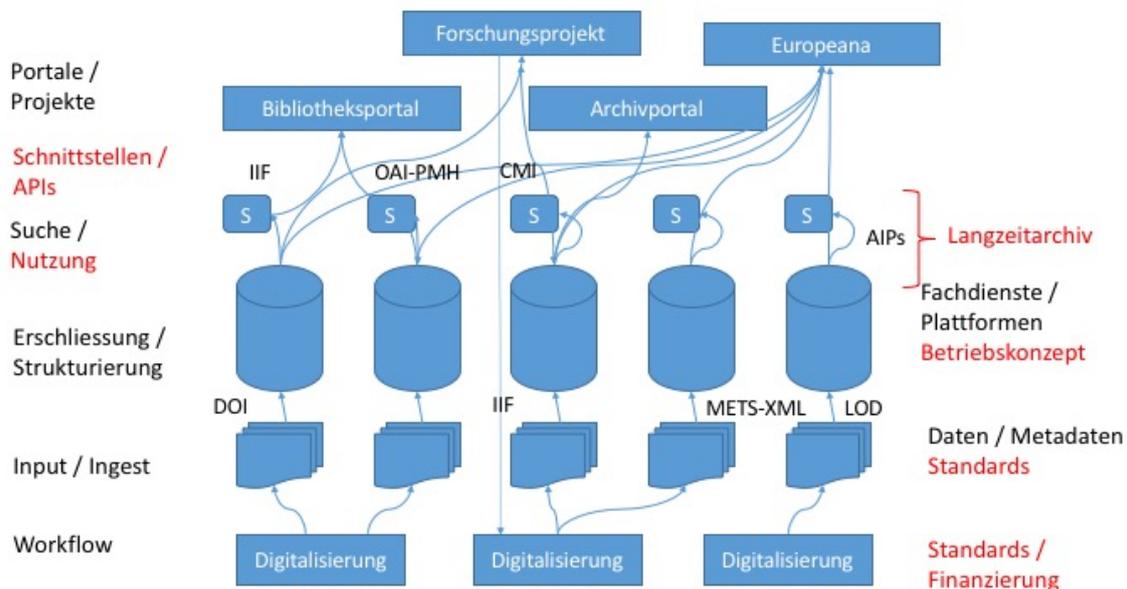


Abbildung: Schema der Architektur eines offenen, dezentralen Systems der Online-Plattformen (eigene Darstellung)

4.3 Europeana

Wie erwähnt, taucht bei der Untersuchung der vorhandenen Plattformen immer wieder die Europeana bzw. eine mögliche Beteiligung an dieser, als Thema auf. Dabei ist zu erwähnen, dass eine solche Beteiligung nicht im Fokus des Programmes "Wissenschaftliche Information" steht, sondern eher der Kulturpolitik des Bundes und der Kantone zuzuordnen wäre. Gleichwohl würde eine solche Beteiligung die Sichtbarkeit der digitalisierten Sammlungen im europäischen Raum erhöhen.

Die Beteiligung könnte für die Plattformen eine denkbare Möglichkeit darstellen (e-codices kooperiert seit 2010). Die Europeana funktioniert als Repository von Metadaten, die einzelnen digitalisierten Objekte, inklusive aller Möglichkeiten, über deren Nutzung zu bestimmen, verbleiben bei den beteiligten Institutionen. Das Modell für die Metadaten, das Europeana Data Model (<http://pro.europeana.eu/page/edm-documentation>), bietet offene Standards und Schnittstellen und somit eine Grundlage auch zur Nutzung der Daten in anderem Kontext (Portale, Forschungsprojekte). Die Europeana stellt zwei Forderungen, damit Einrichtungen sich kontinuierlich an ihr beteiligen können: Die Einhaltung der Standards, inklusive der Lieferung von Metadaten im Europeana Data Model und die Existenz einer nationalen Kontaktstelle, die für die Einrichtungen eines Landes die Transformation der jeweiligen Metadaten in dieses Modell, die Datenlieferung und die notwendige Beratung übernimmt. Diese Kontaktstelle muss vom jeweiligen Land finanziert werden, ansonsten ist die Mitgliedschaft in der EU – obwohl die Europeana von der EU-Kommission gefördert wird – nicht notwendig. Norwegen ist z.B. ein sehr aktives Mitglied der Europeana. (Persönliche Kommunikation mit dem betreffenden Verantwortlichen bei der Europeana, Joris Pekel.)

