

Kontakt

Prof. Dr. Dörte Schmidt
Gesellschaft für Musikforschung
(dschmidt@udk-berlin.de)

Prof. Dr. Holger Simon
Verband Deutscher Kunsthistoriker e.V.
(simon@pausanio.com)

Institution

Akademie der Wissenschaften und der Literatur | Mainz
Prof. Torsten Schrade – Digitale Akademie
Geschwister-Scholl-Str. 2
55131 Mainz

www.nfdi4culture.de
info@nfdi4culture.de



FOKUS- THEMEN und AUFGABEN- BEREICHE

**einer Infrastruktur für
Forschungsdaten zu
materiellen und immateriellen
Kulturgütern**

Version 1.0 vom 8.5.2019

Content

1. Präambel	3
2. Bedarfe	5
3. Fokusthemen	7
3.1 Data & Code Literacy	7
3.2 Datenqualität	10
3.3 Forschungssoftware	15
3.4 Normdaten für Werke	19
3.5 Recht und Ethik der Forschungsdaten (spezifisch)	24
4. Aufgabenbereiche	29
4.1 Digitalisierung materieller und immaterieller Kulturgüter	29
4.2 Datenstandards, Datenqualität und Datenkuratierung	29
4.3 Forschungswerkzeuge und Datendienste	30
4.4 Nachhaltige Bereitstellung und Publikation von Daten (LZA, LZV)	30
4.5 Rechte und Datenethik	31
4.6 Expertise, Beratung, Weiterbildung und Qualifikation	31
Anhang: User Stories	32
Impressum	50

1. Präambel

In der Konsortiumsinitiative „NFDI4Culture“ haben sich Partner*innen aus Forschung, Gedächtnisinstitutionen und Infrastruktureinrichtungen zusammengefunden, die sich dem Aufbau einer dezentralen, fach- und forschungsnahen Infrastruktur für Forschungsdaten aus dem Bereich der **materiellen und immateriellen Kulturgüter** widmen. Die wissenschaftliche Notwendigkeit und zugleich das kulturpolitische Potential von NFDI4Culture ergibt sich aus den Besonderheiten der beforschten, **primär nicht-textuellen Güter**, deren materiale und mediale Dimensionen einen Eigenwert besitzen, der nicht vollständig in einer digitalen Repräsentation aufgeht. Als Abformung eines Kulturguts kann das Digitalisat zudem seinerseits immaterielles Kulturgut werden, z. B. bei Verlust oder Zerstörung des materiellen Objekts, durch die Geschichtlichkeit medialer Objekte u. a. m.

Der bisherige Austausch zwischen den einzelnen Partner*innen von NFDI4Culture dient dabei der Selbstverständigung zu Aufgaben und Zielen eines möglichen NFDI-Konsortiums und profitiert von den spezifischen Perspektiven und Kompetenzen aller Beteiligten. So werden Wissenstransfer und eine gemeinsame Sprechfähigkeit gewährleistet und befördert. Die vorliegenden Texte werden in diesem frühen Stadium des allgemeinen NFDI-Prozesses als Positionsbestimmung vorgelegt und dienen **als Bezugspunkt für Diskussionen** zu den für eine NFDI4Culture charakteristischen Herausforderungen in einem größeren Kontext: sowohl in und mit den Fachcommunities als auch zwischen den verschiedenen Konsortiumsinitiativen.

Aufgabenfelder und Ziele einer NFDI4Culture können, werden und sollen sich entsprechend der in der Diskussion formulierten Bedarfe anpassen und weiterentwickeln. Das vorliegende **„Living Document“** dient dabei als Referenzpunkt, in dem die verschiedenen, dynamisch fortlaufenden Diskussionen an einer Stelle immer wieder transparent gebündelt und dokumentiert werden.

Zum Inhalt

Die der NFDI4Culture zu Grunde liegende wissenschaftliche Community produziert im **Zusammenspiel von Forschung, Gedächtnisinstitutionen (GLAM) und Infrastruktureinrichtungen** über den gesamten Forschungsprozess¹ hinweg verschiedenartige Forschungsdaten zu materiellen und immateriellen Kulturgütern. Diese Daten wurden bisher in einer stark analog ausgerichteten Wissenschaft erhoben und vor allem über das Medium Buch dem Wissenschaftsprozess zur Weiterverwendung zur Verfügung und sichergestellt. Bereits heute und absehbar zunehmend entstehen vor allem digitale Forschungsdaten, für die forschungsnah und wissenschaftsbasierte Lösungen gefunden und einfach zugänglich gemacht werden müssen. Die NFDI4Culture will in enger Kooperation mit allen beteiligten Partner*innen aus Forschung, Gedächtnisinstitutionen und Infrastrukturanbietern ein professionelles Netzwerk etablieren, das an den FAIR-Prinzipien² orientierte Angebote für alle Phasen des **Forschungsdatenlebenszyklusses (DLZ)** von der Datengenerierung und -anreicherung über Datenanalyse und -archivierung bis hin zur Datendistribution und -nachnutzung sicherstellt.

Die Handlungsfelder im Rahmen einer NFDI4Culture ergeben sich in erster Linie aus den Schlagworten **Standardisierung, Qualitätssicherung und Vermittlung** (der generierten Daten, der sie fassenden spezifischen Formate, der verwendeten Software und sie aufnehmenden Dienste) im Kontext einer Forschung zu materiellen und immateriellen Kulturgütern. Eingedenk dieser Schlagworte ergeben sich für eine NFDI4Culture folgende Perspektiven entlang des gesamten DLZ, die anschließend in den Fokusthemen und Aufgabenbereiche (siehe  Verweise) mehrfach adressiert und geschärft werden.

1 Prozess der Informationsverarbeitung / Forschungsprozess:

Ideation (Ideen erzeugen, Fragen stellen, Informationen finden)
Konzeptualisierung (Methoden wählen, Partner*innen suchen und vernetzen, Förderer finden)

Operationalisierung (Techniken und Technologien identifizieren)
Kollektion und Kreation (Daten erzeugen und sammeln, Experimente durchführen)

Interpretation (Daten darstellen und analysieren, Ergebnisse beschreiben, Publikation vorbereiten)

Dissemination (Forschungsergebnisse publizieren)

(vgl. F. Lohmeier, J. Mittelbach u. M. Stöhr: Informationsservices auf Augenhöhe - So können Bibliotheken den Forschungsprozess proaktiv unterstützen, in: Handbuch Informationskompetenz, hg. v. Wilfried Sühl-Strohmenger, Berlin/Boston 2016, S. 289–306, hier S. 292).

2 FAIR = Findable, Accessible, Interoperable, Usable, vgl. online unter https://www.forschungsdaten.org/index.php/FAIR_data_principles

Standards und Prozesse: Wie lassen sich Standards und Prozesse forschungsgeleitet weiterentwickeln, damit sie in den adressierten Disziplinen breitere Anwendung finden? Diese Kernfragestellung aller NFDI-Vorhaben berührt fast alle für die NFDI4Culture relevanten Themen:

So ist die Erstellung standardisierter, kompatibler und interoperabler Forschungsdaten (☞ 3.4) gerade im Bereich materieller und immaterieller Kulturgüter und ihrer digitalen Repräsentationen Voraussetzung für den Zugang zu den damit adressierten Gegenständen. Genauso ist die Abstimmung zu Formatstandards als iterativer Prozess zu verstehen, der die wissenschaftliche Perspektive einer größtmöglichen Spezifik des beschriebenen Gegenstandes ebenso in den Blick nimmt, wie die Erfordernisse der Infrastruktureinrichtungen hinsichtlich einer Normierung im Rahmen der produktiven Langzeitarchivierung (☞ 4.4). Das Thema „Datenqualität“ (☞ 3.2) umfasst im Sinne einer Standardisierung von Daten und ihrer Prozesse sowohl die typischen Bereiche des Information Retrievals als auch eine Art Quellenkritik im Hinblick auf die Zuverlässigkeit, Aktualität, Vollständigkeit und Nutzbarkeit aller zu einer Forschungsfrage verfügbaren Informationen. Gerade jedoch für immaterielle und materielle Kulturgüter gilt es, die Spezifik der Daten in den Metadaten zu verankern und so die Differenz zwischen der digitalen Repräsentation und dem materialen Objekt für Menschen und Maschinen lesbar darzustellen.

Werkzeuge und Dienste: Forschungssoftware hat sich im Zuge des digitalen Wandels in den Wissenschaften zu einem **elementaren Bestandteil von Forschungsprozessen** im Allgemeinen und von datengetriebenen Forschungsaktivitäten im Speziellen entwickelt – gerade auch in den Digital Humanities: So müssen neben der Langzeitverfügbarkeit von Daten auch die zur Erstellung und Präsentation verwendeten Werkzeuge und Dienste in den Blick genommen werden. Dementsprechend ergeben sich folgende Fragen: Welche Softwarewerkzeuge und Datendienste existieren entlang des Datenlebenszyklus für die Arbeit mit Forschungsdaten zu materiellen und immateriellen Kulturgütern? Welche Werkzeuge und Diens-

te müssen noch (weiter)entwickelt werden? Wie lässt sich eine institutionelle und technische Nachhaltigkeit von Werkzeugen und Diensten gewährleisten, ohne die die wissenschaftliche Erschließung und Analyse dieser Kulturgüter sowie die Nachvollziehbarkeit der wissenschaftlichen Ergebnisse oftmals kaum möglich sein wird (☞ 3.3 u. 4.3)?

Bereitstellung und Publikation: Um eine optimale Weiternutzung der Forschungsdaten im Sinne der FAIR-Prinzipien zu erreichen, muss ihre möglichst intuitive Auffindbarkeit durch geeignete Konzepte gewährleistet werden (☞ 4.4). Diese umfassen u. a. Systeme der eindeutigen Adressierung wie Persistent Identifier, die Adressierung und Vernetzung zusammengehöriger Inhalte und Ausformungen über Normdaten, mögliche Visualisierungen, Publikationswege, geeignete Schnittstellen etc. Die Speicherung geschieht bevorzugt in den zertifizierten Langzeitarchivsystemen der Infrastruktureinrichtungen nach frühzeitig abgestimmten Datenmanagementplänen (☞ 3.2). Das umfasst nicht zuletzt auch verteilte Konzepte, bei denen projektspezifische Forschungsdaten aus Wissenschaftseinrichtungen mit Daten zu Objekten aus Sammlungseinrichtungen verknüpft werden. Die etablierten Systeme und Regelwerke der normierten Erschließung müssen zwischen den Akteur*innen diskutiert und – wo nötig – in einem iterativen Prozess den sich neu ergebenden Bedürfnissen der Forschung angepasst werden.

Rechte und Datenethik: In der Forschung zu materiellen und immateriellen Kulturgütern haben wir es mit Werken zu tun, die nicht selten einen rechtlichen Schutz genießen. Rechtssicherheit bei der Ablage, Bereitstellung und Nachnutzung von Forschungsdaten ist insbesondere für die NFDI4Culture mit sehr heterogenen und gleichzeitig oftmals komplexen Untersuchungsgegenständen, die teilweise komplizierten Rechtsverhältnissen unterliegen, eine Herausforderung. Dies wird offensichtlich in der Forschung zu urheberrechtlich geschützten und mit Leistungsschutz- oder Persönlichkeitsrechten belegten AV-Medien oder Fotografien. In verschiedenen Bereichen mit kulturell oder historisch sensiblen Gegenständen (z.B. in der Ethnologie) sind darüber hinaus ethi-

sche Aspekte mitzudenken. Die NFDI4Culture wirkt in dem Prozess zu einer digitalen Forschung auf Lösungen hin, die sowohl eine freie und offene Wissenschaft (Open Access, Open Science, Open Data) sicherzustellen als auch die Interessen der Urheber*innen und der Kultur- und Kreativwirtschaft unterstützen. (☞ 3.5 u. 4.5).

Expertise und Beratung: NFDI4Culture sieht eine vordringliche Aufgabe im Aufbau von geeigneten Strukturen, die die herkömmliche Trennung von Wissenschafts-, Gedächtnis- und Infrastruktureinrichtungen im positiven Sinne aufbrechen, um einen institutionalisierten Austausch zu den formulierten Fokusthemen und Aufgabenbereichen zu ermöglichen. Die einschlägigen Fachinformationsdienste (arthistoricum, Propylaeum, musiconn, Darstellende Kunst etc.) sollen als Ausgangspunkt dienen, um verteilte und am Bedarf der Forschenden orientierte Beratungs- und Expertisenetze innerhalb der Gedächtnisinstitutionen, Wissenschaft und Infrastruktureinrichtungen zu etablieren. Ihr Aufbau ist ein Kernelement der NFDI4Culture und notwendig für die Durchsetzung der hier diskutierten Themen. Zudem ist auf eine Verankerung von Data und Code Literacy in entsprechenden Lehrveranstaltungen hinzuwirken. Auf Basis der in NFDI4Culture vertretenen Expertise sind bedarfsgerechte Ausbildungs-, Weiterbildungs-, Lehr- und Lernkonzepte sowie Beratungs- und Unterstützungsangebote zu bündeln, ist ihr Ausbau zu befördern und ihre Implementierung zu unterstützen. (☞ 3.1 u. 4.6).

Die vorliegende Textsammlung ist als „Living Document“ zu verstehen und wird sukzessive erweitert. Die angerissenen Fragen und Themenkomplexe sind dementsprechend in Tiefe und Schärfe offen für die Diskussion der verhandelten Gegenstände: Alle Überlegungen werden so im Abgleich mit den Fachcommunities sowie weiteren Konsortien verifiziert und inhaltlich ausdifferenziert.

2. Bedarfe

K. Bicher (Universität Paderborn), S. Pittroff (AdW Mainz)

NFDI4Culture orientiert sich an den **Bedarfen der Nutzer*innen**. Sie stehen im Zentrum der geplanten **Angebote und Dienste von NFDI4Culture**. Die angedachte Organisationsstruktur eines möglichen NFDI-Konsortiums bindet daher Forschende sehr eng in den Entscheidungsebenen ein und setzt auf ihre Mitwirkung in themengeleiteten Foren. So kann eine fortwährende Anpassung der Infrastruktur an die durch die Nutzer*innen definierten Notwendigkeiten und die aktuellen technischen Möglichkeiten gewährleistet werden. Damit trägt NFDI4Culture zum allgemeinen Ziel des NFDI-Prozesses bei, technische Lösungen in der Forschung fortwährend zu verbessern und gezielt auf die Bedarfe der Forschung hin zu entwickeln. Ebenso trägt sie dazu bei, zeit- und kostenintensive Parallelentwicklungen zu verhindern und immer da generische Ansätze anzuregen, wo sie projektübergreifend und interdisziplinär denkbar sind. So entwickeln sich die in der vorliegenden Textsammlung angesprochenen Kernfragen bereits aus dem Prozess der Identifikation solcher Bedarfe.

Prototypische Nutzungsszenarien (☞ Anhang User Stories) aus den für die Initiative NFDI4Culture relevanten Forschungsbereichen (z. B. der Musikwissenschaft, Kunstgeschichte, Denkmalpflege, Archäologie, Theater-, Tanz- und Filmwissenschaft) zeigen, welche Bedarfe die Nutzer*innen auch vor dem Hintergrund unterschiedlicher Level an Digitalität im Forschungsprozess formulieren und sollen als Bezugspunkte für die Services eines möglichen NFDI-Konsortiums dienen.

Die exemplarischen User Stories veranschaulichen Fragen und Anforderungen der von NFDI4Culture vertretenen Communities. Sie spiegeln die Perspektive der Forschung auf die Herausforderungen eines nachhaltigen Umgangs mit Forschungsdaten und demonstrieren damit den Gegenpart zu den in den übrigen Texten angesprochenen Themen, die v. a. aus der Sicht der In-

frastrukturangebote verhandelt werden. So wird gewährleistet, dass schon die initialen Überlegungen der Konsortiumsinitiative auf die Bedarfe der Forschung abgestimmt sind. Dies soll sicherstellen, dass der Wissenschaft tatsächlich ein politisch wirksames Gremium zur nachhaltigen Umsetzung ihrer Bedürfnisse erwächst. Das Ziel ist, zukünftige Forschungsfragen zu materiellen und immateriellen Kulturgütern beantworten zu können und deren Analytik wie Ergebnisse in Form von Daten langfristig verfügbar und zugänglich zu halten. Die hier angenommenen Fälle basieren auf zunächst ungerichteten Statements einzelner Interessensvertretungen der verschiedenen in der NFDI4Culture zusammenkommenden Fachcommunities³ und werden im weiteren Verlauf des NFDI-Prozesses systematisch in enger Absprache mit den Disziplinen verifiziert und ergänzt.

Die User Stories nehmen dabei verschiedene Forschungsdatentypen in den Blick, neben spezifischen standardisierten Formaten für Bild- und AV-(Meta-)Daten werden Fragen nach Vernetzungsmöglichkeiten über Normdateien oder zur nachhaltigen Verfügbarkeit von Annotationsdaten gestellt. **Die User Stories beschreiben die Notwendigkeit und den Bedarf an infrastruktureller Unterstützung der Forschung** während des gesamten DLZ: Von der Recherche über Erfassung und Erschließung ggf. in Spezialdatenbanken bis zur Nachnutzung und Publikation von Forschungsdaten werden Bedürfnisse formuliert. Als relevante Kriterien und wesentliche Elemente der Angebote thematisieren die User Stories Fragen der Anschlussfähigkeit, z. B. über standardisierte (Meta-)Datenformate, der Rechtssicherheit und Rechteverwaltung sowie der Nachnutzbarkeit / openness der in verschiedenen Repositorien vorgehaltenen Datenkorpora.

Einige der formulierten Bedarfe lassen sich gut in allgemein-geisteswissenschaftlichen Forschungsdateninitiativen auffangen, andere ge-

äußerte Bedarfe verlangen eine spezifische Antwort aus der Community der NFDI4Culture. Für die NFDI4Culture wird eine Gemeinsamkeit der an sich sehr unterschiedlichen fachspezifischen Herausforderungen offensichtlich: **das besondere Verhältnis, in dem das eigentliche Werk zu seiner digitalen Repräsentation steht**. Der Forschungsgegenstand in der Kunstgeschichte kann im Falle eines mittelalterlichen Schreines etwa ein dreidimensionales Objekt sein, das mehrfach umgearbeitet wurde und dessen Standort, und somit sein örtlicher Kontext, sich im Laufe der Jahrhunderte häufig geändert hat. In der Musikwissenschaft wird hingegen die Aufführung einer Oper untersucht, die eine hohe Komplexität im Zusammenspiel von zeitlich linearen und nonlinearen Kontexten mit sich bringt. Diese Beispiele zeigen, dass ein Schrein zwar fotografiert und eine musikalische Darbietung digitalisiert werden können, hierbei aber nur eine Dimension von vielen relevanten festgehalten wird. Viele unterschiedliche Formate sind also notwendig, um sich dem eigentlichen Werk digital anzunähern. Niemals aber wird es in der Vielzahl seiner Digitalisate und Annotationen vollständig aufgehen können. Die Untersuchung der Digitalisate ermöglicht aber Sichten auf ein Werk, die es ohne jene nicht gegeben hätte. Somit produzieren Forschende digitale Originale, die Rückschlüsse auf ein Werk erlauben, jedoch nicht mit diesem zu verwechseln sind. Die in diesen Zusammenhängen entstehenden Forschungsdaten bedürfen demnach spezifischer Berücksichtigung im NFDI-Kontext und werden in ihrer Spezifik von einer NFDI4Culture bedacht. Die folgenden Texte widmen sich verschiedenen Themenfeldern, in denen diese Spezifik besonders virulent wird und stellen Ausgangspunkte und Ansätze für den Umgang mit Forschungsdaten aus diesen Kontexten vor.

³ Z. B. Positionspapiere der Fachgruppen der Gesellschaft für Musikforschung zum Umgang mit Forschungsdaten, interne Arbeitspapiere. Einige der formulierten Grundsätze sind in das Memorandum der Gesellschaft für Musikforschung zur Schaffung nationaler Forschungsdateninfrastrukturen eingegangen, online unter <https://www.musikforschung.de/index.php/memoranda/schaffung-nationaler-forschungsdateninfrastrukturen-nfdi/langfassung>

3. Fokusthemen

3.1 Data & Code Literacy

*O. Brand (Philipps-Universität Marburg),
T. Schrade (AdW Mainz)*

Gegenstand

Forschungen zu materiellen und immateriellen Kulturgütern finden inzwischen zu einem Großteil auf der Basis digitaler Daten statt (z. B. Bilddaten, Audio- und Videodaten, Metadaten zu kulturellen Objekten). Ein planvoller und reflektierter Umgang mit Forschungsdaten zu materiellen und immateriellen Kulturgütern spielt in allen Phasen des Forschungsprozesses eine wichtige Rolle: Von der Erhebung und Organisation von Datenbeständen über deren Analyse und Interpretation bis hin zur Visualisierung, Publikation und dauerhaften Kuratierung. **Data Literacy** umfasst dabei die **Fähigkeit, planvoll mit Daten umzugehen** und sie im jeweiligen Forschungskontext **bewusst einsetzen und hinterfragen zu können**.¹ Genau wie die Fähigkeit zur Quellenkritik zum grundlegenden Repertoire von geistes- und kulturwissenschaftlich Forschenden zählt, muss in Zukunft auch Data Literacy als neue Kompetenz in den Methodenkanon aufgenommen werden.

