

---

**eHumanities – interdisziplinär:**

**Zwischenbericht des Modellprojektes zum  
Forschungsdatenmanagement in den digitalen Geistes- und  
Sozialwissenschaften**

---

Autoren:

Anna Daudrich<sup>a</sup>, Rainer Gnan<sup>b</sup>, Sonja Kümmer<sup>b</sup>, Stephan Lücke<sup>c</sup>, Markus Putnings<sup>a</sup>, Christian Riepl<sup>c</sup>, Jürgen Rohrwild<sup>a</sup>, Volker Schallehn<sup>b</sup>, Julian Schulz<sup>c</sup>, Konstanze Söllner<sup>a</sup>, Martin Spenger<sup>b</sup>, Alexander Weiß<sup>b</sup> und Florian Zacherl<sup>c</sup>

Projektpartner:

- a) Universitätsbibliothek der Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU)
- b) Universitätsbibliothek der Ludwig-Maximilians-Universität München (LMU)
- c) IT-Gruppe Geisteswissenschaften der Ludwig-Maximilians-Universität München (LMU)

*Bisherige Projektlaufzeit:*

01.01.2018 – 31.12.2018

erstellt am 18.12.2018.

EINLEITUNG	3
AP1: UNTERSUCHUNGEN ZUM METADATENMODELL	3
Aktueller Stand	3
Nächste Schritte	6
Weiteres Vorgehen	6
Neue Ziele	6
AP2: VORSCHLAG EINES TRANSFERIERBAREN DATENMANAGEMENTPLANS	6
Aktueller Stand	6
Vergleich von Werkzeugen zur Erstellung von Datenmanagementplänen	6
Fachspezifische DMP-Templates	8
Umsetzung eines transferierbaren DMP	8
Empfehlung	9
Nächste Schritte	9
AP3: ENTWICKLUNG EINES ABGESTIMMTEN SCHULUNGSPAKETES	10
Aktueller Stand	10
Bedarfsanalyse	10
Festlegung des didaktischen Konzepts	11
Software- bzw. Plattformauswahl	11
Nächste Schritte	12
AP4: BESCHREIBUNG UND ETABLIERUNG VON SERVICES IM FDM DER AUSGEWÄHLTEN FACHDOMÄNEN	13
Aktueller Stand	13
Nächste Schritte	14
AP5: UNTERSTÜTZUNG BEI RECHERCHE UND PUBLIKATION UNTER VERWENDUNG ENTSPRECHENDER TOOLS	15
Aktueller Stand	15
Nächste Schritte	17
AUSBLICK	18
ANHANG	22
A: Evaluationskriterien für DMP	22
B: E-Learning Modulplan	24
C: Projektbeteiligte	26

## EINLEITUNG

Die Entstehung der Digital Humanities (DH) ist die natürliche Reaktion auf die fortschreitende Proliferation digitaler Daten und die Datenverarbeitung in den Geistes- und Sozialwissenschaften. *eHumanities - interdisziplinär* versteht sich als eine regionale Initiative, die Forschungsdaten-bezogene Dienstleistungen von Infrastruktureinrichtungen auf diese Entwicklung hin ausrichtet. Das Projekt gliedert sich in fünf Arbeitspakete (AP), in denen sowohl konkrete technische Lösungen (AP 1, 4, 5) als auch fachspezifische Unterstützungs- und Beratungsangebote (AP 2, 3) für diese Fachdomänen entwickelt werden. Um hierbei möglichst nahe an den Anforderungen der Wissenschaft zu bleiben, fließen sowohl die Bedürfnisse der Forscher<sup>1</sup> als auch die Detailstruktur der Forschungsdaten (FD) einer Vielzahl von DH-Projekten in die Planungen ein.

Nachfolgend werden die Ergebnisse der einzelnen AP beschrieben. Der sich anschließende Ausblick geht auf die zukünftige Ausrichtung von *eHumanities - interdisziplinär* in der nächsten Projektphase und auf Möglichkeiten ein, die Projektergebnisse auf andere Fachdomänen, insbesondere die Natur- und Ingenieurwissenschaften, zu übertragen.

### AP1: UNTERSUCHUNGEN ZUM METADATENMODELL

Ziel des Arbeitspakets ist es, ein auf den [FAIR-Prinzipien](#) basierendes Metadatenmodell zu finden bzw. zu erweitern, mit dem sich Forschungsdaten aus den DH einheitlich beschreiben lassen. Um die Auffindbarkeit der Forschungsdaten in Datenrepositorien und bibliothekarischen Nachschlagewerken und damit ihre Nachnutzung zu befördern, soll das Metadatenmodell nicht nur formale Metadaten enthalten, sondern auch eine tiefe inhaltliche Beschreibung der FD ermöglichen.<sup>2</sup>

Die in diesem Arbeitspaket erzielten Ergebnisse fließen in die anderen Arbeitspakete des Projekts (insb. AP 3 und 5) ein und sollen darüber hinaus zu einer Empfehlung für die DH ausgearbeitet werden.

#### Aktueller Stand

In einem ersten Schritt wurden die umfangreichen, an der IT-Gruppe Geisteswissenschaften (ITG) angesiedelten Datenbestände gesichtet und erschlossen. Unter den für das AP relevanten 66 Datenbeständen sind abgeschlossene und laufende Projekte. Die Daten stammen aus allen sechs geisteswissenschaftlichen Fakultäten der LMU und sind hinsichtlich der verwendeten Daten- und Dateiformate heterogen. Insofern stehen sie exemplarisch für

---

<sup>1</sup> Aus Gründen der Lesbarkeit wird im Text für Personengruppen die männliche Form verwendet, diese umfasst aber Personen aller Geschlechter.

<sup>2</sup> Zur Begriffsbestimmung und als Ausgangspunkt für weiterführende Informationen vgl. RfII - Rat für Informationsinfrastrukturen: Leistung aus Vielfalt. Empfehlungen zu Strukturen, Prozessen und Finanzierung des Forschungsdatenmanagements in Deutschland. Göttingen 2016, Anhang, S. A-18, 2.22 Metadaten.

die ausdifferenzierte geisteswissenschaftliche Forschungslandschaft und eignen sich daher sehr gut als primäre Datenbasis.

In einem zweiten Schritt wurden repräsentative Projekte aus diesem Set formal erschlossen. Als grundlegendes Metadatenschema wurde *DataCite* gewählt, da *DataCite*

- a) an Forschungsdatenzentren weit verbreitet und universell einsetzbar ist<sup>3</sup>
- b) Registrierungsschema für DOIs<sup>4</sup> ist und
- c) an der UB der LMU bereits eingesetzt sowie
- d) von [GeRDI](#) präferiert wird.

Die Auszeichnung der Projekte mit *DataCite* erfolgte in Kooperation mit den projektverantwortlichen Fachwissenschaftlern. Dadurch konnten wertvolle Erkenntnisse darüber gewonnen werden, an welchen Stellen es zusätzlicher Vorgaben bedarf, um eine einheitliche Befüllung der Felder und damit eine bessere Nachnutzbarkeit der Daten sicherzustellen.

Zusätzlich zu der Beschreibung auf Projektebene wurde bei dem DFG-Langfristvorhaben [VerbaAlpina](#) auch der zentrale Datenbestand erschlossen. Dies geschah zunächst auf einer Metaebene, indem die zentralen Datenkategorien von VerbaAlpina – Morpho-lexikalischer Typ, Konzept und Gemeinde<sup>5</sup> sowie die Basiskategorie der Einzelbelege – zu Schichten aggregiert und jeweils mit *DataCite* erschlossen wurden.

*VerbaAlpina* hat in mehrfacher Hinsicht Pilotcharakter: Zum einen mit Blick auf die Herausbildung und Etablierung von Workflows und Schnittstellen (vgl. AP 4), zum anderen hinsichtlich der hier getroffenen Überlegungen zur Erschließungstiefe. Mit der Beschreibung der atomaren Informationseinheiten (etwa 200.000 Datensätze) inklusive DOI-Vergabe schlägt das Projekt bisher wenig begangene Pfade ein.<sup>6</sup> Die Entscheidung für diese Granularität wird dadurch gerechtfertigt, dass die Referenzierbarkeit jedes einzelnen Forschungsdatums per DOI eine zentrale Anforderung der Wissenschaftler ist.

Parallel zur Erschließung mit *DataCite* wurde eine weitere Anforderung seitens der Fachwissenschaftler (vgl. AP 3) aufgegriffen und der Grundstein für eine Bereitstellung der FD über bibliothekarische Nachschlagewerke gelegt: Mit Unterstützung eines Experten aus der Medienbearbeitung der UB der LMU wurde mit der Arbeit an einem konzeptionellen Mapping von *DataCite* zu [MARC](#) begonnen.

---

<sup>3</sup> Zu den Vorteilen von DataCite gegenüber vergleichbaren Metadatenschemata siehe Tobias Weber / Dieter Kranzlmüller: How FAIR can you get? Image retrieval as a use case to calculate FAIR metrics. Garching 2018, S. 5f. <https://arxiv.org/abs/1810.03824>.

Große Datenrepositorien, die *DataCite* nutzen, sind unter anderem [Zenodo](#), [figshare](#), [Pangaea](#); siehe hierzu auch [re3data.org](#).

<sup>4</sup> *Digital Object Identifier*; vgl. <https://www.doi.org/>, siehe auch [VerbaAlpina Methodologie](#).

<sup>5</sup> Siehe dazu [Thomas Krefeld / Stephan Lücke: s.v. "Typisierung", in: VA-de 18/1, Methodologie](#) sowie [Thomas Krefeld: s.v. "Konzept", in: VA-de 18/1, Methodologie](#).

<sup>6</sup> Siehe auch [den Eintrag zum Forschungsdatenmanagement in der VerbaAlpina Methodologie](#).

*DataCite* eignet sich besonders gut als generisches, d.h. fächerübergreifendes Datenformat, da es von inhaltlichen Kriterien abstrahiert. Um jedoch den Anforderungen der FAIR-Prinzipien, insbesondere der Auffindbarkeit, gerecht zu werden, ist es erforderlich, neben formalen Metadaten, wie Autor, Titel, Erscheinungsjahr, auch den Inhalt eines Forschungsdatums präzise zu beschreiben. Da dies in *DataCite* nur rudimentär möglich ist, soll *DataCite* um ein inhaltsbeschreibendes Metadatenmodell ergänzt werden.