Im Rahmen dieser Studie ist die Beteiligung an der Europeana nicht weiter geprüft worden, da dies nicht zum Projektauftrag gehörte. Grundsätzlich wäre in dieser Richtung eine Weiterentwicklung der schweizerischen Infrastruktur für Kulturvermittlung möglich. Es wäre denkbar, eine solche Kontaktstelle einzurichten – weshalb dies bis heute noch nicht geschehen ist, entzieht sich der Kenntnis der Autoren des Berichts. Die Schweizerische Nationalbibliothek arbeitet zwar in der Europeana mit, scheint diese Rolle jedoch nicht zu übernehmen. Es wäre mit den Plattformbetreibern zu klären, ob

eine solche Koordinationsstelle eingerichtet und finanziell unterstützt werden könnte. Dies scheint mit der jetzt schon geleisteten Arbeit in der Standardisierung von Daten relativ einfach möglich zu sein. Gleichzeitig ist es immer auch möglich, ohne Datenlieferung an die Europeana das Europeana-Datenmodell zu nutzen.

5. Fazit

Im Projekt AISOOP (Analyse der Informationsarchitektur, Schnittstellen und Organisation der Online-Plattformen) wurde untersucht, ob die kollaborativ von schweizerischen Bibliotheken betriebenen Plattformen für digitalisierte Medien den Anforderungen der aktuellen Wissenschaften entsprechend aufgestellt sind.

Grundsätzlich sind die Plattformen gut für ihre jetzigen Aufgaben aufgestellt, zu denen insbesondere die Unterstützung des gesamten Workflows von der Digitalisierung bis zur Bereitstellung gehören. Sie sind personell und infrastrukturell so ausgestattet, dass sie gut funktionieren. Einige beklagen zu wenig Personal für eine sinnvolle Weiterentwicklung. Gleichzeitig sind sie finanziell und institutionell gut für ihren Regelbetrieb abgesichert. Sie übernehmen Aufgaben auch für kleinere Einrichtungen. Grundsätzlich sind sie so aufgestellt, dass sie auf zukünftig eventuell auftauchende neue Anforderungen reagieren können. Insgesamt könnten sie auch mit den Anforderungen der datengetriebenen Wissenschaft umgehen, wenn diese einmal real an sie herangetragen werden. Die Plattformen sind an sich offen und an Weiterentwicklungen interessiert, solange diese finanziert wird.

Die Strukturen der Plattformen bauen auf dem föderal organisierten System von Wissenschaftseinrichtungen und Bibliotheken der Schweiz auf. Sie sind nicht zentral, sondern kollaborativ organisiert.

Grundsätzlich scheint es so, als würde die tatsächlich von den Plattformen geleistete Arbeit unterschätzt, insbesondere das Engagement bei der Unterstützung kleinerer Einrichtungen bei Digitalisierungsprojekten. Dafür werden die aktuellen Anforderungen an die Plattformen im Bezug auf das Angebot von Daten, Linked Data und Metadaten tendenziell überschätzt. Wie gesagt, wären die Plattformen grundsätzlich in der Lage, auf diese Anforderungen zu reagieren, wenn sie aufkommen oder wissenschaftspolitisch gefordert würden, solange dafür Ressourcen zur Verfügung gestellt werden. Die Infrastruktur und das Wissen für solche Weiterentwicklungen sind aufgebaut. Ein Potential der Wissenschaftsförderung kann es sein, solche Ressourcen (vor allem finanzieller Art) zur Verfügung zu stellen, wenn die Anforderungen aus der Wissenschaft tatsächlich auftreten. Momentan muss dies noch mit Vorsicht und kritisch betrachtet werden, auch wenn diese Erwartungen und Hoffnungen alenthalben geäußert werden.