Mit dem digitalen Wandel in den Wissenschaften geht jedoch nicht nur eine stärkere Fokussierung auf und kritische Reflexion von digitalen Daten einher, sondern auch eine steigende Anwendung von Algorithmen und Software-Werkzeugen. In den Fachdisziplinen ist der Bedarf für den Erwerb von Kenntnissen in softwarebasierter Methodik und ein Wissen um informatische Denk- und Lösungsmodelle hoch.² Zusammen mit der Data Literacy muss also auch eine grundsätzliche **Code Literacy** sowie die Fähigkeit zum **Computational Thinking** bei den Forschenden aufgebaut werden, selbst wenn diese Begriffe und die hinter ihnen stehenden Konzepte erst in jüngerer Zeit wieder stärker in den Blick geraten sind.³

Da Forschungsprozesse zunehmend datengetrieben ablaufen, ist es für Forschende von entscheidender Bedeutung, auch die **Algorithmen und die Software-Instrumente**, mit denen die Analysen ausgeführt und die Ergebnisse erzielt werden, in

wissenschaftlichen und gesellschaftlichen Zusammenhängen **kritisch einschätzen und reflektieren** zu können.⁴ Code Literacy und Computational Thinking steigern dabei auch die Fähigkeit von digital Forschenden, **computerbasierte Problemlösungsstrategien** im Forschungsprozess anwenden zu können. Dazu gehört beispielsweise die Fähigkeit, komplexe Sachverhalte in einfacher zu lösende Teilprobleme zu gliedern (decomposing), Lösungsmuster aus einem Forschungsbereich in einen anderen zu übertragen (analogisieren), konkrete Probleme stärker zu abstrahieren und dabei auch ein gesteigertes Bewusstsein für eine kontinuierliche Evaluation der erzielten Ergebnisse zu entwickeln.⁵ Aus Perspektive der Geistes- und Kulturwissenschaften müssen **Data Literacy** und **Code Literacy** als **zentrale Kulturtechniken unserer Zeit** angesehen werden. Die **Source Codes** zur Erhebung, Auswertung, Visualisierung und Publikation der Forschungsdaten zu materiellen und immateriellen Kulturgütern stellen darüber hinaus **immaterielle Kulturgüter in sich selbst** dar. Bei ihnen handelt es sich um eine neue kulturwissenschaftliche Quellengattung, deren kritische **Reflexion** zu den **Kernaufgaben des Konsortiums NFDI4Culture** gehört.⁶

Handlungsfelder und Ziele

Der für den digitalen Wandel in den Wissenschaften dringend notwendige Aufbau von Fähigkeiten im Bereich von Data und Code Literacy ist in jüngerer Zeit von zahlreichen wissenschaftlichen Gremien gefordert worden.⁷ Beide Themen gehören zum zentralen **Querschnittsbereich einer NFDI** und müssen dort in den kommenden Jahren konsequent verfolgt und nachhaltig entwickelt werden, wenn eine „digitale Kompetenz bei allen Gruppierungen des wissenschaftlichen Personals“ erfolgreich entstehen soll.⁸ Ausbildungs- und Qualifizierungskonzepte müssen dabei auf drei Ebenen entwickelt werden: (a) auf der übergeordneten Ebene der NFDI insgesamt, (b) auf

4 Vgl. EC EGFD (2018), S. 13.

5 Vgl. Rauscher (2018), online unter <https://excitingedu.de/computational-thinking/>

6 Vgl. Krajewski (2015), online unter <https://www.hsozkult.de/debate/id/diskussionen-2901> und Schrade (2018), online unter <https://metacontext.github.io/2018-programming-historians/>

7 Vgl. u. a. RfII (2016), 36, 49; DFG (2018), 37, 39; Hochschulforum Digitalisierung (2016), (2018); Allianz (2018), HRK (2013). Siehe weiterhin die Informationen des Stifterverbands, online unter <https://www.stifterverband.org/data-literacy-education>, des BMBF, online unter <https://www.bildung-forschung.digital/de/data-literacy---mehr-datenkompetenz-fuer-studierende-2355.html> und der Gesellschaft für Informatik, online unter <https://gi.de/dataliteracy/>

8 Allianz der Deutschen Wissenschaftsorganisationen

1 Vgl. Ridsdale et al. (2015), S. 6.

2 Vgl. Katerbow et. al (2018), S. 426; in den Geisteswissenschaften z. B. durch die Digital Humanities.

3 Erste Denkanstöße hierzu gab es bereits seit den 1960er Jahren, vgl. Papert (1980).

der generischen Ebene zwischen Konsortien und (c) auf der Ebene von NFDI4Culture und den dort vereinten Fachdisziplinen mit ihren konkreten Forschungsgegenständen, Forschungsdaten und digitalen Forschungswerkzeugen.

a) **Auf der übergeordneten Ebene der NFDI** geht es darum, bestehende nationale und internationale Angebote, Projekte und Plattformen mit existierenden Angeboten zu Data und Code Literacy zu sichten (z. B. das PARTHENOS-Projekt⁹), deren Potentiale für die NFDI zu analysieren und diese dann auf Basis abgestimmter Strategien und Workflows in Zusammenarbeit zwischen den NFDI-Konsortien weiter auszubauen. Daneben gilt es, übergeordnete Leitlinien für einen reflektierten Umgang mit wissenschaftlichen Daten und Code zu erstellen¹⁰ und ausgehend von der NFDI zu einer „neuen Datenkultur der Offenheit und des Teilens (Data Sharing)“¹¹ in den Wissenschaften zu gelangen.

b) **Weiterhin gilt es in enger Vernetzung zwischen NFDI-Konsortien** generische Aspekte von Data und Code Literacy mit Blick auf die Geisteswissenschaften zu bearbeiten. Hierzu gehören z. B. die Entwicklung von Ausbildungs- und Weiterbildungskonzepten, die ein grundlegendes **Datenverständnis** und -bewusstsein in die Geisteswissenschaften tragen (Was sind geisteswissenschaftliche Primärdaten, was sind Sekundärdaten?, Welche Rolle spielen Metadaten in den Geisteswissenschaften?). Auch muss der Forschungsdatenlebenszyklus in enger Zusammenarbeit auf geisteswissenschaftliche Forschungsprozesse abgebildet werden. Dabei geht es auch um die Entwicklung von gemeinsamen **Qualitätskriterien**, -standards und -sicherungsprozessen für geistes- und kulturwissenschaftliche Forschungsdaten. Weiterhin müssen übergreifende Lehrkonzepte zur Ausbildung einer **Datenethik**, der **Datensicherheit** sowie ein gemeinsamer Fahrplan für die allgegenwärtige Frage der **rechtlichen Voraussetzungen** für das Teilen und Nachnutzen geisteswissenschaftlicher Forschungsdaten erstellt werden.

c) **Innerhalb von NFDI4Culture** geht es um die Entwicklung von Ausbildungs- und Lehrkonzepten für **Data und Code Literacy auf der Basis von Forschungsdaten zu materiellen und immateriellen Kulturgütern**. Im Zentrum der Bemühungen

steht die Förderung eines gesteigerten Bewusstseins um **fachspezifische Datenstandards** bei den Nutzer*innen und den Beteiligten des Konsortiums (bspw. eine grundsätzliche Expertise in der Anwendung des CIDOC-CRM in der Kunstgeschichte und von MEI in der Musikwissenschaft). Ein spezifischer Aspekt, zu dem NFDI4Culture beitragen kann, ist der bisher wenig untersuchte **performative Aspekt der Erstellung, Annotation und Analyse von Forschungsdaten und Forschungssoftware**. Mit jeder Datenannotation und jeder geschriebenen Codezeile bilden Forschende subjektive Forschungsrealitäten in den Daten und der Software ab, was wiederum erhebliche Auswirkungen auf die Interpretation der Daten und die durch die Software erzeugten Ergebnisse haben kann. Bei einigen Teilen der in NFDI4Culture behandelten Datensets existiert sogar eine **doppelte Performativität** (z. B. bei annotierten Forschungsdaten zu modernem Tanz oder moderner Musik): Zum einen werden in den Daten spezifische Performances erfasst und näher beschrieben, zum zweiten ist die Annotation dieser Daten ein performativer Akt an sich. Dabei ist es bei den in NFDI4Culture behandelten Forschungsdaten nie möglich, über die Daten zum „eigentlichen Objekt“ zu gelangen. **Daten zu materiellen und immateriellen Kulturgütern stellen immer nur eine Repräsentation des untersuchten Gegenstandes dar, nie den Gegenstand selbst**. Für diese besondere Situation gilt es innerhalb des Konsortiums **spezifische Beurteilungskriterien** für einen reflektierten Umgang mit den Forschungsdaten im Sinne der Data Literacy zu finden.

Die an NFDI4Culture beteiligten Akteur*innen werden sich insbesondere um eine **curriculare Verankerung** von Data und Code Literacy im Sinne von **Critical Data Studies und Critical Algorithm Studies** bemühen. Die Grundlagen hierfür sind von den Partner*innen in Lehrangeboten an den beteiligten Universitäts- und Hochschulstandorten gelegt.¹² Auch erfolgreiche **Transferformate in die freie Wirtschaft** existieren bereits.¹³

(2017), S. 2; vgl. auch Heidrich (2018), S. 6.

9 Vgl. <http://www.parthenos-project.eu/>.

10 Dazu gehören u. a. Aspekte wie das Erfassen, Managen, Kuratieren, Analysieren, Visualisieren, Interpretieren, Kontextualisieren oder Anwenden von Forschungsdaten und Forschungssoftware.

11 RfII (2016), S. 52.

12 Bspw. der hochschulübergreifende M.A. Studiengang „Digitale Methodik in den Geistes- und Kulturwissenschaften“ (Universität Mainz / Hochschule Mainz), der M.A. Studiengang „Cultural Data Science“ (Philipps-Universität Marburg), der M.Sc. Studiengang „Audiovisual Arts Computing“ (Musikhochschule Detmold) sowie die B.A./M.A. Studiengänge „Informationsverarbeitung“, „Medienwissenschaft mit Medieninformatik“ und der M.A. Studiengang „Archäologie mit Studienrichtung Archäoinformatik“ (Universität zu Köln).

13 Bspw. der Lehrgang „Digital Officer & Data Science“ der Pausanio Akademie; siehe <https://pausanio.com/akademie/>.

Literatur und Quellen

Allianz der Deutschen Wissenschaftsorganisationen (2017): Den digitalen Wandel in der Wissenschaft gestalten. Die Schwerpunktinitiative „Digitale Information“ der Allianz der deutschen Wissenschaftsorganisationen. Leitbild 2018–2022, hg. v. Steuerungsgremium der Schwerpunktinitiative „Digitale Information“ der Allianz der deutschen Wissenschaftsorganisationen. Online unter <http://doi.org/10.2312/allianzoa.015>

Bundesministerium für Bildung und Forschung (2019): Data Literacy – Mehr Datenkompetenz für Studierende. Online unter <https://www.bildung-forschung.digital/de/data-literacy---mehr-datenkompetenz-fuer-studierenden-2355.html>

DFG - Deutsche Forschungsgemeinschaft (2018): Förderung von Informationsinfrastrukturen für die Wissenschaft. Ein Positionspapier der Deutschen Forschungsgemeinschaft, Bonn 2018. Online unter http://www.dfg.de/download/pdf/foerderung/programme/lis/positionspapier_informationsinfrastrukturen.pdf

EC EGF D - European Commission Expert Group on FAIR Data (2018): Turning FAIR into reality. Final Report and Action Plan from the European Commission Expert Group on FAIR Data, Brüssel 2018. Online unter <https://doi.org/10.2777/54599>

Edmond, J., M. Doran u. G. Nugent-Folan (2018): Is Software Production a Cultural Practice? Position Paper for 2018 Working Group. Trinity College Dublin Centre for Digital Humanities. Online unter <http://dh.tcd.ie/dh/wp-content/uploads/2018/03/Trinity-Centre-for-Digital-Humanities-Software-Culture-Working-Group-Position-Paper.pdf>

Föhr, P. (2017): Historische Quellenkritik im Digitalen Zeitalter. Doctoral Thesis, University of Basel, Faculty of Humanities and Social Sciences. Online unter <http://dx.doi.org/10.5451/uni-bas-006805169>

Gesellschaft für Informatik e.V. (2019): Data Literacy – Kompetenzen in der Hochschule. Online unter <https://gi.de/dataliteracy/>

Heidrich, J., P. Bauer u. D. Krupka (2018): Future Skills: Ansätze zur Vermittlung von Data Literacy in der Hochschulbildung. Arbeitspapier Nr. 37, hg. v. Geschäftsstelle. Hochschulforum Digitalisierung beim Stifterverband für die Deutsche Wissenschaft e.V. Berlin. Online unter https://hochschulforumdigitalisierung.de/sites/default/files/dateien/HFD_AP_Nr37_DALI_Studie.pdf

Heidrich, J., P. Bauer u. D. Krupka (2018): Strukturen und Kollaborationsformen zur Vermittlung von Data-Literacy-Kompetenzen. Arbeitspapier 32, hg. v. Geschäftsstelle. Hochschulforum Digitalisierung beim Stifterverband für die Deutsche Wissenschaft e.V. Berlin. Online unter https://gi.de/fileadmin/GI/Hauptseite/Aktuelles/Meldungen/2018/HFD_AP_Nr32_Data_Literacy_Kompetenzen_Literatur.pdf

Hochschulforum Digitalisierung (2016): The Digital Turn. Hochschulbildung im digitalen Zeitalter. Arbeitspapier Nr. 27, hg. v. Geschäftsstelle. Hochschulforum Digitalisierung beim Stifterverband für die Deutsche Wissenschaft e.V. Berlin. Online unter <https://hochschulforumdigitalisierung.de/sites/default/files/dateien/Abschlussbericht.pdf>

Katerbow, M., M. Royeck u. A. Raabe (2018): DFG-Förderung und der digitale Wandel in den Wissenschaften, in: Informatik Spektrum 41 (2018), S. 421. Online unter <https://doi.org/10.1007/s00287-018-01135-0>

Krajewski, M. (2015): Programmieren als Kulturtechnik, in: H-Soz-Kult, 30.11.2015. Online unter <http://www.hsozkult.de/debate/id/diskussionen-2901>

Manovich, L. (2011): Cultural Software. 2011. Online unter <http://manovich.net/content/04-projects/070-cultural-software/67-article-2011.pdf>

Papert, S. (1980): Mindstorms, New York 1980. <https://docs.google.com/viewer?a=v&pid=sites&srcid=ZGVmYXVsdGRvbWFpbXha2hsYWdoZWZ8Z3g6NzgyOWYxNWNjMjE5ZjVh>

Rauscher, M. (2019): Computational Thinking. Online unter <https://excitingedu.de/computational-thinking/>

RfII - Rat für Informationsinfrastrukturen (2016): Leistung aus Vielfalt. Empfehlungen zu Strukturen, Prozessen und Finanzierung des Forschungsmanagements in Deutschland, Göttingen 2016. Online unter <http://www.rfii.de/?wpdmdl=1998>

Ridsdale, C. et al. (2015): Strategies and best practices for data literacy education. Knowledge synthesis report. Online unter <http://hdl.handle.net/10222/64578>

Schrade, T. (2013): Datenstrukturierung, in: U. Frietsch u. J. Rogge (Hrsg.): Über die Praxis des kulturwissenschaftlichen Arbeitens. Ein Handwörterbuch, Bielefeld 2013, S. 91–97.

3.2 Datenqualität

R. Altenhöner (Staatsbibliothek zu Berlin / Preussischer Kulturbesitz), C. Bracht (Philipps-Universität Marburg), O. Brand (Philipps-Universität Marburg), K. Bulle (Philipps-Universität Marburg), J. Kett (Deutsche Nationalbibliothek), G. Taentzer (Philipps-Universität Marburg)

Gegenstand

Forschende wollen auf **Forschungsdaten** mit **gesichert höchster Qualität** zugreifen können, die genau die Spezifik des Feldes der materiellen und immateriellen Kulturgüter widerspiegelt. Unter Datenqualität ist dabei zu verstehen, dass sowohl relevante Rechercheergebnisse als auch die Zuverlässigkeit, Aktualität, Vollständigkeit und Nutzbarkeit der zu einer Forschungsfrage verfügbaren Informationen gegeben sind, ebenso die Erhebung von solchen Metadaten, die für die dauerhafte Erhaltung und Auffindbarkeit der Forschungsdaten relevant sind.

Zwar existieren bereits allgemeine Kriterien für Datenqualität¹; für Forschungsdaten zu materiellen und immateriellen Kulturgütern jedoch **fehlen bislang anerkannte Qualitätskriterien**, die die Besonderheit dieser Forschungsdaten aufgreifen und qualitativ widerspiegeln: Die erschlossenen oder noch zu erschließenden Daten weisen einige Besonderheiten auf, die sich zum einen in dem spezifischen Verhältnis von digitalem zu analogem Objekt manifestieren. Gleichzeitig weisen die bei NFDI4Culture im Fokus stehenden Forschungsdaten die Besonderheit auf, dass sie nicht selten einen gewissen Grad an Unsicherheit (etwa beim Entstehungszeitraum) beinhalten. Beide Spezifika sollten in der zunehmend digitalisierten Forschung in noch zu entwickelnden Kriterien für die Qualität der Daten ausgedrückt werden, um zu einem Gegenstand der Forschung werden zu können.

Forschende und Datennutzende müssen zu nahezu jedem Zeitpunkt und leicht zugänglich die Qualität der Forschungsdaten überprüfen bzw. sich auf diese verlassen zu können. Wenn Unsicherheiten im Hinblick auf einzelne Aspekte der

Forschungsdaten vorhanden sind, sollte sich dies in den Kriterien zur Datenqualität einfach ausdrücken und nachvollziehen lassen, um bei der Forschung oder Verarbeitung der Daten berücksichtigt werden zu können. Diese Informationen müssen nicht nur für die Nutzer*innen sondern auch maschinenlesbar codiert sein.

Das Thema „Datenqualität“ umfasst, wie oben bereits formuliert, sowohl die typischen Bereiche des **Information Retrievals** (z. B. Recall und Precision, Relevanz, terminologische Kontrolle) als auch die Zuverlässigkeit, Aktualität, Vollständigkeit und Nutzbarkeit der zu einer Forschungsfrage verfügbaren Informationen. Ein weiterer wichtiger Aspekt ist die Erhebung von solchen Metadaten, die für die **langfristige Erhaltung der Forschungsdaten** relevant sind. Diese beziehen sich auf (a) technische Informationen zu Datenformaten, (b) zur technischen Entstehungsumgebung, aber auch (c) zur Kontextualisierung der Entstehungssituation. Auf dieser Basis ist es möglich, die langfristige Verfügbarkeit der Forschungsdaten abzusichern, indem zum Beispiel notwendige vertrauenswürdige Transformationen der Daten durchgeführt werden und/oder geeignete Werkzeuge zur Anzeige und Nutzung bereitgehalten werden.

Die Entwicklung von Qualitätskriterien für Forschungsdaten zu materiellen und immateriellen Kulturgütern ist insofern von hoher Bedeutung, als kulturwissenschaftliche Erschließungsdaten hinsichtlich der Datenmodelle wie auch der Inhalte ausgesprochen heterogen und dynamisch sind. Erst die **Definition und Anwendung von Qualitätsstandards** können die formale und semantische Interoperabilität und Nachnutzbarkeit solcher Daten gewährleisten – und damit die Transparenz und Nachvollziehbarkeit der damit durchgeführten Forschung erhöhen.

Metadaten in den Kulturwissenschaften haben den Zweck, kulturelle Artefakte und ihre digitalen Repräsentationen (Daten) anhand domänenspezifischer syntaktischer oder semantischer Merkmalsgruppen zu beschreiben, dadurch leichter auffindbar zu machen und Beziehungen zwischen den Artefakten, ihren digitalen Repräsentationen (z. B. Bildern) und weiteren Metadaten (z. B. Perso-

¹ Korrektheit, Konsistenz, Zuverlässigkeit, Vollständigkeit, Genauigkeit, Aktualität, Redundanzfreiheit, Relevanz, Einheitlichkeit, Eindeutigkeit, Verständlichkeit, Schlüsseleindeutigkeit, Referenzielle Integrität. Vgl. Apel et al. (2015).

nennormdaten) herzustellen. Die Festlegung und **nachhaltige Pflege domänenspezifischer Merkmalsgruppen** erfolgt vor allem in den jeweiligen Fachdisziplinen und setzt neben der fachbezogenen Spezifik einen gewissen **Standardisierungsgrad** voraus, zum Beispiel durch **domänenspezifische Normdatenkonzepte** (z. B. Werknormdaten in der GND), **Regelwerke** (z. B. das bibliothekarische Regelwerk RDA), Austauschformate (z.B. LIDO) oder **generische Ontologien und Referenzmodelle** (z. B. CIDOC-CRM für den gesamten Bereich des Kulturerbes).

Von diesen beschreibenden Metadaten zu unterscheiden sind **Metadaten über Metadaten**, die den Zweck haben können, wissenschaftlich relevante semantische Aussagen zu treffen, etwa über die Informationsquellen der beschreibenden Metadaten, oder auch darüber, zu welchem Zeitpunkt die Metadaten von welcher/welchem Autor*in bearbeitet wurden oder wo die Informationsressourcen zu finden sind. Insbesondere diese **Metadaten zweiter Ordnung** enthalten somit **wesentliche Aussagen über die Qualität der Daten** und der sie beschreibenden Metadaten erster Ordnung.

Um die syntaktische und semantische Interoperabilität und Nachnutzbarkeit solcher Daten zu gewährleisten, ist es erforderlich, dass über die Definition von Qualitätsstandards und deren Anwendung entlang von passenden Kurationskriterien und damit verbundenen Verfahren zur Datenerhebung ein transdisziplinärer Standardisierungsprozess in Gang gesetzt wird, und zwar im nachhaltigen Dialog sowohl der Forschenden wie auch der Infrastrukturanbieter.

Ziel

Ziel von NFDI4Culture soll es sein, zunächst (1) **Kriterien zur Qualitätsmessung** der sehr unterschiedlich strukturierten Daten aus dem Bereich der nicht-textuellen, materiellen und immateriellen Kulturgüter zu entwickeln. Grundlage der Qualitätskriterien sind ein zu definierender Anwendungskontext und konkrete Anforderungen je Kriterium. Es gilt daher, (2) **spezifische Anwendungsszenarien** und die dafür relevanten Qualitätskriterien zu beschreiben. Erst daraus können (3) **Maßnahmen zur Verbesserung der Datenqualität** abgeleitet werden.

Der Nutzen von Qualitätsstandards ist hoch, da **die Qualität unserer Erkenntnisse** mit digitalen Analysewerkzeugen und Methoden **direkt von der Qualität unserer Daten abhängt**. Anhand der Kriterien wird es zum Beispiel mit Hilfe automatisierter Verfahren möglich sein zu beurteilen, ob Daten schlicht unvollständig sind oder tatsächlich unsicheres Wissen die Ursache für fehlende eindeutige Aussagen ist. Dies ist eine wesentliche Information, die den Forschenden im Forschungsprozess leicht zugänglich und maschinenlesbar zur Verfügung stehen muss, damit Transparenz und Nachvollziehbarkeit ihrer Forschung auch bei der Verwendung algorithmischer Methoden gegeben sind. Eine Maßnahme zur Verbesserung der Qualität ist beispielsweise, die **Unsicherheit des Wissens** in diesem Bereich explizit in den Daten auszudrücken. Ein typischer Anwendungskontext aus dem Bereich von NFDI4Culture ist die Bestanderschließung und -verwaltung von Sammlungen materieller und immaterieller Kulturgüter. Weil die Datenerfassung hier nie abgeschlossen ist, sondern kontinuierlich aktualisiert wird, ergeben sich besondere Anforderungen für die gezielte Versionierung und Zitierbarkeit der Daten.

Der Anspruch von NFDI4Culture ist es, für den weiten Bereich der Forschung zu materiellen und immateriellen Kulturgütern, basierend auf repräsentativen Anwendungskontexten, ein Set von Kriterien zur Beurteilung und Verbesserung der **Datenqualität in Forschungsdatenrepositorien** sowie der **Qualitätssicherungsprozesse zu etablieren** – wünschenswerterweise als ISO-Standard. Dabei bleiben spezifische Herausforderungen bestehen, so etwa die Skalierbarkeit und das Spannungsfeld zwischen der Fokussierung auf sinnfällige Schnittmengen innerhalb des breiten Spektrums der beteiligten Forschungsbereiche auf der einen Seite und Herausarbeitung der spezifischen Anforderungen je Bereich auf der anderen Seite.

Handlungsfelder und Umsetzungsschritte

a. Vernetzung bestehender Initiativen im Hinblick auf Datenqualität zu materiellen und immateriellen Kulturgütern

Es befassen sich bereits mehrere Akteur*innen damit, wie die Qualität von Forschungsdaten gesichert werden kann. Diese sind zu vernetzen und der dazugehörige Diskurs ist zu gestalten, um gemeinsam die besonderen Herausforderungen von Forschungsdaten zu materiellen und immateriellen Kulturgütern (im Vergleich zu Forschungsdaten anderer Fächer) bezüglich ihrer Qualität herauszuarbeiten. Die **Vernetzung dieser Akteur*innen** ist notwendig, um fachspezifische Anforderungen hinreichend fundiert ableiten zu können. Potentielle Partner für diese Vernetzung finden sich bei den Beteiligten an NFDI4Culture, d.h. bei den Forschenden, sowie bei den beteiligten Infrastruktureinrichtungen. Zentral einzubeziehen sind etwa die folgenden Anbieter*innen und Projekte: *Arbeitsgemeinschaft kunsthistorischer Bildarchive und Fototheken (AKBF)*; *Corpus der barocken Deckenmalerei in Deutschland (CbDD)*; *museum4punkt0*; *KultSam*; *KONDA*; *GND4C*, *prometheus - Das verteilte Bildarchiv für Forschung und Lehre*. Dabei sollen Kriterien und Grundlagen für eine konkretisierte Anwendung der FAIR-Prinzipien für konkrete Projekte bzw. Repositorien zu entwickelt werden².

b. Anlaufpunkt und Servicestelle zur Zertifizierung von Datenqualität zu materiellen und immateriellen Kulturgütern

Für die Etablierung von Qualitätsstandards und ihre Integration in bestehende Workflows zur Erstellung und Bearbeitung von Forschungsdaten sollen **Zertifizierungsverfahren** entwickelt werden. Um den hohen Informations- und Beratungsbedarf seitens der Datennutzenden und der digitalisierenden Infrastruktureinrichtungen (GLAMs) zu bedienen, sollen Beratungsstellen (Helpdesk) und Anlaufstellen für Zertifizierungsverfahren angeboten werden. Um die Nutzung der Zertifizierung zu erhöhen, sind Incentives für Zertifizierung zu entwickeln und zu bewerben: etwa indem 'von höherer Ebene' (Politik, DFG) die Zertifizierung eine Voraussetzung für Förderung ist und die Zertifizierung auch entsprechend nachgefragt wird.