Die Entwicklung dieses Modells verläuft überwiegend bottom-up-getrieben, auf Grundlage der Datenmodelle ausgewählter Projekte. Die dort vorkommenden Entitäten (und Relationen) werden extrahiert und auf ihre Ähnlichkeit hin überprüft. In einem nächsten Schritt werden existierende Datenmodelle wie *EDM* und *CIDOC-CRM* daraufhin untersucht, ob sie die identifizierten Entitäten(typen) abbilden können. Darauf aufbauend lassen sich bezüglich des weiteren Vorgehens drei Optionen identifizieren: Im besten Fall kann auf ein bereits etabliertes Datenmodell zurückgegriffen werden. Ist die Abdeckung jedoch zu gering, ist zu überlegen, ob für das inhaltsbeschreibende Datenmodell ein bestehendes Format adaptiert oder auf einem Merging der ähnlichen Entitäten aufgesetzt werden soll.

Die Arbeiten im AP 1 erfolgen im kontinuierlichen Austausch mit anderen Akteuren im Bereich Forschungsdatenmanagement (FDM), um die Anschlussfähigkeit der Projektergebnisse zu gewährleisten. In einer Videokonferenz wurde mit den Metadatenexperten von *GeRDI* (TU Dresden, ZBW Kiel) sowie mit der Projektkoordinatorin der Hessischen FDM-Landesinitiative (*HeFDI*) über das *GeRDI*-Metadatenmodell (*DataCite* mit generischer Erweiterung) und über Möglichkeiten einer domänenübergreifenden Inhaltserschließung gesprochen.

Über den Arbeitskreis „Digitale Geistes- und Sozialwissenschaften in München“ - kurz *dhmuc* - fließen Anforderungen aus dem Bereich „GLAM“ (*Galleries, Libraries, Archives and Museums*) in die Entwicklung des Metadatenmodells ein.

Außerdem wurde eine Arbeitsgruppe mit Vertretern des Leibniz-Rechenzentrums der Bayerischen Akademie der Wissenschaften (LRZ) und der Technischen Universität München (TUM) gebildet. Die informelle Arbeitsgemeinschaft dient dem interdisziplinären Austausch über grundlegende (technische) Fragen, wie z.B. Datenmanagement und Metadaten.

Die bisherige Beschäftigung mit den ausgewählten Pilotprojekten hat gezeigt, dass von Seiten der ITG und der UB ein hoher Betreuungsaufwand erforderlich ist. Workflows lassen sich bis zu einem gewissen Grad standardisieren, doch wenn Projektspezifika zum Tragen kommen, beispielsweise bei der Datentransformation, wird der Zeit- und Ressourcenaufwand auch mit Fortschreiten des Projekts relativ hoch bleiben.

Jedoch lässt sich die Arbeit, konkret der Dateneingest in Open Data LMU, erleichtern, wenn es gelingt, den Datenproduzenten vom Zeitpunkt der Projektplanung an die Notwendigkeit einer strukturierten, normierten sowie plattform- und softwareunabhängigen Arbeitsweise zu vermitteln. Gleichzeitig gilt es, das Bewusstsein unter den Forschenden für den Einsatz von Normdaten und kontrollierten Vokabularen zu stärken, da dieser fundamentale Bedeutung für die weitere Verarbeitung der Daten hat.

## Nächste Schritte

### *Weiteres Vorgehen*

Nachdem der Prozess der formalen Erschließung im Pilotprojekt *VerbaAlpina* weitgehend abgeschlossen werden konnte, liegt in der sich anschließenden Projektphase der Schwerpunkt auf der Ausarbeitung des inhaltsbeschreibenden Metadatenschemas. Hierzu wird die Analyse des ITG-Datenbestands weiter vorangetrieben. Sobald ein Entwurf für eine inhaltsbeschreibende Ergänzung zu *DataCite* erarbeitet wurde, wird dieser der [DataCite Metadata Working Group](#) zur Diskussion vorgelegt.

Parallel dazu wird der Prozess der Erschließung auf Datensatzebene an einem weiteren ITG-Projekt getestet und optimiert. Um das Anforderungsspektrum auch im Hinblick auf bisher noch nicht berücksichtigte Datentypen möglichst breit zu gestalten, soll ein Projekt aus dem Bereich der Kunstwissenschaften gewählt werden.

### *Neue Ziele*

Es wird angestrebt, den Erfassungs-Workflow von Metadaten dahingehend zu automatisieren, dass die Eingaben in das DMP-Tool (siehe AP 2) zu Beginn des Projekts an zentraler Stelle abgelegt und nach Abschluss des Projekts in die Metadatenerzeugung einfließen können („DMP 2.0“). Hierfür wurde ein erster Test mit RDMO durchgeführt, dessen Ergebnisse als vielversprechend zu bewerten sind (vgl. AP 2). Im weiteren Projektverlauf sollen neben den formalen und inhaltlichen Metadaten auch strukturelle und technische Metadaten Gegenstand des Arbeitspakets sein. Die gewonnenen Metadaten sollen langfristig auch als *Linked-Open-Data* angeboten werden.

## AP2: VORSCHLAG EINES TRANSFERIERBAREN DATENMANAGEMENTPLANS

Datenmanagement-Pläne (DMP) dokumentieren den geplanten Umgang mit FD während eines Forschungsprojektes und darüber hinaus. Unerwartete und unvorhersehbare Entwicklungen sind ein natürlicher Bestandteil des Forschungsprozesses. Ein ausführlicher DMP minimiert jedoch das Risiko vorhersehbarer Komplikationen beim Datenhandling und sollte im Idealfall bereits in der Projektgenese erstellt werden.

Im Rahmen von *eHumanities - interdisziplinär* werden verschiedene Software-Werkzeuge zur Erstellung von DMP miteinander und mit generischen Anforderungen verglichen. Das Ziel ist ein auf die DH abgestimmter, nachnutzbarer DMP, dessen Konzept von anderen Disziplinen und Fächern übernommen werden kann.

## Aktueller Stand

### *Vergleich von Werkzeugen zur Erstellung von Datenmanagementplänen*

Zunächst wurde ein Katalog mit Anforderungen an eine DMP-Softwarelösung zusammengestellt (siehe Anhang A). Um die Nutzerseite besser berücksichtigen zu können, wurde auf die langjährigen Erfahrungen der ITG bei der Erstellung von DMP für die

Forschenden und auf entsprechende Bedarfsumfragen an der Philosophischen Fakultät der FAU zurückgegriffen (siehe AP 3). Dabei wurde unter anderem die Möglichkeit, eine lokale Installation betreiben zu können, als wichtiges Kriterium identifiziert.<sup>7</sup>

Nach Erstellung des Kriterienkataloges wurden verfügbare DMP-Tools recherchiert und soweit möglich in Form einer lokalen Testinstallation aufgesetzt. Die gesichteten Tools sind: [DMPTool](#), [DMPonline](#)<sup>8</sup>, [CLARIN-D Wizard](#), [roadmap](#), [RDMO](#), DMP-Tool der Universität Bielefeld, [IEDA DMPT](#) und [GFBio DMPT](#).<sup>9</sup>

Von den untersuchten Werkzeugen wurden *roadmap* und *RDMO* als vielversprechende Optionen identifiziert.<sup>10</sup> Beide ermöglichen leicht umsetzbare lokale Installationen und bieten weitgehende Freiheiten in der Konfiguration und Erstellung von neuen Planvorlagen (Templates). Gleichzeitig werden die Werkzeuge mit Blick auf „machine-actionability“ von größeren Teams weiterentwickelt – also auf die Möglichkeit hin, die Informationen aus dem DMP automatisch in andere Systeme zu exportieren und zu integrieren.

*roadmap* bietet standardmäßig Optionen für Beispielantworten und wiederverwendbare Hilfetexte (sogenannte „Themes“). Die Mehrheit der vielen verfügbaren Vorlagen ist hierbei explizit auf die Anforderungen einzelner Förderorganisationen zugeschnitten. Demgegenüber erlaubt *RDMO*, mit Hilfe von Wenn-Dann-Bedingungen komplexere Vorlagen zu gestalten. Dies wird in den öffentlichen *RDMO*-Templates genutzt, um den gesamten Planungsprozess detailreich abzubilden, auch wenn sie auf einen Förderer wie den Schweizerischen Nationalfonds zugeschnitten sind. In der Praxis zeigte sich, dass das *roadmap*-Backend jedoch eine merklich flachere Lernkurve hat und deutlich schnelleres Erstellen von Vorlagen ermöglicht.

Beide Werkzeuge sind vollumfänglich für den Einsatz zur Forschungsunterstützung geeignet. Es ist jedoch zu erwarten, dass die primäre Schwäche von *RDMO*, der relativ komplexe Erstellungsprozess für Vorlagen und der damit verbundene Arbeitsaufwand auf administrativer Seite, langfristig durch transferierbare DMP innerhalb der in Deutschland bereits etablierten Community kompensiert werden wird.

In Anbetracht der wachsenden Zahl verfügbarer DMP-Tools und den weitgehenden Konfigurationsmöglichkeiten der bereits etablierten Tools wurde auf den Versuch einer Eigenentwicklung eines DMP-Tools verzichtet. Insbesondere mit Blick auf die Nachnutzbarkeit

---

<sup>7</sup> Zum einen gibt es in der heutigen kompetitiven Forschungslandschaft Vorbehalte gegenüber Systemen, die von Dritten betrieben werden. Zum anderen besteht der Wunsch nach Ansprechpartnern vor Ort, die bei Problemen direkt in das System eingreifen können.

<sup>8</sup> Im März 2018 wurden DMPTool und DMPonline auf *roadmap* umgestellt. Die getrennte Weiterentwicklung der beiden ursprünglichen Tools ist damit de facto eingestellt; hiermit war eine weitere Evaluation nicht mehr zielführend, da es bereits mittelfristig nur noch eine Softwarelösung geben wird.

<sup>9</sup> Noch nicht untersucht wurde beispielsweise [TUB-DMP](#) der TU Berlin, dessen Sourcecode erst im Juni 2018 frei zugänglich war oder der [UQ Research Data Manager](#) der University of Queensland, dessen DMP-Tool nur eine Komponente eines komplexen Datenmanagementsystems darstellt.