Eine Öffnung der Schnittstellen und die Bereitstellung der Inhalte als Open Data/Open Linked Data scheinen aber grundsätzlich möglich und würden auch die Sichtbarkeit der Bestände und die Einzelnutzung erleichtern. Der Aufwand dafür sollte von den Plattformen abgeschätzt werden, ebenso sollten die Plattformen bei eigenem Interesse dafür Projekte entwickeln. Eine Förderung der Weiterentwicklung der Plattformen, z.B. aus dem Programm "Wissenschaftliche Information" (2017-2020 P-5), erscheint als durchaus sinnvoll und empfehlenswert.

Aktuell besteht die Nachfrage eher in der finanziellen Unterstützung bei der Digitalisierung von Inhalten für Forschungszwecke, die dann in eine der bestehenden Plattformen integriert werden können.

Die „Strategie Digitalisierung“ des Programms "Wissenschaftliche Information", insbesondere aber das in diesem Rahmen geförderte Projekt DLCM sowie das von der Schweizerischen Akademie der Geistes- und Sozialwissenschaften geförderte DaSCH beschäftigen sich mit dem Management und der Archivierung von Forschungsdaten selber. Bislang ist dies keine Anforderung, die an die in dieser Studie untersuchten Plattformen herangetragen wurde. Vielmehr scheint sich – wie sich auch in der Literatur zu tatsächlich durchgeführten Forschung in den Digital Humanities zeigte – in der Forschungspraxis eine Differenz zwischen Sammlungen von digitalen Objekten (bzw. digitalisierten Medien), die als Datengrundlage genutzt werden können, und Forschungsdaten, die im Forschungsprozess entstehen, zu zeigen. Konkret zeigt sich dies bei den digitalen Editionen, die in der Regel Digitalisate aus Archiven und Bibliotheken nutzen, diese aber in eigenen Systemen bearbeiten, annotieren und kommentieren und schliesslich publizieren. Denkbar wäre es, diese Anwendungen technisch

zusammenzuführen, aber nur, wenn die Plattformen dabei finanziell massiv unterstützt würden. Dann wäre, wie in DaSCH für Daten aus Forschungsprozessen angedacht, eine sehr flexible Unterstützung von Forschungsdaten in unterschiedlichen Formaten und Strukturen nötig. Ausser der theoretischen Möglichkeit scheint aber wenig dafür zu sprechen, dass solche Services sinnvoll sind, besonders, wenn sie sich – wie es z.B. erklärtes Ziel des Programms “Wissenschaftliche Information“ ist – mittelfristig selbst tragen sollen. Die Forschenden scheinen sie nicht einzufordern. Sie benötigen „nur“ direkten und offenen Zugang zu den digitalen Objekten, um sie selbständig und in ihren Systemen be- und verarbeiten zu können.

Als Potential der Plattformen ist die bessere Dokumentation der eigenen Angebote nach aussen zu nennen. Die schon vorhandenen Schnittstellen, digitalen Inhalte, Unterstützungsleistungen und Betriebskonzepte werden zumeist erst genannt, wenn die Plattformen direkt danach gefragt werden. Gerade für Forschende, die bislang weder mit Digitalisaten noch mit Big Data umgehen, wäre es sinnvoll, auch ohne solche Anfragen auf die Möglichkeiten der Plattformen hingewiesen zu werden, z.B. durch bessere Dokumentation auf der jeweiligen Homepage oder durch Öffentlichkeitsarbeit in den unterschiedlichen Communities.

Neue Plattformen, die in Zukunft als Teil der nationalen Forschungsinfrastruktur gefördert werden könnten – z.B. für andere Datenformen als Texte –, sollten von Beginn an auf Fragen der Interkonnektivität und Offenheit achten. Es ist nach den bisherigen Entwicklungen nicht davon auszugehen, dass heute mit absoluter Sicherheit ein in Zukunft durchgängig geltendes Metadatenformat für diese Interkonnektivität angegeben werden kann. Das Europeana Data Model wäre aber für solche Plattformen ein unbedingt zu prüfendes Format.

