

Gefördert durch



Projektnummer 441958017

Deutsche  
Forschungsgemeinschaft



---

nationale  
Forschungsdaten  
Infrastruktur  
*for CULTURE*

---

# Handreichung für ein FAIRes Management kulturwissenschaftlicher Forschungsdaten

Abstract:

Um optimal nachnutzbar zu sein, sollen Forschungsdaten gemäß den FAIR-Prinzipien aufbereitet werden. Diese Handreichung erklärt, was diese Leitlinien bedeuten und wie sie in den Kulturwissenschaften und in Kulturerbe-Sammlungen umgesetzt werden können.

Angela Kailus

<https://doi.org/10.5281/zenodo.7716941>

NFDI4Culture Permalink: <https://nfdi4culture.de/go/E3625>

Metadaten: <https://nfdi4culture.de/id/E3625>

Datenpublikation: <https://doi.org/10.5281/zenodo.7689172>

Version: 1.0.1

# Inhaltsverzeichnis

## [Einleitung](#)

### [1. Zielgruppe dieser Handreichung](#)

### [2. Besonderheiten kulturwissenschaftlicher Daten](#)

### [3. Vier grundlegende Prämissen zum Verständnis der FAIR-Prinzipien](#)

#### [3.1. Sowohl Menschen als auch Maschinen sind als „Datenauswerter“ vorgesehen](#)

##### [Die Rolle von Datenstandards](#)

#### [3.2. Die FAIR-Prinzipien gelten sowohl für Daten als auch für Metadaten](#)

##### [Digitale Objekte als Organisationseinheiten](#)

#### [3.3. Bei den FAIR-Prinzipien geht es nicht unbedingt um offene Daten](#)

#### [3.4. Die FAIR-Prinzipien sind keine Regeln oder Normen](#)

### [4. Akteure und ihre Rollen im FAIR-Prozess](#)

#### [Datenproduzierende](#)

#### [Betreibende von Datenplattformen](#)

### [5. Findable – Auffindbar](#)

#### [F1. \(Meta-\)Daten wird ein global eindeutiger und persistenter Identifikator zugewiesen](#)

##### [PID-Systeme](#)

##### [Forschungsdaten zitieren](#)

##### [Persistente Identifikatoren für Personen und Institutionen verwenden](#)

#### [F2. Daten werden mit umfangreichen Metadaten \(vgl. R1\) beschrieben](#)

#### [F3. Metadaten enthalten eindeutig und explizit den Identifikator der Daten, die sie beschreiben](#)

#### [F4. \(Meta-\)Daten werden in einer durchsuchbaren Ressource registriert oder indiziert](#)

### [6. Accessible – Zugänglich](#)

#### [A1. \(Meta-\)Daten sind über ihren Identifikator mithilfe eines standardisierten Kommunikationsprotokolls abrufbar](#)

##### [A1.1 Das Protokoll ist offen, kostenlos und universell implementierbar](#)

##### [A1.2 Das Protokoll unterstützt bei Bedarf Verfahren zur Authentifizierung und Rechteverwaltung](#)

#### [A2. Metadaten bleiben verfügbar, auch wenn die zugehörigen Daten nicht \(mehr\) verfügbar sind](#)

### [7. Interoperable – Interoperabel](#)

#### [I1. \(Meta-\)Daten nutzen eine formale, zugängliche, gemeinsam genutzte und breit anwendbare Sprache für die Wissensrepräsentation](#)

##### [Zukunftssichere Dateiformate festlegen und verwenden](#)

##### [Interoperable Datenstandards festlegen und verwenden](#)

##### [Metadatenschemata dokumentieren](#)

##### [Prozesse zur Verbesserung der Metadatenqualität einrichten](#)

[12. \(Meta-\)Daten enthalten Vokabulare, welche den FAIR-Prinzipien folgen](#)  
[Offene, gut definierte Vokabulare verwenden](#)

[13. \(Meta-\)Daten enthalten qualifizierte Verweise auf andere \(Meta-\)Daten](#)

## [8. Reusable – Nachnutzbar](#)

[R1. \(Meta-\)Daten sind detailliert beschrieben und enthalten präzise, relevante Attribute](#)

[R1.1. \(Meta-\)Daten enthalten eine eindeutige, zugreifbare Angabe einer Nutzungslizenz](#)

[Lizenzieren Sie Daten für die Nachnutzung](#)

[R1.2. \(Meta-\)Daten enthalten detaillierte Provenienz-Informationen](#)

[Benennungskonventionen einhalten](#)

[Daten systematisch dokumentieren](#)

[Datenintegrität bewahren](#)

[R1.3. \(Meta-\)Daten entsprechen den fachgebietsrelevanten Community-Standards](#)

## [9. FAIR-Prinzipien umsetzen – Hinweise zur Projektplanung](#)

[9.1 FAIRification benötigt Ressourcen und Expertise](#)

[9.2 Projektziel Nachnutzung: bei der Planung die Weichen stellen](#)

[9.3 Aktives Projektmanagement mit dem Datenmanagementplan](#)

[Tipps zum Umgang mit dem Datenmanagementplan](#)

[Tipps zur konkreten Projektplanung](#)

[Wählen Sie eine vertrauenswürdige Datenplattform](#)

## [10. Empfehlungen für Dateiformate](#)

[Text](#)

[Musik](#)

[Bild \(Rastergrafiken\)](#)

[Bild \(Vektorgrafiken\) und CAD](#)

[Audio](#)

[Video/Film](#)

[3D](#)

[Tabellen](#)

[Datenbanken](#)

[Geoinformationssysteme \(GIS\)](#)

[Webseiten](#)

## [11. Empfehlungen für Rahmenwerke und Referenzmodelle](#)

[Grundlegende Semantic-Web-Technologien](#)

[Referenzmodelle](#)

## [12. Empfehlungen zu Metadatenstandards](#)

[12.1 Empfehlungen für Metadateninhalte](#)

[Bibliotheken](#)

[Archivgut](#)

[Materielles und visuelles Kulturgut in Museen, Bauwerke, Monumente, archäologische Stätten](#)

[Annotationen](#)

## 12.2 Empfehlungen für Metadatenstrukturen

fachübergreifende Verbreitung

Bibliotheken

Archivgut

Materielles und visuelles Kulturgut in Museen, Bauwerke, Monumente, archäologische Stätten

Textuelle Quellen

Quellen der Musik

Metadaten in Bilddateien

Annotationen

3D

Geodaten

Forschungsmetadaten in Repositorien

Langzeitsicherung

## 12.3 Empfehlungen für Vokabulare, Normdaten und Anwendungsontologien

Kontrollierte Vokabulare / Normdaten

Personen

Körperschaften

Geografika

Werke, materielle Kulturerbe-Objekte

Bauwerke, Monumente

Sachschlagworte

Rollen, Beziehungen

Kodierungen

Kodierungen für Zeit- und Datumsangaben

Kodierungen für Sprachen

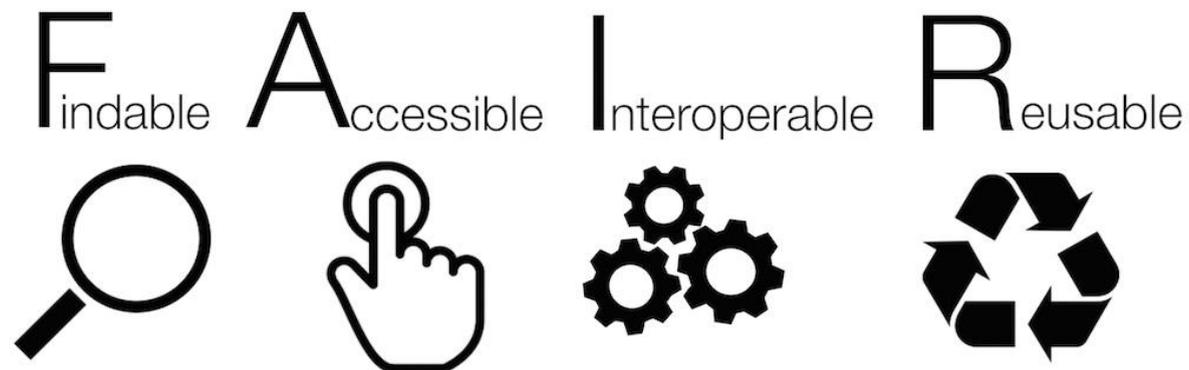
Anwendungsontologien

13. Verwendete Literatur

14. Weiterführende Literatur

Danksagung

# Einleitung



Sangya Pundir, CC BY-SA 4.0, via Wikimedia Commons

Die FAIR-Prinzipien enthalten Leitlinien für eine Datenverwaltungspraxis, die eine bessere Nachnutzbarkeit von Forschungsdaten unterstützt. Die Abkürzung FAIR bedeutet Findable (auffindbar), Accessible (zugänglich), Interoperable (interoperabel) und Reusable (wiederverwendbar). Jeder Buchstabe verweist auf ein Kernkriterium, das Nachnutzbarkeit ermöglicht. Jedem sind einige der insgesamt 15 Grundsätze zugeordnet, die genauer beschreiben, mit welchen Maßnahmen die FAIRness von Daten erreicht werden kann, um Forschungsdaten sowohl für Menschen wie auch Maschinen in bestmöglicher Weise verständlich und so offen wie möglich zugänglich zu machen.