² Dies erfolgt anhand der 27 konkretisierenden Empfehlungen der Expert Group on FAIR Data, online unter https://ec.europa.eu/info/sites/info/files/turning_fair_into_reality_1.pdf sowie anhand der Checkliste der TIB Hannover, online unter <https://blogs.tib.eu/wp/tib/2017/09/12/die-fair-data-prinzipien-fuer-forschungsdaten>

Ansatzweise nimmt die Deutsche Digitale Bibliothek (DDB) mit ihren spartenspezifischen Fachstellen die oben beschriebenen Aufgaben bereits wahr, indem sie Kultur- und Wissenschaftseinrichtungen bei Erstellung, Lieferung, Clearing und Veröffentlichung der Daten unterstützt. Diese Fachstellen könnten als Startpunkt für die o.g. Anlaufstellen genutzt werden.³

c. Interoperabilität durch Mapping auf einheitliche Austauschformate verbessern

Durch die Transformation bestehender Datenstrukturen auf einheitliche Austauschformate werden die Stärken und Schwächen der zugrundeliegenden Datenmodelle deutlich. Auf diese Weise lassen sich konkrete Anforderungen zur strukturellen Verbesserung der Datenqualität, d. h. der Erfassung und Aufbereitung von Forschungsdaten ableiten. Ansatzpunkt könnte zunächst LIDO als anerkanntes Austauschformat sein; dies ist im Rahmen von NFDI4Culture jedoch für andere Fachdisziplinen und Datentypen zu prüfen und ggf. weiterzuentwickeln.

d. Anreicherung vorhandener Metadaten mit Normdaten / Data Augmentation

Für die Disambiguierung und als Voraussetzung für die Integration von Forschungsdaten ins Semantic Web ist die Verknüpfung des verwendeten Vokabulars mit Normdaten essentiell. Dies ist am Anfang des Datenlebenszyklus noch relativ einfach, nachträglich jedoch nur mit hohem manuellen und intellektuellen Aufwand möglich. Einen Ausweg bietet hier die **Data Augmentation** als Bereich maschinellen Lernens. Im Rahmen eines möglichen NFDI-Konsortiums könnte im Zusammenspiel mit anderen Konsortien die Data Augmentation für Forschungsdaten erprobt werden.

e. Langzeitverfügbarkeit/Langzeitarchivierung

In Abstimmung mit bestehenden Initiativen und Arbeitsgruppen, insbesondere der DINI/nestor-AG Forschungsdaten, sind die Anforderungen zur **Langzeitarchivierung** für den Bereich des materiellen und immateriellen Kulturguts zu prüfen und ggf. weiterzuentwickeln.

f. Datenqualität und Privatheit

Die Qualitätssicherung kann unter Umständen Konsequenzen bezüglich der Privatheit haben.

³ Vgl. Positionspapier der KLA: Archive als Informationsdienstleister und Infrastruktureinrichtungen (Dezember 2018), online unter <https://www.bundesarchiv.de/DE/Content/Downloads/KLA/positionspapier-forschungsdateninfrastruktur.pdf>

Automatische Datenvervollständigungen können zum Beispiel leichtere Rückschlüsse auf Personen ermöglichen. Die Autorschaft von Forschungsdaten ist ein Kriterium für die Bewertung der Daten und muss daher nachvollziehbar sein. Inhaltliche Bezüge zu Personen, zum Beispiel in Interviews, oder Angaben zur Provenienz bedürfen kontextabhängig jedoch der Anonymisierung. Hierfür gilt es pragmatische Lösungen zu finden. Näheres erfolgt in Verbindung mit dem Fokusthema Rechte und Datenethik (☞ 3.5).

Literatur und Quellen

Aggarwal, C.C., P. S. Yu (2009): A Survey of Uncertain Data Algorithms and Applications, in: IEEE Transactions on Knowledge and Data Engineering 21/5 (2009), S. 609–623. Online unter <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=4626956>

Apel, D., W. Behme, R. Eberlein u. C. Merighi (2015): Datenqualität erfolgreich steuern: Praxislösungen für Business-Intelligence-Projekte, 3. überarb. u. erw. Aufl., Heidelberg 2015.

Arendt, T., S. Kranz, F. Mantz, N. Regnat u. G. Taentzer (2011): Towards Syntactical Model Quality Assurance in Industrial Software Development: Process Definition and Tool Support, in: R. Reussner, M. Grund, A. Oberweis u. W. Tichy (Hrsg.): Software Engineering 2011 – Fachtagung des GI-Fachbereichs Softwaretechnik, Bonn 2011, S. 63–74. Online unter <https://dl.gi.de/handle/20.500.12116/19847>

Arendt, T., G. Taentzer (2013): A tool environment for quality assurance based on the Eclipse Modeling Framework, in: Automated Software Engineering 20/2 (2013), S. 141–184.

Arts, D. G. T., N. F. de Keizer u. G.-J. Scheffer (2002): Defining and Improving Data Quality in Medical Registries: A Literature Review, Case Study, and Generic Framework, in: Journal of the American Medical Informatics Association 9/6 (2002), S. 600–611.

Baca, M. (Hrsg.) (2016): Introduction to Metadata, Los Angeles 2008, 3. Aufl. 2016. Online unter <http://www.getty.edu/publications/intrometadata/>

Basili, V., G. Caldiera u. D.H. Rombach (1994): The goal question metric approach, in: Encyclopedia of Software Engineering, New York 1994.

Binder, F., B. Entrup, I. Schiller u. H. Lobin (2014): Uncertain about Uncertainty: Different ways of processing fuzziness in digital humanities data. Digital Humanities 2014. Online unter https://ids-pub.bsz-bw.de/frontdoor/deliver/index/docId/7638/file/Binder_Entrup_Schiller_Lobin_Uncertain_about_Uncertainty_2014.pdf

Bowers, S., B. Ludäscher (2004): An Ontology-Driven Framework for Data Transformation in Scientific Workflows, in: E. Rahm (Hrsg.): Data Integration in the Life Sciences. DILS 2004 (= Lecture Notes in Computer Science 2994), Berlin/Heidelberg 2004, S. 1–16, 2004.

Coburn, E., R. Light, G. McKenna, R. Stein u. A. Vitzthum (2010): LIDO – Lightweight Information Describing Objects Version 1.0., 2010. Online unter <http://www.lido-schema.org/schema/v1.0/lido-v1.0-specification.pdf>

Crofts, N., M. Doerr, T. Gill, S. Stead u. M. Stiff (Hrsg.) (2011): Definition of the CIDOC Conceptual Reference Model – Version 5.0.4, 2011. Online unter http://www.cidoc-crm.org/sites/default/files2/cidoc_crm_version_5.0.4.pdf

Ehrig, H., K. Ehrig, C. Ermel, F. Hermann u. G. Taentzer (2007): Information Preserving Bidirectional Model Transformations, in: Fundamental Approaches to Software Engineering 2007, S. 72–86.

English, L.P. (1999): Improving Data Warehouse and Business Information Quality, New York 1999.

Famelis, M., R. Salay, A. Sandro u. M. Chechik (2013): Transformation of Models Containing Uncertainty, in: A. Moreira et. al. (Hrsg.): Model-Driven Engineering Languages and Systems (= Programming and Software Engineering 8107), Berlin/Heidelberg 2013, S. 673–689.

Genero, M., M. Piattini, u.C. Calero (2005): A survey of metrics for UML class diagrams, in: Journal of Object Technology 4/9 (2005), S. 59–92. Online unter http://www.jot.fm/issues/issue_2005_11/article1

Karsai, G., H. Krah, C. Pinkernell, B. Rumpe, M. Schneider u. S. Völkel (2009): Design Guidelines for Domain Specific Languages, in: Proceedings of the 9th OOPSLA Workshop on Domain-Specific Modeling (2009), S. 7–13.

Kontokostas, D, P. Westphal, S. Auer, S. Hellmann, J. Lehmann, R. Cornelissen u. A. Zaveri (2014): Test-driven evaluation of linked data quality, in: Proceedings of the 23rd international conference on World wide web (2014). ACM, S. 747–758.

Konferenz der Leiterinnen und Leiter der Archivverwaltungen des Bundes und der Länder (KLA) (2018): Positionspapier der KLA: Archive als Informationsdienstleister und Infrastruktureinrichtungen (Dezember 2018). Online unter <https://www.bundesarchiv.de/DE/Content/Downloads/KLA/positionspapier-forschungsdateninfrastruktur.pdf>

Laranjero, N., S. Nur Soydemir u. J. Bernardino (2015): A Survey on Data Quality: Classifying Poor Data, in: 21st Pacific Rim Internat. Symp. on Dependable Computing (PRDC), IEEE 2015.

Lindenthal, J. (2016): Datenqualität und Retrieval. Vorschläge zur Verbesserung der Suche in der Deutschen Digitalen Bibliothek. Online unter http://jl.balilabs.de/DDB/DQ/DDB_Datenqualitaet%3%A4t_Retrieval_1.0.pdf

Mohagheghi, P., V. Dehlen u. T. Neple (2009): Definitions and approaches to model quality in model-based software development - a review of literature, in: Journal Information and Software Technology 51/12 (2009), S. 1646–1669.

Rula, A., A. Zaveri (2014): Methodology for Assessment of Linked Data Quality, 1st. Workshop on Linked Data Quality, 2014. Online unter <http://ceur-ws.org/Vol-1215/paper-04.pdf>

Salay, R., M. Famelis u. M. Chechik (2012): Language independent refinement using partial modeling, in: J. de Lara, A. Zisman (Hrsg.): Fundamental Approaches to Software Engineering. FASE 2012 (= Lecture Notes in Computer Science

7212), Berlin/Heidelberg 2012, S. 224–239.

Sarma, A. Das, O. Benjelloun, A. Halevy, S. Nabar u. J. Widom (2009): Representing uncertain data: models, properties, and algorithms, in: The VLDB Journal 18/5 (2009), S. 989–1019.

Scheven, E. (2013): Die neue Thesaurusnorm ISO 25964 und die GND. Online unter http://www.isko-de.org/data/uploads/2013_Scheven.pdf

Stviliai, B., L. Gasser, M. B. Twidale, and L. C. Smith (2007): A framework for information quality assessment, in: Journal of the American Society for Information Science and Technology 58/12 (2007), S. 1720–1733.

Sunyé, G., D. Pollet, Y. Le Traon u. J.- M. Jézéquel (2001): Refactoring UML models, in: Modeling Languages, Concepts, and Tools. UML 2001 (= Lecture Notes in Computer Science 2185), Berlin/Heidelberg 2001, S. 134–148.

Taentzer, G., T. Kehrer, C. Pietsch u. U. Kelter (2018): A Formal Framework for Incremental Model Slicing, in: A. Russo, A. Schürr (Hrsg.): Fundamental Approaches to Software Engineering. FASE 2018 (= Lecture Notes in Computer Science 10802), Berlin/Heidelberg 2018, S. 3–20.

Wilkinson, M. D., S. Sansone, E. Schultes, P. Doorn, L. Olavo Bonino da Silva Santos, M. Dumontier (2018): A design framework and exemplar metrics for FAIRness, in: Scientific Data 5/180118 (2018). Online unter <https://doi.org/10.1038/sdata.2018.118>

Zaveri, A., A. Rula, A. Maurino, R. Pietrobon, J. Lehmann u. S. Auer (2016): Quality Assessment for Linked Open Data: A Survey, A systematic literature review and conceptual framework, in: Semantic Web 7/1 (2016), S. 63–93. Online unter <http://www.semantic-web-journal.net/system/files/swj556.pdf>

3.3 Forschungssoftware

*D. Röwenstrunk (Universität Paderborn),
T. Schrade (AdW Mainz)*

Gegenstand

Forschungssoftware ist im Zuge des digitalen Wandels in den Wissenschaften zu einem **elementaren Bestandteil von Forschungsprozessen** im Allgemeinen und von datengetriebenen Forschungsaktivitäten im Speziellen geworden: „Komplementär zur stetig wachsenden Datenmenge ergibt sich die ebenfalls fachübergreifende Notwendigkeit der Nutzung von Forschungssoftware, um die verfügbaren Forschungsdaten zu analysieren, kommentieren, annotieren, sie letztlich adäquat nutzen sowie [...] überhaupt per Berechnung, Modellierung, Simulation erzeugen zu können“¹. Die Anwendungsszenarien und Typen von **Forschungssoftware in der Domäne von Forschungsdaten zu materiellen und immateriellen Kulturgütern** sind dabei äußerst vielseitig: Sie reichen von der täglichen Recherche- und Schreibaarbeit mit Webbrowsern und Textverarbeitungsprogrammen bis hin zum Einsatz von Spezialsoftware für die datentechnische Codierung von Musik und Tanz (immaterielles Kulturgut) oder die Annotation und digitale Rekonstruktion von 3D-Daten zu kulturellen Objekten (materielles Kulturgut).² Um eine adäquate **Softwarequalität und -nachhaltigkeit** zu gewährleisten, muss solche Spezialsoftware von entsprechend ausgebildeten, professionellen **Research Software Engineers (RSEs)** konzipiert und **unter (geistes-)wissenschaftlicher Durchdringung der kulturellen Anwendungsdomäne implementiert** werden.

Ausgangssituation

Auch wenn längst internationale Standards (bspw. ISO/IEC 25000)³ und einschlägige Vorgehensmodelle⁴ für die nachhaltige und qualitätsorientierte Entwicklung von Software existieren, finden die **Entwicklungsprozesse von Forschungssoftware häufig noch wenig organisiert, viel zu selten dokumentiert und meist isoliert** innerhalb einzelner Projekte statt. Insbesondere für das mit der Softwareentwicklung betraute wissenschaftliche Personal ergeben sich in diesem Arbeitsumfeld große Herausforderungen: Trotz der fachwissenschaftlichen und informationstechnischen Doppelqualifikation von RSEs werden die tariflichen Eingruppierungen häufig niedriger angesetzt, als bei fachwissenschaftlichem Personal. Gleichzeitig müssen RSEs in vielen Fällen unter sehr großem **zeitlichen Druck** arbeiten, da ihre Entwicklungsleistung meist eine Erfolgsvoraussetzung für softwarebasierte Forschungsprojekte ist. Diese Leistung bringt ihnen im gegenwärtigen Wissenschaftssystem aber **keine akademische Anerkennung** und sie müssen sich darüber hinaus „traditionell“ wissenschaftlich qualifizieren, um den akademischen Karriereweg sinnvoll verfolgen zu können.⁵

Eine im Jahr 2018 durchgeführte Umfrage⁶ der AG Research Software Engineering in den Digital Humanities (DH-RSE)⁷ zu den Arbeits- und Karrierebedingungen von RSEs im geistes- und kulturwissenschaftlichen Forschungsfeld ergab, dass ca. 80% der Personen alleine oder höchstens zu zweit in Projekten entwickeln, **kein separates Personal für Qualitätssicherungsmaßnahmen** (bspw. QA Testing) verfügbar ist, formale Testprozesse eher nicht vorgenommen werden und in über 70% der Fälle auch **kein Übergabeplan** für die Forschungssoftware im Falle eines Personalwechsels existiert. Die daraus resultierende, **generelle Softwarequalität** wurde von den teilnehmenden Personen **mittel bis schlecht** beurteilt.⁸ Zwar ist das Anse-

3 Vgl. ISO/IEC 25000:2014, Systems and software engineering – Systems and software Quality Requirements and Evaluation (SQuaRE) – Guide to SQuaRE (2014), online unter <https://www.iso.org/standard/64764.html>

4 Vgl. dazu jüngst Martin (2018).

5 Vgl. Brett (2017), S. 22 und Busch, Rettinghaus, Schrade, Schulte (2018), online unter <https://dh-rse.github.io/workshop/dhd2018/karriere/anerkennung/arbeitsbedingungen/2018/08/20/aktuelle-situation-der-rses-karriere-welche-karriere.html>

6 Czmiel et al. (2018). Ergebnisse online unter <https://dh-rse.github.io/dhd-workshop-2018-presentation/#/step-4>

7 Siehe <https://dh-rse.github.io/>

8 Vgl. die Diagramme unter <https://dh-rse.github.io/dhd-workshop-2018-presentation/#/step-20>

1 Katerbow (2018), S. 426.

2 Vgl. Czmiel et al. (2018), S. 56.

hen von RSEs in geistes- und kulturwissenschaftlichen Forschungsprojekten nach Selbsteinschätzung eher hoch, gleichzeitig gibt es aber auch **keine wirklichen Beförderungs- oder Karriere-möglichkeiten**. Dies mag auch daran liegen, dass die wissenschaftliche Attribution für den Beitrag der RSEs zum Ergebnis eines Projektes zwar in Form von Erwähnungen bei Konferenzeinreichungen oder Vorträgen gegeben ist, die Software als akademische Leistung aber in den wenigsten Fällen über wissenschaftliche Nachweissysteme zitierfähig verfügbar gemacht wird.⁹ Somit hat die schon länger existierende Redewendung „Applications are like fish, Data is like wine“¹⁰ für den wissenschaftlichen Bereich nach wie vor ihre Gültigkeit.

Die dringend notwendige **Optimierung der Nachhaltigkeits- und Qualitätsaspekte von Forschungssoftware** und die **Verbesserung der Arbeits- und Karrieremöglichkeiten von Research Software Engineers** sind in jüngerer Zeit auf europäischer Ebene vom Expertennetzwerk Knowledge Exchange (KE)¹¹ aufgegriffen worden, in dem sich auch JISC (Großbritannien), CNRS (Frankreich) und die DFG (Deutschland) engagieren. Das Software Sustainability Institute¹² in London gehört in diesem Umfeld zu den führenden Forschungseinrichtungen und hat neben der Herausgabe verschiedener Studien zu Forschungssoftware auch maßgeblich an der Formierung der RSE-Bewegung mitgewirkt. Hieraus ist in Deutschland jüngst die aktive DE-RSE-Community hervorgegangen.¹³ Die Allianzinitiative der deutschen Wissenschaftsorganisationen hat zudem erste Leitlinien zum Thema Forschungssoftware erarbeitet.¹⁴

Handlungsfelder und Ziele

Der Rat für Informationsinfrastrukturen schreibt in seinem zweiten Diskussionsimpuls zur Ausgestaltung der NFDI: „Der Ressourcenbedarf der NFDI wächst mit ihrem schrittweisen Aufbau und umfasst vor allem Fachpersonal für Entwicklung und Betrieb der Dienste-Portfolios (**Investition in Köpfe**)“.¹⁵ Diese Aussage trifft in besonderer Weise auf das **Fokusthema Forschungssoftware** zu. Ein nachhaltiger Umgang mit Forschungsdaten lässt sich auf allen Ebenen des Forschungsprozesses nur durch eine nachhaltige Softwareinfrastruktur erzielen. **Nachhaltigkeit von Software entsteht niemals aus der Software selbst, sondern nur aus den dauerhaft verfügbaren Kompetenzen von und der kontinuierlichen Pflege und Weiterentwicklung durch wissenschaftliche Softwareentwickler*innen**. Zentrales Anliegen der NFDI muss es also sein, die akademischen Ausbildungs- und Karrieremöglichkeiten und die spezifischen programmiererischen Fähigkeiten von wissenschaftlichem Personal im Bereich der Softwareentwicklung zu **verbessern** und zu **fördern**. **Darüber hinaus sollten** Angebote entwickelt werden, die die wissenschaftliche Softwareentwicklung **in Einklang mit Standards aus Industrie und Wirtschaft bringen**. Innerhalb der NFDI ist der Bereich Forschungssoftware auf einer generellen, einer interkonsortialen und einer konsortialen Ebene anzugehen. Gleichzeitig müssen zur Erzielung von Innovation und Fortschritt **Optimierungsprozesse** auf der **Managementebene** (leitende Wissenschaftler*innen und Wissenschaftsmanager*innen) der **Entwicklungsebene** (wissenschaftliche Softwareentwickler*innen) und der **Infrastrukturbene** (Anbieter von wissenschaftlicher Software) angestoßen werden.¹⁶

a) **Auf der übergeordneten Ebene der NFDI** sind fachübergreifende Konzepte dafür zu entwickeln, wie die wissenschaftliche Leistung von RSEs in den Fächern anerkannt und mit akademischen Karrierepfaden versehen werden kann: „until there is a career path into which these experts can be hired, academia will continue to experience serious problems attracting and retaining software experts“¹⁷. Eine **adäquate wissenschaftliche Veröffentlichungsform für Software** einschließlich entsprechender **Zitationsstandards und -empfehlungen** zu finden, gehört ebenfalls zu den übergeordneten Aufgaben der NFDI. Hierbei unterscheidet sich die Veröffentlichung von For-

9 Vgl. die Folien 21-23, online unter <https://dh-rse.github.io/dhd-workshop-2018-presentation/#/step-21>

10 Governor (2017), Why Applications Are Like Fish and Data is Like Wine, online unter <https://redmonk.com/jgovernor/2007/04/05/why-applications-are-like-fish-and-data-is-like-wine/>

11 Siehe <http://www.knowledge-exchange.info/>

12 Siehe <https://www.software.ac.uk/>

13 Siehe <https://rse.ac.uk/> und <https://www.de-rse.org/de/index.html>

14 Vgl. Katerbow et al. (2018), online unter <http://doi.org/10.5281/zenodo.1172970>

15 RFII (2018), S. 3.

16 Vgl. zu dieser Aufteilung auch Katerbow et al. (2018).

17 Hettrick (2016), S. 15.

schungssoftware nicht von der Veröffentlichung von Forschungsdaten, beides muss zusammen gedacht und behandelt werden: „Research software should become a citable scientific deliverable of equivalent value to researchers as that of a publication“¹⁸. Und letztlich müssen auch entsprechende Ausbildungs- und Qualifizierungsmöglichkeiten für akademische Softwareentwickler*innen realisiert werden. Hierbei muss es darum gehen, Einstiegshürden so niedrig wie möglich zu halten und **Entwicklungskompetenzen möglichst früh in die akademische Ausbildung zu integrieren**: „Many researchers know how to code, but few understand the wider set of skills that are needed to develop reliable, reproducible and reusable software. [...] software engineering should be incorporated [...] at the very start of a research career“¹⁹.

b) **Zwischen den NFDI-Konsortien** sollten Kriterien für nachhaltige und qualitätsorientierte Forschungssoftware in den Fachdisziplinen erarbeitet werden. Hierzu gehört die gemeinsame **Diskussion und Analyse von gängigen Anwendungsszenarien** (bspw. für Annotation, Analyse, Visualisierung, Publikation von Forschungsdaten), aber auch die **Fixierung funktionaler Merkmale geisteswissenschaftlicher Software** in Bereichen wie bspw. der Benutzbarkeit (im Sinne einer guten Usability), der Interoperabilität (im Sinne eines einfachen Datenaustauschs) und der Zuverlässigkeit (im Sinne von Fehlertoleranz).²⁰ Darüber hinaus geht es um die **Ermöglichung und Förderung eines kontinuierlichen Wissenstransfers zwischen geisteswissenschaftlichen Entwicklungsteams**. Hierzu müssen zwischen Konsortien **gemeinsame Entwicklungsforen** für den gegenseitigen Erfahrungsaustausch zu aktuellen Entwicklungsmethoden (bspw. agile Methoden, Test Driven Development, Behaviour Driven Development etc.), Vorgehensmodellen, Entwurfs- und Architekturmustern sowie Qualitätssicherungsmaßnahmen gebildet werden.

c) **Auf der Ebene von NFDI4Culture** werden die Spezifika für die Entwicklung von nachhaltiger und qualitativ hochwertiger Software für die Erforschung von materiellen und immateriellen Kulturgütern in den Blick genommen. Zunächst müssen **Nachweissysteme für Forschungssoftware** aufgebaut und **stabile Entwicklungscommu-**

nities um die nachgewiesenen Softwareprodukte herum etabliert werden, um sie dauerhaft institutionell abzusichern. Gleichzeitig müssen **Ausbildungswege** für Softwareentwickler*innen geschaffen werden, die besonderes Gewicht auf die **spezifischen Charakteristika der Anwendungsdomäne von NFDI4Culture** legen (bspw. softwarebasierter Umgang mit codierter Musik, softwarebasierter Umgang mit 3D-Objekten etc.). Für die Nutzer*innen sind analog dazu anwenderfreundliche Workshopformate im Stile der Software Carpentry²¹ zu entwickeln, die eine grundsätzliche (auch codetechnische) **Befähigung zur richtigen Verwendung von Software** zur Erforschung von materiellen und immateriellen Kulturgütern gewährleisten.