Für die bisher bestehenden Plattformen wäre zu empfehlen, dass diese sich untereinander auf ein Format einigen würden, das ihren jeweiligen Ansprüchen genügt, bei dem in den Institutionen klar ist, wie es generiert und wie auf Daten in diesem Format zugegriffen werden kann und das gleichzeitig gemeinsam genutzt werden kann. Die OAI-PMH-Schnittstellen sind dafür eine gute Voraussetzung, ebenso ist das Europeana-Datenmodell auch für diese Aufgabe ein unbedingt zu prüfendes Format. Es hätte den Vorteil, dass es von einer aktiven Community aus Einrichtungen mit den gleichen Herausforderungen wie den Institutionen, welche die schweizerischen Plattformen betreiben, umgehen müssen, weiterentwickelt wird.

IIIF ist als Modell im Blickfeld zu halten, kann aber erst für alle Plattformen empfohlen werden, wenn es auch standardisiert Video- und Audioformate unterstützt.

Literatur

- Antonijević, Smiljana: *Amongst digital humanists : an ethnographic study of digital knowledge production*. Basingstoke; Hampshire : Palgrave Macmillan, 2015
- Beer, Nikolaos ; Herold, Kristin ; Kolbmann, Wibke ; Kollatz, Thomas ; Romanello, Matteo ; Rose, Sebastian ; Walkowski, Niels-Oliver: *Interdisciplinary Interoperability*, DARIAH-DE working papers (Nr. 3). Göttingen, 2014
- Bender, Michael: *Forschungsumgebungen in den Digital Humanities: Nutzerbedarf - Wissenstransfer - Textualität, Sprache und Wissen*. Berlin ; Boston : Walter de Gruyter, 2016
- Bron, Marc ; Van Grop, Jasmijn ; de Rijke, Maarten: Media Studies Reserach in the Data-Driven Age: *How Reserach Questions Evolve*. In: Journal of the Association for Information Science and Technology Bd. 67 (2016), Nr. 7, S. 1535–1554
- Burdick, Anne ; Drucker, Johanna ; Lunenfeld, Peter ; Presner, Todd ; Schnapp, Jeffrey: *Digital Humanities*. Cambridge ; London : The MIT Press, 2012
- Clivaz, C. ; Gregory, A. ; Hamidović, D. (Hrsg.): *Digital Humanities in Biblical, Early Jewish and Early Christian Studies*, Scholarly Communication. Leiden ; Boston : Brill, 2014
- ETH-Bibliothek (Hrsg.): *Jahresbericht 2015 der ETH-Bibliothek*. Zürich: ETH-Bibliothek, 2016
- Ganascia, Jean-Gabriel: *The Logic of the Big Data Turn in Digital Literary Studies*. In: *frontiers in Digital Humanities* (2015)
- Gardiner, Eileen ; Musto, Ronald G.: *The Digital Humanities: A Primer for Students and Scholars*. New York : Cambridge University Press, 2015
- Graham, Shawn ; Milligan, Ian ; Weingart, Scott: *Exploring Big Historical Data: The Historian's Macroscopic*. London : Imperial College Press, 2015 — ISBN 978-1-78326-637-1
- Griffin, G. ; Hayler, M. (Hrsg.): *Research Methods for Reading Digital Data in the Digital Humanities*, Research Methods for the Arts and Humanities. Edinburgh : Edinburgh University Press, 2016
- Hartsell-Gundy, A. ; Braunstein, L. ; Golomb, L. (Hrsg.): *Digital Humanities in the Library: Challenges and Opportunities for Subject Specialists*. Chicago : American Library Association, 2015
- Humbel, Marco: *Die Umsetzung von Open Data an Wissenschaftlichen Bibliotheken der Schweiz: Eine qualitative Untersuchung*. Bachelor-Thesis HTW Chur, 2016
- Kaplan, Frédéric: *A map for big data research in digital humanities*. In: *frontiers in Digital Humanities* (2015)
- Kurz, Susanne: *Digital Humanities: Grundlagen und Technologien für die Praxis*. Wiesbaden : Springer Fachmedien, 2015
- Pierazzo, Elena: *Digital Scholarly Editing: Theories, Models and Methods*. Farnham ; Burlington : Ashgate, 2015
- Presner, Todd ; Shepard, David ; Kawano, Yoh: *HyperCities: Thick Mapping in the Digital Humanities*. Cambridge ; London : Harvard University Press, 2014
- Programm „Wissenschaftliche Information: Zugang, Verarbeitung und Speicherung“ (SUK 2013-2016 P-2): *Hauptstossrichtung Publikationen, Umsetzungsmassnahme EP-10, Digitalisierung. Angepasste Strategie und Umsetzungsmassnahmen*. Zürich, 2015
- Programm „Wissenschaftliche Information: Zugang, Verarbeitung und Speicherung (SUK 2013-2016 P-2): *White Paper for a Swiss Information Provisioning and Processing Infrastructure 2020*. Zürich, 2015
- Reicher, Ruth ; Becker, Rainer ; Bender, Michael ; Munson, Mathew ; Schmunk, Stefan ; Schöch, Christof: *Verfahren der Digital Humanities in den Geistes- und Kulturwissenschaften*, DARIAH-DE working papers (Nr. 4). Göttingen, 2014
- Ridolfo, Jim: *Digital Samaritans: Rhetorical Delivery and Engagement in the Digital Humanities, Digital Rhetoric Collaborative*. Michigan : University of Michigan Press, 2015
- Sabharwal, Arjun: *Digital Curation in the Digital Humanities: Preserving and promoting archival and special collections*, Chandos Information Professional Series. Amsterdam ; Boston ; Cambridge ; Heidelberg ; London ; New York ; Oxford ; Paris ; San Diego ; San Francisco ; Singapore ; Sydney ; Tokyo : Chandos, 2015 — ISBN 978-0-08-100143-1