Die FAIR-Prinzipien wurden erstmals im Jahr 2016 veröffentlicht<sup>1</sup>. Obwohl sie ursprünglich aus den Lebenswissenschaften stammen, sollen sie auch in anderen Forschungsdisziplinen angewandt werden. Innerhalb kurzer Zeit wurden die Richtlinien breit rezipiert und implementiert. Sie bilden eine wichtige Grundlage etwa für Datenservices, fachspezifische Leitlinien und Forschungsinfrastrukturen. Förderprogramme der Europäischen Union und Förderinstitutionen wie die Deutsche Forschungsgemeinschaft (DFG) sowie zahlreiche Hochschulen haben die FAIR-Prinzipien in ihre jeweiligen Leit- und Förderrichtlinien aufgenommen. Sie sind grundlegender Bestandteil guter wissenschaftlicher Praxis und wurden in langfristig angelegten EU-Projekten wie etwa der European Open Science Cloud (EOSC)<sup>2</sup> fest

<sup>1</sup> <https://doi.org/10.1038/sdata.2016.18>

<sup>2</sup>

[https://ec.europa.eu/info/research-and-innovation/strategy/strategy-2020-2024/our-digital-future/open-science/european-open-science-cloud-eosc\\_en](https://ec.europa.eu/info/research-and-innovation/strategy/strategy-2020-2024/our-digital-future/open-science/european-open-science-cloud-eosc_en)

implementiert. Die Etablierung der FAIR-Prinzipien in übergeordneten Selbstverpflichtungen, Förderprogrammen und Infrastrukturprojekten hat eine enorme Wirkung entfaltet. Mit der EOSC in Verbindung stehen eine Reihe von Initiativen, Institutionen und Projekten, die die Etablierung eines FAIRen Forschungsdatenmanagements auf verschiedenen Ebenen unterstützen. Besonders sind hier die Research Data Alliance (RDA)<sup>3</sup>, die Go FAIR-Initiative<sup>4</sup>, FAIRsFAIR<sup>5</sup>, FAIRsharing<sup>6</sup> und als deutsche Initiative die Nationale Forschungsdateninfrastruktur (NFDI)<sup>7</sup> zu nennen. Die FAIR-Prinzipien sind daher in den letzten Jahren zur bestimmenden Leitformel für nachhaltiges Forschungsdatenmanagement geworden. Die Akteure reagieren damit auf eine kritische Situation, in der fehlende Standards, offene rechtliche Fragen, Handlungsunsicherheiten und auch wettbewerbliche Vorbehalte den umfassenden und einfachen Zugang zu Daten blockieren. Die FAIR-Prinzipien bieten eine Reihe von Lösungen, um die enormen Potenziale auf dem Gebiet vernetzter Forschungsdateninfrastrukturen möglichst uneingeschränkt zur Geltung kommen zu lassen.

Beispiel für die Integration in die Förderrichtlinien der DFG

Leitlinien zur Sicherung guter wissenschaftlicher Praxis – Kodex, 2019<sup>8</sup>, S. 19 und Materialien auf dem DFG-Portal Wissenschaftliche Integrität<sup>9</sup>

## 1. Zielgruppe dieser Handreichung

Diese Handreichung wendet sich insbesondere an die vom Konsortium NFDI4Culture angesprochenen Fachgemeinschaften, aber auch weit darüber hinaus. Die Leserschaft im Umfeld des Konsortiums sollte fachliche Kenntnisse in den Architektur-, Kunst-,

---

<sup>3</sup> <https://www.rd-alliance.org/>

<sup>4</sup> <https://www.go-fair.org/>

<sup>5</sup> <https://www.fairsfair.eu/>

<sup>6</sup> <https://fairsharing.org/>

<sup>7</sup> <https://www.nfdi.de/>

<sup>8</sup> <https://doi.org/10.5281/zenodo.6472827>

<sup>9</sup> [https://wissenschaftliche-integritaet.de/kommentar\\_tags/fair-prinzipien/](https://wissenschaftliche-integritaet.de/kommentar_tags/fair-prinzipien/)

Musik-, Theater-, Tanz-, Film- oder Medienwissenschaften und berufliche Erfahrungen im Bereich digitale Strategien oder Datenmanagement mitbringen. Mitarbeiter:innen in den Projekten, Forschungs- und Kulturerbe-Einrichtungen, die Daten produzieren und verwalten, erfahren nun hier, wie sie diese Forschungs- und Sammlungsdaten bestmöglich nachnutzbar machen können.

Im Hochschulbereich sind vor allem Wissenschaftler:innen angesprochen, die im Rahmen eines oft zeitlich begrenzten Projektvorhabens bestimmte Ressourcen oder Sammlungsbestände unter bestimmten Forschungsfragen bearbeiten und die im Rahmen der Projektplanung oder -durchführung mit der Entwicklung von Datenmanagementplänen befasst sind. Nicht selten sind über bestehende Kooperationen mit Infrastruktureinrichtungen auch die dort angesiedelten Forschungsdatenmanager:innen eingebunden. Auch sie sind gehalten, die Fachspezifik der Daten, ihrer Entstehungskontexte und Anwendungsszenarien zu berücksichtigen.

In bestandshaltenden Sammlungen wie Museen, Archiven und Bibliotheken, aber auch in den Fachämtern der Denkmalpflege, befassen sich wissenschaftliche Kurator:innen, Dokumentar:innen oder Vertreter:innen der Fachinformatik mit dem Aufgabenfeld. Hier gehört die Entwicklung und Pflege von digitalen Strategien oft ebenso wie die von Sammlungskonzepten zu den Kernaufgaben der Häuser. Sie betreffen daher nicht nur Projekte, die in ihren Zielen und verfügbaren Zeit- und Personalressourcen begrenzt sind, sondern Daueraufgaben im Bereich des Sammlungsmanagements. So gibt es häufig die Anforderung, vorhandene, kontinuierlich kuratierte Datenbestände für neue Nutzungsszenarien anzupassen oder die Daten bereits abgeschlossener Projekte weiter zugänglich zu machen.

Betreiber:innen von Repositorien und Datenplattformen haben in der Regel einschlägige professionelle Kenntnisse im Datenmanagement. Sie gehören daher nicht zu den primären Adressaten dieser Handreichung. Da es aber für Datenproduzierende wichtig ist zu verstehen, wie die FAIR-Prinzipien im Zusammenspiel zwischen ihnen und den Publikationsplattformen umgesetzt wird, werden auch die Aufgaben der letzteren beleuchtet. Mehr dazu im Abschnitt Akteure und ihre Rollen im FAIR-Prozess.

## 2. Besonderheiten kulturwissenschaftlicher Daten

Es gibt einige besondere Rahmenbedingungen im Umgang mit kulturwissenschaftlichen Forschungsdaten. In globalen Anwendungsbezügen betreffen sie das Menschenrecht auf kulturelle Teilhabe und wissenschaftlichen Fortschritt, das zusammen mit dem Recht auf Bildung über den Weg des UN-Sozialpakts seit 1968 im deutschen Bundesrecht fest verankert ist.

Neben den vom Völkerrecht abgeleiteten nationalstaatlichen Grundrechten gilt es beim Management kulturwissenschaftlicher Daten, auch einige wissenschaftsimmanente Paradigmen zu berücksichtigen: Durch eine grundlegende Ausrichtung der Kulturwissenschaften auf Fragestellungen zur lebensweltlichen beziehungsweise historischen Kontextualisierung genießt das materielle oder immaterielle, meist nicht-digitale Untersuchungsobjekt eine Sonderstellung als authentischer Informationsträger, als Quelle wissenschaftlichen Wissens. Solche Gegenstände können materielle Objekte wie barocke Deckengemälde, chinesische Theatermasken, Programmzettel von Wanderkinos des europäischen Films, oder Lagepläne von im Krieg zerstörter Gebäude sein, aber auch immaterielle Kulturgüter wie der Musikinstrumentenbau in Sachsen oder die ebenfalls als UNESCO Kulturerbe eingetragene Deutsche Orchester- und Theaterlandschaft. Es kann sich um Aufzeichnungen menschlicher Aktivitäten handeln, die als filmische Mitschnitte von Performances oder Phonographenwalzen mit Gesangsaufnahmen vorliegen, oder auch um biografische Daten. Daher sind digitale Daten zu materiellen und immateriellen Kulturgütern stets nur eine mediale Repräsentation des beforschten Gegenstandes und nicht der beforschte Gegenstand selbst. Auch wenn immer größere Anteile kultureller Werke in Zukunft in digitaler Form vorliegen werden, entstehen weiterhin Daten, die sich mit den nicht-digitalen Werken als Informationsträgern beschäftigen und daher in ihrer wissenschaftlichen Bezugnahme über den Bereich der Digitalität hinausweisen.

Der Begriff „Forschungsdaten“<sup>10</sup> im Bereich der Kulturwissenschaften schließt daher die Daten, die die Bestände von Gedächtnisinstitutionen wie Bibliotheken, Archiven, Museen und Denkmalfachämtern erschließen, explizit ein. Digitale Daten machen die Sammlungsobjekte, die Gegenstand der Forschung sind, meist erst auffindbar und

---

<sup>10</sup> <https://nfdi4culture.de/go/E3508>

eindeutig adressierbar. In den Beschreibungsdaten zu historischen materiellen Objekten steckt bereits die Leistung einer historischen und materiellen Bestandsaufnahme und Kontextualisierung – sie vermittelt, warum ein Gegenstand als Kulturerbeobjekt bedeutsam und als Forschungsgegenstand verwendbar ist. Forschungsdaten vermitteln daher bereits grundlegende Forschungsergebnisse und bieten Grundlage und Ausgangspunkt weiterführender Forschung, oft auch überhaupt erst die Bedingung ihrer Möglichkeit.