Abschließend festzuhalten ist, dass **Software selbst einen komplexen immateriellen Charakter besitzt**²² und ihre Verwendung immer auch **Rückwirkungen** auf die **Forschungsmethodik** sowie die damit erzielten **Forschungsdaten** und **Ergebnisse** hat.²³ Eine Hermeneutik und Reflexion dieses Vorgangs gehört neben den technologischen Aspekten zu den fachlich-inhaltlichen Anliegen von NFDI4Culture.

18 Hettrick (2016), S. 11. Aktuelle Ansätze zur wissenschaftlichen Zitation von Software bei Katz (2018), online unter https://doi.org/10.1007/978-3-319-96418-8_34 sowie Druskat (2017), online unter <https://doi.org/10.5281/zenodo.1120389>

19 Hettrick (2016), S. 14.

20 Vgl. hierzu auch ISO/IEC 25000:2014.

21 Siehe <https://software-carpentry.org/>.

22 „Unlike buildings, which may be made of bricks, concrete, wood, steel, and glass software is made of software. Large software constructs are made of smaller software components [...]“, Martin (2018), S. XVI.

23 Vgl. van Dyck-Hemming et al. (2018).

Literatur und Quellen

Allianz der Deutschen Wissenschaftsorganisationen (2017): Den digitalen Wandel in der Wissenschaft gestalten. Die Schwerpunktinitiative „Digitale Information“ der Allianz der deutschen Wissenschaftsorganisationen. Leitbild 2018–2022, hg. v. Steuerungsgremium der Schwerpunktinitiative „Digitale Information“ der Allianz der deutschen Wissenschaftsorganisationen. Online unter <http://doi.org/10.2312/allianzoa.015>

Ayelt, K. et. al (2017): Status Quo Agile 2016/17. Studie der Hochschule Koblenz in Kooperation mit der GPM Deutschen Gesellschaft für Projektmanagement e. V. Koblenz 2017. Online unter https://www.gpm-ipma.de/fileadmin/user_upload/GPM/Know-How/Studie_Status_Quo_Agile_2017.pdf

Brett, A., M. Croucher et. al. (2017): Research Software Engineers: State of the Nation Report 2017, Zenodo. Online unter <http://doi.org/10.5281/zenodo.495360>

Busch, H., K. Rettinghaus, T. Schrade u. S. Schulte (2018): Aktuelle Situation der RSEs: Karriere - welche Karriere?, in: Blog der AG DH-RSE im Verband Digital Humanities im deutschsprachigen Raum, 20.08.2018. Online unter <https://dh-rse.github.io/workshop/dhd2018/karriere/anerken-nung/arbeitsbedingungen/2018/08/20/aktuelle-situation-der-rses-karriere-welche-karriere.html>

Czmiel, A., S. Druskat u. T. Schrade (2018): Research Software Engineering und Digital Humanities. Reflexion, Kartierung, Organisation, in: Konferenzabstracts der DHd 2018, Kritik der digitalen Vernunft, Köln 2018, S. 56–59. Folien online unter <https://dh-rse.github.io/dhd-workshop-2018-presentation>

Hettrick, Simon (2016): Research Software Sustainability: Report on a Knowledge Exchange Workshop, Edinburgh, The Software Sustainability Institute. Online unter http://repository.jisc.ac.uk/6332/1/Research_Software_Sustainability_Report_on_KE_Workshop_Feb_2016_FINAL.pdf

Katerbow, M., M. Royeck u. A. Raabe (2018): DFG-Förderung und der digitale Wandel in den Wissenschaften, in: Informatik Spektrum 41 (2018), S. 421. Online unter <https://doi.org/10.1007/s00287-018-01135-0>

Katerbow, M., G. Feulner (2018): Handreichung zum Umgang mit Forschungssoftware. Zenodo. Online unter <http://doi.org/10.5281/zenodo.1172970>

Katz, D.S., N.P. Chue Hong (2018): Software Citation in Theory and Practice, in: J. Davenport, M. Kauers, G. Labahn, J. Urban (Hrsg.) Mathematical Software – ICMS 2018. ICMS 2018 (= Lecture Notes in Computer Science 10931), Cham 2018. Online unter https://doi.org/10.1007/978-3-319-96418-8_34

Martin, R. C. (2018): Clean Architecture: A Craftsman's Guide to Software Structure and Design, Boston 2018.

Rat für Informationsinfrastrukturen (2018): Zusammenarbeit als Chance. Zweiter Diskussionsimpuls zur Ausgestaltung einer Nationalen Forschungsdateninfrastruktur (NFDI) für die Wissenschaft in Deutschland, Göttingen 2018. Online unter www.rfii.de/?wpdmdl=2529

Schrade, T. (2017): Nachhaltige Softwareentwicklung in den Digital Humanities. Konzepte und Methoden, in: Konferenzabstracts der DHd2017, Bern 2017, S. 168–171. Folien online unter <https://digicademy.github.io/2017-dhd-sustainable-software/>

van Dyck-Hemming, A., A. Notroff, C. Odebrecht u. N. Reichert (2018): Institutionelle Kulturen und Workflows der Softwareentwicklung: Gemeinsamkeiten und Unterschiede, in: Blog der AG DH-RSE im Verband Digital Humanities im deutschsprachigen Raum, 20.08.2018. Online unter <https://dh-rse.github.io/workshop/dhd2018/kultur/community/workflows/2018/07/30/institutionelle-kulturen-und-workflows-der-softwareentwicklung-gemeinsamkeiten-und-unterschiede.html>

3.4 Normdaten für Werke

K. Bicher (Universität Paderborn), A. Hammes (Universität Paderborn), A. Kailus (Philipps-Universität Marburg), J. Kett (DNB), G. Schelbert (HU Berlin), A.-V. Walzel (SLUB Dresden)

Gegenstand

Normdateien sind wichtiger Bestandteil der Geisteswissenschaften im Zusammenspiel von Forschenden, Gedächtnisinstitutionen und Infrastruktureinrichtungen. Im Kontext einer zunehmend digitalisierten und vernetzten Wissenschaft erfüllen sie verschiedene Funktionen und dienen als:

- **Referenzpunkt** und Identifikator für die bezeichnete Entität
- **Knotenpunkt** innerhalb eines semantischen Netzes, über den unabhängig voneinander aufbewahrte Informationen zu einzelnen Entitäten miteinander verknüpft werden können
- **Speicher** für Kontextinformationen¹, die als disambiguierende Elemente dem Identifikator zur eindeutigen Identifizierung mitgegeben werden

Durch konsequente Anwendung von Normdaten werden im Forschungskontext relevante Entitäten wie Objekte, Personen, Orte, Sachschlagworte oder Körperschaften mit eindeutigen Identifikatoren (URIs) versehen, womit umfassende Verknüpfungen in einer definierten, maschinenverwertbaren Form ermöglicht werden. Die Auseinandersetzung mit Wesen und Potentialen von Normdateien ist für eine NFDI deshalb von zentraler Bedeutung. Gemäß dem FAIR-Prinzip² der Interoperabilität sollen Forschungsdaten so aufbereitet werden, dass sie in einer (semi-)automatisierten Weise kombiniert werden können. Mit Hilfe von Linked Data und darauf basierenden Semantic Web-Technologien kann so die qualifizierte Verlinkung zwischen dezentralen Repositorien über Grenzen von Systemen und Domänen hinweg verbessert werden. Als Referenzpunkt und Daten-Drehscheibe fungieren in diesem Kontext

Normdateien und Fachvokabulare³, die möglichst mit einer Lizenz für die freie Nachnutzung ausgestattet sind und bereits in den einschlägigen Institutionen genutzt werden.

Für die Initiative NFDI4Culture ist es naheliegend, zunächst Linked Data-Nutzungen für jene Gegenstände anzustreben bzw. voranzutreiben, deren Identifizierung und Referenzierbarkeit als primäre Herausforderung aller durch das Konsortium angesprochenen Disziplinen gelten kann: Dies sind die **Werke** der Musik, Architektur, bildenden und performativen Künste, Archäologie und Medienkünste. Eine NFDI4Culture nimmt dementsprechend **Besonderheiten von Normdateien für immaterielle und materielle Kulturgüter** in den Blick. Die Modellierung dieser Werknormdaten berücksichtigt dabei die Spezifika der Forschungsgegenstände und die daraus resultierenden domänenspezifischen Anforderungen. Entsprechende Ansätze für einen konzeptionellen und quantitativen Ausbau gibt es bereits, z. B. im Rahmen der GND und Wikidata, sie werden aktuell jedoch noch nicht umfassend und flächendeckend genutzt.

Handlungsfelder und Ziele

a) Auf der übergeordneten Ebene der NFDI sind Normdaten aufgrund ihrer Rolle für die Umsetzung des FAIR-Kriteriums der Interoperabilität in der aktuellen Diskussion bereits ein wichtiges Querschnittsthema. Absehbar werden die sich daraus ableitenden allgemeinen Aufgaben eher fokussiert an zentralen Stellen und **konsortiumsübergreifend** umgesetzt: die **Anpassung und Weiterentwicklung von Normdatenangeboten an die Anforderungen des Web of Data**, ihre nachhaltige Bereitstellung, die übergreifende Koordinierung und Harmonisierung von Normdaten verschiedener Anbieter sowie die Einbettung in generische Dienste und technische Entwicklungen. International agierende Einrichtungen wie die DNB positionieren sich zunehmend⁴ in vernetzten Zusammenhängen: Im Rahmen des DFG-Projekts *GND*

1 So fungieren Normdateien einerseits als Hilfs- und Arbeitsmittel für die Forschung, in sie gehen aber auch Forschungsergebnisse ein.

2 Auf die zentrale Bedeutung der Standardisierung von Forschungsdaten nach den sog. FAIR-Prinzipien weist auch die Bund-Länder-Vereinbarung zu Aufbau und Förderung einer Nationalen Forschungsdateninfrastruktur (NFDI) vom 26. November 2018 in ihrer Präambel hin (<https://www.gwk-bonn.de/fileadmin/Redaktion/Dokumente/Papers/NFDI.pdf>). Vgl. auch  3.2.

3 Anders als Normdateien dienen Fachvokabulare der strukturierten Beschreibung eines Objekts und seiner Thematik, nicht jedoch seiner eindeutigen Identifizierung. Entsprechend werden im Bereich der Kunst beispielsweise AAT und Iconclass genutzt. Für den Bereich der Musik werden entsprechende Konzepte zur Zeit diskutiert.

4 Weitere Vernetzungsaktivitäten: Aufbau des Virtual International Authority File (VIAF, 2003-2011) als Aggregator verschiedener, internationaler Normdatenangebote; Deutsche Digitale Bibliothek; zunehmende Beteiligung auch nicht-bibliothekarischer Akteur*innen an den Arbeitsgruppen und Gremien des Standardisierungsausschusses der DNB.

für *Kulturdaten (GND4C)*⁵ wird aktuell ein neues Kooperationsmodell für die GND entwickelt, das die Öffnung der GND für die Anforderungen einer spartenübergreifenden Partizipation von Anwendergemeinschaften aus Wissenschaft, Gedächtnis- und Infrastruktureinrichtungen gleichermaßen vorsieht. Die DNB erkundet außerdem Wege der Kooperation zwischen GND und Wikidata und sondiert dynamische Austausch- und Anreicherungsverfahren zwischen GND und Fachrepositorien auf der Grundlage von Wikibase⁶. In diesem Feld wollen auch andere Datenzentren und die Fachinformationsdienste aktive Rollen übernehmen, so dass der Vernetzungsgrad deutlich zunehmen wird.

Um die Engführung dieser generellen Entwicklungen mit der Situation in den jeweiligen Forschungscommunities zu gewährleisten, übernimmt NFDI4Culture die Aufgabe, die **fachspezifischen Sichten und Anforderungen** zu sammeln, zu bündeln und an geeigneter Stelle innerhalb des Normdaten-Kompetenznetzes als Bedarfe einzubringen. Dabei sollten sich die verschiedenen NFDI-Initiativen fortlaufend zu Strategien und Vorgehensweisen abstimmen, um die Öffnung und Nachnutzung von Forschungsdaten über Normdatenreferenzierung nach allgemeinen Prinzipien zu fördern. Die Aufgabe impliziert konkreter, sich über die Bereitstellung und Nutzung übergreifender Frameworks oder Formate zu verständigen, um die einfache Zugänglichkeit der Daten mit Semantic-Web-Technologien zu unterstützen. Damit ist der grundsätzliche Anspruch verbunden, an der Schnittstelle von Wissenschaft und Bibliothek tätig zu werden (☞ 1). Mit Blick auf die möglichst umfassende Nachnutzbarkeit von Normdaten gilt es – auch politisch – offene Lizenzierungsmodelle (z. B. CC0) zu unterstützen.

Umgekehrt sehen es die Beteiligten der NFDI4Culture als wichtige Aufgabe an, die übergeordneten Entwicklungen und Angebote an die jeweiligen Forschungscommunities zurück zu vermitteln und die konkrete Umsetzung innerhalb der Applikationskontexte der Forschung zu befördern.

b) Weiterhin gilt es in enger Vernetzung zwischen NFDI-Konsortien, die strategische Rolle von Normdaten für Kulturgüter als gemeinsame Herausforderung **geisteswissenschaftlicher Disziplinen** und der **bestandsbewahrenden Kulturerbe-Institutionen** zu begreifen. Das Spezifikum

besteht darin, dass es bei geisteswissenschaftlichen Daten meist um digitale Repräsentations- und Verarbeitungsformen von kulturellen Objekten geht, die ganz wie bei traditionellen Museen, Archiven und Bibliotheken vom Prinzip her zeitlich unbegrenzt bereitgestellt werden sollen. Die bewahrenden Institutionen, deren Bestände Grundlage und Gegenstände wissenschaftlicher Fragestellungen bilden, sind daher aktiv einzubeziehen.⁷ Unter Berücksichtigung der FAIR-Kriterien sollten auf einer generischen Ebene zwischen den künftigen NFDI-Konsortien übergreifende Leitlinien für die relevanten Stadien des DLZ (bes. Erzeugung, Aufbereitung, Nutzung, Nachnutzung) formuliert werden. Diese Leitlinien sollten dann Anwendung sowohl im Bereich der Normdaten aber auch bei den geisteswissenschaftlichen Forschungsdaten insgesamt finden.

Werke als geistige Schöpfungen und damit Werknormdaten sind auch für andere geisteswissenschaftliche Disziplinen relevant, darunter die Literaturwissenschaften, die Archäologie oder die Geschichtswissenschaften. In Zusammenarbeit mit Vertreter*innen dieser Fächer prüft NFDI4Culture in einem dynamischen Prozess die Anforderungen an Werknormdaten gegen übergreifende Standards und Qualitätskriterien, bringt Impulse zu deren Weiterentwicklung ein und befördert die konkrete Umsetzung dieses Rahmenwerks beim zukünftigen quantitativen und qualitativen Ausbau von Werknormdatenbeständen. Es werden bestehende Angebote, Projekte und Plattformen (allgemeine wie GND, VIAF oder Wikidata, aber auch fachbezogene wie RISM⁸, Census⁹, CONA¹⁰) gesichtet, deren Potentiale für den anvisierten Geltungsbereich analysiert. Auf Basis abgestimmter Strategien und Workflows in Zusammenarbeit zwischen den NFDI-Konsortien ist der weitere Ausbau der fachbezogenen Angebote und ihr Anschluss an die übergreifenden Angebote über Koreferenzierung anzustreben.

7 Vgl. auch Positionspapier des Verbandes der Historiker und Historikerinnen Deutschlands (VHD) zur Schaffung nationaler Forschungsdateninfrastrukturen (NFDI), online unter <https://www.historikerverband.de/verband/stellungnahmen/positionspapier-zur-schaffung-nationaler-forschungsdateninfrastrukturen-nfdi.html>

8 Répertoire International des Sources Musicales, <http://www.rism.info/de/startseite.html>

9 Census of Antique Works of Art and Architecture Known in the Renaissance, <http://www.census.de/>

10 Cultural Objects Name Authority, <http://www.getty.edu/research/tools/vocabularies/cona/>

5 GND für Kulturdaten (GND4C), online unter <https://wiki.dnb.de/x/dlf9Bw>

6 Kooperation GND – Factgrid, online unter <https://blog.factgrid.de/archives/1475>

c) Innerhalb von NFDI4Culture

Die Konsortiumsinitiative unterstützt und befördert **domänenspezifische Aktivitäten zur Weiterentwicklung konzeptioneller Rahmenwerke für Werknormdaten** im Zusammenschluss von Forschenden, Gedächtnisinstitutionen und Infrastruktureinrichtungen. Teilaspekte sind Definition und Funktion von Norm- und Referenzdaten in verschiedenen Kontexten (z. B. übergeordneten, moderierten, u.U. offenen Norm- und Referenzdateien, z. B. Wikidata, oder projektspezifischen Dateien) in Bezug auf Erschließungstiefe und Anschlussfähigkeit, Nachnutzungsoptionen und Zugänglichkeit (Lizenzen). Sie berücksichtigt dabei historisch bedingte Spezifika bestehender Normdaten. So ist das in Bibliotheken verwendete, aus dem LRM-/FRBR-Datenmodell¹¹ abgeleitete Stufenmodell¹² noch unterdefiniert für Werke, die sich nicht seriell produzierten Ressourcen zuordnen lassen. Auf Grundlage der zu analysierenden Konzepte für Werke im Fokus der Initiative NFDI4Culture wird deshalb ein Anforderungs- und Kriterienkatalog entwickelt, auf dessen Basis Werknormdaten in den bestehenden Repositorien supplementiert und angereichert werden können. Dabei ist zu erwarten, dass sich die Anforderungen an die Modellierung von Werkentitäten in den Teildisziplinen unterscheiden.¹³

Solche Werknormdaten wären der **Referenzpunkt für den Nachweis von digitalen und analogen Ressourcen**, die sich mit diesen Werken befassen oder Forschungsmaterialien zu ihnen bereitstellen. Sie können außerdem dazu dienen, in Beziehung stehende verschiedene Ausprägungen von Werken (Skizzen, Entwürfe, Aufführungen, Aufnahmen, Modelle, Kopien, Nachbildungen, Quellenmaterial, historische graphische Bildquellen, fotografische Abbildungen oder 3D-Modelle) qualifiziert auffindbar zu machen. Außerdem werden Anforderungen an die Granularität für den Nachweis komplexer Werke ebenso diskutiert wie Fragen nach Art und Ausführlichkeit der für eine Identifikation und Nachnutzbarkeit notwendigen Kontextinformationen.

11 FRBR: https://www.ifla.org/files/assets/cataloguing/frbr/frbr_2008.pdf

LRM: <https://www.ifla.org/publications/node/11412>

12 Modellierung gemäß der Werkstufen Werk - Expression - Manifestation - Item.

13 Ein Werknormdatensatz muss zunächst nur so viele Informationen zu seinem Gegenstand enthalten, dass dieser unverwechselbar zu identifizieren ist. Zusätzlich soll er seinerseits so viele normdatenreferenzierte Kerninformationen enthalten, dass der Normdatensatz selbst als Knoten in einem semantischen Netz dienen kann.

Für die Erarbeitung und Kommunikation eines solchen Rahmenwerkes setzt sich die NFDI4Culture für eine „Normdatenagentur“ mit Beratungs- und Service-Angeboten zur Unterstützung teilnehmender Einrichtungen, Projekte und Einzelpersonen ein. Diese bündelt die erforderliche Expertise, die für die kontinuierliche strategische Weiterentwicklung und den faktischen Support notwendig ist. Dabei hat die NFDI4Culture neben der Diskussion konzeptioneller Fragen die Einbindung resp. den Ausbau übergreifender Werknormdateien in verschiedenen Forschungskontexten im Blick. Auf diese Art und Weise etablieren sich Werknormdateien zunehmend als Knotenpunkte im semantischen Netz.

Konkret sollte die Agentur folgende Aufgaben übernehmen:

- **Entwicklung und Ausbau von Standards** in der Beschreibung von immateriellen und materiellen Werken der Musik, performativen Künste, Architektur- und Kunstgeschichte, Archäologie und Medienkünste auf der Grundlage von Modellierungsstandards und Ontologien wie CIDOC-CRM¹⁴, FRBRoo¹⁵, IFLA-LRM
- Unterstützung von Initiativen, die dem **quantitativen Ausbau** von Werknormdaten in den relevanten **Normdatenhubs** dienen (z. B. durch Massendatenimporte von Datensets aus Museen, Archiven, Denkmalpflege, Forschungsprojekten)
- Entwicklung und Ausbau von einfach nachnutzbaren, auf Standardschnittstellen basierenden **Angeboten zur Anbindung von Normdaten** und zur normdatenbasierten Vernetzung von Forschungsdatenbeständen
- Bereitstellung von leicht anpassbaren **Werkzeugen** (Software), mit denen eine (nachträgliche) **Normdatenreferenzierung** von Forschungsdaten vorgenommen oder unterstützt werden kann oder aus bestehenden strukturierten Datenbeständen geeignete neue Normdatenbestände extrahiert und zur Einbringung in die Normdatenrepositorien vorbereitet werden können
- Vernetzung, Koordinierung und Bündelung von **Normdaten-Initiativen** und Projekten, auch für Anfragen anderer geisteswissenschaftlicher Fächer (im Sinne des Transfers, ☞ 1)
- Vermittlung von Kompetenzen bei der Erstellung, Verwendung und Entwicklung von

14 <http://www.cidoc-crm.org/>

15 <https://www.ifla.org/publications/node/11240>

Normdaten im Fachkontext und Anwendung von rechnerbasierten Methoden zum Auslesen von Normdaten (XQuery, SPARQL etc.) sowie die Einrichtung geeigneter Infrastrukturen zur Erleichterung des Normdateneinsatzes in Forschungszusammenhängen (☞ 3.1)

Literatur und Quellen

Association of Research Libraries / Wikimedia Foundation (2019): ARL White Paper on Wikidata Opportunities and Recommendations, 18. April 2019. Online unter <https://www.arl.org/publications-resources/4751-arl-white-paper-on-wikidata-opportunities-and-recommendations>

Bicher, K., B. Wiermann (2018): Normdaten zu „Werken der Musik“ und ihr Potenzial für die digitale Musikwissenschaft, BIBLIOTHEK – Forschung und Praxis, 2018, AR 3218.

Bicher, K. (2018): FRBR-Ebenen und Normdaten zu Werken der Musik für Musikwissenschaft und Musikbibliotheken, Sessionspace bei der GNDCon 2018, 4.12.2018, Frankfurt a. Main. Online unter <https://wiki.dnb.de/display/GNDCON2018/Dokumentation+der+Sessions+der+GNDCon+2018>

Bracht, C. (2016): Warum eine sinnvolle Dokumentation von Kunstwerken ohne die Bibliotheken nicht mehr länger denkbar ist. Neue Strategien am Bildarchiv Foto Marburg, Beitrag zur VDA-Fortbildung für Fachreferent*innen er Kunstwissenschaft, Heidelberg 2016, 6.10.2016. Folien online unter https://www.vdb-online.org/veranstaltung/718/bracht_foto_marburg.pdf

Bund-Länder-Vereinbarung zu Aufbau und Förderung einer Nationalen Forschungsdateninfrastruktur (NFDI) vom 26. November 2018. Online unter <https://www.gwk-bonn.de/fileadmin/Redaktion/Dokumente/Papers/NFDI.pdf>

Div. Autor*innen (o.J.): Wikidata:WikiProject Cultural heritage. Online unter https://www.wikidata.org/wiki/Wikidata:WikiProject_Cultural_heritage

Effinger, M. (o. J.): Werktitelnormdaten für Graphik, arthistoricum. Online unter <https://www.arthistoricum.net/netzwerke/graphik-vernetzt/werktitelnormdaten/>

Effinger, M., G. Knaus, F. Ehrl (2018): Anforderungen an Werknormdaten für druckgraphische Werke aus kunsthistorischer Perspektive, Sessionspace bei der GNDCon 2018, 3.12.2018, Frankfurt a. Main. Online unter <https://wiki.dnb.de/pages/viewpage.action?pageId=143294739>

GND-Kooperative / J. Kett (2017): Initiative für Normdaten und Vernetzung: GND-Entwicklungsprogramm 2017–2021, Stand 06/2017. Online unter https://wiki.dnb.de/download/attachments/125418325/GND_Entwicklungsprogramm17-21_2017-06.pdf

Kailus, A., R. Stein (2018): Besser vernetzt: Über den Mehrwert von Standards und Normdaten zur Bilderschließung. Informationszugang, semantische Interoperabilität, Linked Open Data, Normdaten, Koreferenzierung, in: Piotr Kuroczyński, Peter Bell, Lisa Dieckmann (Hrsg.), Computing Art Reader. Einführung in die digitale Kunstgeschichte, Heidelberg 2018, S. 119–139. Online unter <http://vg07.met.vgwort.de/na/78b0cc67822c4fe197cde9ed9029783c?l=https://books.ub.uni-heidelberg.de/arthistoricum/reader/download/413/413-17-83322-3-10-20190111.pdf>

Kailus, A. (2018): Spartenübergreifende Öffnung und Weiterentwicklung der Gemeinsamen Normdatei: Das Projekt GND für Kulturdaten (GND4C), in: Eva Emenlauer-Blömers, Andreas Bienert, James R. Hemsley (Hrsg.) Konferenzband EVA Berlin 2018. Elektronische Medien & Kunst, Kultur und Historie. 25. Berliner Veranstaltung der internationalen EVA-Serie Electronic Media and Visual Arts, Heidelberg 2018, S. 241–245. Online unter <http://vg07.met.vgwort.de/na/7def22c8796a49e593a66b29d112c0a6?l=https://books.ub.uni-heidelberg.de/arthistoricum/reader/download/442/442-17-83288-1-10-20181206.pdf>

Kailus, A., M. Rosenkötter (2018): GND für Bauwerke: Ausbau zum Knotenpunkt im semantischen Kulturdatennetz, Sessionspace bei der GNDCon 2018, 4.12.2018, Frankfurt a. Main. Online unter <https://wiki.dnb.de/display/GNDCON2018/Dokumentation+der+Sessions+der+GNDCon+2018>

Niggemann, E., R. Goebel, R. Kretzschmar, C. Bracht u. J. Ahlers (2017): GND Für Kulturdaten (GND4C). Beschreibung des Vorhabens – Projektanträge im Bereich „Wissenschaftliche Literaturversorgungs- und Informationssysteme“ (LIS), 2017. Online unter https://wiki.dnb.de/download/attachments/134055796/B_eRT_GND4C_public.pdf?version=1&modificationDate=1518198808000&api=v2

Patton, G. E. (2010): Funktionale Anforderungen an Normdaten: Ein konzeptionelles Modell, Berlin 2010.