<sup>10</sup> IEDA DMPT wurde nicht genauer untersucht, da es durch ezDMP ersetzt werden wird.

von an verschiedenen Forschungseinrichtungen bereits vorhandenen Hilfestellungen und Vorlagen für DMP scheint eine weitere Fragmentierung der verwendeten Plattformen nicht zielführend zu sein.

### *Fachspezifische DMP-Templates*

Ein Template besteht in der Regel aus einem Fragenkatalog, der entweder allgemein gehalten ist oder auf Anforderungen von Fachgesellschaften und Förderern zugeschnitten sein kann, sowie entsprechend zugeordneten Hilfestellungen, Beispielantworten und Zusatzinformationen. Um ein möglichst modulares und leicht transferierbares Template zu gewährleisten, hat sich in der praktischen Projektarbeit eine Aufteilung in fachspezifische, organisationsspezifische und allgemeingültige Elemente als vorteilhaft erwiesen.

Die Trennung dieser Aspekte erleichtert die Nachnutzung allgemeiner und fachspezifischer Informationen an anderen Einrichtungen, ohne Veränderungen an den Templates vornehmen zu müssen. Lediglich die lokalen Gegebenheiten, etwa die institutionelle FD-Policy und Kontaktstellen, müssen angepasst werden. Gerade für kleinere Hochschulen, die nur wenige Personalressourcen haben, um umfangreiche Templates an lokale Besonderheiten anzupassen, ist dies hilfreich. Ferner gestattet die Aufteilung in generische und disziplinspezifische Informationen die Erweiterung des Templates auf andere Fachbereiche.

*roadmap* erlaubt eine grobe Trennung nach diesem Vorbild auf Basis der sogenannten „[DCC themes](#)“, d. h. von Themenkomplexen, die jeder DMP abdecken sollte und zu denen verschiedene allgemeine und fachspezifische Hilfetexte bereitgestellt werden können, die der Nutzer auswählen kann. In *RDMO* kann eine strikte (fachliche) Aufgliederung auf Ebene der Fragenkataloge nur durch Konditional-Klauseln erreicht werden. Allerdings besteht seitens einiger Forscher der Wunsch, im interdisziplinären Feld der DH stets auf alle Hilfestellungen zugreifen zu können. Daher wurden die *RDMO*-Fragebögen modifiziert, um verschiedene fachspezifische Informationen anbieten zu können.

### *Umsetzung eines transferierbaren DMP*

Zunächst wurden in den lokalen Testinstanzen der beiden Tools Vorlagen für Datenmanagementpläne erstellt und mit fach- und einrichtungsspezifischen Informationen angereichert. Dabei wurden jeweils zwei verschiedene Templates eingerichtet: ein ausführliches, generisches Template auf Basis des Standard-*RDMO*-DMP und eines für das Horizont-2020-Programm der EU. Hierfür wurden zunächst vier Fachrichtungen (Alturumswissenschaften & (Kunst-)Geschichte, Sozialwissenschaften, Erziehungswissenschaften & Psychologie sowie Sprach- und Literaturwissenschaften) festgelegt und danach entsprechende fachliche Hilfestellungen und Beispiele entwickelt. Entsprechende Texte für die DCC Themes finden sich auf [fdm-bayern.org](http://fdm-bayern.org).<sup>11</sup>

---

<sup>11</sup> [Die Hilfetexte in Form der vom DCC etablierten Themes können über die Webseite eingesehen werden.](#)

In der Praxis zeigt sich, dass die eingeführte Aufteilung der Fachdisziplinen in vier Gruppen noch zu grob ist. Dies ist besonders spürbar in der Gruppe Literatur- und Sprachwissenschaften sowie bei Fachbereichen, die in anderen Geisteswissenschaften wenig verbreitete Methodologien einsetzen, wie etwa der Psychologie.<sup>12</sup> Daher ist eine noch feinere fachliche Abstufung von Vorteil. Dies ist auch mit Blick auf die Nachnutzbarkeit erstrebenswert, da so eine Abstimmung der Templates auf den lokalen fachlichen Zuschnitt der Hochschulen erleichtert wird.

### Empfehlung

Grundsätzlich sind sowohl *RDMO* als auch *roadmap* für den Einsatz als konfigurierbares, disziplinübergreifendes DMP-Werkzeug gut geeignet. *roadmap* bietet momentan noch mehr Funktionen als *RDMO*, ein niedrigschwelliges Backend und eine große internationale Community. *RDMO* kompensiert dies allerdings durch eine bereits etablierte deutschsprachige Nutzergemeinschaft, erweiterte Funktionalitäten beim Konstruieren von Templates und ein mittelfristig größeres Angebot an nachnutzbaren, deutschsprachigen Vorlagen. Welches Tool eingesetzt werden sollte, ist daher auch von den institutionellen Gegebenheiten abhängig. Für Hochschulen mit engen Kooperationen mit anglo-amerikanischen Einrichtungen kann beispielsweise die weite Verbreitung von *roadmap* in den USA ausschlaggebend sein.

### Nächste Schritte

Wie in AP 1 dargestellt, sind viele der von Forschenden in ein DMP-Tool eingebrachten Informationen auch für die Metadaten-Generierung relevant. Eine enge Verzahnung von lokalem Datenrepositorium und DMP-Tool kann daher Doppelarbeit für die Forschenden vermeiden. Exemplarisch wurde hierfür bereits in *RDMO* eine Export-Option (via einer sogenannten *DataCite-View*) eingerichtet, die über die einzelnen Datensätze vorhandene Informationen in das *DataCite*-Metadatenchema abbildet. In einem nächsten Schritt wird an einer Umsetzung über die API des DMP-Werkzeugs gearbeitet. Im Zuge dessen ist geplant, das Tool im nächsten Schritt auch an der LMU einzurichten.

Weiter wird eine feinere fachliche Aufgliederung der Hilfetexte angestrebt. Zwar setzt die teils sehr interdisziplinäre Arbeitsweise in den DH der Granularität gewisse Grenzen. Doch scheint eine größere Anzahl von fachlichen Texten die Transferierbarkeit der Templates zu verbessern, da einzelne Einrichtungen die für sie relevanten Fächer auswählen können. Schließlich soll eine einfachere Möglichkeit entwickelt werden, mit der die institutionsspezifischen Anpassungen an den DMP-Templates direkt über eine Eingabemaske

---

<sup>12</sup> Literaturwissenschaften und Linguistik verwenden sehr unterschiedliche Datentypen/-strukturen (etwa digitale Editionen oder linguistische Textkorpora), Analysemethoden und Datenquellen. Daher sollten sie bei der Erstellung von Hilfetexten als eigenständig betrachtet werden. Die Psychologie setzt Methoden, die in der Medizin verbreitet sind, ein (beispielsweise fMRT-Aufnahmen) und forscht an und mit Patienten. Damit entstehen spezielle Anforderungen an das Datenmanagement, insbesondere bei den Themenbereichen Datenschutz und Ethik, die noch deutlich komplexer sind als beispielsweise in den Erziehungswissenschaften. Die Psychologie kann daher von eigens auf sie abgestimmten detaillierten Hilfestellungen profitieren.

vorgenommen werden können. Hierzu ist geplant, die wichtigsten, wiederholt auftretenden institutionellen Informationen (beispielsweise Kontaktstelle für Datenschutzfragen, institutionelles Repositorium oder FD-Policy) in den DMP-Vorlagen durch Platzhalter darzustellen und die Inhalte über eine separate Eingabeoberfläche von Anwendern des Tools zu erfragen. Die DMP-Vorlagen können damit auf einfache Weise angepasst werden – ohne manuellen Eingriff in oder tieferes Wissen über die Struktur der XML-Vorlage.

### AP3: ENTWICKLUNG EINES ABGESTIMMTEN SCHULUNGSPAKETES

Im Rahmen dieses APs werden modulare Lehrinhalte zum FDM entwickelt. Es wird ein auf die Bedürfnisse der DH angelegtes zeit- und ortsunabhängiges sowie nachnutzbares E-Learning-Angebot aufgebaut. Das Online-Schulungspaket soll sowohl für die Lehre als auch in der Fortbildung einsetzbar sein und neben fachunabhängigen Lerninhalten gleichermaßen auch fachspezifische Themen berücksichtigen. Im Probelauf werden die entwickelten Module zunächst an der FAU und im Anschluss an die darauffolgende Evaluation und Verbesserung auch an der LMU angewandt. Eine breite institutionsübergreifende Nachnutzbarkeit der Lehrangebote als Open Educational Resources ist dabei ein zentrales Ziel.

Aktueller Stand

#### *Bedarfsanalyse*

Die ersten Arbeitsschritte galten der Erarbeitung einer empirisch belegten Grundlage, die als Ausgangspunkt und erste Orientierung bei der Konzeption der Module fungieren wird. Ausgehend von der Zielsetzung, ein zielgruppen- und bedarfsorientiertes Lernangebot zu entwickeln, wurde eine Umfeldanalyse zum gegenwärtig praktizierten FDM in den DH durchgeführt. Mittels einer Online-Umfrage und ergänzenden Leitfaden-Interviews wurden dabei Wissenschaftler der Philosophischen Fakultät und des Fachbereichs Theologie der FAU befragt. In diesem Zusammenhang wurden Informationen zum aktuellen Umgang mit FD sowie zum Bedarf an entsprechender Schulung und Beratung ermittelt bzw. abgeschätzt.<sup>13</sup>

Anhand der durchgeführten Umfeldanalyse konnte zunächst positiv festgehalten werden, dass das Bewusstsein für den Mehrwert der nachhaltigen Verfügbarkeit und interdisziplinären Nachnutzbarkeit von FD in den DH an der FAU stark ausgeprägt ist. Die grundsätzliche Bereitschaft, eigene FD anderen zur Verfügung zu stellen, ist weitgehend vorhanden. Ein großer Bedarf an Sensibilisierung und Beratung besteht allerdings in Fragen der konkreten Handhabung der Daten über ihren gesamten „Lebenszyklus“<sup>14</sup> hinweg. Die Wissenslücken reichen von der mehrheitlich unzureichenden Planung des Datenmanagements über mangelnde Kompetenzen bei der Auswahl der Datenformate, Software und

---

<sup>13</sup> Ausführliche Evaluation siehe hier: <https://www.fdm-bayern.org/files/2018/11/forschungsdatenmanagement-in-den-geisteswissenschaften-an-der-fau-umfrage.pdf>

<sup>14</sup> Vgl. <https://www.fdm-bayern.org/startseite/lebenszyklus-von-forschungsdaten/>

Metadatenstandards bis hin zu Unsicherheiten im Umgang mit rechtlichen Themen sowie mit Repositorien und Datenarchiven.