Die Kulturwissenschaften haben digitale Methoden bislang eher zögerlich entwickelt, im Unterschied etwa zu den eher mit metrischen Daten arbeitenden Natur- Lebens- und Sozialwissenschaften. Dadurch, dass die Kulturwissenschaften eher mit nicht-metrischen Informationen arbeiten, werden selbst digitale Daten, die zum jeweiligen Forschungsobjekt vorliegen, häufig weiterhin mit nicht-digitalen Methoden ausgewertet.

Kulturwissenschaftler:innen, die digital vorliegende Informationsbestände verwenden, arbeiten allerdings zunehmend multimodal: Sie bearbeiten ihre Forschungsfragen durch gemeinsame Auswertung von Daten zu textuellen, visuellen oder auditiven Objekten. Dabei verwenden sie häufig hybride, nicht miteinander integrierte Software und technische Umgebungen. Eine effektive Verwaltung, Archivierung und Nachnutzung der in solchen Arbeitsprozessen entstehenden Forschungsdaten ist dadurch erheblich erschwert.

Auch können die in Rücksicht auf bestimmte heuristische Perspektiven ausgerichtete Forschungsdaten in den Kulturwissenschaften sehr individuell sein: Forschende bereiten ihre Quellen häufig so auf, dass sie für das jeweilige Forschungsziel des Projekts unmittelbar nützlich sind. Was noch für eine spezifische Forschungsfrage und die zugehörigen epistemischen Instrumente funktioniert, mag für andere Forschergruppen nicht oder nur sehr eingeschränkt verwendbar sein.

Ein Datenmanagement im Einklang mit den FAIR-Prinzipien unterstützt den Übergang zu Daten, die nicht nur von Menschen, sondern auch von Maschinen interpretiert werden können und erfordert in „interpretierenden“ Disziplinen der Kulturwissenschaften ein erhebliches Maß an methodologischer Reflexion und informationstechnischer Perspektivierung. Die FAIRe Bereitstellung von

kulturwissenschaftlichen Forschungsdaten zusammen mit einer offenen Lizenzierung eröffnet aber viele neue Möglichkeiten für eine wissenschaftlich fruchtbare Weiterverwendung der Daten. Auf europäischer und nationaler Ebene gewinnt die Unterstützung für die gemeinsame Datennutzung in allen Disziplinen an Schwung und wird bereits zur Regel für die öffentlich finanzierte Forschung und Digitalisierung. Auch die für die Kulturwissenschaften relevanten Förderprogramme werden daher in den Leitlinien und einzureichenden Datenmanagementplänen zunehmend auf die FAIR-Prinzipien ausgerichtet.

Zum Verhältnis von Untersuchungsgegenständen und Daten in den Geisteswissenschaften

Schöch, Christof: Big? Smart? Clean? Messy? Data in the Humanities, in: Journal of Digital Humanities Volume 2, Issue 3, 2013, S. 2-13<sup>11</sup>

### 3. Vier grundlegende Prämissen zum Verständnis der FAIR-Prinzipien

Um die FAIR-Prinzipien richtig zu verstehen, sollten Sie sich über die vier folgenden Aspekte im Klaren sein.

#### 3.1. Sowohl Menschen als auch Maschinen sind als „Datenauswerter“ vorgesehen

Die Menge der Veröffentlichungen im digitalen Raum nimmt immer mehr zu, dazu kommt eine Vielzahl von heute vorhandenen digitalen Werkzeugen, Aufbereitungs- und Auswertungsmethoden, die neue Wege in der Entdeckung und Analyse relevanter Datenbestände erlauben. Menschen sind deshalb immer mehr auf die Unterstützung

---

<sup>11</sup> <https://nbn-resolving.org/urn:nbn:de:bvb:20-opus-129492>

durch Computer angewiesen, um die für sie jeweils relevanten Daten zu finden und auszuwählen. Die FAIR-Prinzipien sollen daher die optimale Nachnutzbarkeit für Mensch und Maschine gleichermaßen gewährleisten. „Maschinentauglichkeit“ ist hier die Fähigkeit von Computersystemen, Daten ohne oder mit nur minimalem menschlichem Eingriff zu finden, darauf zuzugreifen, sie zu integrieren und wiederzuverwenden.

„Maschinell verwertbar“ ist ein digitales Objekt, wenn es Informationen bereitstellt, die es einem autonom agierenden, computergestützten Algorithmus ermöglichen,

- die Art des digitalen Objekts sowohl in Bezug auf die Struktur als auch auf die Zweckbestimmung zu identifizieren,
- zu bestimmen, ob es im Kontext der aktuellen Aufgabe nützlich ist, indem er Metadaten und/oder Datenelemente (Bitsequenzen) abfragt,
- zu bestimmen, ob es in Bezug auf die Lizenz oder andere Zugänglichkeits- oder Nutzungsbeschränkungen nutzbar ist,
- es ähnlich nachzunutzen, wie es ein Mensch mit vergleichbarer Fragestellung tun würde und
- die Herkunft der Daten sorgfältig aufzuzeichnen, um die gesammelten Daten angemessen zitieren zu können.

Eine gute Forschungsdatenverwaltung gemäß den FAIR-Prinzipien ermöglicht ein Netz von Daten und Diensten, die sich gegenseitig finden, miteinander kommunizieren und für die Wiederverwendung verfügbar bleiben. Grundlegend dafür sind Linked Data-Technologien. Sie basieren darauf, dass digitale Objekte und mit ihnen verbundene Ressourcen über Uniform Resource Identifier (URIs) global eindeutig identifizierbar sind. Das ist die Voraussetzung, dass sie über eine Verbindung mit Identifikationssystemen (z. B. Ontologien) typisiert, kategorisiert und damit auch von Maschinen „verstanden“ werden können. In gleicher Weise werden die Beziehungen zwischen Ressourcen mit URIs ausgewiesen. Die führenden übergreifenden Standards zur Kodierung von Semantik sind RDF (Resource Description Framework)<sup>12</sup>, das die Syntax für Datenaustausch definiert, OWL (Web Ontology Language)<sup>13</sup>, eine formale Beschreibungssprache zur Erstellung, Publikation und zum Austausch von Ontologien,

---

<sup>12</sup> <https://www.w3.org/RDF/>

<sup>13</sup> <https://www.w3.org/OWL/>

und SKOS (Simple Knowledge Organization System)<sup>14</sup>, eine formale Sprache zur Kodierung von Dokumentationssprachen wie Thesauri, Klassifikationen oder anderen kontrollierten Vokabularen. Die eingebettete, mit den Daten verbundene Semantik bietet erhebliche Vorteile bei der qualifizierten Auswertung der Daten und dem Umgang mit inhaltlich heterogenen Datenquellen.

## Die Rolle von Datenstandards

Die Grundfunktion von Datenstandards ist es, Informationen durch einheitliche Kodierung oder Gleichmäßigkeit der Beschreibung besser auswertbar zu machen. Ihre Funktionsfähigkeit ist nicht in der Vorgabe begründet, sondern vielmehr im gemeinsamen Gebrauch. Es sind die am weitesten verbreiteten Systeme und Konventionen innerhalb einer Anwendergemeinschaft, die festlegen, welche Arten von Informationen für jedes Informationsobjekt in einer Datensammlung auf welche Weise erfasst werden. Sie sind in der Regel gut dokumentiert, und sie haben aktive Anwendergemeinschaften, die kontinuierlich an ihrer inhaltlichen und technischen Weiterentwicklung und an den Softwaresystemen arbeiten, um sie an die aktuellen und zukünftigen Herausforderungen anzupassen. Ein Beispiel für eine solche Adaption ist das Dokumentenformat Text Encoding Initiative (TEI)<sup>15</sup>, das, ausgehend von seinem ursprünglichen Anwendungskontext in spezialisierten Forschungsbibliotheken, heute international in einer Vielzahl fachlich breit ausgerichteter Texteditionsprojekte eingesetzt wird.

Die konsequente Nutzung von Standards ist bei der Erstellung der Metadaten einschließlich ihrer Quellen und Erhebungsbedingungen besonders wichtig. Bei einer maschinellen Nachnutzung der Metadaten ist dies Voraussetzung für ihre Auswertbarkeit. Standards sichern die mittel- und längerfristige Verständlichkeit und Nachnutzbarkeit der Daten ab und ermöglichen, dass Personen mit Daten weiterarbeiten können, die sie nicht selbst erstellt haben. Wenn kaum verbreitete oder schlecht dokumentierte Formate, Schemata oder Modelle verwendet werden oder die Daten in Software eingebettet sind, die proprietär ist oder wegen mangelnder Pflege oder Dokumentation nicht (mehr) zugänglich ist, ist die Verständlichkeit der Daten häufig bereits nach wenigen Jahren selbst für Menschen nicht mehr gewährleistet.

---

<sup>14</sup> <https://www.w3.org/TR/skos-reference/>

<sup>15</sup> <https://tei-c.org/>

Welche Standards jeweils befolgt werden sollten, hängt von den Gepflogenheiten der Fachdomäne im Umgang mit der jeweiligen Materialart, dem Dokumentationsinteresse des Forschungsprojekts und von der Art der zu erzeugenden Daten ab.