- Schreibman, S. ; Siemens, R. ; Unsworth, J. (Hrsg.): *A new Companion to Digital Humanities, Blackwell Companions to Literatur and Culture*. Chichester, West Sussex : Wiley Blackwell, 2016
- Steggle, Matthew: *Digital Humanities and the Lost Drama of Early Modern England: Ten Case Studies, Studies in Performance and Early Modern Drama*. Farnham ; Burlington : Ashgate, 2015
- Terras, Melissa ; Nyhan, Julianne ; Vanhoutte, Edward: *Defining Digital Humanities: A Reader*. Farnham ; Burlington : Ashgate, 2013 — ISBN 9781409460636
- Thompson Klein, Julie: *Interdisciplining digital humanities: boundary work in an emerging field, Digital Humanities*. Ann Arbor : University of Michigan Press, 2015
- Wanger, R.; Ehrismann, M.: *E-Periodica: die Plattform für digitalisierte Schweizer Zeitschriften*. In: *Schriften des Forschungszentrums Jülich Reihe Bibliothek / Library Bd. 22* (2016), S. 161-175

Anhang A: Umfrage

Ziele des Projekts AISOOP

- Analyse des IST-Zustandes bezüglich Informationsarchitektur, eingesetzter Systeme, Schnittstellen, Metadatenstandards sowie der Organisationsstruktur im Hinblick auf eine Öffnung und Verstetigung der Plattformen.
- Analyse der Erhebung der Nutzung der Plattformen (verwendete Methoden, Kennzahlen und Standards)
- Bedarfsanalyse seitens der Wissenschaft: hinsichtlich Inhalten, Formaten und Funktionen von Online-Plattformen
- Definition des SOLL-Zustandes bezüglich Standards, Informationsarchitektur, Schnittstellen, Metadaten, Funktionen und Kennzahlen für die Nutzung.
- Analyse des GAPs zwischen IST und SOLL und Beschreibung möglicher Handlungsfelder.

Fragen

An Betreiber von Online-Plattformen

Name der Plattform:

URL:

Verantwortliche Person:

A Technische Aspekte

A.1 Wie sieht die Informationsarchitektur der Plattform aus (Hardware, Betriebssystem/e, Software)?