Es ist davon auszugehen, dass im Rahmen der vorhandenen Ressourcen nicht alle ermittelten Wissenslücken durch entsprechende E-Learning-Module innerhalb der Projektlaufzeit vollständig abgedeckt werden können. Die Priorisierung der Themenwahl wird sich bei der Entwicklung einzelner Lerneinheiten nach primären Bedürfnissen der Wissenschaft richten.

#### *Festlegung des didaktischen Konzepts*

Wie aus der Bedarfserhebung hervorging, ist sich die Mehrheit der Forschenden der besagten Defizite bewusst und befürwortet ausdrücklich die Durchführung entsprechender Maßnahmen zur Kompetenzvermittlung. Der Wunsch nach einem praktischen Training zur Sicherstellung des Wissenstransfers in den Forschungsalltag wurde dabei vielfach geäußert.

Ausgehend von diesem konkreten Bedarf wird, anknüpfend an Lehrmethoden des *Constructive Alignment*,<sup>15</sup> ein handlungs- und ergebnisorientiertes Lernen den didaktischen Baustein der E-Learning-Module bilden. Im Unterschied zu einem inhaltsorientierten Ansatz, in dessen Rahmen die Lerninhalte über anwendungsspezifische Interaktionen im Vordergrund stehen, konzentriert sich das ergebnisorientierte Lernen auf den tatsächlichen Wissenserwerb und das Erreichen der Lernziele. Letzteres kann durch eine oder mehrere Prüfungs- bzw. Kontrollaufgaben (z.B. in Form von Selbsttests, Quizzes oder auch durch Feedback in einer Präsenzveranstaltung) gemessen werden. Ziel ist es, die zu Beginn jedes Moduls kommunizierten Lernziele in den Testaufgaben systematisch widerzuspiegeln. Individuelle Lernfortschritte werden dadurch sichtbar, was wiederum zur Steigerung der Lernmotivation führen könnte.

#### *Software- bzw. Plattformauswahl*

Für das geplante E-Learning standen die Lernplattformen *ILIAS* und *Moodle* in der engeren Auswahl. Die beiden gehören aktuell zu den verbreitetsten Learning Management Systemen (LMS) an deutschen Hochschulen und stellen eine Vielzahl an notwendigen Funktionalitäten und Werkzeugen zur Verfügung. Auch wenn sich beide LMS für das Vorhaben des Projektes grundsätzlich eignen, stellt *ILIAS* im Rahmen der FAU – nicht zuletzt aus pragmatischen Gründen – die bessere Alternative dar: Die *ILIAS*-basierte Lernumgebung *StudOn*<sup>16</sup> ist die zentrale Lehr- und Lernplattform der Universität; eine Unterstützung durch zuständige Mitarbeiter ist daher vor Ort gewährleistet. Zudem dürfte die den Forschenden der FAU vertraute Lernplattform die Akzeptanz des zu entwickelnden E-Learnings an der Universität begünstigen.

---

<sup>15</sup> Vgl. <https://www.e-teaching.org/didaktik/konzeption/constructive-alignment/constructive-alignment/>; siehe auch John Biggs / Catherine Tang: *Teaching for Quality Learning at University: What the Student Does*. 4. Aufl. Maidenhead 2011: Open University Press.

<sup>16</sup> Vgl. <https://www.studon.fau.de/studon/>

Dennoch werden die Möglichkeiten einer möglichst reibungslosen Importierbarkeit der Module aus *ILIAS* in *Moodle* stets berücksichtigt, sodass die anschließende Erprobung an der LMU, die *Moodle* als zentrale Lernplattform betreibt, ohne Verzögerungen erfolgen kann.

#### Nächste Schritte

Ausgehend von der beschriebenen Strategie ist die Konzeption und Erprobung der E-Learning-Module in der anschließenden Projektphase vorgesehen. Die Arbeiten erfolgen unter Beratung des Instituts für Lerninnovation (ILI) der FAU. In diesem Zusammenhang wurde die Empfehlung ausgesprochen, das E-Learning zunächst schrittweise einzuführen und seine Anwendbarkeit zu prüfen, um frühzeitig Feedback von zukünftigen Anwendern zu erhalten und somit aufwändige Korrekturen zu minimieren.

Hierfür empfiehlt sich die Entwicklung eines zunächst komplexitätsreduzierten, jedoch funktionsfähigen prototypischen Moduls, das in einem ersten Schritt auf die Anforderungen eines eingeschränkten Anwenderkreises zugeschnitten sein wird. Als Hauptbenutzergruppe werden Angehörige des wissenschaftlichen Nachwuchses fungieren, zu dem Masterstudierende sowie Promovierende zu zählen sind. Diese Zielgruppe generiert – im Gegensatz zu Bachelorstudierenden – bereits eigene Forschungsdaten, benötigt allerdings häufig allgemeine FDM-Kompetenzen, um diese nachhaltig zu verwalten. Fortgeschrittene Forschende – wie etwa wissenschaftliche Mitarbeiter oder Postdoktoranden – wünschen dagegen in der Regel die Beratung zu konkreten Fragestellungen, die besser individuell geklärt werden können.

Das Thema für das Prototyp-Modul wurde folglich anhand der ermittelten primären Bedürfnisse und Wissenslücken der adressierten Fokusgruppe festgelegt. Diese umfassen im Wesentlichen folgende thematische Aspekte: Datentypen und Datenformate, Metadaten, ethische und rechtliche Aspekte sowie Repositorien und Datenarchive. Dem letzten Themenbereich werden sich die ersten prototypischen Lerninhalte widmen, wobei das konkrete Unterthema „Forschungsdaten suchen und nachnutzen“<sup>17</sup> im Fokus stehen wird. Dieser Schwerpunkt ist dadurch zu begründen, dass er sowohl für Masterstudierende als auch für Promovierende eine hohe Relevanz besitzt, während das Teilthema „Forschungsdaten archivieren und publizieren“ in der Regel erst gegen Ende der Promotionsphase an Aktualität gewinnt. Ein weiteres gewichtiges Argument für die Auswahl des genannten Themenbereiches stellen bestehende Präsenzs Schulungen der UB FAU zur fachspezifischen und fachübergreifenden Datenbankrecherche dar. Diese können als erste Orientierung bei der Konzeption der Lerninhalte dienen und bieten zudem die Möglichkeit, das E-Learning-Modul im Kontext eines Blended-Learnings zu erproben.

---

<sup>17</sup> Das Grobkonzept des Prototyp-Moduls befindet sich im Anhang B.

#### AP4: BESCHREIBUNG UND ETABLIERUNG VON SERVICES IM FDM DER AUSGEWÄHLTEN FACHDOMÄNEN

Ziel des Arbeitspakets 4 ist es, Dienstleistungen im Bereich FDM zu entwickeln und zu etablieren, die den gesamten Lebenszyklus von Forschungsdaten abdecken. Modellcharakter hat dabei die an der LMU durch die Projektarbeit entstandene, enge Kooperation zwischen der ITG, einer Einrichtung, die eng mit der Fachcommunity verzahnt ist, und der UB, einer wissenschaftlichen Bibliothek. Da sich bereits abzeichnet, dass sich diese Organisationsform zu einem Erfolgsmodell entwickeln könnte, sollen im Rahmen von AP4 die Rollenverteilung zwischen ITG und UB, sowie die im Entstehen begriffenen Workflows dokumentiert werden, um von anderen Akteuren im Bereich FDM adaptiert werden zu können.

##### Aktueller Stand

Die Aufgaben, die sich je nach Projektphase (ab ovo, in vita, post mortem) auf die ITG und die UB verteilen, orientieren sich am Grundsatzpapier<sup>18</sup> der AG Datenzentren im [DHd](#). So sind Beratung, Projektbegleitung und Bereitstellung von IT-Infrastruktur zentrale Handlungsfelder der ITG (Schwerpunkt ab ovo und in vita). An der UB (Schwerpunkt post mortem) ist die Erschließung und die langfristige Verfügbarhaltung der Forschungsdaten angesiedelt, ferner die Vergabe von DOIs (bedeutsam v. a. für die von den FAIR-Prinzipien geforderte Interoperabilität), Rechtemanagement sowie die Konfiguration von Zugriffskontrollen.

Der von den Datenproduzenten in den Fachdisziplinen ausgehende Datenfluss über die ITG zur UB und von dort zu überregionalen Aggregatoren setzt sich aus mehreren Arbeitspaketen bzw. Prozessen zusammen, die jeweils über einen Workflow gesteuert werden sollen. Um sicherzustellen, dass diese Prozessabläufe praxistauglich sind und nach einer Testphase von anderen Institutionen genutzt werden können, werden sie anhand des Pilotprojekts *VerbaAlpina*<sup>19</sup> erarbeitet und erprobt. Die Wahl fiel auf *VerbaAlpina*, da zu Projektstart bereits eine Kooperation zwischen *VerbaAlpina* und *GeRDI* bestand. Diese wurde im Zuge des Projekts um die UB erweitert. Die ITG ist seit Beginn maßgeblich an Planung und Umsetzung einschließlich Hosting von *VerbaAlpina* beteiligt.

Eine Besonderheit von *VerbaAlpina* ist, dass sich das Projekt regelmäßiger Versionierungen bedient, um den Zugriff auf und die Zitierfähigkeit von Forschungsdaten bereits vor Abschluss des Forschungsvorhabens zu ermöglichen bzw. zu garantieren. Diese Versionierungen stellen eine Herausforderung für das Forschungsdatenmanagement dar, da nicht nur der Datensatz als solcher, sondern auch seine verschiedenen Versionen zu Gegenständen des Forschungsdatenmanagements werden. Konkret bedeutet dies, dass mit jeder neuen Version

---

<sup>18</sup> Zu Begriffsbestimmung, Abgrenzung und Leistungsspektrum von Datenzentren sowie Projektphasen und Aufgaben siehe DHd AG Datenzentren: Geisteswissenschaftliche Datenzentren im deutschsprachigen Raum – Grundsatzpapier zur Sicherung der langfristigen Verfügbarkeit von Forschungsdaten (Version 1.0), S. 5, 10, 12 und 16-19. Zenodo, 03.02.2018. Link: <http://doi.org/10.5281/zenodo.1134760>.