Bestandserschließungsprojekte in Kulturerbe-Institutionen orientieren sich in der Regel daran, was in in ihrer Sparte – sei es Bibliothek, Archiv, Museum oder Denkmalpflege – verbreitet verwendet wird. Für Digitalisierungsprojekte in bestandshaltenden Einrichtungen sind die DFG-Praxisregeln „Digitalisierung“<sup>16</sup> seit vielen Jahren eine weit über die Spartengrenzen und den ursprünglich intendierten Kontext der DFG-Antragstellung hinaus anerkannte und bewährte Good-Practice-Empfehlung zur Qualitätssicherung mit Hilfe von Standards. Sie wurden 2023 in einer aktualisierten Fassung<sup>17</sup> vorgelegt.

Mit Blick auf zukünftige Nachnutzungsmöglichkeiten der Forschungsdaten sollten Datenproduzierende aber immer auch Standards sondieren, die in einer Fachdomäne vielleicht noch nicht verbreitet eingesetzt werden, aber der Fragestellung angemessen erscheinen und viel Potential für eine gute Interoperabilität und Wiederverwendbarkeit der Daten haben. Dies gilt insbesondere für die Vokabularstandards.

Die für Forschungsdaten relevanten Datenstandards beziehen sich auf mehrere Geltungsbereiche. Empfehlungen für die von NFDI4Culture vertretenen Fachgebiete finden Sie auf den hier verlinkten Seiten dieser Handreichung.

- **Standard-Dateiformate** gewährleisten in hohem Maß, dass die Dateien zu einem späteren Zeitpunkt und von anderen Akteuren verwendet werden können. Hier finden Sie Empfehlungen für Dateiformate.
- **Referenzmodelle (top-level ontologies)** sind komplexe Sets von Repräsentationselementen (Klassen, Attribute und Beziehungen), mit denen ein Wissens- oder Diskursbereich modelliert werden kann. Die Definitionen der Elemente enthalten Informationen über ihre Bedeutung und Beschränkungen für ihre logisch konsistente Anwendung. Sie ermöglichen es den Maschinen, Metadaten zu „verstehen“, indem auf der Grundlage von Ableitungsregeln aus vorhandenen Daten neue Fakten erschlossen werden können. Hier finden Sie Empfehlungen für Rahmenwerke und Referenzmodelle.

---

<sup>16</sup> [https://www.dfg.de/formulare/12\\_151/](https://www.dfg.de/formulare/12_151/)

<sup>17</sup> <https://doi.org/10.5281/zenodo.7435724>

- **Metadatenstandards** legen Kriterien, Methoden und Prozesse fest, nach denen die Dokumentation in bestimmten Fachcommunities erfolgen soll, um Daten später gemeinsam nutzen zu können. Gemäß dem Zweck, für den sie entwickelt wurden, untergliedern sie sich in drei Bereiche:
  - **Daten-Inhaltsstandards** oder Katalogisierungsregeln legen fest, wie bestimmte Ressourcen, Objekte oder Sachverhalte beschrieben werden sollen. Sie empfehlen, wie Informationen strukturiert und formatiert werden sollen und welches Vokabular verwendet werden soll. Hier finden Sie [Empfehlungen für Metadateninhalte](#).
  - **Daten-Strukturstandards** beziehen sich auf **Metadaten-Elementsets**. Sie geben die Struktur, den Inhalt, die Semantik und den Umfang von Metadaten in Form eines Kategorienschemas vor. Liegen sie in Form eines **Metadatenschemas** vor, sind sie auch technisch verarbeitbar. Hier finden Sie [Empfehlungen für Metadatendatenstrukturen](#).
  - **Standards für Datenwerte** beziehen sich auf **kontrollierte Vokabulare**, deren Verwendung als Datenwert in den Elementen des Metadatenschemas empfohlen wird. Kontrollierte Vokabulare gewährleisten, dass Begriffe („units of thought“<sup>18</sup>) definiert, mit eindeutigen Bezeichnungen versehen und in Beziehung zueinander gesetzt werden können. Ihre Verwendung ist die Voraussetzung für vollständige und präzise Suchergebnisse und die richtige Verknüpfung vergleichbarer Daten. Hier finden Sie [Empfehlungen für Vokabulare, Normdaten und Anwendungsontologien](#).

## 3.2. Die FAIR-Prinzipien gelten sowohl für Daten als auch für Metadaten

Die 15 Grundsätze der FAIR-Prinzipien beziehen sich auf die „(Meta-)Daten“. Was bedeutet das? Sie gelten einerseits für die eigentlichen Datenelemente (Bitsequenzen). Beispiele hierfür sind ein Text, eine Datei, ein Bild, Quellcode einer Software, ein 3D-Modell, ein Dienst oder Zeitreihen (Audiosignale). Es kann sich aber auch um ein Aggregat vieler Einheiten handeln, die ihrerseits separat adressiert werden können, wie z. B. eine Datenbank, ein digitalisiertes Buch, die digitalisierten

---

<sup>18</sup> <https://www.w3.org/TR/2009/REC-skos-reference-20090818/#concepts>

Materialien eines Nachlasses, ein Aufzeichnungssystem, eine Software oder eine Forschungsdatenpublikation mit mehreren Komponenten.

Metadaten beschreiben die Eigenschaften anderer Daten. Sie geben Auskunft über ihren Inhalt, ihre Eigenschaften oder ihre Struktur. Sie geben an, in welchen Kontexten sie stehen oder wie sie genutzt werden können.

Im Prozess der Umsetzung der FAIR-Prinzipien nehmen Daten und Metadaten unterschiedliche Funktionen ein. Die eingesetzten Methoden, um Daten und die zugehörigen Metadaten FAIR zu machen, können sich unterscheiden. Auch der Grad der FAIRness kann für beide Komponenten unterschiedlich sein. Daher ist es zuerst notwendig, die beiden Komponenten zu identifizieren und sie als digitale Objekte (DO) zu strukturieren.

## Digitale Objekte als Organisationseinheiten

Ein digitales Objekt ist zunächst „ein Objekt, das sich aus einer Reihe von Bitsequenzen zusammensetzt“<sup>19</sup>. Das bedeutet, dass jede Datei als digitales Objekt betrachtet werden kann. Einige digitale Objekte können einfach strukturiert sein, z. B. eine Textdatei. Ein Video, das sich aus mehreren Elementen zusammensetzt (Videospur, Audiospur, Containerdatei und möglicherweise weitere), kann als ein komplexes digitales Objekt betrachtet werden.

Um von maschinellen Agenten interpretiert werden zu können, muss das digitale Objekt adressierbar sein, es muss strukturiert und typisiert werden. Zu diesem Zweck wird der Bitsequenz ein Identifikator, möglichst ein global eindeutiger und persistenter Identifikator (PID), und eine Beschreibung seiner Eigenschaften in Form einer Metadateneinheit zugewiesen. Bitsequenz, PID und Metadaten sind über eindeutige Relationen zu einem erweiterten digitalen Objekt verbunden, das sich als adressierbare und auswertbare Wissensseinheit darstellt.

---

<sup>19</sup> [https://files.dnb.de/nestor/materialien/nestor\\_mat\\_16.pdf](https://files.dnb.de/nestor/materialien/nestor_mat_16.pdf)

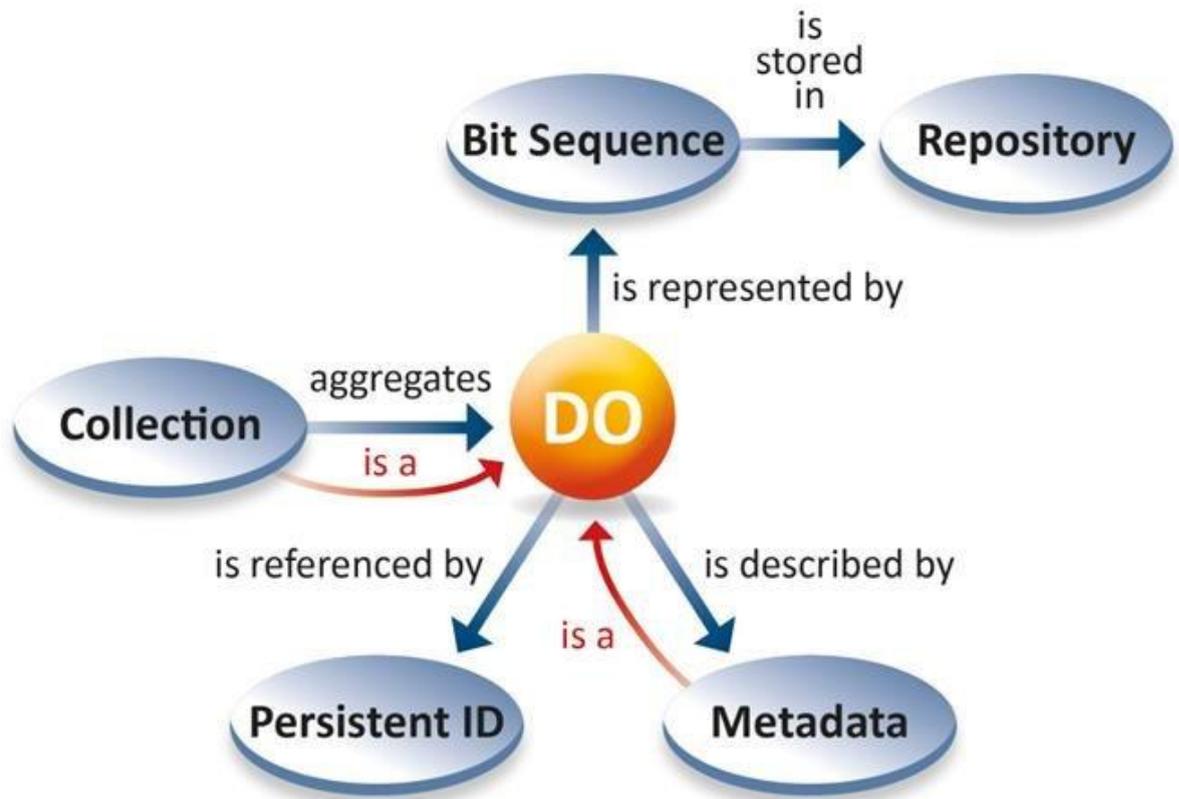


Abb. 1: Beziehungen zwischen Bestandteilen eines digitalen Objekts, seinem Aggregator und Repositorium. Smedt, K. / Koureas, D. / Wittenburg, P.: FAIR Digital Objects for Science: From Data Pieces to Actionable Knowledge Units, in: Publications, 2020, 8. 21., fig. 4. CC-BY 4.0