Poulter, M. (2017): Wikidata: the new hub for cultural heritage. Online unter <https://blog.wikimedia.org/2017/01/wikidata-the-new-hub-for-cultural-heritage/>

Schelbert, G. (2017): ... warum nicht gleich Wikidata?!, in: M. Stolz (Hrsg.), Konferenzabstracts DHd 2017 Bern. Digitale Nachhaltigkeit (13.–18. Februar 2017), S. 287–288. Online unter https://wikis.hu-berlin.de/mediathek/...Warum_nicht_gleich_Wikidata%3F

Simon, H. (2006): Normierung und Standardisierung der Sacherschließung? Ein Plädoyer für die Heterogenität von Sammlungsbeschreibungen (Vortrag auf Tagung Electronic imaging and the visual arts (EVA), Berlin, 10.–12. November 2004), in: Jörn Sieglerschmidt (Hrsg.): Museums-Informationssystem (MusIS), Konstanz 2006, S. 100–105.

Verband der Historiker und Historikerinnen Deutschlands (VHD) (2017): Positionspapier des Verbandes der Historiker und Historikerinnen Deutschlands (VHD) zur Schaffung nationaler Forschungsdateninfrastrukturen (NFDI). Online unter <https://www.historikerverband.de/verband/stellungnahmen/positionspapier-zur-schaffung-nationaler-forschungsdateninfrastrukturen-nfdi.html>

Van der Graaf, M., L. Waaijers (2014): Authority files: Breaking out of the library silo to become signposts for research information, 2014. Online unter http://repository.jisc.ac.uk/6224/1/Authority_files_-_Breaking_out_of_the_library_silo.pdf

Wiermann, B. (2018): Bibliothekarische Normdaten und digitale Musikwissenschaft, in: Die Musikforschung 71/4 (2018), S. 338–357 .

3.5 Recht und Ethik der Forschungsdaten (spezifisch)

O. Brand (Philipps-Universität Marburg), T. Hartmann (FIZ Karlsruhe), S. Pittroff (AdW Mainz), G. Schelbert (HU Berlin), D. Schmidt (UdK Berlin), H. Simon (Universität zu Köln)

Gegenstand

In der Forschung zu materiellen und immateriellen Kulturgütern haben wir es mit Werken zu tun, von denen einige rechtlichen Schutz genießen. Hinzu kommen Rechte, die im Reproduktions- und/oder Digitalisierungsprozess entstehen können (Urheberrechte von Fotograf*innen, Tonmeister*innen etc.) oder das Haus- oder Persönlichkeitsrecht betreffen. Die dem NFDI-Prozess zugrundeliegenden FAIR-Prinzipien tragen der Anforderung aus der Wissenschaft Rechnung, dass Forschungsdaten für die Forschung offen und leicht zugänglich sein müssen. Ansonsten ist eine freie Wissenschaft nicht möglich. Zugleich ist für die Urheber*innen (Künstler*innen, Autor*innen, Interpret*innen) wie für die Kultur- und Kreativwirtschaft (wie z.B. Ausstellungswesen, Verlage, Konzertveranstalter*innen) die Verwertung der Werke notwendig und Grundlage für die Produktion neuer Kunst und Kulturangebote. Forscher*innen, Urheber*innen wie Kreativwirtschaft sind eng verbunden, genuiner Teil der Kulturlandschaft und damit Gegenstand der Forschung.

Durch den digitalen Wandel in Wissenschaft, Kunst und Kultur entstehen in der heutigen Praxis vor allem im Verhältnis von Forschungs- und Verwertungsinteressen immer wieder rechtliche Herausforderungen, für die bisher keine zufriedenstellenden Lösungen gefunden wurden. Darüber hinaus kann es im Bereich der Kulturgüter ethische Gründe für Zugangsbeschränkungen geben: So werfen beispielsweise Publikationen aus ethnologischen Sammlungen neben rechtlichen auch ethische Fragen im Umgang mit historisch oder kulturell sensible Kulturgüter historisch oder kulturell sensible Kulturgüter auf, da die Auswertung der topografischen Metadaten zu archäologischen Funden führen und Raubgrabungen befördern können.

Im Folgenden soll daher die Ausgangssituation differenziert aufgeschlüsselt werden werden, um eine interne Diskussion in der Konsortiumsinitiative NFDI4Culture über die anschließend Handlungsempfehlungen und Lösungen zu befördern.

Ausgangssituation Recht

Aus rechtlicher Sicht sind für den Bereich der Forschungsdaten zu unserem materiellen und immateriellen Kulturerbe vor allem folgende Felder relevant:

1. Urheberrecht

1.1. Urheberpersönlichkeitsrechte §12-14 UrhG

In einer NFDI wird das Urheberpersönlichkeitsrecht immer dann virulent, wenn urheberrechtlich geschützte Werke digitalisiert oder im Rahmen der Forschung als Forschungsdaten weiterverarbeitet werden.

Einzig Urheber haben das Recht, sowohl über die Nennung des eigenen Namens zu entscheiden als auch über die Art der Veröffentlichung und sich gegen Entstellungen oder andere Beeinträchtigungen ihrer Werke zu verwahren, die ihren berechtigten Interessen zuwiderlaufen. Während die Nennung des Namens einer allgemein üblichen wissenschaftlichen Praxis die „Autor*innen“ zu nennen folgt, entscheidet das Recht auf Veröffentlichung über den Zugang zum Werk. Inwieweit das Recht auf Entstellung auch auf digitale Methoden zutrifft, bei denen zur digitalen Erforschung der Werke der Binärcode z.B. von Bildern verändert und interpretiert wird, ist eine in diesem Zusammenhang zu klärende Frage.

Urheberrechte hängen immer an natürlichen Personen und werden vererbt (§28 UrhG). Werke werden erst 70 Jahre nach dem Tod des*der Urheber*in gemeinfrei und Lichtbilder 50 Jahre nach der Veröffentlichung. Urheber*in und Erben*innen nehmen die Verwertungsrechte selbst wahr. Sie können diese auch an Dritte abtreten (z.B. an Verwertungsgesellschaften) oder verkaufen (z.B. an Verlage).

1.2. Verwertungsrechte (§15-24 UrhG)

Bei den Verwertungsrechten werden im Hinblick auf Forschungsdaten vor allem das Vervielfältigungsrecht (§16 UrhG), das Bearbeitungsrecht (§ 23 UrhG) und das Recht der öffentlichen Zugänglichmachung (§19a) tangiert.

1.2.1. zum Vervielfältigungsrecht (§16 UrhG)

Das Vervielfältigungsrecht betrifft die „Übertragung des Werkes auf Vorrichtungen zur wiederholbaren Wiedergabe“ (§16 UrhG Abs. 1). In der analogen Welt ist es den Forscher*innen ohne weitere Rechtklärung erlaubt z.B. Bilder, Filme,

Tonträger oder Texte zu vergleichen, in dem man diese nebeneinander legt oder die Texte, Tonträger und Bilder für wissenschaftliche oder private Zwecke kopiert. Die Privatkopie ist eine bereits vorhandene Schranke im UrhG. Sie erlaubt nach §53 Abs. 1 Vervielfältigungen, die zum einen unentgeltlich und auf einem materiellen Träger erfolgen und zum anderen nur kleine Teile eines Werkes umfassen. Für digitale Datenträger (etwa Film- oder Audio-Daten) gilt dies nur für Archivierungszwecke. Die graphische Aufzeichnungen von Werken der Musik bedarf nach § 53 Abs. 4 stets der Zustimmung der Rechteinhaber. Weil jede digitale Verwendung von z.B. Bildern oder Musik eine Kopie und damit eine Vervielfältigung nach §16 UrhG darstellt, bedürfen digitale Methoden der Datenanalyse und Datenanreicherung von Bild und Ton nach heutiger Rechtsgrundlage stets der Erlaubnis des*der jeweiligen Urhebers*in. Die im März 2018 neu eingeführte Forschungskopie (§ 60c UrhG) erlaubt in Abs 2 Vervielfältigungen von bis zu 75 Prozent eines Werkes. Für die Erforschung der immateriellen und materiellen Kulturgüter benötigt die Forschung aber auch den Zugriff auf das vollständige Werk.

1.2.2. zum Bearbeitungsrecht (§ 23 UrhG)

Das Bearbeitungsrecht betrifft die Bearbeitung oder Umgestaltung eines Werkes. Diese Werke können Werke der Kunst und Kultur sein und auch Datenbankwerke.

Im digitalen Forschungsprozess sind z.B. bei Bild-, Objekt und Klanganalysen Bearbeitungen und Umgestaltungen der digitalen Repräsentationen von Kunstwerken notwendig. Zugleich finden im Rahmen der Anreicherungsprozesse von Daten für die Datenanalyse Bearbeitungen und Umgestaltung von Datenbankwerken statt.

Das Bearbeitungsrecht wird von den Verwertungsgesellschaften nicht wahrgenommen und kann somit heute nicht lizenziert werden.

1.2.3. zum Recht der öffentlichen Zugänglichmachung (§19a)

Das Recht der öffentlichen Zugänglichmachung betrifft jede drahtgebundene oder drahtlose Zugänglichmachung in der Öffentlichkeit. Die Forschungsgemeinschaft und der universitäre Raum ist Teil dieser Öffentlichkeit.

In der analogen Welt war es erlaubt, Bücher, Noten, analoge Ton-, Film-, Bild- oder digitale Datenträger zu kaufen und an andere Personen weiterzugeben. Die Bibliothek macht sie auf dieser Grundlage der bisherigen Forschung zugänglich und sichert somit eine freie Wissenschaft. Jedes

digitale Archiv, Bild- oder Musikportal, welches Daten an andere Wissenschaftler*innen zur Verfügung stellt, tangiert in der digitalen Welt damit sowohl das Recht auf öffentlichen Zugänglichmachung, weil es die Werke einer großen Gruppe zugänglich macht, und das Vervielfältigungsrecht, weil es stets digitale Kopien erstellt. Solche Portale sehen sich, soweit sie nicht auf „rechtfreie“ Werke zugreifen, die zum Beispiel einer CC0 Lizenz unterstellt sind (vgl. Deutsche Digitale Bibliothek, Europeana), einer komplexen und teilweise widersprüchlichen Rechtslage in Bezug auf die Lizenzierung von Leistungsschutzrechten gegenüber. Eine Lizenzierung über die Verwertungsgesellschaften, auf die die Urheber per Gesetz für die Wahrnehmung ihrer Rechte verwiesen sind, ist etwa in der Musik prinzipiell durch die GVL oder die VG Musikedition möglich, für Lichtbilder aber nicht, weil die zuständige VG Bild-Kunst die Verwertung der Leistungsschutzrecht von den Fotografen nicht bekommt. Digitales Forschen zu zeitgenössischer Kunst und Kultur ist nach dem heutigen Rechtsstand zumeist ohne komplexe Rechtereklärung nicht erlaubt.

2. Persönlichkeitsrecht (Art. 1 und 2 I GG)

Die Persönlichkeitsrechte werden aus einem Bündel von Rechten, vor allem aber aus dem ersten beiden Artikeln des GG, die den Schutz der Persönlichkeit und seine Grundrechte sicherstellen. Auch wenn hier nur ein Teil der Gesamtheit materiellen und immateriellen Kulturguts betroffen ist, so sind die Forschungen auch zu diesen Gegenständen unverzichtbar.

Quantitative digitale Forschung an materiellen und immateriellen Kulturgütern tangiert das selten, weil die Daten (Foto, Gesprochenes, Vorgetragenes) anonymisiert gespeichert und analysiert werden oder aber für eine solche Auswertung eigens anonymisiert werden können. Bei der Veröffentlichung muss aber sichergestellt werden, dass die Daten nicht auf eine bestimmte Person zurückzuführen sind. Speziell für die Forschungsdaten zu materiellen und immateriellen Kulturgütern müssen Werkzeuge zu einer rechtskonformen Anonymisierung entwickelt werden, die für die Forschung keine Qualitätsverluste bedeuten.¹ Dabei liegt die Herausforderung im Einzelfall, da eine Anonymisierung nicht immer ohne weiteres möglich ist. So kann man etwa bei bekannteren Personen auch anhand von

1 Vgl. z.B. https://www.forschungsdaten.org/index.php/Datetei:Horst-Ladewig_MultimodaleDaten-ForschungsfeldViadri-na.pdf

Schrift, Stimme oder bestimmten Inhalten auf ihre Identität schließen. Damit wird der Umgang mit Persönlichkeitsrechten auch zu einem Aufgabenfeld für ein mögliches NFDI-Konsortium.

3. Eigentums- und Hausrecht

Das Eigentums- und Hausrecht tangiert die digitale Erforschung von materiellen und immateriellen Kulturgütern, weil in allen für dieses Feld relevanten Bereichen private Händler*innen, Sammler*innen sowie privat getragene Institutionen aktiv sind. Verhindern Eigentümer den Zugang zu den Objekten stehen sie für Forschung nicht zur Verfügung.

Wichtig ist hier eine Differenzierung der Rechtspersonen Eigentümer*in und Besitzer*in. Dem / Der Eigentümer*in, dem / der eine Sache gehört, darf zwar über den Zustand des Werkes (z.B. Kölner Dom) befinden, darf es zerstören oder erhalten. Der / Die Eigentümer*in hat aber kein Urheberrecht und auch keinen rechtlichen Einfluss auf die Erforschung seiner Werke. (Erzbistum Köln etwa muss bei Forschungen über den Kölner Dom nicht um Erlaubnis gefragt werden.) Den Besitzer*in wiederum gehört das Werk zwar nicht, sondern es befindet sich nur in seinem*ihren Besitz. Dadurch hat er*sie aber die Möglichkeit, über z.B. sein*ihr Hausrecht den Zugang zu den Werken zu beschränken (z.B. Fotografier- und Aufzeichnungsverbot in Archiven, Museen, Theater oder Konzert, aber auch bei privaten Sammler*innen oder Händler*innen).

Diese Rechtslage ist nicht typisch für die digitale Forschung, sondern war bereits früher ein Problemfeld. Die große Reichweite digitaler Forschungsdaten, beispielsweise hochauflösender Bildaufnahmen, sensibilisiert viele Besitzer*innen, erhöht ihr Unbehagen und mindert ihre Bereitschaft, Zugang zum Erstellen solcher Aufnahmen und Aufzeichnung zu gewähren. Vor diesem Hintergrund gilt es aufzuklären und das Forschungsinteresse deutlich zu machen.

4. Archivrecht

Ein nicht unerheblicher Teil der Kulturgüter einschließlich deren Digitalisate sind in (digitalen) Archiven, Bibliotheken, Sammlungen und anderen Gedächtniseinrichtungen aufbewahrt. Der Zugang und die Nutzungsmöglichkeiten für Forschende können bei diesen Kulturgütern auch von Archivgesetzen abhängen, die z. B. besondere Schutz- und Aufbewahrungsvorgaben mit sich bringen. Nicht selten sind zudem ethische und wissenschaftshistorische Überlegungen bei der Bereitstellung von (digitalen) Archivalia einzubeziehen.² Ergänzt wird der besondere Rechtsrahmen durch vielfältige Vertragsbestimmungen, z. B. bei Nachlassverfügungen oder Leihgaben. Wir werden daher in Kooperationen mit den Archiven und Sammlungen die rechtlichen, ethischen und wissenschaftshistorischen Voraussetzungen klären, um den Forschenden Zugang und Nutzungsbefugnisse zu vermitteln. Daran anknüpfend wollen wir möglichst auch Schnittstellen zu digitalen Archiv- und Sammlungsbeständen entwickeln und in der Forschung zu materiellen und immateriellen Kulturgütern etablieren.

5. Kulturgüterschutz

Der Kulturgüterschutz wird vor allem über die Haager Konvention zum Schutz von Kulturgut von 1954 und unterschiedlicher UNESCO-Abkommen durchgesetzt. Es zeichnen sich hier keine rechtlichen Konflikte ab, vielmehr dient die Digitalisierung in vielen Fällen der Bewahrung der materiellen und immateriellen Kulturgüter und ihrer Vielfalt.

² Vgl. Dossier zum interdisziplinärem Forschungsseminar, online unter: <http://edoc.hu-berlin.de/18452/14299>

Ausgangssituation Ethik

Aus ethischer Sicht könnte ein NFDI-Konsortium für das materielle und immaterielle Erbe vor allem folgende Rechtsbereiche tangieren:

1. Kulturgüter aus historisch oder kulturell sensiblen Kontexten

Der Umgang mit historisch und kulturell sensiblen Objekten in Deutschland ist ein hoch politisches Thema. Hier geht es nicht nur um Eigentumsrechte und die damit verbundene Rückgabe von Sammlungsbeständen, Grabungsobjekten, Fundstücken etc. an die entsprechenden Länder. Auch die ethischen Kriterien für den Umgang mit und die öffentliche Zugänglichmachung von Überresten, Riten und Bräuchen sind vor dem breiteren Horizont der aktuellen Neubewertung des Unrechtskontextes des Kolonialismus in der Diskussion und erfordern besondere Aufmerksamkeit.³ In diesem Themenfeld ist eine ethische Abwägung von wissenschaftsgeleiteten Veröffentlichungen etwa von Grabungen, ethnologischen Sammlungen oder aus anderen Gründen historisch oder kulturell sensiblen Sammlungsbeständen im Verhältnis zu ihren Kontexten geboten.

Aus ethnologischer Perspektive zeigt sich etwa, dass es zwischen verschiedenen Rechtsauffassungen zu vermitteln gilt, wenn es um Besitz-, Persönlichkeits- und Urheberrechte geht. Z.B. können Fotos mit nach unserem Rechtssystem ausgelaufenen Persönlichkeitsrechten der Gemeinschaft, der die Abgebildeten angehören, der die Abgebildeten angehören, als verletzend betrachtet werden.

Auch das europäische Urheberrechtsdenken, das auf dem Individualitätsprinzip beruht und Gemeinschaftskonzepte außen vorlässt (so müssen z.B. bei Musikwerken natürliche Personen als Urheber, aber nicht z.B. das Dorf oder eine Gemeinschaft genannt werden) erweist sich im transkulturellen Bereich als vermittlungsbedürftig. Ähnliche Fragen stellen sich für die europäischen Konzepte der Kreativität und der Provenienz.

2. Raubgräberei und Zerstörung von Kulturgütern

In der digitalen Forschung werden Daten durch andere Daten automatisch angereichert, so dass z.B. sehr leicht eine Topografie von archäologischen Gräbern oder Fundorte von anderen Kulturgütern erstellt werden kann. Für die Forschung bringen solche Datenanreicherungen sinnvolle und notwendige Erkenntnisse, die aber auch von Raubgräbern oder politischen Gruppen zur Zerstörung von Kulturgütern genutzt werden können. Hier ist eine ethische Sensibilität bei der Abwägung von wissenschaftsgeleiteter Veröffentlichung solcher Daten im Verhältnis zu ihrer möglichen Verwendung geboten.

3. Open Access und FAIR-Prinzipien

Eine freie Wissenschaft basiert auf den freien Zugang zum Wissen. Bibliothek, Archiv und Museum sind in der analogen Kultur zur zentralen Institutionen geworden, über die das – zumeist schriftlich codierte – Wissen wie – vorwiegend materielle – Kulturgüter aufbewahrt und den Wissenschaftler*innen bis heute zugänglich gemacht wird. Das Urheberrecht aus dem 19. Jahrhundert hat diese Funktion gestützt, in dem es die Verwertung von Wissen und Kulturgütern sichergestellt und zugleich die Zurverfügungstellung über Gedächtnisinstitutionen erlaubt hat.

In einer digitalen Wissenschaftskultur kommt das bestehende Urheberrecht (vgl. oben) an seine Grenzen und es scheint einen solchen offenen Wissenstransfer zu verhindern. Open Access ist eine Antwort des Wissenschaftssystems auf dieses Dilemma, dem sich alle Wissenschaftsorganisationen verpflichtet haben. Die Durchsetzung von Open Access in der Forschung setzt allerdings voraus, dass der / der Autor*in im Besitz der Rechte über alle Inhalte ist, die er veröffentlicht werden sollen. Dieses Prinzip stößt im Bereich der Erforschung von materiellen und immateriellen Werken auf Grenzen, weil es hier zu einem Konflikt mit den Urheberrechten an den Forschungsgegenständen kommt. Wissenschaftler*innen sollen etwa über zeitgenössische Musik forschen, sie dürfen sie aber nicht als ganzes Werk digital zugänglich machen.

Ähnliche Herausforderungen ergeben sich auch bei den FAIR-Prinzipien der Forschungsdaten, die accessible und reusable zur Verfügung gestellt werden sollen. Das wissenschaftsethische Gebot stößt auch hier bei den materiellen und immateriellen Kulturgütern an Grenzen des Urheberrechts.

³ Vgl. Deutscher Museumsbund (2018), online unter <https://www.museumsbund.de/wp-content/uploads/2018/05/dmb-leitfaden-kolonialismus.pdf>

Handlungsfelder und Ziele

Die Ausgangssituation macht deutlich, dass die Themenbereiche Recht und Ethik in Bezug auf die Forschungsdaten zu materiellen und immateriellen Kulturgütern zentrale Herausforderungen darstellen, die Forschung behindern oder ermöglichen kann.

Auf der Ebene eines möglichen NFDI-Konsortiums für materielle und immaterielle Kulturgüter soll über das Helpdesk eine kompetente Rechtsberatung zur Verfügung stehen, die im Überblick über die verschiedenen rechtlichen Aspekte der unterschiedlichen Datentypen jenseits von Einzelfallgutachten und dennoch flexibel zu Rechtssicherheit im Umgang mit Forschungsdaten verhilft. Zudem sollen konsortial vertragliche Lösungen entwickelt werden, um die spätere Freischaltung zunächst rechteebewährter Daten vorzubereiten und die sich verändernde Rechtesituation immer wieder an die Daten rückzukoppeln. Dazu ist eine personell unterlegte Struktur nötig, die Nutzer*innen wie Anbieter*innen dauerhaft bei der Entwicklung fach- und gegenstandsspezifischer Lösungen unterstützt, hierfür Standards entwickelt und diese auch immer wieder an die aktuelle Rechtslage anpasst.