<sup>19</sup> Das Projekt wird von zehn Mitarbeitern aus unterschiedlichen Disziplinen betreut und zusätzlich von über 20 externen und ehemaligen Mitarbeitern unterstützt. Seit Oktober 2014 (mit Perspektive auf Verlängerung bis 2025) wird das Projekt durch die DFG gefördert. (Link: <https://www.verba-alpina.gwi.uni-muenchen.de/>).

von *VerbaAlpina* nicht nur neue Datensätze mit Metadaten ausgezeichnet und einer DOI versehen werden, sondern auch bereits in der Vorgängerversion enthaltene Datensätze, die sich in der Zwischenzeit geändert haben.

Die Komplexität, die hieraus erwächst, erfordert eine enge Abstimmung zwischen Datengeber und -nehmer, in diesem Fall der ITG und UB. Die Vorteile dieser Kooperation kommen besonders auch beim konzeptionellen Mapping, d.h. der Auszeichnung der Forschungsdaten mit Metadaten, zum Tragen, da hierbei sowohl Fachwissen als auch Erfahrung im Bereich Metadatenmanagement erforderlich sind.

Nach Abschluß des konzeptionellen Mappings beginnen an der UB<sup>20</sup> die Arbeiten am technischen Mapping, d. h. es werden auf Basis des konzeptionellen Mappings Metadaten im *DataCite-XML*-Format erzeugt. Anschließend werden die Forschungsdaten mit den zugehörigen Metadaten in das Forschungsdatenrepositorium der LMU (*Open Data LMU*) eingespielt und jeweils mit einem persistenten Link (DOI) versehen.

Um die FD auch über andere Sucheinstiege auffindbar zu machen, betreibt die UB der LMU eine OAI-PMH-Schnittstelle. Über diese können *GeRDI*, *Google Dataset Search*, *BASE (Bielefeld Academic Search Engine)* und weitere Plattformen die Metadaten der Forschungsdaten harvesten und in ihre Suchindizes einbinden.<sup>21</sup> Mit den *GeRDI*-Vertretern am LRZ steht das Projektteam in engem Austausch, um die Kompatibilität und damit eine bestmögliche Nachnutzbarkeit der Daten zu gewährleisten. In regelmäßigen Treffen werden das Metadatenschema, die Belegung der Felder sowie zusätzliche Meta-Informationen, wie z. B. die Angabe einer File-Schnittstelle, über die direkt auf die FD zugegriffen werden kann, diskutiert.

#### Nächste Schritte

Mit *VerbaAlpina* wurde ein Pilotprojekt gewählt, das sich bereits in einer fortgeschrittenen Projektphase befindet. Um die Prozesse in der ITG stärker auf FDM auszurichten und schließlich zu standardisieren, sollen im weiteren Projektverlauf gezielt Forschungsprojekte in der Antragsphase in die Projektarbeit einbezogen werden.

Aufgrund der Komplexität des skizzierten Datenflusses und der Beteiligung verschiedener Akteure ist es notwendig, die sich derzeit bei der Beschäftigung mit *VerbaAlpina* herausbildenden Arbeitsabläufe zu standardisieren. Angestrebt werden transparente, zwischen ITG, UB und *GeRDI* abgestimmte Workflows, die von nachfolgenden Projekten sowie von anderen Institutionen nachgenutzt werden können. Dies soll unterstützt werden durch die Konzeption einer Muster-Kooperationsvereinbarung und die Erstellung anwenderspezifischer Handreichungen.

---

<sup>20</sup> Die technische Infrastruktur befindet sich momentan im Aufbau (siehe AP5) und der hier beschriebene Workflow befindet sich in einer explorativen Phase, es zeichnet sich aber bereits eine konkrete Architektur ab.

<sup>21</sup> Eine umfassende Übersicht zu Forschungsdaten-Repositorien ist auf [re3data](https://re3data.org/) zu finden.

Die bisherigen Untersuchungen und konkreten Erfahrungen deuten darauf hin, dass gerade die enge Verzahnung zwischen den forschungsdatenproduzierenden Fachwissenschaften, einem domänenspezifischen Kompetenzzentrum wie z. B. der ITG und der UB ein ideales Organisationsmodell sein könnten, das ein Maximum an Effizienz entfalten dürfte. Unverzichtbar erscheint vor allem der enge persönliche und damit vertrauensbildende Kontakt zwischen den Vertretern aller beteiligten Instanzen, für den auch und gerade die räumliche Nähe die wichtigste Voraussetzung darstellt. In der nächsten Projektphase soll daher diese Kooperation zwischen ITG und UB weiter vertieft und längerfristig in ein Datenzentrum überführt und damit institutionalisiert werden. Durch eine aktive Teilnahme an der AG Datenzentren der DHd kann hierbei auf bestehende Strukturen aufgebaut werden.

#### AP5: UNTERSTÜTZUNG BEI RECHERCHE UND PUBLIKATION UNTER VERWENDUNG ENTSPRECHENDER TOOLS

Arbeitspaket 5 strebt die Optimierung sowohl der Publikations-, als auch der Recherche-Möglichkeiten insbesondere in Bezug auf Auffindbarkeit, Nachnutzbarkeit und Zugänglichkeit von Forschungsdaten auf technischer Ebene an. Diese Optimierung basiert sowohl auf den Ergebnissen einer durchgeführten Anforderungsanalyse bezüglich eines geeigneten Recherche-Instruments, als auch auf Empfehlungen aus AP 1. Dies betrifft beispielsweise die Erweiterung der [OAI-PMH-Schnittstelle](#) des Publikationsservers *Open Data LMU* um weitere geeignete Metadaten schemata.

#### Aktueller Stand

Unter dem **Aspekt des Publizierens** wurde von Seiten des AP1 eine Empfehlung für das Metadaten schema *DataCite* ausgesprochen. Dieser Empfehlung folgend implementierte die UB der LMU ein *EPrints*-Plugin (*DataCite.pm*), welches den Funktionsumfang der *Open Data LMU*-OAI-PMH-Schnittstelle dergestalt erweitert, dass ein Harvesten über einen öffentlichen Zugang auch im *DataCite*-Format möglich ist, was durch das *GeRDI*-Projekt bereits produktiv genutzt wird. Es ist beabsichtigt, dieses *EPrints*-Plugin auch anderen Institutionen zur Nachnutzung bereitzustellen.

Um die Datensätze auf *Open Data LMU* auch in der seit Kurzem von Google angebotenen [Dataset Search](#) auffindbar zu machen, wurden alle Datensätze auf der Weboberfläche entsprechend den Empfehlungen von Google mit Metadaten nach dem Vokabular von [schema.org](#) ausgezeichnet.

Unter dem **Aspekt der Recherche** steht die Verbesserung der Auffindbarkeit und damit auch die Nachnutzbarkeit dezentraler und schwer zugänglicher Forschungsdaten im Vordergrund.

*Yewno* als proprietärer Discovery Service verspricht, mit Hilfe von Künstlicher Intelligenz und maschinellem Lernen englischsprachige Volltexte aus unterschiedlichen Disziplinen semantisch zu erschließen und zueinander in Relation zu setzen, um damit letztlich auch disziplinäre Grenzen bei der Recherche zu überschreiten.

*Apache Solr* als Open-Source-Suchmaschinentechnologie mit seinen vielfältigen Features und Plugins verspricht hohe Flexibilität nicht nur in Bezug auf die Anbindung unterschiedlicher Datenquellen. Vor dem Hintergrund der nun folgenden Fragestellungen wurden die beiden genannten Software-Produkte auf die Tauglichkeit im gegebenen Forschungsdatenkontext hin überprüft:

1. Verbesserung der Auffindbarkeit durch den Einsatz von künstlicher Intelligenz

In Gesprächen mit dem Anbieter von *Yewno* hat sich gezeigt, dass eine Ausweitung des Angebots auf FD auf absehbare Zeit nicht vorgesehen ist. Es gäbe zwar die Option, die Software selbständig mit Daten zu trainieren, jedoch ist dies aufgrund von mangelndem Datenvolumen und fehlenden Personalressourcen keine Option, die verfolgt werden kann.

2. Welchen Datenmodellen entsprechen die zu indexierenden Daten? Wie müssen die entsprechenden Indexschemata in *Solr* demnach modelliert werden?

Mögliche Kandidaten für das zu erwartende Datenmodell stellen aktuell sowohl das Metadatenschema *DataCite* (Empfehlung AP 1), als auch das Linked-Data-Datenformat RDF dar.

3. Welche Schnittstellen bietet *Solr*, um verschiedene Datenhaltungssysteme zur Indexierung anzubinden?

*Solr* bietet genau zwei unterschiedliche Schnittstellen zur Indexierung von Datenquellen. Auf der einen Seite steht der sogenannte *DataImportHandler*, welcher es erlaubt, neben JDBC-basierten Datenbanken auch HTTP-basierte Datenquellen oder Dateisysteme zu indexieren. Diese Variante erspart etwaige ressourcenintensive Vorverarbeitungsschritte wie Datenexport, Datentransformation, usw. Auf der anderen Seite existiert der sogenannte *UpdateRequestHandler*. Über diese Schnittstelle ist es möglich, schemakonforme Dateien im XML-, JSON- oder auch CSV-Dateiformat im Index abzulegen.

In Anbetracht der Tatsache, dass nicht nur im Falle von *VerbaAlpina* (vgl. AP 4), sondern auch in vielen weiteren Fällen *MySQL*-Datenbanken als Datenquellen vorgehalten werden, erschien es zunächst sinnvoll, anhand einer einfachen *Solr*-Instanz die Anbindung und Indexierung einer exemplarischen *MySQL*-Datenbank mit einigen wenigen Datensätzen über den *DataImportHandler* zu erproben.