Im Bereich der Gedächtnisinstitutionen (Bibliotheken, Archive, Museen) wird der Begriff „Metadaten“ in der Regel für die Zusatzinformationen verwendet, mit denen die dort verwahrten, häufig physischen Objekte, Sammlungen und Ressourcen geordnet, beschrieben und verwaltet werden. Es ist daher durchaus üblich, dass Metadaten nicht-digitale Entitäten behandeln, die materiell oder immateriell, konkret wie abstrakt sein können. Statt auf eine Bitsequenz können sich Metadaten z. B. auf ein Gemälde, ein Libretto, einen Quelltext, eine Person, einen geografischen Ort, ein Ereignis oder einen Begriff der Ideengeschichte beziehen. Wenn Normdaten zu diesen nicht-digitalen Entitäten vorhanden sind, kann man sie über Normdaten-PIDs adressieren. Das digitale Objekt enthält in diesem Fall als Bitsequenz einen Datensatz in einem strukturierten Format, der seinerseits Metadaten zu einem nicht-digitalen Gegenstand enthält.

Um die Funktion von Metadaten besser zu verstehen, ist es hilfreich, sie in verschiedene Kategorien zu unterteilen. Für die FAIRness der Daten spielen sie alle eine Rolle.

- Administrative Metadaten sind Daten über eine Bitsequenz oder einen Gegenstand, die für ihre Verwaltung relevant sind. Dazu gehören Angaben über ihre Art und ihre Zugehörigkeit zu einem bestimmten Projektkontext, z. B. Projekt-/Ressourceneigentümer:innen, Forschungsleitung, Projektmitarbeiter:innen, Finanzierung, Projektzeitraum, Entstehungsdatum der Bitsequenz. Zu dieser Gruppe von Metadaten gehören auch
  - technische Metadaten – Informationen, die zur Verwendung und Verarbeitung von Dateien erforderlich sind (Dateiformat, Bildauflösung, Kompressionsrate, Hardware- und Softwareanforderungen, Authentifizierungs- und Sicherheitsdaten wie z. B. Passwörter),
  - rechtliche Metadaten – Informationen zu geistigem Eigentum und Nutzungsrechten,
  - Provenienz-Metadaten zur Herkunft von Informationen und den Modifikationen, die die Bitsequenz oder der Metadatensatz selbst erfahren haben. Sie benennen z. B. Akteur:innen, Zeitpunkte und die Art und Methoden der Bearbeitung oder die zugrunde liegenden Informationsquellen.
  - Archivierungsmetadaten – Informationen, die für die langfristige Verwaltung und Archivierung digitaler Bestände erforderlich sind. Sie gewährleisten die Integrität eines digitalen Objekts während der gesamten Aufbewahrungszeit. Ein gängiges Modell hierfür ist PREMIS (Preservation Metadata: Implementation Strategies)<sup>20</sup>.
- Deskriptive Metadaten sind Daten über eine Bitsequenz oder einen Gegenstand, die es menschlichen und maschinellen Akteur:innen ermöglichen, sie zu finden, zu identifizieren und zu zitieren, z. B. zu Autorschaft oder Herstellungskontexten, Titel, Themenschlagwörter, persistente Identifikatoren, verwandte Publikationen und Objekte.
- Strukturelle Metadaten beschreiben die interne Organisation komplexer digitaler Objekte oder Gegenstände und stellen den Zusammenhang zwischen ihren Bestandteilen oder zu weiteren digitalen Objekten her. Über METS oder

---

<sup>20</sup> <http://www.loc.gov/standards/premis/>

TEI-Auszeichnungen stellen sie z. B. ein digitales Inhaltsverzeichnis für einen digitalisierten frühneuzeitlichen 120-seitigen Druck bereit und ermöglichen so die sinnvolle Organisation der Untereinheiten für die Nachnutzung.

Die für das digitale Objekt relevanten Metadateninformationen können je nach Datenformat und angestrebtem Nutzungskontext unterschiedlich sein. Es ist in der Regel davon auszugehen, dass Metadaten nicht nur in der für sie angelegten Dokumentation (z. B. einer Datenbank oder Tabelle), sondern auch auf der Ebene von Software- und Systemkonfigurationen oder der Prozesskontrolle (z. B. Logdateien) zu lokalisieren sind und von dort extrahiert werden müssen.

Beispiel einer born digital Bilddatei (Bitsequenz) und ihrer Metadaten

Administrative Metadaten

- technisch: Dateiformat, Größe, Auflösung, Farbtiefe, Kompressionsrate
- rechtlich: Lizenzierung, Nutzungsrechte, Rechteinhaber
- Provenienz: Entstehungsinformationen der Primärdatei (Kamerahersteller:in und -modell, Belichtungszeit, Geokoordinaten des Aufnahmeorts), Veränderungshistorie (Konvertierung und Bildbearbeitung unter Benennung der verwendeten Software, Verfahren, Formate, technischen Parameter, Zeitpunkte und Akteur:innen)

Deskriptive Metadaten

- Fotograf:in/Fotograf, Aufnahmezeitpunkt, Identifikation der dargestellten Entität (z. B. Person, Objekt, Ereignis, Ort), Kontext, Motivation zur Anfertigung der Bilddatei, Referenz auf andere per PID adressierbare Versionen und Nutzungskontexte

Der PID des digitalen Objekts führt in der Regel auf eine Landing Page, eine HTML-Seite, die Metadaten über das digitale Objekt zeigt. Diese muss für Menschen und Maschinen ausreichend Informationen enthalten, um das digitale Objekt zu identifizieren zu können. Die Landing Page muss ihnen außerdem den Zugang zur Bitsequenz selbst (Text, Bild, Video) wie auch zu eventuell vorhandenen weiteren Metadaten ermöglichen. Die Maschineninterpretierbarkeit der Metadaten kann

gewährleistet werden, indem sie über [schema.org](http://schema.org)<sup>21</sup> in die Header der HTML-Seite eingebettet werden.

### 3.3. Bei den FAIR-Prinzipien geht es nicht unbedingt um offene Daten

Eine möglichst offene Lizenzierung der Daten im Sinne des Open-Access-Paradigmas wird empfohlen, da sie die umfassendsten Nachnutzungsmöglichkeiten eröffnet und zunehmend als gute wissenschaftliche Praxis angesehen wird. Ein Meilenstein ist die Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities<sup>22</sup> von 2003, sie ist bis heute von mehreren hundert Institutionen unterzeichnet worden.

Es ist aber auch möglich, auf FAIRe Art und Weise mit Daten zu arbeiten, die nicht uneingeschränkt bereitgestellt werden können. Zugriffsbeschränkungen sind mit den FAIR-Prinzipien vereinbar, solange die Bedingungen und Wege zum Zugang menschen- und maschineninterpretierbar dokumentiert sind. Für Daten und die zugehörigen Metadaten können unterschiedliche Grade der Zugänglichkeit gelten.

Daten können FAIR sein, wenn sie nicht für Dritte zugänglich sind, wenn sie nur für eine bestimmte Gruppe von Personen zugänglich sind oder wenn sie für jeden zugänglich sind (offene Daten). Unterschiedlich restriktive Zugänge können auch für verschiedene Stadien des Datenlebenszyklus<sup>23</sup> oder für unterschiedliche Nutzungszwecke gelten. So können beispielsweise neue, experimentelle Daten zunächst nur für den Erzeugenden und seine Gruppe zugänglich sein, dann für Konsortialpartner, wenn die Ergebnisse weiterentwickelt worden sind, und schließlich für die Öffentlichkeit bei der Publikation. Persönliche sensible Daten dürfen niemals öffentlich zugänglich und nutzbar sein.

Bei der Digitalisierung von Sammlungsobjekten ergeben sich oftmals komplexe Fragestellungen rund um die Urheberrechte und Nutzungsrechte an den Sammlungsobjekten und den Digitalisaten von ihnen. Die Rechtklärung und eine umfassendere Bereitstellung von Digitalisaten im Open Access wurden seit 2021 durch die Anpassung des deutschen Urheberrechts vereinfacht.