Ziele und Handlungsfelder des komplexen Bereichs der Rechte und Ethik, die im Rahmen der NFDI4Culture verfolgt werden, sind zudem Gegenstand eines lebhaften Diskurses. Sie will den Diskurs für die Wissenschaften versachlichen, um pragmatische und auf Ausgleich bemühte Lösungen zu finden. Die Initiative NFDI4Culture wirkt in dem Diskurs darauf hin, sowohl eine freie und offene Wissenschaft im Sinne des Open Access, Open Science und Open Data sicherzustellen, als auch die Rechte der Urheber*innen zu respektieren und das Verwertungsinteresse der Kultur- und Kreativwirtschaft zu unterstützen.

Literatur und Quellen

Deutscher Museumsbund (Hrsg.) (2018): Leitfaden zum Umgang mit Sammlungsgut aus kolonialen Kontexten, Berlin 2018. Online unter <https://www.museumsbund.de/wp-content/uploads/2018/05/dmb-leitfaden-kolonialismus.pdf>

Hartmann, T. (2018): Urheberrecht abgelaufen, trotzdem abgemahnt? Wikimedia kämpft vor Gericht für Gemeinfreiheit. Netzpolitik.org am 01.11.2018. Online unter <https://netzpolitik.org/2018/urheberrecht-abgelaufen-trotzdem-abgemahnt-wikimedia-kaempft-vor-gericht-fuer-gemeinfreiheit/>

Hartmann, T. (2017): Open Access rechtlich absichern – warum es ein Opt-in braucht, in: Konstanze Söllner, Bernhard Mittermaier (Hrsg.): Praxishandbuch Open Access, Berlin 2017, S. 45–52. Autorenversion online unter <http://hdl.handle.net/11858/00-001M-0000-002D-6F33-E>

Hartmann, T. (2014): Urheberrecht in der Bildungspraxis. Leitfaden für Lehrende und Bildungseinrichtungen, Bielefeld 2014.

Hartmann, T. (2013): Zur urheberrechtlichen Schutzfähigkeit der Forschungsdaten, in: Jürgen Taeger (Hrsg.): Law as a Service (LaaS) – Recht im Internet- und Cloud-Zeitalter, Bd. 1, Edewecht 2013, S. 505–515. Zweitveröffentlichung / Volltext online unter <http://hdl.handle.net/11858/00-001M-0000-0014-1208-E>

Klimpel, P. (2015): Eigentum an Metadaten? Urheberrechtliche Aspekte von Bestandsinformationen und ihre Freigabe, in: E. Euler, M. Hagedorn-Saupe, G. Maier, W. Schweibenz u. J. Sieglerschmidt (Hrsg.): Handbuch Kulturportale, Online-Angebote aus Kultur und Wissenschaft, S. 57–65, Berlin 2015.

Klimpel, P., E.-M. König (2015): Urheberrechtliche Aspekte beim Umgang mit audiovisuellen Materialien in Forschung und Lehre. Gutachten für die Gesellschaft für Medienwissenschaft und den Verband der Historiker und Historikerinnen Deutschlands, Berlin September 2015.

Klimpel, P., F. Rack u. J. H. Weitzmann (2017): Neue rechtliche Rahmenbedingungen für Digitalisierungsprojekte von Gedächtnisinstitutionen. Handreichung. 4. gänzlich neu bearbeitete Auflage, Berlin 2017. Online unter <http://dx.doi.org/10.12752/2.0.002.3>

Simon, H. (2007): prometheus und justitia – Bildarchive der Kunst- und Kulturwissenschaften im Spannungsfeld des medialen Umbruchs hin zu einer digitalen Informationsgesellschaft, in: K.-N. Peifer, G. Gersmann (Hrsg.): Forschung und Lehre im Informationszeitalter – zwischen Zugangsfreiheit und Privatisierungsanreiz (= Schriften zum Europäischen Urheberrecht 4), Berlin 2007, S. 65–86.

Universitätsbibliothek Mainz (Hrsg.) (2019): Handout: Rechtliche Fragen bei der Bereitstellung von Forschungsdaten, Mainz 2019. Online unter <https://www.digitale-bibliotheksdienste.uni-mainz.de/files/2019/02/Handreichung-Forschungsdatenbereitstellung.pdf>

4. Aufgabenbereiche

Die Fokusthemen (☞ 3.1 bis 3.5) beleuchten die Themen der Initiative NFDI4Culture grundlegend und im Querschnitt zu den sechs Schlüsselfragen, die bereits in einem Extended Abstract bei der DFG eingereicht wurden. Zu ihnen sollen Angebote und Lösungen speziell für die Arbeit mit Forschungsdaten zu materiellen und immateriellen Kulturgütern entwickelt werden.

4.1 Digitalisierung materieller und immaterieller Kulturgüter

Use Case: *In einer Gedächtnisinstitution soll eine Sammlung kultureller Objekte digitalisiert und die Digitalisate webbasiert für die Forschung zur Verfügung gestellt werden.*

Viele Forschungs- und Gedächtnisinstitutionen haben bereits seit den 1990er Jahren mit einer rudimentären datenbankgestützten Dokumentation begonnen, die jedoch nur in einigen großen Häusern über eine reine Inventarisierung hinausgekommen ist. In wenigen Fällen werden die Digitalisate bisher in Form webbasierter Informationsangebote und -infrastrukturen für die Forschung zur Verfügung gestellt (z. B. Museum Kunst und Gewerbe Hamburg, Städel Museum Frankfurt, UB Heidelberg). Die Digitalisierung erfolgt an einigen Stellen momentan nach **unzureichenden Standards** und mit **unterschiedlichem technischem Know-how**. Insbesondere kleineren Häusern fehlt meist das technische und praktische Wissen für eine Umsetzung. Dadurch steht bislang nur ein Bruchteil des kulturellen Erbes in digitaler Form für die Forschung zur Verfügung. Zudem werden Digitalisierungskampagnen häufig **noch zu wenig an den Bedarfen der Forschung ausgerichtet**. Somit existiert bereits auf der Ebene der Digitalisierung ein **sehr hoher Abstimmungsbedarf** zwischen den Institutionen für **forschungsgeleitete Digitalisierungsverfahren** in den Bereichen 2D-, 3D-, Video- und Audio-Digitalisierung.

4.2 Datenstandards, Datenqualität und Datenkuratierung

Use Case: *Forscher*innen verschiedener Disziplinen möchten auf Basis von Digitalisaten Merkmale mittelalterlicher Glasmalereien oder Musiknotate codieren und annotieren. Welche Standards können verwendet werden?*

Eine **generelle Interoperabilität** von Forschungsdaten zu materiellen und immateriellen Kulturgütern **ist in der Praxis bislang noch nicht gegeben**. Gleichwohl existieren gute Ansätze in Form von **Standards**. So hat das International Council of Museums mit dem **CIDOC-CRM** ein Referenzmodell entwickelt, mit denen sich Kulturgüter präzise beschreiben lassen. Mit dem **IFLA-LRM** steht ein Entity-Relationship-Modell zur (bibliographischen) Beschreibung verschiedener Ausprägungen immaterieller und materieller Entitäten zur Verfügung. Mit der Music Encoding Initiative (**MEI**) existiert ein einschlägiger Standard für die Codierung notierter Musik. Über **LIDO** (Lightweight Information Describing Objects) können Metadaten zwischen Repositorien ausgetauscht werden. Das International Image Interoperability Framework (**IIIF**) stellt standardisierte APIs bereit, die einen übergreifenden Austausch digitaler Objekte ermöglichen und hat eine native Kompatibilität mit dem Annotation Data Model des W3C. **Semantic Web-Technologien** ermöglichen die verbesserte Auffindbarkeit von verteilten Forschungsdaten und die qualifizierte Verlinkung zwischen den Datenrepositorien. Verknüpfungen verweisen auf **Normdateien und Fachvokabulare** wie z. B. die Gemeinsame Normdatei (**GND**), den Art and Architecture Thesaurus (**AAT**), den Thesaurus of Geographic Names (**TGN**) und zunehmend auch Wikidata. Die Verwendung, Weiterentwicklung und gegenseitige Verschränkung dieser und weiterer Standards muss in den Institutionen und bei den Forschenden angeregt werden. Es besteht ein großer Bedarf für die Entwicklung **übertragbarer Datenmanagement-, Qualitätssicherungs- und Kuratierungsprozesse**.

4.3 Forschungswerkzeuge und Datendienste

*Use Case: Forscher*innen benötigen zur Untersuchung kultureller Schaffensprozesse in spezifischen historischen Epochen Zugriffs- und Analysemöglichkeiten über eine möglichst breite Anzahl entsprechender digitaler Kulturgüter.*

Es gibt einen hohen **Bedarf für verteilte Nachweis- und Recherchesysteme**, die einen einfachen Zugang zu und eine leistungsstarke Analyse auf großen Datenbeständen zu Kulturgütern ermöglichen. Da die Digitalisate und prozessualen Daten in diesem Feld sehr unterschiedlichen Rechten unterliegen, müssen im Rahmen des Konsortiums **föderierte Datendienste** etabliert werden, mit denen die **Datensouveränität** der bestands-erhaltenden Institutionen gewährleistet werden kann. Plattformen für bestimmte Typen von Daten existieren bereits (z. B. prometheus - Das verteilte Bildarchiv für Forschung- und Lehre, das Répertoire International des Sources Musicales (RISM), das Fachrepositorium media/rep/, im weiteren Sinne auch die Europeana und die DDB). Der effiziente föderierte Zugriff auf dezentrale heterogene Datenrepositorien erfordert offene standardisierte Metadaten basierend auf offenen Schnittstellen. In der Forschung ist der **Etablierungsgrad** solcher Systeme aber noch sehr unterschiedlich. Die Initiative NFDI4Culture möchte sich darum bemühen, die Zugänglichkeit von Forschungsdaten und verteilten Archiven zu fördern sowie algorithm- und softwarebasierte Forschungsmethoden (bspw. KI, Computer Vision) zu integrieren. Damit soll dem **hohen Bedarf an digitalen Forschungswerkzeugen** entsprochen und Synergien für eine nachhaltige Nutzung der Softwarewerkzeuge z. B. als Open Source-Anwendungen erzielt werden. Bestehende Forschungszentren wie das Zentrum Musik – Edition – Medien (ZenMEM), das Cologne Center for eHumanities (CCeH), die Digitale Akademie der Mainzer Akademie sowie die Fachinformationsdienste werden eine höhere **Software Sustainability** gewährleisten.

4.4 Nachhaltige Bereitstellung und Publikation von Daten (LZA, LZV)

*Use Case: Nachwuchswissenschaftler*innen möchten und müssen digitale Daten und Ergebnisse aus ihren Untersuchungen im Open Access dauerhaft und zitierfähig publizieren.*

Angebote für die **professionell unterstützte, nachhaltige und unkomplizierte Open Access Publikation** von Forschungsdaten werden in den Geisteswissenschaften immer stärker nachgefragt. Im Bereich der generischen Lösungen haben sich neben privatwirtschaftlich agierenden Plattformen (z. B. Academia) auch öffentlich geförderte Lösungen wie z. B. Zenodo (CERN) und RADAR (FIZ Karlsruhe) etabliert. **Fachnahe Lösungen**, die ein an wissenschaftlichen Standards ausgerichtetes Publizieren von Forschungsdaten (bspw. unter Einbeziehung DataCite oder ORCID) und Ergebnissen in Form von Publikationen ermöglichen und eine wissenschaftliche Sichtbarkeit in der eigenen Community sicherstellen, sind bisher vor allem durch die Fachinformationsdienste entwickelt worden, **aber noch nicht in der Breite bekannt**. So wurden an der UB Heidelberg (heiUP) und an der SLUB Dresden (musiconn.publish) Lösungen entwickelt, die den digitalen Workflow von der Publikation von Forschungsdaten über eBooks und online Publikationen bis hin zu einer Printausgabe ermöglichen. **Diese Lösungen gilt es auszubauen und in die Breite der teilnehmenden Fachdisziplinen hineinzutragen**. Für Forschungsdaten aus Forschungsprojekten sollen **fachspezifische Repositorien** die Datenpublikationen auch jenseits von Text und zweidimensionalem Bild etwa in (annotierten) 3D-Formaten ermöglichen und dauerhaft bereitstellen.

4.5 Rechte und Datenethik

Use Case: *Eine Institution möchte Digitalisate und prozessuale Daten zu künstlerischen Artefakten moderner Künstler*innen rechtskonform für die Forschung bereitstellen und benötigt dafür ein funktionierendes Rechtmanagement.*

Forschungsdaten zu materiellen und immateriellen Kulturgütern unterliegen – genau wie die Kulturgüter selbst – meist einer komplexen Rechts-situation. Bei der Publikation von Daten müssen Aspekte der Datenethik beachtet werden, z. B. im Bereich von Daten zu Kulturgütern aus kolonialem Erbe oder bei Forschungsdaten, deren Veröffentlichung zu einem Verlust der eigentlichen Objekte führen könnte (z. B. durch Raubgräberei). Es wird eine zentrale Aufgabe sein, durch Beratung **Rechtssicherheit für die Anbieter*innen und Nutzer*innen** von digitalen Angeboten mit Blick auf die bestehende Gesetzgebung herzustellen. Auch müssen effiziente und dynamische **Authentifizierungs- und Zugangslösungen** implementiert werden, die digitale Forschung zum einen ermöglichen und zum anderen Rechtssicherheit gewährleisten. Dieser Implementierungsprozess endet nicht mit der Überführung von Forschungsdaten in ein Repository sondern er muss im Zuge der Verfügbarhaltung kontinuierlich juristisch überprüft und angepasst werden. Darüber hinaus sollten die Erfahrungen und Anforderungen an den Gesetzgeber weitergegeben werden, um diesen bei einer wissenschaftsfreundlichen **Gesetzgebung zu beraten**.

4.6 Expertise, Beratung, Weiterbildung und Qualifikation

Use Case: *Forscher*innen möchten ein digitales Forschungsprojekt unter Einbeziehung materieller oder immaterieller Kulturgüter aufsetzen und benötigen Beratung und Expertise im Datenmanagement, im Datenumgang und in softwarebasierten Forschungsmethoden.*

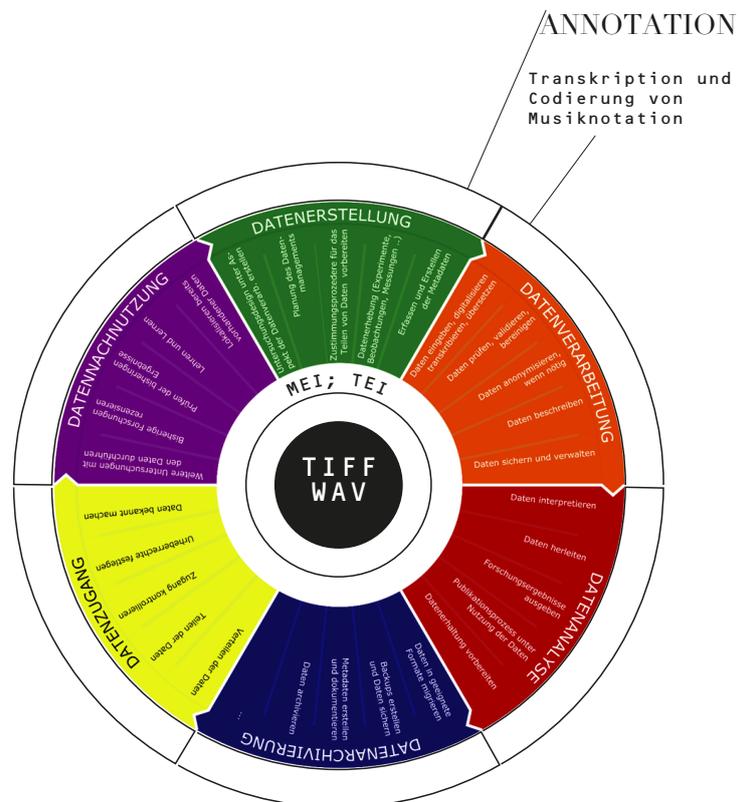
Forschungsdaten zu materiellen und immateriellen Kulturgütern stellen immer nur eine Repräsentation des Gegenstandes und nie den Gegenstand selbst dar. Für diese besondere Situation gilt es innerhalb des Konsortiums **spezifische Beurteilungskriterien für einen reflektierten Umgang** im Sinne der **Data Literacy** zu entwickeln. Da Forschungsprozesse zunehmend datengetrieben ablaufen, ist es für Forschende von entscheidender Bedeutung, auch die Algorithmen und die **Software-Werkzeuge** in wissenschaftlichen und gesellschaftlichen Zusammenhängen **kritisch einschätzen und reflektieren** zu lernen (☞ 3.1). Es besteht ein großer Bedarf an bundesweit verfügbaren **Beratungs- und Qualifizierungsangeboten**, die auf allen Stufen eines Antrags- oder Forschungsprozesses kompetent ansetzen. NFDI4Culture möchte sich auch um eine stärkere **curriculare Verankerung** von Data und Code Literacy im Sinne von **Critical Data Studies und Critical Algorithm Studies** bemühen.

Anhang: User Stories

K. Bicher (Universität Paderborn), A. Hammes (Universität Paderborn), S. Pittroff (AdW Mainz), Franziska Voß (Universitätsbibliothek Frankfurt a. M.), A.-V. Walzel (SLUB Dresden)

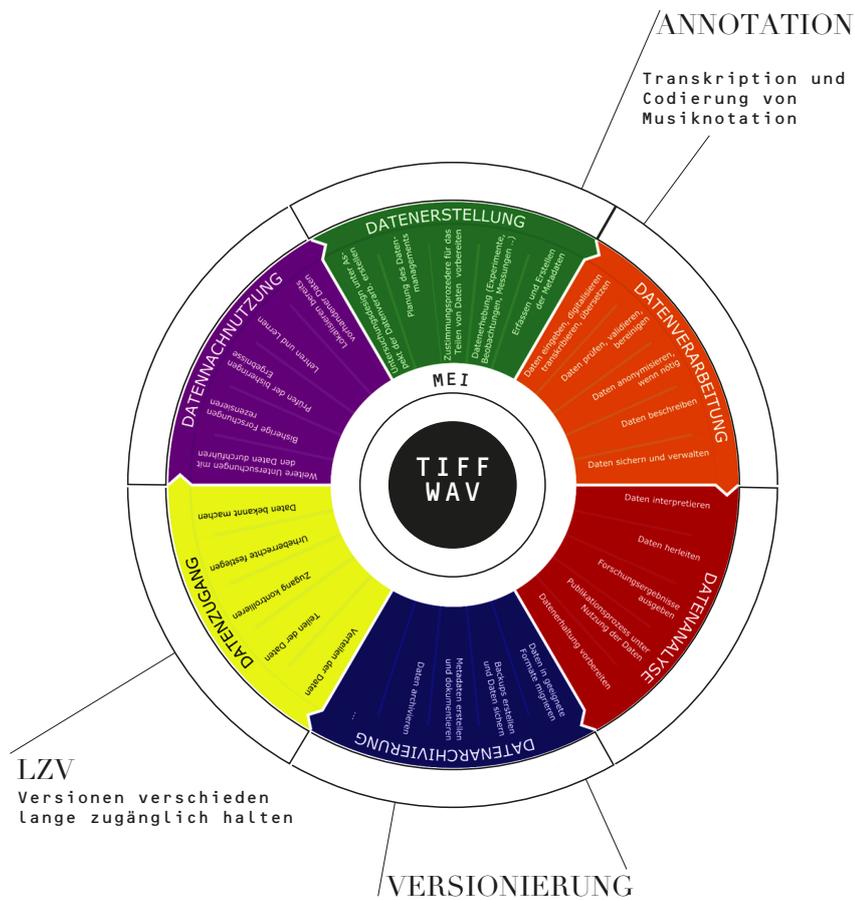
Grafik **Der Forschungsdatenzyklus und verwandte Themengebiete**: CC-BY 4.0, SLUB Dresden; basierend auf Darstellungen von forschungsdaten.info und dem UK Data Archive

Grafik **User Stories**: CC-BY-SA 4.0, Sarah Pittroff, AdW Mainz



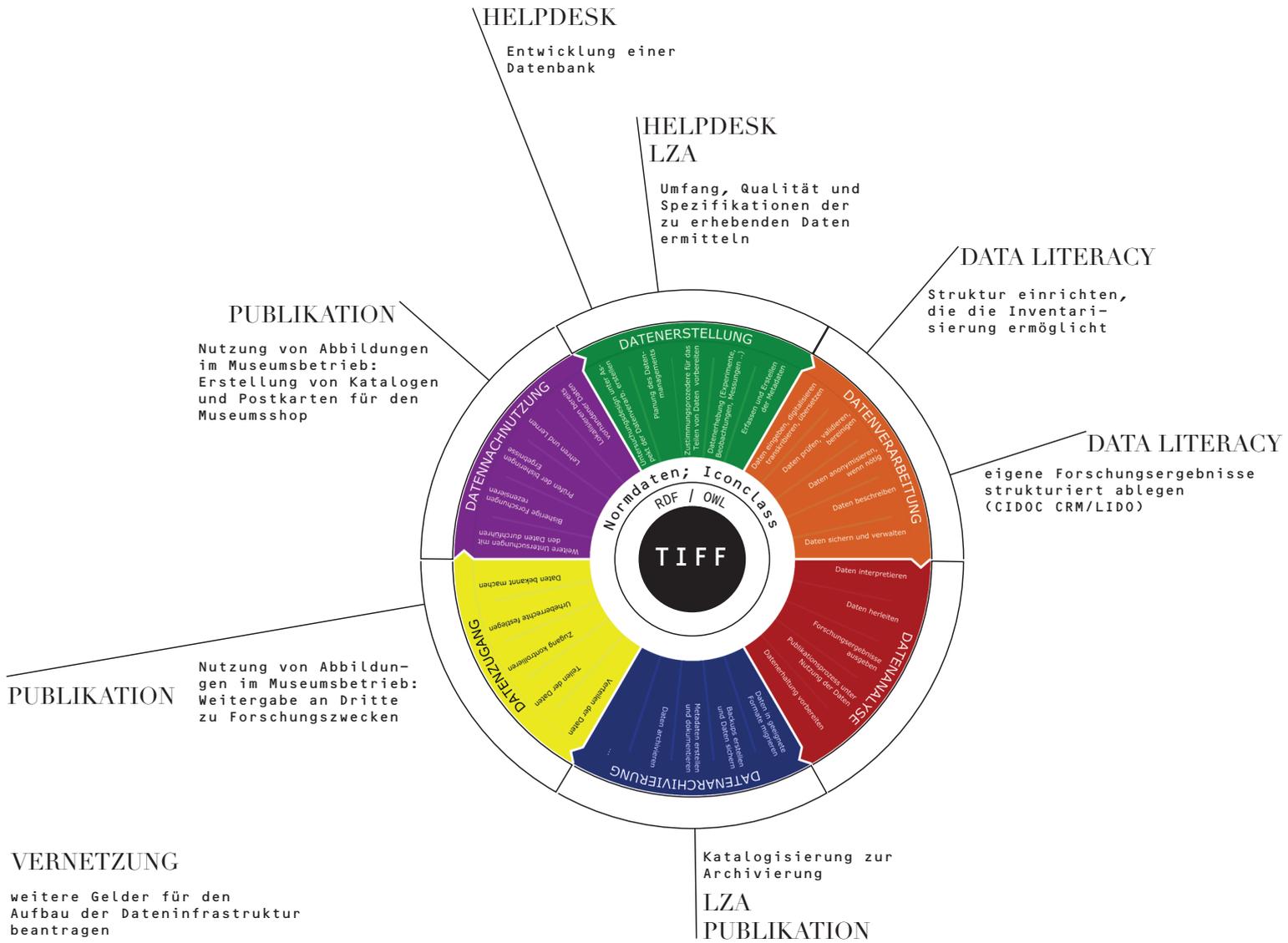
NFDI4Culture User Story II

Als Musikwissenschaftlerin möchte ich musikalische Quellen annotieren (mit Text, ggf. auch wiederum Musiknotation), um sie weiter zu analysieren.



NFDI4Culture User Story IIa

Als Editor einer Musikedition möchte ich eine MEI-basierte Notenausgabe erstellen und verschiedene Versionen verschieden lange zugänglich halten, um a) Entscheidungen erst später zu treffen und/oder b) Entscheidungen nachvollziehbar zu halten.