4. Wie skaliert *Solr*?

*Solr* skaliert sehr gut. Für den Fall, dass die Speicherkapazität einer Maschine für den von *Solr* benötigten Index nicht mehr ausreichen sollte, kann einfach durch das Anbinden einer oder auch mehrerer weiterer Maschinen die Gesamtspeicherkapazität erhöht werden.

Zudem ist bei der Verwendung von virtuellen Maschinen die Möglichkeit gegeben, die Leistungsfähigkeit der Maschinen beispielsweise durch die Hinzunahme weiterer CPUs einfach zu steigern.

## 5. Wie ist die Usability von Solr zu bewerten?

Die Überprüfung der Usability bezieht sich in diesem Zusammenhang ausschließlich auf *Solr* als Backend. *Solr* bietet Software-Entwicklern und Administratoren ein Webinterface, um eine Vielzahl von Operationen ohne anfallende Programmierarbeit durchführen zu können:

- Betrachten von Konfigurationsdateien
- Abschicken von Request an gegebene Indexe
- Analysetools für das Finetuning der Konfigurationen
- Anzeige der index-spezifischen (z.B. bei *Shards*) Java-Properties
- Schema-Browser
- Browse-Interface
- und vieles weitere mehr

Ein erster allgemeiner Überblick verspricht eine umfangreiche, relativ intuitive und leicht zu erschließende Oberfläche und Navigationsstruktur. Eine besonders interessante Funktion stellt das Analyse-Interface dar, da es einerseits die Analyse eines Strings X während der Suchzeit simuliert, andererseits aber auch die Analyse eines Strings Y während der Indexierung. Dieses Feature ist besonders relevant, wenn man am Beginn eines Projektes steht. Hiermit kann beispielsweise ein für den jeweils zu verarbeitenden Datentyp ideales linguistisches Verarbeitungsmodul (z.B. *Stemming*) gefunden und für den produktiven Einsatz ausgewählt werden.

Neben dem Analyse-Interface liefert *Solr* standardmäßig auch ein vollständiges prototypisches User-Interface mit, welches auf [Apache Velocity](#) basiert, modular aufgebaut ist, leicht zu konfigurieren ist und die Möglichkeit zulässt, bereits ohne konkretes Frontend Funktionalitäten wie Facettierung, Highlighting oder auch AutoSuggest zu testen.

## 6. Wie kann die Anbindung des Solr-Index über eine OAI-PMH-Schnittstelle erfolgen?

Vor dem Hintergrund, dass eine Vielzahl von Projekten als Zielsystem-Schnittstelle das OAI-PMH-Protokoll präferiert und der Funktionsumfang von *Solr* mit Hilfe des Plugins [oai4solr](#) entsprechend erweitert werden kann, soll in einem weiteren Schritt dieses Plugin in *Solr* integriert werden.

### Nächste Schritte

In Bezug auf die Datenmodelle soll einerseits ein konzeptionelles Kategorienmapping von *DataCite-XML* nach *Solr-XML* als Implementierungsgrundlage für einen entsprechenden Konverter und ein zu modellierendes Indexschema entworfen werden. Andererseits steht eine Evaluierung des Plugins [SolrRDF](#) unter Hinzuziehung des Datenbestands des Verbundkatalogs (BVB) über den LinkedOpenData-Service des [B3Kat](#) an, um die Tauglichkeit von *Solr* bezüglich des Umgangs mit Linked-Data-Technologien erproben und bewerten zu können. Zusätzlich sollen die Vor- bzw. Nachteile einer schema-basierten gegenüber einer schema-losen Indexierung analysiert und vor dem Hintergrund der zu erwartenden Einsatzbereiche (z.B. Entwicklungsphase vs. Produktiveinsatz) abgewogen werden. Den

erfolgreichen Verlauf der Anbindung einer exemplarischen MySQL-Datenbank an eine einfache *Apache Solr*-Instanz vorausgesetzt, stehen nun die Anbindung und vollständige Indexierung der *VerbaAlpina*-Datenbank mit Hilfe einer bereits im Rahmen dieses Arbeitspaketes aufgebauten komplexen *SolrCloud*-Infrastruktur im Vordergrund. In diesem Zusammenhang wird ein auf maschinellem Lernen basierendes Ranking-Verfahren (LTR) zum Einsatz kommen. Auf dieser technologischen Basis soll zudem durch das Befüllen des Indexes mit dem *VerbaAlpina*-Datenbestand die oben genannten Skalierungsarten getestet und bestätigt werden. Während der Aufbauphase sollen auch die bezüglich der Usability genannten Features zur Optimierung der Suchmaschine und damit des Recherche-Instruments gänzlich erschlossen und erprobt werden.

#### AUSBLICK

In einem kurzen Ausblick sollen die Übertragbarkeit der bisher erzielten Projektergebnisse auf andere Fachdomänen sowie die Einordnung in eine künftige Nationale Forschungsdateninfrastruktur (NFDI) betrachtet werden.

Dies betrifft zunächst die Weiterverwendbarkeit der betroffenen FD. *VerbaAlpina* als prototypischer Anwendungsfall wird über die Bereitstellung in *GeRDI* mit Hilfe von *DataCite* interdisziplinären Nutzungskontexten zugeführt. Das Vorgehen bei *VerbaAlpina* ist grundsätzlich auf ähnliche Datenbanken übertragbar. Der laufend stattfindende Austausch mit den *GeRDI*-Ansprechpartnern aus TUM und LRZ dient dabei der interdisziplinären Vernetzung und Abstimmung über die gemeinsame Verwendung von Standards und die aufeinander abgestimmte Weiterentwicklung von Modellen.

Auch beim Einsatz der überprüften DMP-Tools lassen sich Ergebnisse auf andere Fachkontexte übertragen. Durch das Trennen von fachspezifischen und allgemeinen Informationen und Hilfetexten lassen sich DMP-Vorlagen verhältnismäßig leicht für andere Fachbereiche erweitern. Allerdings gibt es Themenbereiche, die in den DH eine geringe Rolle spielen, für die Natur- und Ingenieurwissenschaften aber zentral sind. Ein Beispiel hierfür sind patentrechtliche Erwägungen und Industriekooperationen. Daher ist es empfehlenswert, neben generischen Vorlagen auch fachspezifische Templates zu entwerfen, in denen diese Fragenkomplexe im Detail abgedeckt werden. Im DMP-Tool *RDMO* lässt sich die bestehende Vorlage relativ einfach dahingehend erweitern.

Das Metadatenmodell (siehe unten Abb. 1) berücksichtigt zum einen unterschiedlich gewichtete Metadaten (verpflichtend – optional), zum anderen verschiedene Ebenen des Anreicherungsprozesses. Projekte arbeiten zunächst domänenspezifisch und sollten bei einer strukturierten, plattform- und softwareunabhängigen Datenablage frei in der Wahl des Schemas sein. Von dort aus erfolgt grundsätzlich ein Mapping auf die von *GeRDI* geforderten, allgemein anerkannten, verpflichtenden, formalen Metadatenkategorien von *DataCite*. Optional können in der *DataCite*-Extension inhaltliche, durch Normdaten und kontrollierte Vokabulare standardisierte Metadaten für eine Tiefenerschließung ergänzt werden. Eine

Transformation in andere Metadatenmodelle (re3data, Dublin, KDSF oder MARC) ist ebenfalls möglich. Die Standardisierung im Bereich der inhaltlichen Metadaten wäre der Anknüpfungspunkt und das Aufgabenfeld für domänenspezifische Konsortien (in ihrer Funktion in etwa vergleichbar mit dem Unicode-Konsortium). Die dort erarbeiteten Richtlinien und Standards im FDM umzusetzen, ist Aufgabe eines örtlichen Kompetenz- und Datenzentrums, das die Fachdisziplinen beim Forschungs- und Anreicherungsprozess begleitet. Die Metadatenanreicherung sollte möglichst in die Projektarbeit integriert werden. Auch dazu sind in Förderanträgen personelle Ressourcen vorzusehen. Wir gehen davon aus, dass sich die Brauchbarkeit des hier vorgeschlagenen Metadatenmodells und Anreicherungsverfahrens im weiteren Projektverlauf bewähren wird und auf andere Wissenschaftsdomänen übertragbar ist.