---

<sup>21</sup> <http://schema.org/>

<sup>22</sup> <https://openaccess.mpg.de/Berliner-Erklaerung>

<sup>23</sup> <https://forschungsdaten.info/themen/informieren-und-planen/datenlebenszyklus/>

Mehr Informationen zum Thema Rechtklärung und Lizenzierung finden Sie im Abschnitt R1.1. (Meta-)Daten enthalten eine eindeutige, zugreifbare Angabe einer Nutzungslizenz.

Vorteile des Open-Access-Publizierens

Open-Access.Network – Informieren – Open-Access-Grundlagen – Gründe und Vorbehalte<sup>24</sup>

### 3.4. Die FAIR-Prinzipien sind keine Regeln oder Normen

Die FAIR-Prinzipien sollen nicht als Regeln oder Standards verstanden werden, die streng formalisiert umzusetzen sind. Vielmehr sollen sie Leitlinien sein, an denen man sich beim Umgang mit Forschungsdaten in jeder Phase des Datenlebenszyklus orientieren sollte. Die FAIR-Richtlinien hängen natürlich zusammen, sie sind aber technisch gesehen voneinander unabhängig und können in jeder beliebigen Kombination schrittweise umgesetzt werden. So kann das Einstiegsniveau für Datenproduzierende zunächst niedrig sein. Das ist nicht selten der Fall, wenn man mit Daten umgeht, die bereits vor längerer Zeit erstellt wurden. Im Rahmen des Datenqualitätsmanagements sollte man anstreben, den Umsetzungsgrad für die einzelnen FAIR-Kriterien schrittweise zu erhöhen.

Die Übernahme der FAIR-Prinzipien wird oft eine schrittweise Anpassung der Arbeitsabläufe sein – sie kann aber im Rahmen eines Umbruchs geschehen, wenn eine Art von Infrastruktur durch eine andere ersetzt wird.

Zweck dieses Dokuments ist es daher nicht, eine bestimmte technologische Umsetzung für die FAIR-Prinzipien vorzuschlagen, sondern vielmehr die Merkmale zu beschreiben, die Datenressourcen, -werkzeuge und -infrastrukturen aufweisen sollten, um als FAIR gelten zu können. Die FAIRness von Daten kann mit einem breiten Spektrum von Technologien und Verfahren erreicht werden.

---

<sup>24</sup> <https://open-access.network/informieren/open-access-grundlagen/gruende-und-vorbehalte>

Die Data FAIRport Initiative hat ein Ebenen-Modell zur Einstufung der FAIRness digitaler Objekte entwickelt, das als Leitlinie für den stufenweisen und prozesshaften Ausbau der FAIRness von Datenbeständen dienen kann. Die FAIR-Anforderungen können für jede der drei Komponenten des digitalen Objekts separat geprüft und unterschiedlich umgesetzt werden. Eine Bilddatei kann z. B. anderen Zugangsbeschränkungen unterliegen als die zugehörigen Metadaten.

Digitale Objekte können unterschiedliche FAIR-Ebenen enthalten:

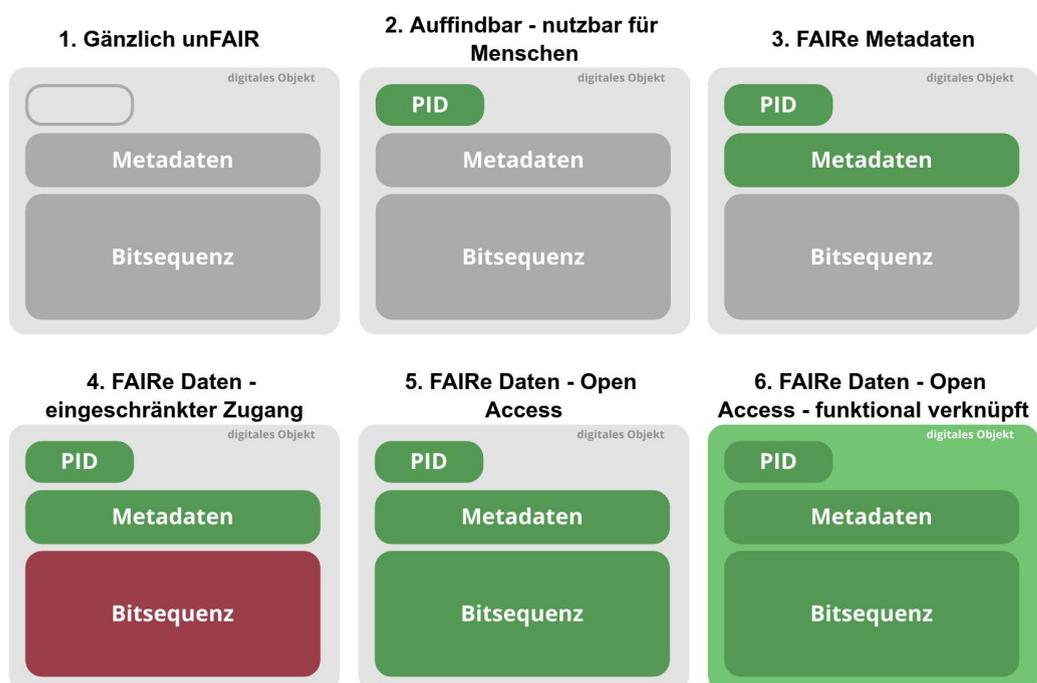


Abb. 2: Daten als zunehmend FAIRe digitale Objekte. FORCE 11. Guiding Principles for Findable, Accessible, Interoperable and Re-usable Data Publishing version b1.0, CC BY-SA 4.0

1. Metadaten sind der Bitsequenz zugeordnet, das digitale Objekt verfügt aber nicht über einen persistenten Identifikator. Die Metadaten sind nicht oder nur eingeschränkt maschineninterpretierbar, und damit ist auch die Bitsequenz nur eingeschränkt maschinell zu verarbeiten (grau).
2. Jedes digitale Objekt hat einen PID. Die Metadaten wurden aber noch nicht für eine maschinelle Auswertung optimiert, das digitale Objekt ist daher in erster

Linie für Menschen verwendbar. Man kann noch nicht von FAIRen  
Nachnutzungsmöglichkeiten sprechen.

3. Jedes digitale Objekt verfügt über umfassende, maschinenauswertbare FAIRe Metadaten (grün). Es fehlen aber noch wichtige Informationen zur Bitsequenz, beispielsweise zur Lizenzierung, zur Datenprovenienz oder zur Strukturierung von komplexen digitalen Objekten.
4. Die Bitsequenzen in den digitalen Objekten sind technisch ebenfalls FAIR, aber nicht frei zugänglich und nicht uneingeschränkt wiederverwendbar, etwa aufgrund rechtlicher Einschränkungen, oder weil es sich um proprietäre Daten handelt (rot). Da die Informationen zu diesen Beschränkungen maschinell auswertbar sind, ist dieses digitale Objekt FAIR.
5. Sowohl die Metadaten als auch die Bitsequenzen selbst sind vollständig FAIR. Sie stehen unter einer genau definierten offenen Lizenz. Damit ist eine umfassende Nachnutzung dieses FAIRen digitalen Objekts möglich.
6. Die digitalen Objekte sind zusätzlich mit Datenbeständen funktional verknüpft, die zu ihnen in Beziehung stehen und die über semantische Technologien in die Auswertung integriert werden können (grüner Hintergrund). Daten liegen als RDF-Statements vor, als granulare Aussagen über die Ressource in Form von Tripeln (Subjekt - Prädikat - Objekt). Es wird außerdem gewährleistet, dass die Informationen zur Datenprovenienz auch für maschinelle Auswertungsmethoden greifbar sind.

Wenn das digitale Objekt nach den FAIR-Prinzipien bereitgestellt wird, gewährleistet es als FAIRes digitales Objekt (FDO) die zuverlässige Interpretation und Verarbeitung der durch das Objekt repräsentierten Daten für Menschen und Computer. FDOs sind außerdem die Basis für ein in Zukunft aufzubauendes FAIR Digital Object Framework (FDOF)<sup>25</sup>, das die maschinelle Verwertbarkeit FAIRer Daten deutlich verbessern wird. Dieses Rahmenwerk will ein vorhersehbares Auflösungsverhalten für Identifikatoren, einen Mechanismus zum Abrufen der Metadaten eines Objekts und ein System zur Typisierung von Objekten definieren.

---

<sup>25</sup> <https://www.go-fair.org/today/fair-digital-framework/>

## 4. Akteure und ihre Rollen im FAIR-Prozess

Es gibt zwei Gruppen von Akteuren, denen unterschiedliche, aber eng verzahnte Rollen bei der Umsetzung der FAIR-Prinzipien zukommen: Datenproduzierende und Betreibende von Datenplattformen und Repositorien. Ein Personenkreis oder eine Institution kann in beiden Rollen aktiv werden.

Auch wenn sich diese Handreichung primär an die Datenproduzierenden wendet, wird die Rolle der Repositorien und Datenplattformen beleuchtet, um den Adressat:innen das Zusammenspiel der beiden Gruppen besser zu erläutern. Sie können dann besser einschätzen, welche Aufgaben zur Unterstützung von FAIR auf Seiten der Datenplattformen liegen. Betreibende von Datenplattformen erhalten Hinweise, wie sie ihr Angebot optimieren und klarer vermitteln können.

### Datenproduzierende

Die Gruppe der Datenproduzierenden schließt alle ein, die für die Forschung relevante Daten erzeugen, bearbeiten und zur Verfügung stellen. Dazu zählen Forschende im engeren Sinne genauso wie die Akteure, die in Gedächtnis- und Infrastruktureinrichtungen mit der Erstellung und dem Management von Daten zu Ressourcen oder Objekten befasst sind.