NFDI4Culture User Story III

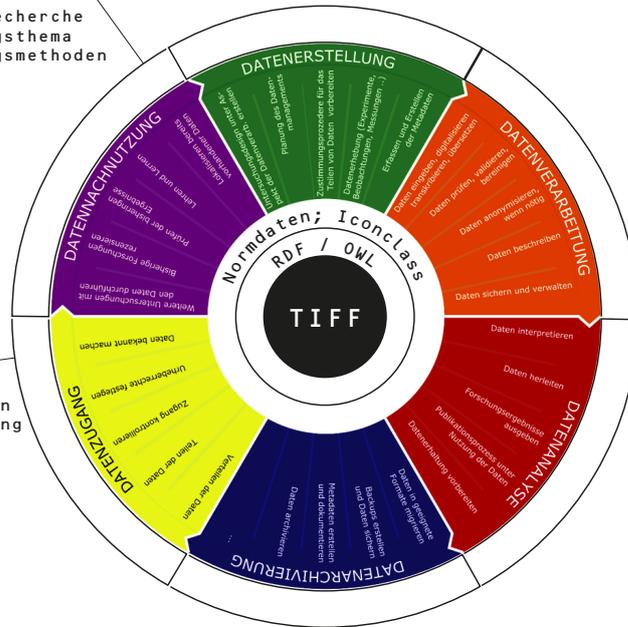
Als Kustodin möchte ich die Gemäldesammlung des Landesmuseums erstmals digital erfassen, um sie zu inventarisieren, eigene Forschungsergebnisse strukturiert abzulegen, zu publizieren und zur Nachnutzung bereitzustellen.

DATA LITERACY

(Internet-)Recherche zum Forschungsthema und Forschungsmethoden

HELPDESK PUBLIKATION

Beratung zur Form der Veröffentlichung von Forschungsergebnissen der Dissertation

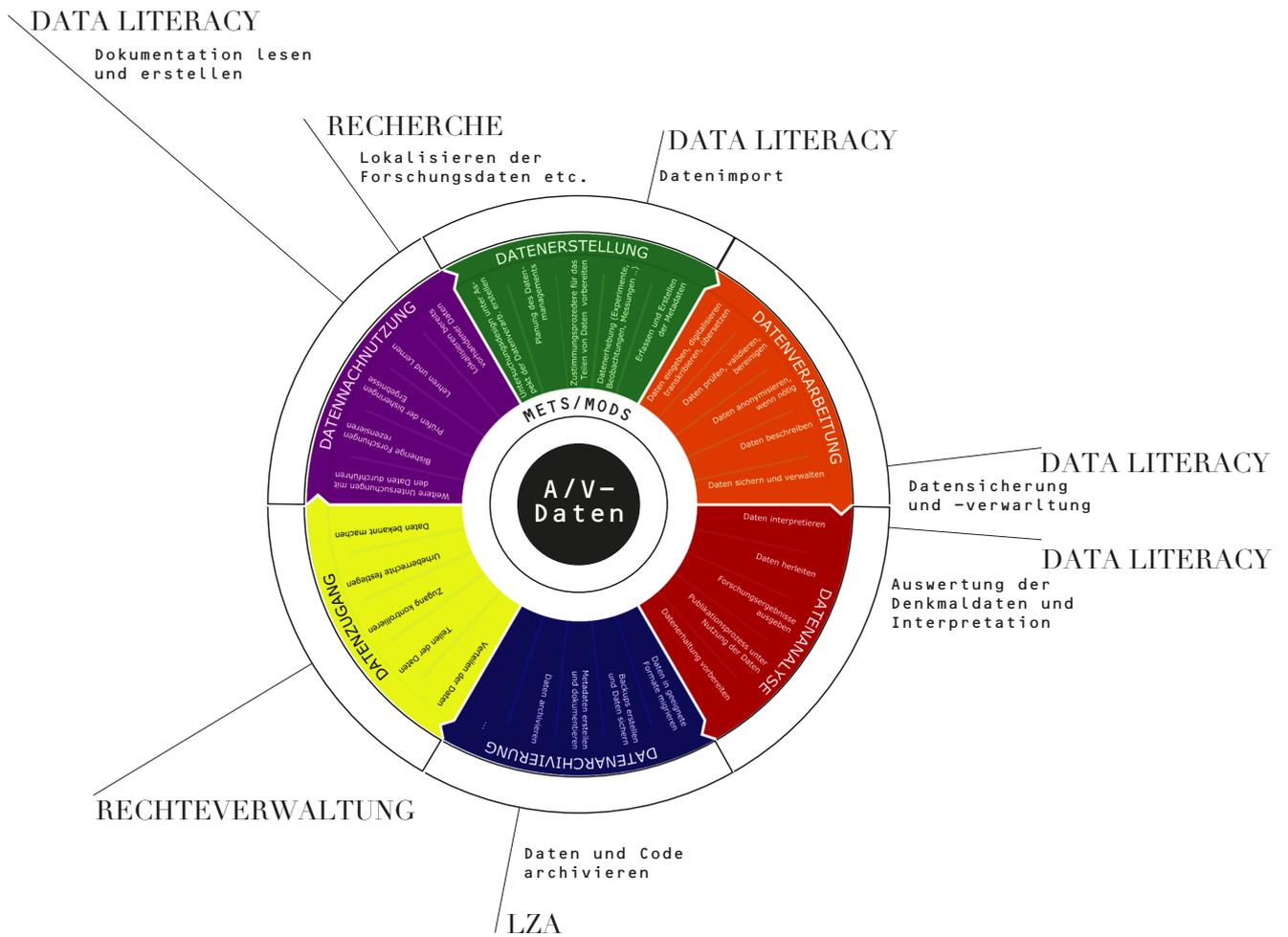


Untersuchung, Einordnung und Vergleich der Forschungsgegenstände mit digitalen Methoden

HELPDESK

NFDI4Culture User Story IV

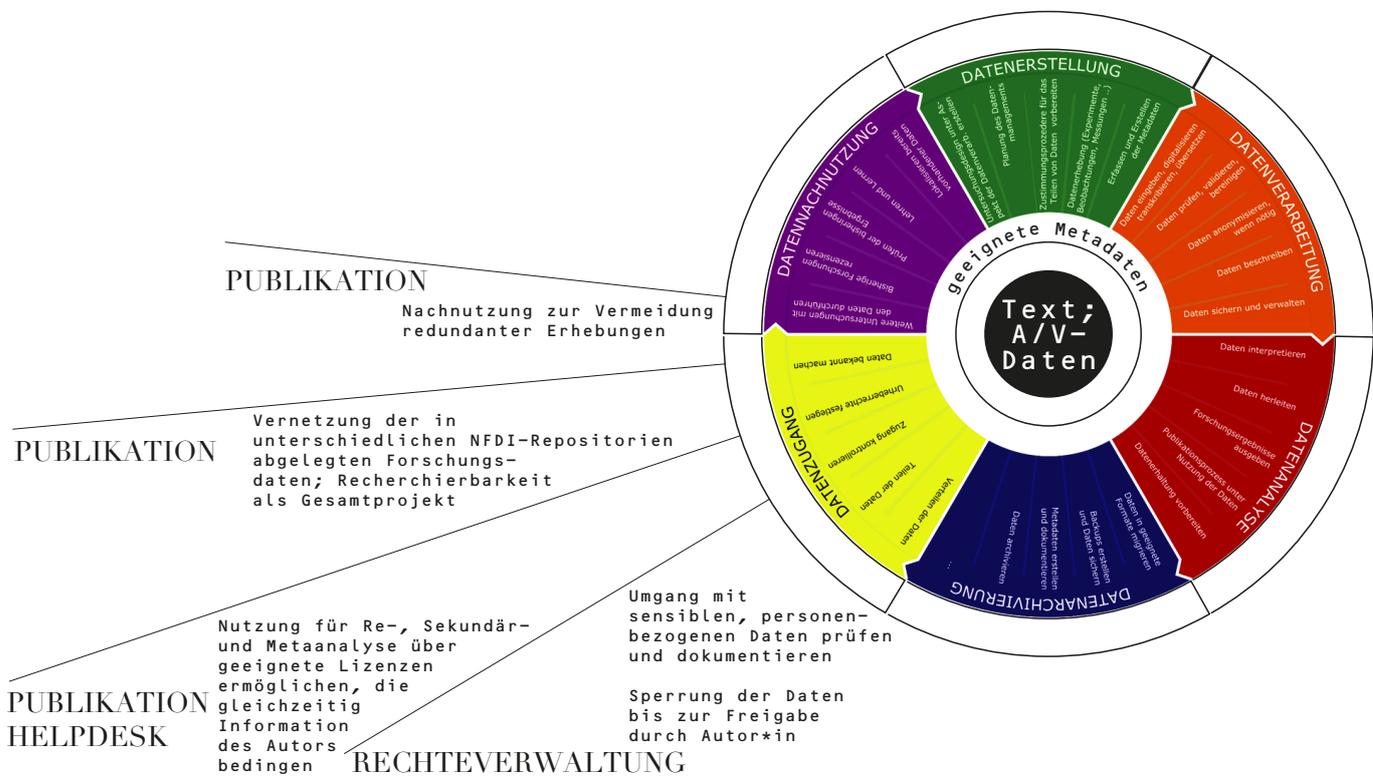
Als Promovendin im Bereich Kunstgeschichte recherchiere ich zum Thema Herrscherportraits. Dazu suche ich Informationen im Internet, um meine Recherche zu komplettieren und zu optimieren. Meine Promotion werde ich digital publizieren.



NFDI4Culture User Story VII

Als Musikethnologin bin ich verantwortlich, während meiner Forschung entstandene Audio-/Videodateien (resp. -Korpora) gemäß den vorher verabredeten rechtlichen Vereinbarungen zugänglich zu halten.

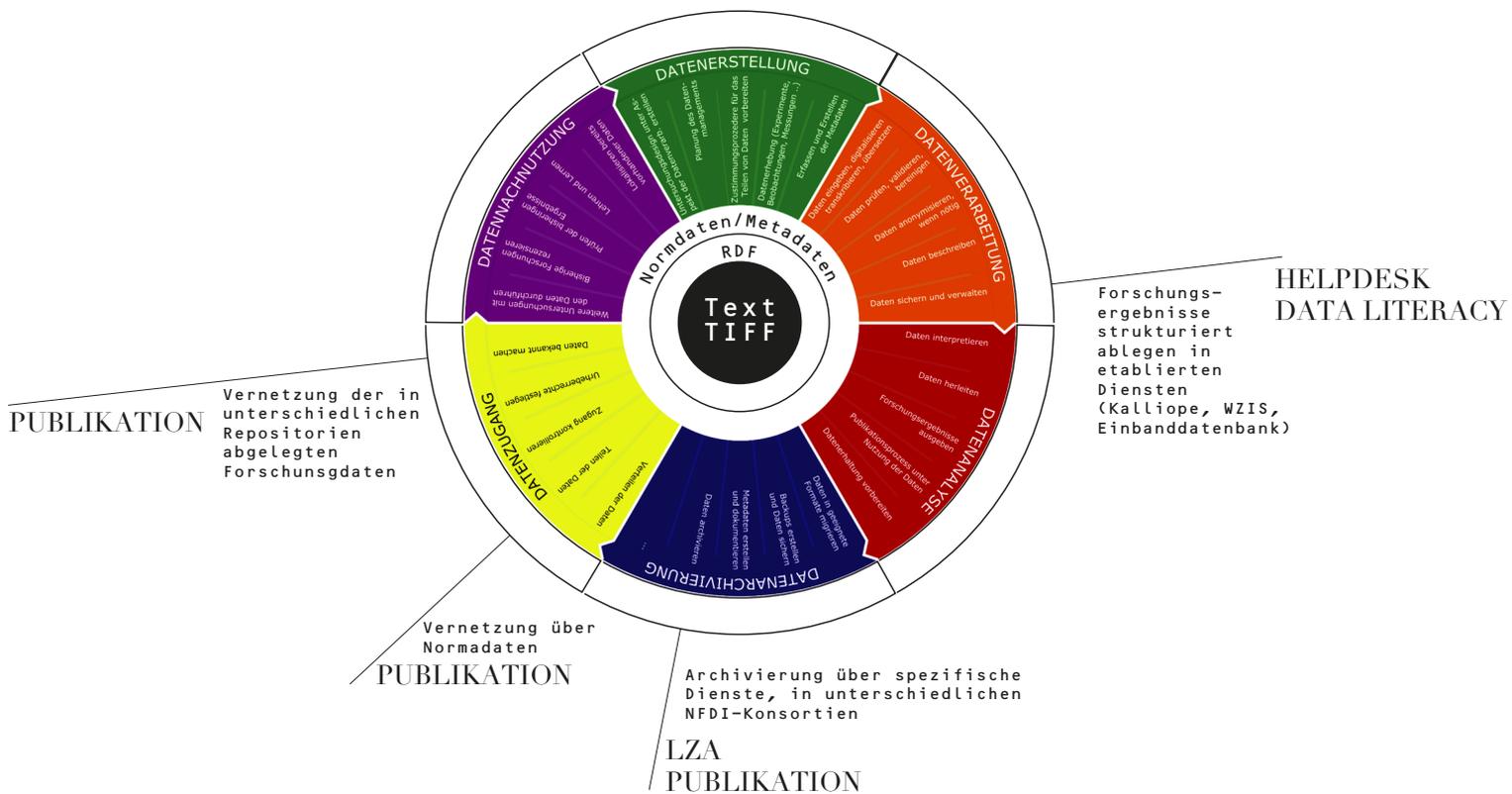
VII a: Als Kustodin der zeitgenössischen Sammlung eines großen Museums möchte ich das Konvolut Video-Art und die Dokumentation der Performances für die Forschung zugänglich machen.



NFDI4Culture User Story VIII

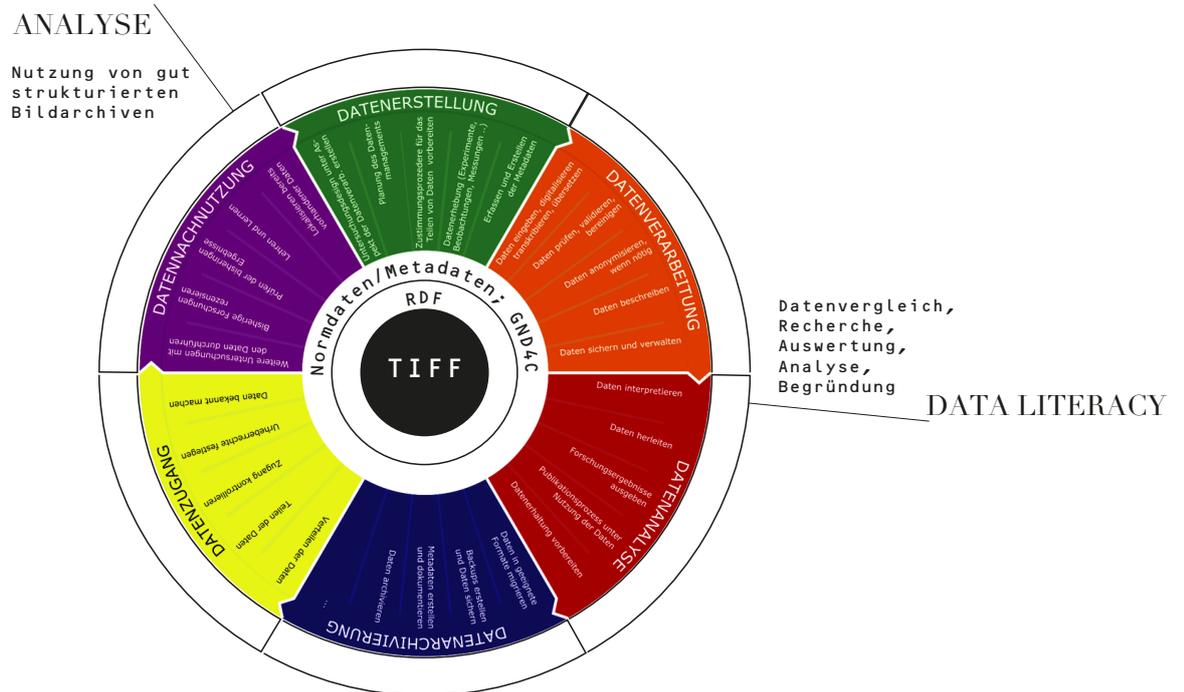
Als Musiksystematiker handiere ich mit verschiedenen quantitativen und qualitativen Datenformaten, unter anderem Fragebögen, Interviews, audio-/visuellen Aufzeichnungen, neuro-/physiologischen Messungen, psychologischen Testverfahren, Online-Surveys oder Experimental-Settings sowie unterschiedlichen Programmcodes.

Der jeweiligen Methodik entsprechend weisen die Datensätze zudem sehr unterschiedliche Umfänge auf. Diese Daten sollen zum Zwecke der Nachprüfbarkeit der wissenschaftlichen Befunde (Qualitätssicherung), zum Zwecke der späteren Nutzung für Re-, Sekundär- und Metaanalysen (Optimierung des Erkenntnisgewinns) sowie zum Zwecke der Vermeidung redundanter Erhebungen (Maximierung des Kosten-Nutzen-Verhältnisses) zur Verfügung gestellt werden.



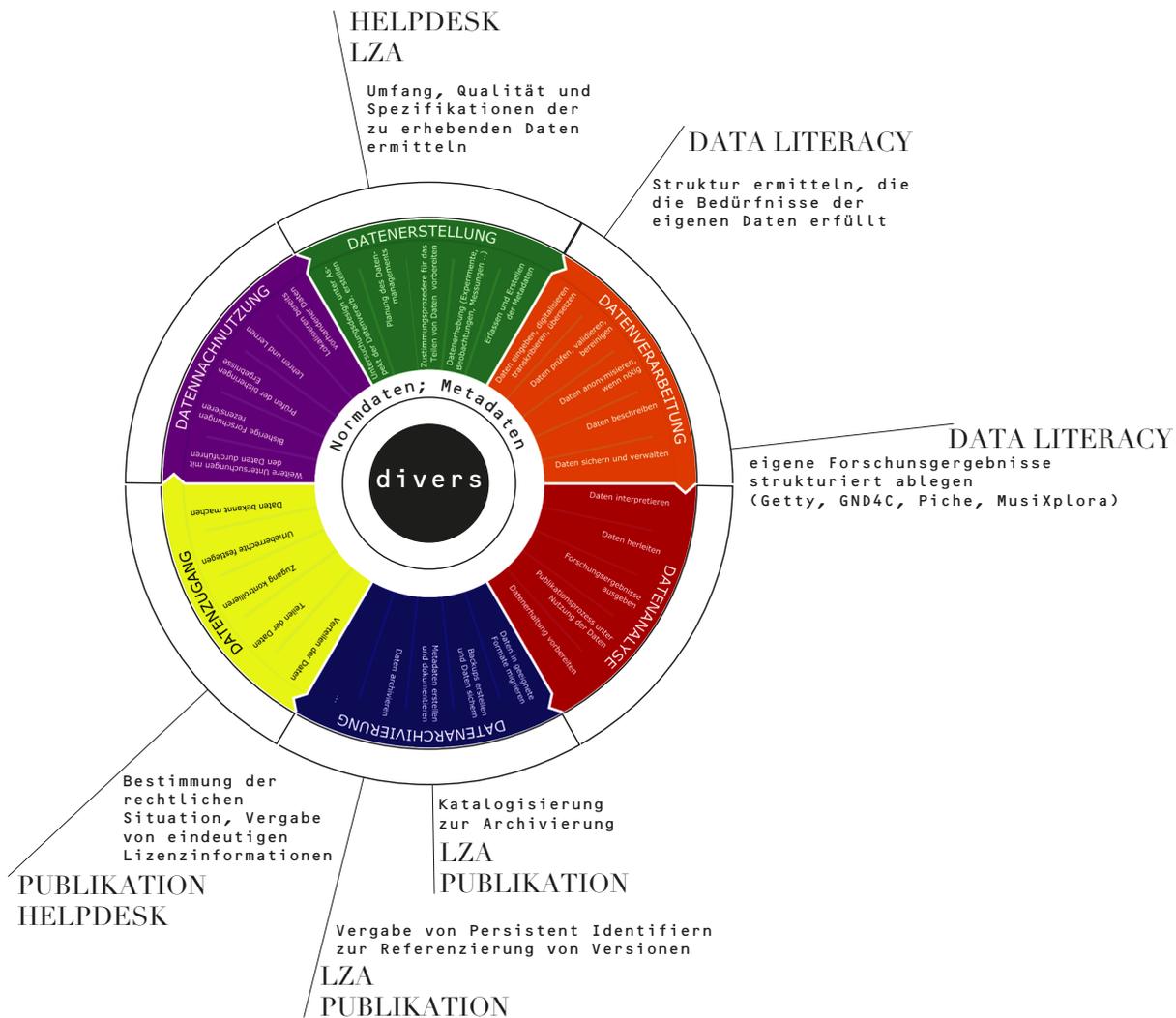
NFDI4Culture User Story X

Als Editor möchte ich die entstehenden Kontextinformationen (zu Wasserzeichen, Schreibern, Schreibprozessen, Überlieferung, Verlagswesen, Notenstich, Sammlungen, Quellenwanderungen, Antiquariatshandel etc., auch Korrespondenz mit Herausgebern/Verlagen, Wissenschaftler*innen/Musiker*innen, die das Spezialwissen der Einrichtungen für Anfragen in Anspruch nehmen) für andere zugänglich und nachnutzbar machen, um Doppelarbeit zu vermeiden.



NFDI4Culture User Story XII

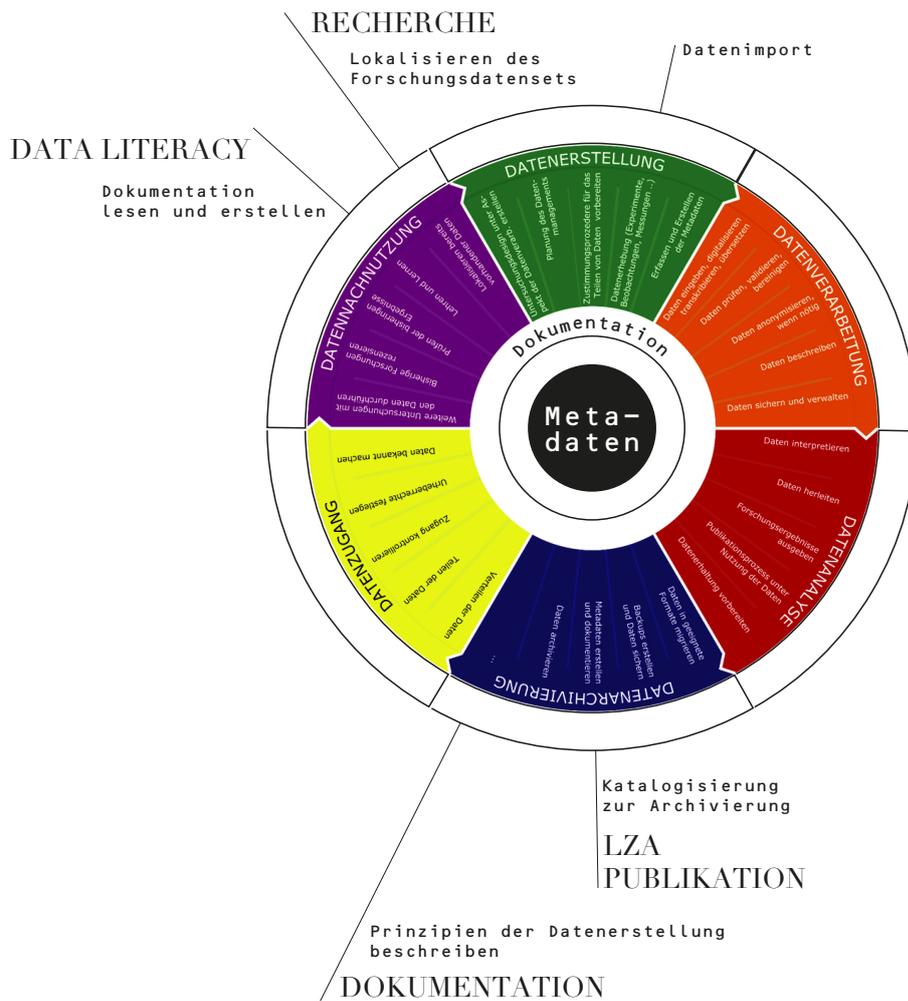
Als Denkmalpfleger suche ich nach bundesweiten Vergleichsbauten zum Stuttgarter Hauptbahnhof, um seine singuläre baukünstlerische Alleinstellung zu belegen, vor Gericht den Denkmalwert zu begründen und somit das Gebäude vor dem Abriss zu bewahren.