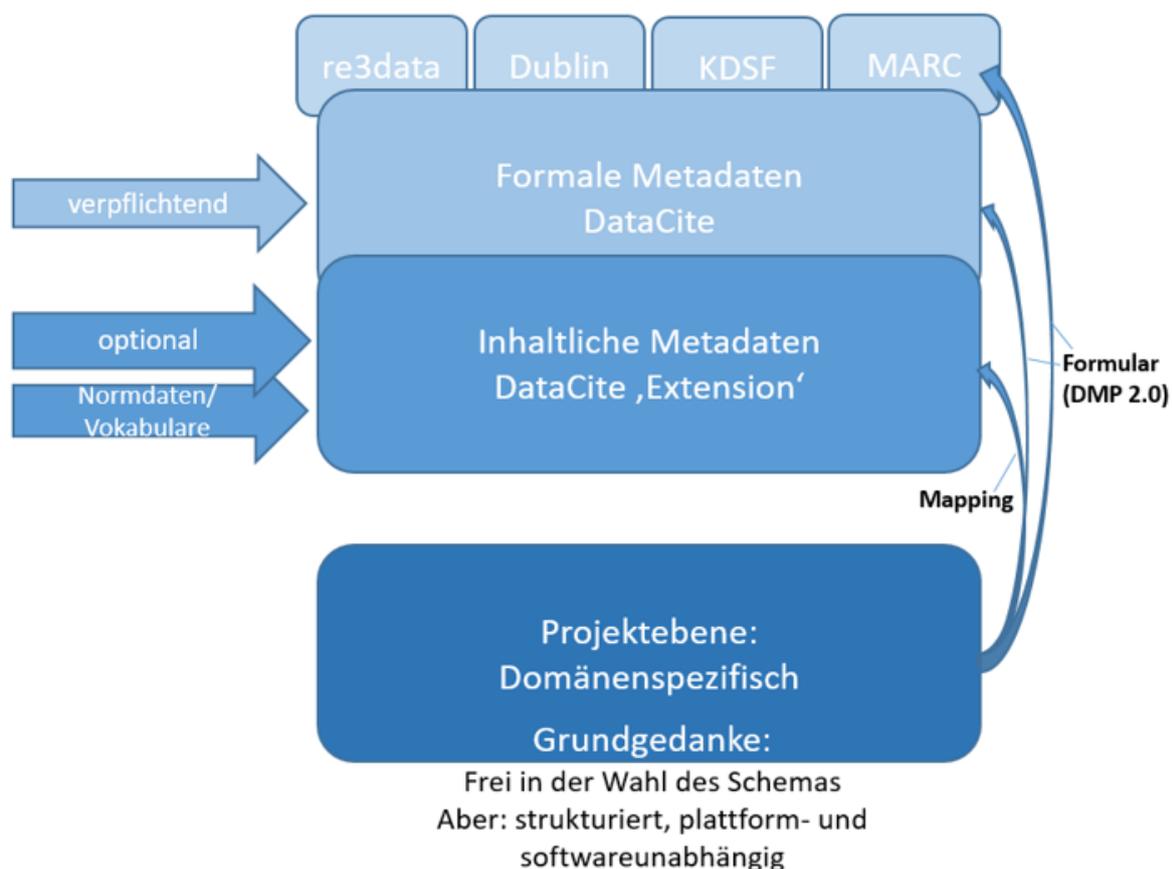


Abb. 1: Metadatenmodell und Prozess der Metadatenanreicherung im Rahmen des FDM

Eine breite institutionsübergreifende Nachnutzbarkeit der Lehrangebote als Open Educational Resources wurde als zentrale Zielsetzung des Schulungspaketes formuliert. Insbesondere fachunabhängige Lerneinheiten eignen sich für die Verwendung über die DH hinaus. Generische Themen können beim konkreten Versuch der Übernahme der Module in weitere Schulungs- bzw. Fortbildungskontexte durch jeweilige disziplinspezifische Schwerpunkte und lokale Bedürfnisse ergänzt werden. Die Herausforderung besteht hier

jedoch darin, Lernmaterialien aus unterschiedlichen Provenienzen und von unterschiedlichen Autoren zu einem homogenen Gesamtangebot zusammenzuführen.

Im Projektverlauf wurden mehrfach auch Fragen des institutionellen Datenmanagements diskutiert. Die Kooperation zwischen der UB der LMU und einer eng an die Fachdisziplinen angebotenen Institution wie der ITG beschreibt aus Sicht aller Projektpartner bereits jetzt einen erfolgsversprechenden Weg für das institutionelle FDM. Das in AP 4 entworfene Konzept eines domänenspezifischen Datenzentrums – hier für die Digital Humanities – könnte Modellcharakter für andere Universitätsstandorte erlangen.<sup>22</sup>

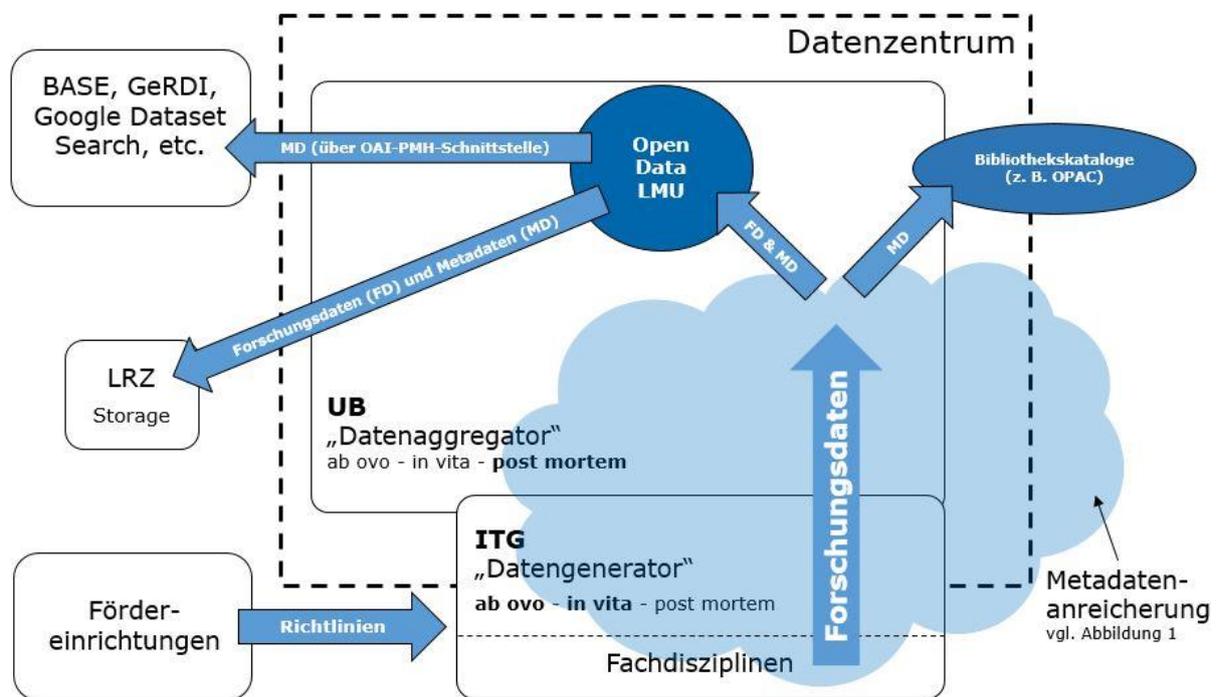


Abb. 2: Rollenverteilung beim FDM am Beispiel der LMU als Prototyp für ein Datenzentrum

Im Laufe der bisherigen Projektzeit hat *eHumanities – interdisziplinär* eine Reihe von Kooperationen neu geknüpft bzw. intensiviert. Ein zentrales Thema in allen Kontexten ist der Aufbau der künftigen Nationalen Forschungsdateninfrastruktur (NFDI). Die NFDI soll nach dem Willen von Bund und Ländern ein sich nachfrageorientiert entwickelndes System sein, das auf die Integration möglichst vieler Projekte in ihrer jeweiligen Fachdomäne angelegt ist. Institutionell und personell soll die NFDI ein Kompetenznetzwerk mit klarem Dienste-Portfolio darstellen. Für die Integrationsfähigkeit in die neuen NFDI-Infrastrukturen dürfte dabei die Beteiligung der datengenerierenden Wissenschaftler ein zentrales Moment darstellen. *eHumanities – interdisziplinär* bringt durch seine enge Anbindung an die Fachwissenschaft die communitybezogenen Voraussetzungen für den Einstieg in die NFDI mit. Dabei steht

<sup>22</sup> Die erweiterte Universitätsleitung der FAU stimmte in ihrer Sitzung am 5.12.2018 dem Vorschlag der UB der FAU zu, nach dem Vorbild der LMU ebenfalls ein Datenzentrum zu begründen.

*eHumanities* – *interdisziplinär* beispielhaft für den Beitrag der Hochschulen im Rahmen einer bestimmten Fachdomäne.

Aus Projektsicht und aus der Perspektive einer künftigen NFDI kann man festhalten, dass weitere Projekte aus anderen Fachdomänen mit einschlägiger Erfahrung im Umgang mit digitalen (und ggf. zu verknüpfenden analogen) FD identifiziert werden sollten, um die gewonnenen Ergebnisse und Kompetenzen in einen anderen Fachkontext zu übertragen und die Anknüpfungsfähigkeit an die NFDI auch in diesen Fachdomänen sicherzustellen. Dies erscheint besonders wichtig, da auch die vorgesehenen NFDI-Konsortien offenbar fachdomänenspezifisch organisiert werden sollen. Die Projektpartner von *eHumanities* sichern dabei jetzt schon ihre Unterstützung für diesen Prozess zu.

## ANHANG

### A: Evaluationskriterien für DMP

Eine ausführliche Diskussion von Bewertungskriterien von DMP findet sich in [Heinrich, M. und Heger, M., Datenmanagementpläne – Eine Bestandsübersicht, Projektarbeit FH Potsdam, 2016](#). Die dort betrachteten Kriterien sind allerdings nicht deckungsgleich mit den hier aufgeführten.

Nummer	Anforderung	Gewichtung	Wunschsituation
Allgemeine Kriterien			
1.1	dauerhafte Trägerorganisation	3	staatliche Einrichtung mit festen Personalmitteln
1.2	Größe der Entwicklercommunity	2	international, dutzende Entwickler aus mehreren Einrichtungen
1.3	Verbreitung des Tools	2	weitverbreitet in Deutschland und Europa (gemeinsame Anforderungen an Datenschutz, usw.). idealerweise auch international im Einsatz
1.4	Softwarelizenz	3	freie Lizenz: GNU PL, MIT, Apache, BSD, ... quellcode-offen
1.5	aktive Weiterentwicklung	3	regelmäßige Updates, neue Features
1.6	Feedback an Entwickler	2	jederzeit zu erreichen, Rücksicht auf Community-Wünsche bei Weiterentwicklung
1.7	Dauer der bisherigen Verfügbarkeit	1	seit 10 Jahren auf dem Markt
2 Technische Kriterien			
2.1	Hosting-Optionen	3	Web-Applikation mit lokaler Server-Installation und dezentrale Installation für kleine Einrichtungen
2.2	Konfiguration der Nutzeroberfläche	3	frei konfigurierbar
2.3	Branding	2	Gestaltung der Web-Oberfläche frei möglich
2.4	Installation vs. Webapplikation	3	Webapplikation ohne lokale Installation am PC, weltweiter Zugriff
2.5	Authentifizierungsoptionen	2	Shibboleth-Authentifizierung und freie Account-Erstellung durch Admin
2.6	Schnittstellen (CRIS, LMU Open Data, ...)	1	Schnittstellen zu Repositorien (insbes. RADAR) und FIS/CRIS-Systemen, eigene API
2.7	Backup-Funktionalitäten	2	Backup der Datenbank automatisiert täglich
2.8	Aufwand für Updates	1	Updates überschreiben keine lokalen Einstellungen (Branding, ...). Update migriert DB automatisch.