Kulturwissenschaftliche Forschung nutzt sehr häufig Informationen und Ressourcen, die von anderen, z. B. Bibliotheken oder Repositorien, bereitgestellt werden. Daher erzeugen Forschende nicht nur Daten, sondern sie sind selbst Nachnutzende. Sie sind auf die Verfügbarkeit von Material („Daten“) angewiesen, auf dessen Grundlage Analysen durchgeführt und Schlussfolgerungen gezogen werden können. Sie müssen beurteilen können, ob sich das verfügbare Datenmaterial für ihre Forschung eignet. Das Material kann vielfältige Formen haben: Es kann Bild-, Ton- oder Filmmaterial, historische Quellen, Texte, softwarebasierte Aufbereitungen oder Analyse-Outputs oder Umfrageergebnisse umfassen und schließt auch die Ergebnisse der Forschungsaktivitäten anderer ein. Als Nutzende profitieren die Datenproduzent:innen von der FAIRness dieser Daten.

## Betreibende von Datenplattformen

Ein Datenrepositorium ist ein Dienst, der Forschungsdaten aufbewahrt und den Zugang zu ihnen ermöglicht. Es ist ein zukunftssicheres Instrument, das sicherstellt, dass die Daten langfristig zugänglich und nutzbar bleiben. Mit Blick auf die in den Kulturwissenschaften essentiellen Datenangebote von Bibliotheken, Museen und Archiven ist es aber geboten, die Gruppe der „Datenplattformen“ über die Forschungsdatenrepositorien hinaus auszuweiten. Zu ihr gehören auch fachliche oder fachübergreifende Online-Bestandsangebote, Webportale oder Verbundsysteme, unter Einschluss ihrer schnittstellenbasierten Datenangebote.

Einem Datenspeicher kommen eine Reihe von spezifischen FAIR-Aufgaben zu:

- aktive und langfristige Aufbewahrung der Daten
- Veröffentlichung maschinenauswertbarer Metadaten, um die webbasierte Suche zu ermöglichen, auch angebots- und domänenübergreifend
- Zuweisung von dauerhaften, eindeutigen Identifikatoren zur Gewährleistung der Zitierfähigkeit
- Verwaltung des Online-Zugangs zu den Daten, so dass sie von anderen menschlichen und maschinellen Akteuren genutzt werden können
- Anwendung von Lizenzbedingungen, Kenntlichmachung von Nutzungsbedingungen für die Daten
- optional: Kontrolle der Qualität der Datensätze und Verbesserung der Metadaten, z. B. durch die Anwendung von Standardvokabularen

Bei den Datenplattformen gibt es auf dem Weg zu einem FAIRen Angebot unterschiedliche Strategien, Schwerpunkte und Umsetzungsgrade. Während Forschungsdatenrepositorien im engeren Sinne (z. B. RADAR, heiDATA, Zenodo) die FAIR-Anforderungen häufig schon weitgehend umgesetzt haben, befinden sich auch Bestandsdaten-Plattformen im Prozess der Ausrichtung ihrer Angebote auf transparente, Linked Open Data-basierte Nutzungsszenarien.

So sind die in den folgenden Kapiteln unter „Die Rolle des Betreibers der Datenplattformen“ aufgeführten Hinweise ebenfalls als Empfehlungen für diese Gruppe zu verstehen.

Den Datenproduzent:innen sollen sie als Richtschnur bei der Projektplanung dienen, welche Informationen zu den Vorgaben und Leistungen in Frage kommender Publikationsplattformen man einholen sollte. Eine frühzeitige Abklärung erleichtert es, Daten gleich so zu strukturieren, dass sie von der Publikationsplattform gut entgegengenommen und präsentiert werden können. Nicht in jedem Fall sind diese Informationen auf den Websites der Plattformen einfach auffindbar, so dass sich auch eine persönliche Kontaktaufnahme empfiehlt.

Datenplattformen: FAIR in der Praxis

Hier beschreiben die für die Repositorien Verantwortlichen, wie sie die FAIR-Prinzipien umsetzen: [RADAR<sup>26</sup>](#), [Zenodo<sup>27</sup>](#)

## 5. Findable – Auffindbar

Zu jedem der Hauptziele – Daten auffindbar, zugänglich, interoperabel und nachnutzbar zu machen – nennen die FAIR-Prinzipien jeweils ein Set von Bedingungen, die umgesetzt werden sollen. Diese werden im Folgenden einzeln vorgestellt, erläutert und um Empfehlungen für Akteure der NFDI4Culture-Communities ergänzt.

Der erste Schritt zur Wiederverwendung von Daten durch Menschen und Computer besteht darin, sie leicht finden zu können. Maschinenlesbare umfangreiche Metadaten sind für das automatische Auffinden von Daten unerlässlich. Daten und Metadaten benötigen einen weltweit eindeutigen und dauerhaften Identifikator und werden in einer durchsuchbaren Ressource registriert oder indiziert.

---

<sup>26</sup>

[https://radar.products.fiz-karlsruhe.de/sites/default/files/radar/docs/info/RADAR\\_FAIR\\_Principles.pdf](https://radar.products.fiz-karlsruhe.de/sites/default/files/radar/docs/info/RADAR_FAIR_Principles.pdf)

<sup>27</sup> <https://about.zenodo.org/principles/>

## F1. (Meta-)Daten wird ein global eindeutiger und persistenter Identifikator zugewiesen

Die eindeutige Adressierbarkeit und Lokalisierung von Daten und Metadaten ist eine notwendige Voraussetzung für alle weiteren FAIR-Schritte vom Zugriff bis zur Wiederverwendung. Um auffindbar zu sein, sollte jedes digitale Objekt und jeder Datensatz eindeutig mit einem persistenten Identifikator (PID) versehen sein.

PIDs werden in das weltweit verwendete System zur Identifizierung von Ressourcen im Web, den Uniform Resource Identifier (URI), eingebunden. Global eindeutige, instanziierebare Ressourcenkennungen ermöglichen nicht nur die eindeutige Adressierung von digitalen Objekten, sondern können sich auf alle identifizierbaren, auch auf nicht-digitale Entitäten beziehen: Personen, Organisationen, Dokumente, physische Objekte oder abstrakte Begriffe. Die Verwendung von URIs beseitigt daher auch Uneindeutigkeiten in der Beschreibung von Daten, wenn jeder in den Metadaten genannten Entität ein eindeutiger Identifikator zugewiesen wird.

Einem unikalen physischen Museumsobjekt wie einer Vase kann ein PID zugewiesen werden. Dieser PID kann zu einer Webseite mit der Beschreibung des physischen Objekts führen, er identifiziert aber nicht die Webseite oder eine andere digitale Repräsentation des Objekts (z. B. ein Datensatz oder ein Digitalbild). Letzteren sollen vielmehr eigene PIDs zugewiesen werden. In den digitalen Repräsentationen oder ihren Metadaten wird ein thematischer Bezug zum physischen Objekt hergestellt, der dessen PID einschließt.

In diesem Zusammenhang bestehen Identifikatoren aus einem Internet-Link, der zu einer Webseite führt, die das Informationsobjekt definiert, wie z. B. einen Forschungsdatensatz, ein bestimmtes Sachschlagwort, eine Person, ein Geografikum. Identifikatoren können anderen Menschen helfen, genau zu verstehen, was Sie meinen, und sie ermöglichen es maschinellen Agenten, Ihre Daten bei der Suche oder automatischen Integration auf sinnvolle Weise zu interpretieren. Identifikatoren sind für die Zusammenarbeit zwischen Mensch und Maschine unerlässlich. Darüber hinaus helfen Identifikatoren anderen, Ihre Arbeit bei der Wiederverwendung Ihrer Daten korrekt zu zitieren.

Ein persistenter Identifikator muss weltweit eindeutig sein – jemand anderes kann den selben Identifikator nicht wiederverwenden oder neu zuweisen, ohne sich auf Ihre Daten zu beziehen. Global eindeutige Bezeichner werden von Registrierungsdiensten vergeben, deren Algorithmen die Eindeutigkeit neu geprägter Identifikatoren garantieren.

Persistente Identifikatoren müssen dauerhaft gültig sein. Das gilt auch dann, wenn sich die Webadresse (URL) einer Ressource ändert. Die Registrierungsdienste garantieren die Auflösbarkeit dieses Links auch in Zukunft, zumindest bis zu einem bestimmten Grad. Es kostet Zeit und Geld, Weblinks aktiv zu halten. Daher müssen teilnehmende Institutionen sich für eine in der Regel kostenpflichtige Mitgliedschaft registrieren. Registrierungsdienste fungieren als zentrale Vergabestellen, welche den teilnehmenden Einrichtungen eindeutige Präfixe für „ihre“ PIDs zuweisen. Viele dieser Datenplattformen erzeugen automatisch weltweit eindeutige und dauerhafte Identifikatoren für die bei ihnen hinterlegten Datensätze.

Anbieter von PIDs sollten eine klare Richtlinie veröffentlichen, die beschreibt, zu welchen Konditionen die dauerhafte Auflösung des Identifikators zum richtigen Speicherort und zum richtigen Informationsobjekt garantiert wird. Es liegt auf der Hand, dass lokal verwendete Identifikatoren (beispielsweise die automatisch vergebenen Datensatz-IDs innerhalb einer lokalen Datenbank), die nicht automatisch auf gemeinschaftlich getragene und öffentlich geteilte Identifikationssysteme referenziert werden können, nicht persistent sind. Ein Datenanbieter, der ein „proprietäres“ Identifizierungsschema wählt, muss daher geeignete und korrekte Zuordnungen zu öffentlichen Identifizierungsschemata bereitstellen, um als FAIR zu gelten.