A.2 In welchen Formaten sind die Metadaten gespeichert?

A.3 In welchen Formaten sind die Daten gespeichert? Welche Dokumententypen kommen auf der Plattform vor?

A.4 In welchen Formaten können die Metadaten von Externen bezogen werden?

A.5 In welchen Formaten können die Daten/Dokumente von Externen bezogen werden?

A.6 Welche Arten von Schnittstellen stehen Externen für den Bezug der Metadaten und Daten zur Verfügung? Wie sind diese Schnittstellen dokumentiert?

A.7 Welches Potenzial wird dem IIIF (International Image Interoperability Framework) beigemessen?

A.8 Auf welche Art/en müssen sich Externe vor dem Bezug von Metadaten und Daten registrieren und authentifizieren?

A.9 Unter welchen Lizenzen/Bedingungen sind die Daten nutzbar?

A.10 Wie wird aufgezeichnet, wer wann welche Daten bezogen hat?

A.11 Sind die Inhalte über andere Plattformen zugänglich? Wie werden sie in diese Plattformen übertragen?

A.12 Gibt es interne Schnittstellen? Wenn ja, welche und wozu dienen sie?

A.13 Wie ist der Workflow organisiert (Digitalisierung, Erschliessung, Ingest, Speicherung, Backup, digitale Langzeitarchivierung, Qualitätssicherung, usw.)?

A.14 Wird OCR (und entsprechend Volltextsuche) angeboten?

A.15 Ist die Publikation der Metadaten in einem zentralen Katalog/Portal vorgesehen (z.B. auf openresearchdata.ch)?

A.16 Sind Bestrebungen im Gange, die Metadaten und/oder die Daten in einen Linked Open Data Graph einzubinden (z.B. über Europeana)?

A.17 Welche weiteren Entwicklungen sind geplant?

B Nutzungsaspekte

B.1 Wer nutzt die Plattform? Welche Zielgruppen werden angesprochen?

B.2 Wird die Plattform von Wissenschaftlern genutzt? Weiss man von welchen und wofür?

B.3 Wurde der Bedarf von Forschenden erhoben?

B.4 Wie wird die Nutzung gemessen? Welche Kennzahlen/Standards? Gibt es eine Empfehlung für einen Standard?

B.5 Wie hat sich die Nutzung in den letzten Jahren entwickelt?

B.6 Sind die Daten für eine Big-Data-Analyse durch externe Forscher zugänglich?

B.7 Können die Inhalte von einem Forschenden bearbeitet/annotiert werden? Wäre das wünschenswert? Ist das geplant?

B.8 Welchen Mehrwert erwarten Sie als Institution, wenn Sie Ihre Daten auf solchen Plattformen anbieten?

C Organisatorische Aspekte

C.1 Wie ist der Betrieb der Plattform organisiert?

C.2 Wie ist die Plattform finanziert? Wie sieht das Betriebskonzept aus? Welche Beiträge leisten die Partner?

C.3 Welche Mitwirkungsmöglichkeit haben andere Institutionen?

C.4 Welche Verfahren und Kooperationen sind mit Wikimedia bekannt oder werden von Ihnen angewendet?

C.5 Wie kann sich eine interessierte Institution beteiligen, die eigene Digitalisate veröffentlichen möchte?

C.6 Welche Entwicklungen sind geplant? Wie schätzen Sie die Veränderungen durch SLSP ein?

Prof. Dr. Rudolf Mumenthaler

Pulvermühlestrasse 57, 7004 Chur

rudolf.mumenthaler@htwchur.ch

081 286 37 19

Dr. Karsten Schuldt

Pulvermühlestrasse 57, 7004 Chur

karsten.schuldt@htwchur.ch

081 286 37 18

Prof. Bruno Wenk

Pulvermühlestrasse 57, 7004 Chur

bruno.wenk@htwchur.ch

081 286 24 45

Bearbeitungszeitraum: April-November 2016

Publikationsdatum: Februar 2017