3	Administrative Kriterien	
3.1	Umfangreiche Rechtverwaltung (Nutzer)	3 Nutzer können Anderen Lese- und / oder Schreibrechte an DMPs geben. DMPs können öffentlich gemacht werden.
3.2	Umfangreiche Rechtverwaltung (Admin)	3 Admin können Nutzer mit unterschiedlichen Rechten ausstatten und feingranulare Nutzergruppen (Lehrstühle, SFBs, ...) anlegen.
3.3	einfaches Erstellen neuer Vorlagen	1 Vorlagen im WYSIWYG-Verfahren. Einfaches Anpassen / Importieren bestehender Vorlagen.
3.4	verschiedene Unterstützungsebenen	2 Hilfetexte, Beispiellantworten, Auswahlmenüs (Textbausteine, Datentypen, ....)
3.5	Verknüpfung mit Normdateien (GND,...)	1 Schlagwörter werden mit GND-Nummer versehen
3.6	Sprachen	3 Mehrsprachig (Deutsch + mindestens Englisch)
3.7	Kosten	2 freies Produkt, nur Kosten für Server, Betreuung
3.8	vorhanden DMP-Vorlagen	2 kuratierte nachnutzbare Vorlagen für Fächer und Förderer
3.9	Statistiken	1 Nutzungsstatistiken möglich
3.10	Metadaten	1 DMPs können selbst mit MD versehen werden
4	Nutzerforderungen	
4.1	Paralleles Arbeiten	3 paralleles Arbeiten an Plänen möglich, Nachrichten zwischen Autoren können ausgetauscht werden, Änderungsnachverfolgung vorhanden
4.2	Kommentarfunktionen	1 s.o. Kommentare und Nachrichten können eingestellt werden
4.3	Versionierung der Pläne	1 Plan kann leicht versioniert werden. Die alte Version ist danach statisch.
4.4	Kontaktfunktion zu Ref. FD/UB/ITG	3 Kontaktformular für Fragen integriert
4.5	einfache, intuitive Bedienung	3 selbsterklärend, WYSIWYG, Auto-Save-Funktion, ...
4.6	doi-Vergabe / Publikationsfunktion	2 DMPs können publiziert werden, DOI-Vergabe via Registrar möglich
4.7	Verwalten von mehreren Projekten	2 beliebige Anzahl von Projekten parallel möglich, eigene Gruppen / Rechte für jedes einzelne
4.8	Projektmanagementfunktion	1 Tool erlaubt Datenarchivierung, Projektplanung (Terminplan), Zuweisen von Aufgaben, Diskussion / Wiki-Funktionalitäten
4.9	Benachrichtigungen	1 erlaubt "Ankündigungen" (Update / Downtime / ...) und "Benachrichtigungen" an User, wenn sein/ihr DMP geändert wurde (bei Ko-Op-Arbeit)
4.10	WYSIWYG/Editor	2 wie Google-Docs

Die Gewichtung gibt die relative Bedeutung der einzelnen Kriterien an; von wichtigen (3) zu nachrangigen / optionalen Kriterien (1).

**Grobkonzept des Prototyp-Moduls “Forschungsdaten suchen und nachnutzen”**

Das Konzept basiert auf dem didaktischen Modell des *Constructive Alignment*. Die Lehr- und Lernziele wurden anhand der Bedarfserhebung priorisiert und können im Fall einer Nachnutzung von Lerninhalten präzisiert, modifiziert und ergänzt werden.

Lehr- und Lernziele	Lernaktivitäten	Lehr- und Lernzielkontrolle
<p><b>Die Lernenden können die für ihre wissenschaftliche Arbeit relevanten Forschungsdaten finden.</b></p> <p>→ Taxonomiestufe<sup>23</sup>: Analysieren</p>	<ul style="list-style-type: none"> <li>● <b>Die Lernenden erarbeiten selbständig Informationen zu verschiedenen Suchquellen (Repositorien, Datenbanken, Suchmaschinen usw.) sowie ihren essentiellen Vor- und Nachteilen, um die Relevanz einzelner Suchquellen für ihr eigenes Forschungsbereich einschätzen zu können.</b> → Mögliche Vermittlungsmedien und Präsentationsformen: Text, Bilder/Grafiken</li> <li>● <b>Die Lernenden erarbeiten selbständig Informationen zu gängigsten Suchstrategien.</b> → Mögliche Vermittlungsmedien und Präsentationsformen: Text, Bilder/Grafiken.</li> <li>● <b>Fallbeispiele werden präsentiert.</b> Mögliches Vermittlungsmedium: Video</li> </ul>	<p><b>Mögliche Kontrollaufgabe:</b></p> <p>Die Lernenden wählen ein konkretes Thema aus ihren Forschungsbereichen aus und suchen nach relevanten Forschungsdaten.</p> <p><b>Überprüfung der Ergebnisse:</b></p> <p><b>Möglichkeit 1:</b> Ergebnisse (u.a. ausgewählte Suchquellen und -strategien) werden in einer Präsenzveranstaltung vorgestellt und diskutiert.</p> <p><b>Möglichkeit 2:</b> Ergebnisse werden online (z.B. über Google Docs oder Wiki) eingereicht. Das Feedback wird ebenfalls online erfolgen.</p>
<p><b>Die Lernenden können bewerten, ob gefundene Forschungsdaten bestimmte Qualitätskriterien</b></p>	<ul style="list-style-type: none"> <li>● <b>Die Lernenden erarbeiten selbständig verschiedene Qualitätskriterien wie technische Aspekte (z.B. Lesbarkeit der Daten),</b></li> </ul>	<p><b>Mögliche Kontrollaufgabe:</b></p> <p>Aufbauend auf die vorherige Aufgabe (Forschungsdaten zu einem bestimmten Thema suchen) müssen die</p>

<sup>23</sup> Für Details zur Zuordnung von Lernzielen unterschiedlicher Taxonomiestufen siehe hier: [https://paedagogics.ch/wp-content/uploads/2014/10/Taxonomiestufen\\_Bloom.pdf](https://paedagogics.ch/wp-content/uploads/2014/10/Taxonomiestufen_Bloom.pdf); siehe auch Benjamin Samuel Bloom: Taxonomie von Lernzielen im kognitiven Bereich. 4. Aufl. Weinheim 1972: Beltz.

<p>erfüllen, um "nachnutzbar" zu sein.</p> <p>→ Taxonomiestufe: Analysieren</p>	<p><b>inhaltliche Aspekte (z.B. Qualität der Datendokumentation) und allgemeine wissenschaftliche Qualität (z.B. Zitationsfähigkeit, Rankings).</b></p> <p>→ Mögliche Vermittlungsmedien und Präsentationsformen: Text, Bilder/Grafiken</p> <ul style="list-style-type: none"> <li>● <b>Worst- und best-case Beispiele werden präsentiert.</b></li> </ul> <p>→ Mögliche Vermittlungsmedien und Präsentationsformen: Video, Text, Bilder/Grafiken</p>	<p>Lernenden die Qualitätskriterien der aus der Trefferliste ausgewählten Forschungsdaten analysieren und ggf. auch erklären, warum andere Ergebnisse aus der Trefferliste diesen Qualitätskriterien nicht entsprechen.</p> <p><b>Überprüfung der Ergebnisse:</b></p> <p><b>Möglichkeit 1:</b> Im Rahmen einer Präsenzveranstaltung</p> <p><b>Möglichkeit 2:</b> Online-Feedback</p>
<p><b>Die Lernenden können verstehen, was beachtet werden soll, um Forschungsdaten anderer WissenschaftlerInnen rechtskonform nachzunutzen.</b></p> <p>→ Taxonomiestufe: Verstehen</p>	<ul style="list-style-type: none"> <li>● <b>Die Lernenden erarbeiten selbständig Inhalte zu Urheber-/Lizenz- und Nutzungsrechte.</b> → Mögliches Vermittlungsmedium: Text</li> <li>● <b>Worst- und best-case Beispiele werden präsentiert.</b></li> </ul> <p>→ Mögliche Vermittlungsmedien und Präsentationsformen: Text, Bilder/Grafiken</p>	<p><b>Mögliche Kontrollaufgabe:</b></p> <p>Aufbauend auf die ersten zwei Aufgaben (Forschungsdaten zu einem bestimmten Thema suchen und bewerten) müssen die Lernenden erklären, welche rechtlichen Risiken mit der Nachnutzung der ausgewählten Forschungsdaten verbunden werden können.</p> <p><b>Überprüfung der Ergebnisse:</b></p> <p><b>Möglichkeit 1:</b> Im Rahmen einer Präsenzveranstaltung</p> <p><b>Möglichkeit 2:</b> Online-Feedback</p>
<p><b>Die Lernenden können Forschungsdaten richtig</b></p>	<ul style="list-style-type: none"> <li>● <b>Die Lernenden erarbeiten selbständig Inhalte zu Prinzipien der Zitierung von</b></li> </ul>	<p><b>Mögliche Kontrollaufgabe:</b></p> <p>Aufbauend auf die ersten</p>

<p><b>zitieren</b></p> <p>→ Taxonomiestufe: Anwenden</p>	<p><b>Forschungsdaten.</b></p> <p>→ Mögliches Vermittlungsmedium: Text</p> <p>● <b>Beispiele werden präsentiert.</b> Mögliches Vermittlungsmedium: Text</p>	<p>drei Aufgaben (Forschungsdaten zu einem bestimmten Thema suchen, bewerten und rechtliche Risiken erklären) müssen die Lernenden 1-2 der ausgewählten Forschungsdaten zitieren.</p> <p><b>Überprüfung der Ergebnisse:</b></p> <p><b>Möglichkeit 1:</b> Im Rahmen einer Präsenzveranstaltung</p> <p><b>Möglichkeit 2:</b> Online-Feedback</p>
--	---	--

C: Projektbeteiligte

UB der FAU:

- Anna Daudrich, M.A.
- Markus Putnings, Dipl.-Wirtsch.-Inf.
- Dr. Jürgen Rohrwild
- Konstanze Söllner, Dipl.-Math., Dipl.-Theol.

UB der LMU:

- Rainer Gnan, M.A.
- Dipl.-Bibl. Sonja Kümmer, M.A.
- Volker Schallehn, M.A.
- Martin Spenger, M.A. (LIS)
- Alexander Weiß, M.A.

ITG der LMU:

- Dr. Stephan Lücke
- Dr. Christian Riepl
- Julian Schulz, M.A.
- Florian Zacherl, M.Sc.