NFDI4Culture User Story XIII

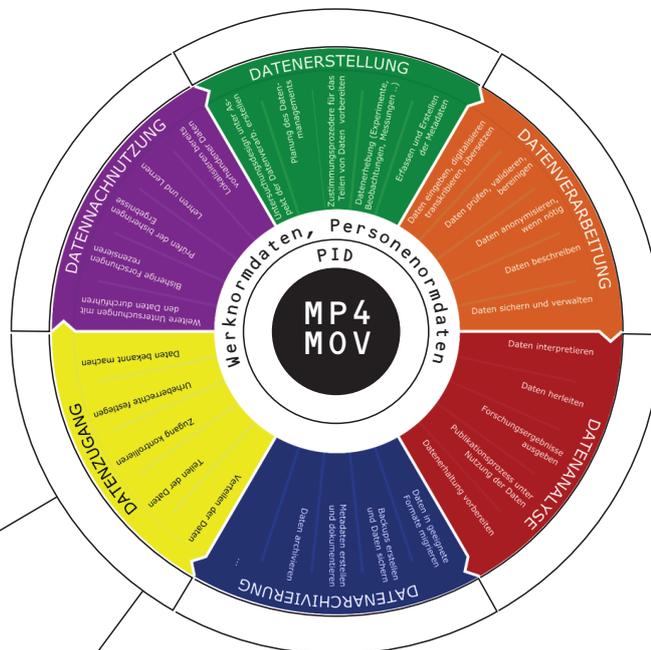
Als Organologe produziere ich medial vielfältige Forschungsdaten primär hinsichtlich materieller und digitaler Objekte (Musikinstrumente, Toninformations-/Tonträger, Modelle/Simulationen, AR/VR-Artefakte), sekundär jedoch auch für deren Kontexte (etwa für Personen, Körperschaften, Orte, Ereignisse, Sachen, Werke oder Titel) – so beispielsweise (textbasierte Daten,

Annotationen, Bild-, Audio-/Video-Medien, Maßreihen, Steuerinformationen, Visualisierungen, Augmented/Virtual Reality-Applications, beschreibende Metadaten). Mit der voranschreitenden Digitalisierung zeigen die Artefakte eine wachsende Komplexität (etwa durch die multiple Kombination verschiedenartiger Digitalisate). Für einen effizienten Wissenstransfer benötige ich eine zugängliche Datenbank mit der Möglichkeit, persistent in verschiedenen Stadien meines Forschungsprozesses zu referenzieren.



NFDI4Culture User Story XIV

Als Forscher möchte ich Forschungsdaten aus vorherigen Projekten nachnutzen, kann sie aber nicht interpretieren.



RECHTE

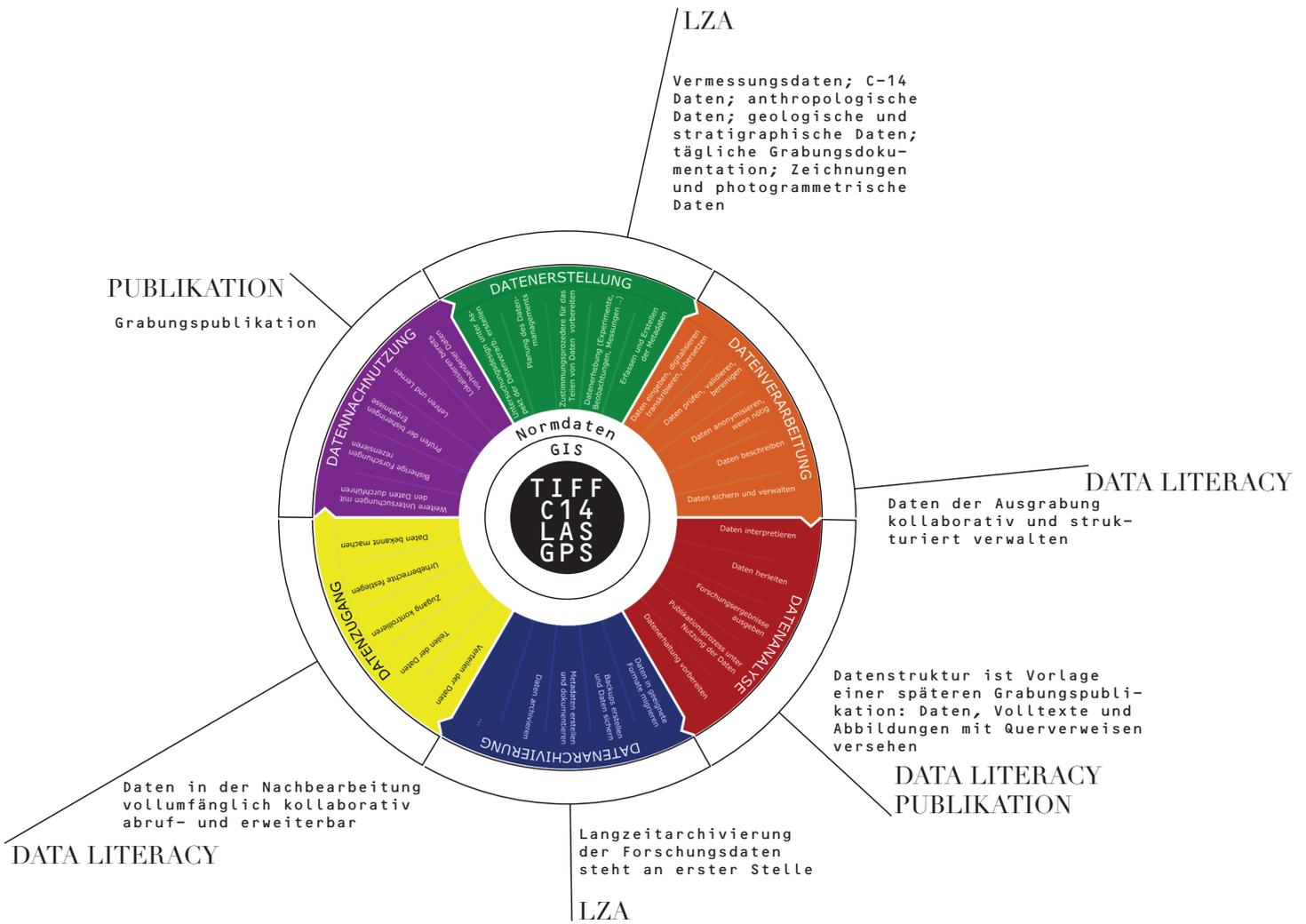
das digitale Medienarchiv kann nur einer ausgewählten Nutzergruppe zur Verfügung gestellt werden.

LZV

Streaming Option des Medienarchivs

NFDI4Culture User Story XIX

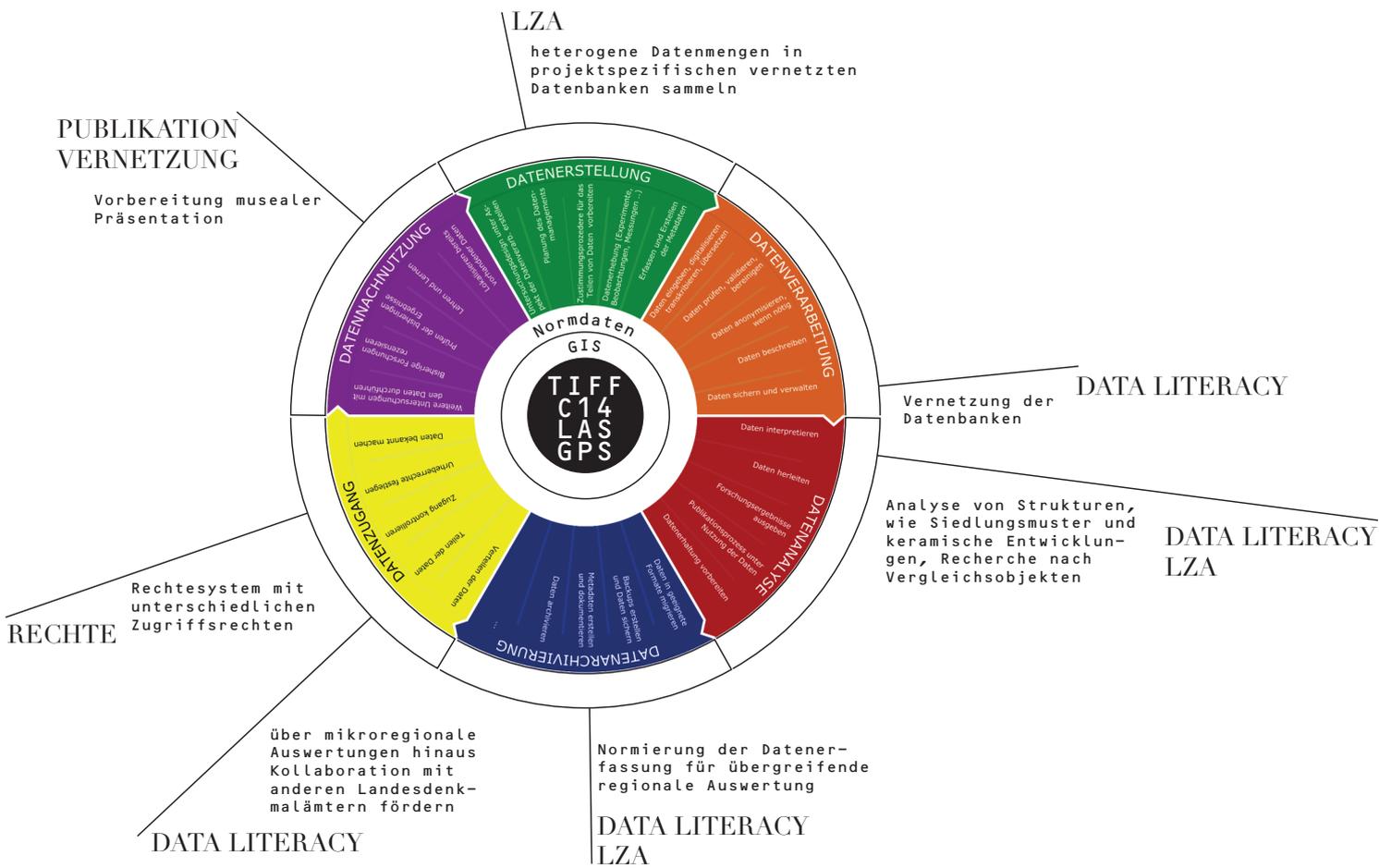
Als Universitätsprofessorin der Theater- und Tanzwissenschaft und Theaterpraktikerin möchte ich ein digitales Medienarchiv mit Streaming Option und einer Zugangsverwaltung / Nutzermanagement aufbauen.



NFDI4Culture User Story XV

Als Universitätsprofessorin leite ich ein Feldforschungsprojekt im europäischen Ausland. Die enormen Datenmengen, die im Zuge meiner Ausgrabungen täglich gesammelt werden (Vermessungsdaten; C-14 Daten; anthropologische

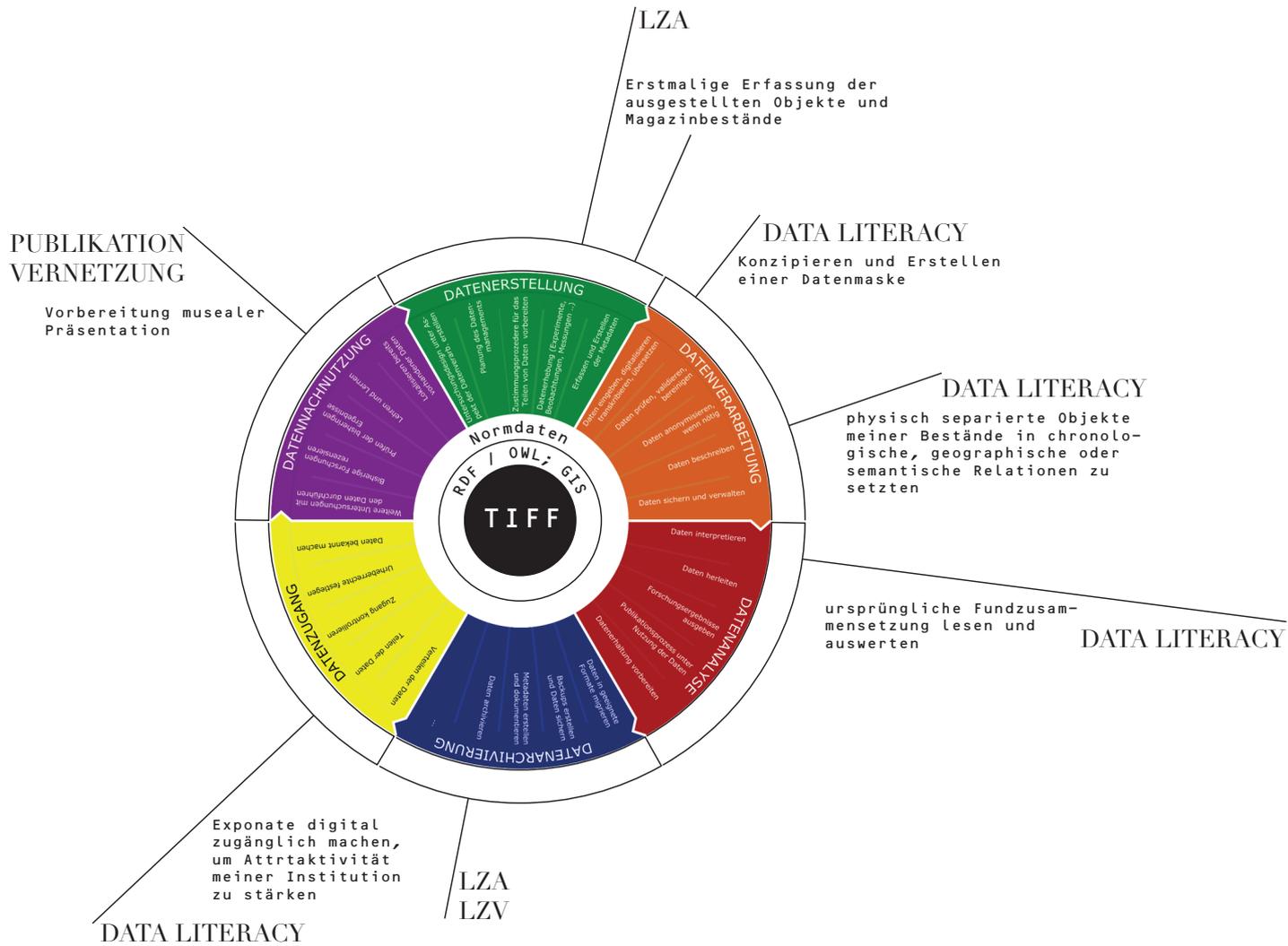
Daten; geologische und stratigraphische Daten; tägliche Grabungsdokumentation; Zeichnungen und photogrammetrische Daten), möchte ich in strukturierter Form digital verwalten.



NFDI4Culture User Story XVI

Als Leiterin eines Landesdenkmalamtes arbeite ich an mehreren Projekten gleichzeitig: Bauaufnahmen, Grabungsprojekte, Archivierung, Bestandsaufnahmen, Survey-Projekte und Feldbegehung, Auswertung, Katalogisierung und Aufarbeitung der Magazinbe-

stände, Vorbereitung musealer Präsentationen. Obwohl die gesammelten Datenmengen heterogener Natur sind, möchte ich alle Daten in projektspezifischen Datenbanken sammeln und eine Vernetzung derselben untereinander ermöglichen.



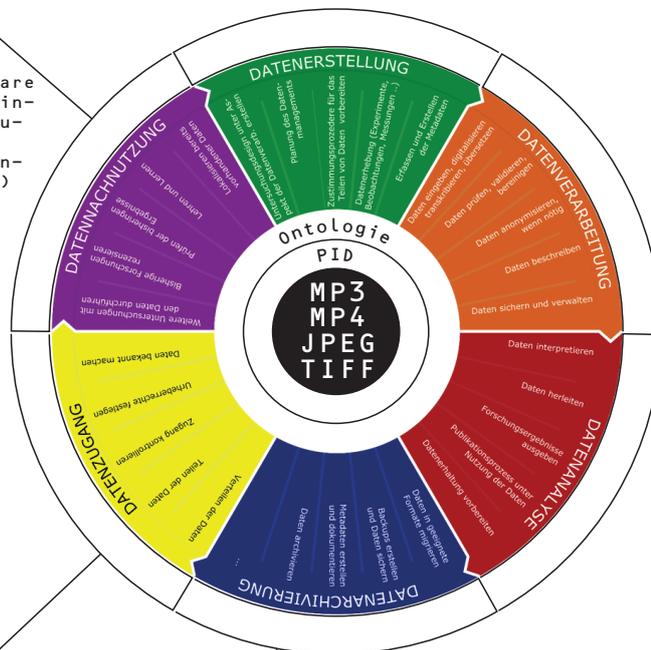
NFDI4Culture User Story XVII

Als Leiterin eines Landesdenkmalamtes arbeite ich an mehreren Projekten gleichzeitig: Bauaufnahmen, Grabungsprojekte, Archivierung, Bestandsaufnahmen, Survey-Projekte und Feldbegehung, Auswertung, Katalogisierung und Aufarbeitung der Magazinbestände,

Obwohl die gesammelten Datenmengen heterogener Natur sind, möchte ich alle Daten in projektspezifischen Datenbanken sammeln und eine Vernetzung derselben untereinander ermöglichen.

PUBLIKATION

Die Annotationssoftware kann von Ausbildungsinstitutionen (Hochschule(n) für Tanz) und Praktikern (Tänzer*innen, Performer*innen) nachgenutzt werden.



Die Ergebnisse, in Form der annotierten Tanzpartituren, stehen der Forschung frei zur Verfügung

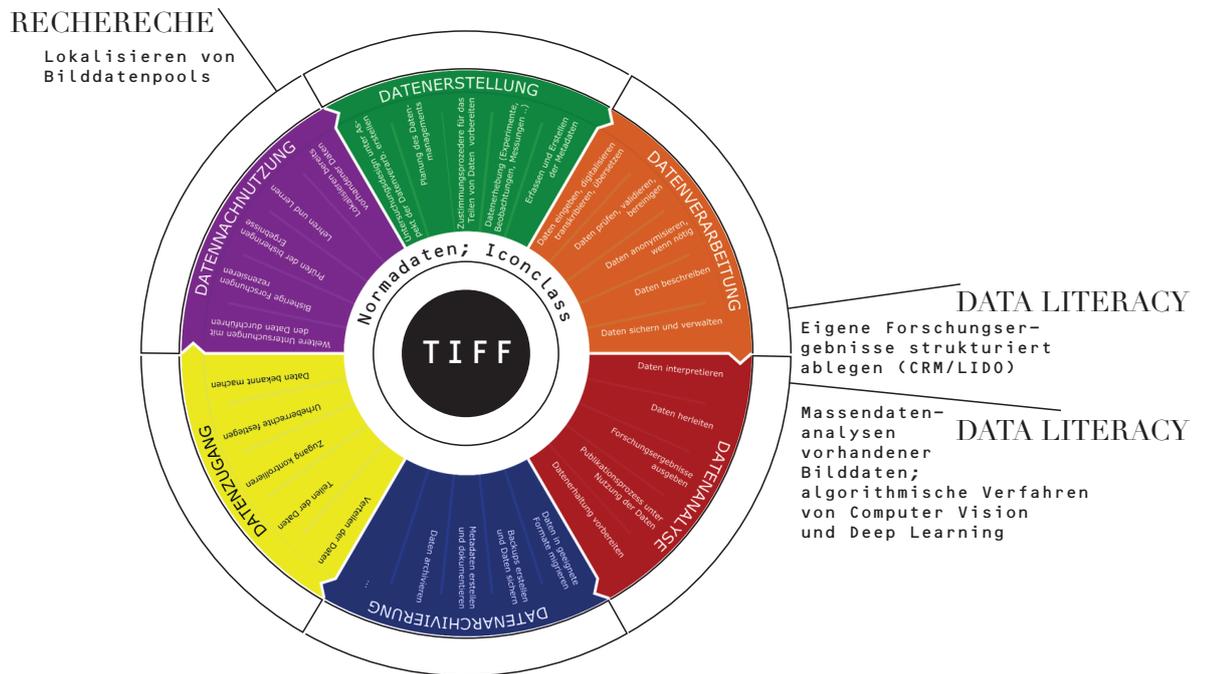
Software wird versioniert und archiviert

LZA

DATA LITERACY

NFDI4Culture User Story XVIII

Als Projektleiter / Koordinator / Entwickler einer open source Software zur digitalen Annotation von choreographischen Partituren suche ich nach Abschluss des Projektes einen Partner, der meine gesammelten Daten langzeitverfügbar aufbewahrt und verfügbar hält.



NFDI4Culture User Story XX

Als ForscherIn möchte ich Massendatenanalysen von mehreren 100.000 Gemälden durchführen, um die Einflüsse von Individualstilen auf die Ausbreitung von Regionalstilen zu untersuchen.

Impressum

Herausgeberin:

NFDI4Culture

<https://nfdi4culture.de/>

Sprecher*innen der Initiative:

Prof. Dr. Dörte Schmidt
Gesellschaft für Musikforschung

Prof. Dr. Holger Simon
Verband Deutscher Kunsthistoriker e.V.

Geplante Hosting Institution:

Akademie der Wissenschaften
und der Literatur | Mainz

Prof. Torsten Schrade
Digitale Akademie
Geschwister-Scholl-Str. 2
55131 Mainz

Redaktion:

Katrin Bicher
Universität Paderborn

Ortrun Brand
Philipps-Universität Marburg

Andrea Hammes
Universität Paderborn

Torsten Schrade
Akademie der Wissenschaften
und der Literatur | Mainz

Sarah Pittroff
Akademie der Wissenschaften
und der Literatur | Mainz

Annika-Valeska Walzel
Sächsische Landesbibliothek – Staats- und Universitätsbibliothek Dresden

Autor*innen und Beteiligte:

Reinhard Altenhöner
Staatsbibliothek zu Berlin / Preußischer Kulturbesitz

Katrin Bicher
Universität Paderborn

Christian Bracht
Philipps-Universität Marburg

Ortrun Brand
Philipps-Universität Marburg

Ina Blümel
Technische Informationsbibliothek Hannover

Klaus Bulle
Philipps-Universität Marburg

Maria Effinger
Universitätsbibliothek Heidelberg

Andrea Hammes
Universität Paderborn

Thomas Hartmann
FIZ Karlsruhe - Leibniz-Institut für
Informationsinfrastruktur

Angela Kailus
Philipps-Universität Marburg

Jürgen Kett
Deutsche Nationalbibliothek

Sarah Pittroff
Akademie der Wissenschaften
und der Literatur | Mainz

Daniel Röwenstrunk
Universität Paderborn

Georg Schelbert
Humboldt-Universität Berlin

Dörte Schmidt
Universität der Künste Berlin,
Gesellschaft für Musikforschung

Torsten Schrade
Akademie der Wissenschaften und
der Literatur | Mainz

Holger Simon
Universität zu Köln,
Verband deutscher Kunsthistoriker e.V.

Gabriele Taentzer
Philipps-Universität Marburg

Joachim Veit
Universität Paderborn

Franziska Voß
Universitätsbibliothek Frankfurt a. M.

Annika-Valeska Walzel
Sächsische Landesbibliothek – Staats- und
Universitätsbibliothek Dresden

Barbara Wiermann
Sächsische Landesbibliothek – Staats- und
Universitätsbibliothek Dresden

Satz:

Sarah Pittroff
Akademie der Wissenschaften
und der Literatur | Mainz

Digital Object Identifier:

Die Onlineversion dieser Publikation finden Sie unter
<https://doi.org/10.5281/zenodo.2763576>

Lizenz:

Alle Texte dieser Veröffentlichung
(ausgenommen Zitate) sind lizenziert unter einer
Creative Commons Attribution-ShareAlike 4.0
International (CC BY-SA 4.0) Lizenz
<https://creativecommons.org/licenses/by-sa/4.0/>