PID-Systeme

PIDs werden im Rahmen verschiedener Identifikationssysteme angeboten.

## Digital Object Identifier (DOI)

DOI<sup>28</sup> kombiniert ein Metadatenmodell mit dem Handle-System (siehe unten) als Auflösungsinfrastruktur (d. h. DOIs sind Handles). Das System wurde mit Unterstützung der International DOI Foundation (IDF) eingeführt und wurde 2012 zur offiziellen ISO-Norm (ISO 26324). Das DOI-System baut auf CNRI-Handles auf. DOI-Registrierungsagenturen sind für die Zuweisung von Identifikatoren verantwortlich. Sie haben jeweils ihr eigenes kommerzielles oder nicht-kommerzielles Geschäftsmodell, um die damit verbundenen Kosten zu tragen. Das DOI-System selbst wird von der IDF gepflegt und weiterentwickelt, die ihrerseits von den Mitgliedern der Registrierungsagenturen kontrolliert wird. Unter dem Handle-System gibt es einen zentralen, kostenlosen und weltweiten Auflösungsmechanismus für DOI-Namen. DOI-Namen von jeder Registrierungsagentur können weltweit in jedem Handle-Server aufgelöst werden; DOIs sind also eigenständig und ihre Auflösung ist nicht von einer einzelnen Agentur abhängig. Für jeden DOI-Namen wird ein Standard-Metadaten-Kernel definiert. Für die Vergabe von DOI-Namen ist eine Lizenzgebühr zu entrichten, ihre Auflösung ist jedoch kostenlos.

Das DOI-System ist im wissenschaftlichen Publikationsprozess weit verbreitet.

## Uniform Resource Name (URN)

URN<sup>29</sup> ist ein Standard der Internet Engineering Task Force (IETF). Es gibt keine zentrale Verwaltung oder Resolver-Infrastruktur. Große Nationalbibliotheken in Europa haben ihre eigene Untergruppe von URN, die URN:NBN, gegründet und betreiben eine gemeinsame Infrastruktur für die

---

<sup>28</sup> <https://www.doi.org/>

<sup>29</sup> <https://www.dnb.de/urnservice>

Namensauflösung. Internationale Standard Buch Nummern (ISBN) für Bücher sind Teil des URN-Systems.

Es fallen keine Lizenzkosten für die Vergabe von URNs an, aber eine URN-Registrierungsagentur muss eine Infrastruktur für die Vergabe und die Auflösung einrichten. Da keine gemeinsame Auflösungsinfrastruktur oder Arbeitsabläufe für URNs existieren, abgesehen von Teilbereichen wie URN:NBN, ist es unmöglich, eine generelle Interoperabilität mit URNs herzustellen.

Der URN-Service der Deutschen Nationalbibliothek<sup>30</sup> für öffentliche Institutionen ist an bestimmte Bedingungen geknüpft, aber kostenlos.

## Handle

Handle<sup>31</sup> ist ein nichtkommerzielles dezentrales System zur Auflösung von Identifikatoren, das von der Corporation for National Research Initiatives (CNRI) betrieben wird. Es wird von vielen anderen übergeordneten Systemen verwendet, z. B. DOI. Verschiedene Initiativen verwenden kommerzielle Handle-Lizenzen, um lokale Handle-Systeme einzurichten, so das European Persistent Identifier Consortium (EPIC). Viele Repositorien betreiben derzeit ihr eigenes lokales Handle-System.

## Archival Resource Key (ARK)

ARK<sup>32</sup> ist kein formaler Standard, aber alle ARKs folgen der gleichen Struktur und den gleichen Arbeitsabläufen. Es gibt keinen zentralen Resolver

---

<sup>30</sup> [https://www.dnb.de/DE/Professionell/Services/URN-Service/urn-service\\_node.html](https://www.dnb.de/DE/Professionell/Services/URN-Service/urn-service_node.html)

<sup>31</sup> <https://www.handle.net/>

<sup>32</sup> <https://arks.org/>

(Verzeichnisdienst zur Namensauflösung) – Organisationen können sich anmelden, um Name Assigning Authority Numbers (NAANs) zu werden und ihre eigene Infrastruktur für die Auflösung von ARKs zu betreiben. Das System wird von der California Digital Library mit Dutzenden von NAANs weltweit über eine kombinierte ARK/DOI-Infrastruktur EZID betrieben. Diese EZID Infrastruktur ermöglicht die Interoperabilität zwischen ARKs und DOI-Namen unter dem Dach von DataCite. Im deutschen Sprachraum spielt ARK eine untergeordnete Rolle.

### Persistent Uniform Resource Locator (PURL)

PURLs sind Webadressen, die angesichts einer dynamischen und sich verändernden Webinfrastruktur als dauerhafte Identifikatoren fungieren. Anstatt direkt auf Webressourcen zu verweisen, verweisen PURLs über einen zwischengeschalteten Resolver auf den tatsächlichen Standort der angeforderten Webressource. Diese Funktion gewährleistet die Kontinuität der Adressierung der Ressourcen, die von Server zu Server transferiert werden können, ohne dass dies negative Auswirkungen auf die Systeme hat, die von ihnen abhängen.

Nach langer Trägerschaft durch OCLC hat das Internet Archive 2016 die Verwaltung des PURL-Dienstes übernommen. PURL-Server werden von verschiedenen Organisationen betrieben. Eine formelle Mitgliedschaft ist notwendig. Der Quellcode steht frei zur Verfügung.

### **Die Rolle der Datenproduzierenden**

Nutzen Sie PIDs für die in den Daten behandelten oder erwähnten Entitäten, dies verbessert ihre Auffindbarkeit und Nachnutzbarkeit deutlich. Differenzieren Sie sorgfältig zwischen PIDs, die sich auf physische und auf digitale Objekte beziehen.

Publizieren Sie auf einer Plattform, die Ihren Daten PIDs zuweist. Klären Sie mit Ihrer Datenplattform, ob eine digitale Ressource, die in unterschiedlichen Formaten publiziert werden soll, separate PIDs benötigt. Bestimmen Sie, mit welcher Granularität die zu veröffentlichenden Elemente Ihrer Forschungsdaten über PIDs separat adressierbar sein sollen.

Benötigen Daten eine größere Menge an PIDs, wie es bei Museums- oder Archivbeständen der Fall sein kann, kann es notwendig werden, selbst Mitglied der PID-Anbieterorganisation zu werden oder eine Partnerorganisation zu finden, die PIDs vergeben kann. Die Zuweisung von PIDs kann mit Kosten verbunden sein. Eine Bedingung ist häufig, dass Datenpakete nach ihrer Einlieferung nur noch in (meist rechtlich begründeten) Ausnahmefällen verändert werden dürfen. Kontinuierlich weiterentwickelte Metadaten zu Sammlungsbeständen benötigen eigentlich eine Versionierung der PIDs und damit auch der zitierbar vorgehaltenen Stände der zugehörigen Daten. Einfacher umsetzbar ist diese Anforderung für die qualitativ hochwertigen Digitalisate der Sammlungsbestände. Beispielhaft<sup>33</sup> ist dies beim Bildarchiv der ETH-Zürich realisiert, das bislang über 770.000 digitale Objekte seines Bestands mit DOIs ausgestattet hat.

Verweisen Sie in Ihren Forschungsergebnissen auf den PID, der Ihrem Datensatz zugewiesen wurde. Verweisen Sie auch auf PIDs von Datensätzen anderer, die Sie verwendet haben und verwenden Sie PIDs zur Referenzierung der Entitäten, auf die sich Ihre Metadaten beziehen. Mehr Informationen dazu im Abschnitt 13. (Meta-)Daten enthalten qualifizierte Verweise auf andere (Meta-)Daten.

### **Die Rolle des Betreibenden der Datenplattform**

Wählen Sie die geeignete Form eines dauerhaften Identifikationssystems und weisen Sie jeder Ressource einen PID zu. Machen Sie die Richtlinien für den Umgang mit dem Identifizierungsschema öffentlich zugänglich, wie es z. B. Zenodo<sup>34</sup> tut.

---

<sup>33</sup> <http://doi.org/10.3932/ethz-a-000674874>

<sup>34</sup> <https://help.zenodo.org/#policies>

Weiterführende Informationen zu PIDs

[THOR – Technical and Human Infrastructure for Open Research. Persistent Identifier Platform](#)<sup>35</sup>

[Koster, Lukas: Persistent identifiers for heritage objects. Code4Lib Journal, 47, 2020](#)<sup>36</sup>

[Arnold, Eckart / Müller, Stefan: Wie permanent sind Permalinks? Informationspraxis Bd. 3 Nr. 1, 2017](#)<sup>37</sup>

PIDs für Sammlungsobjekte: Teilprojekt [HeritagePIDs](#)<sup>38</sup> des britischen Programms [Towards a National Collection \(TaNC\)](#)<sup>39</sup>

## Forschungsdaten zitieren

Wenn Forschungsdaten einen dauerhaften Identifikator haben und in Übereinstimmung mit den Community-Standards zitiert werden, sind die entsprechenden digitalen Objekte oder Datensätze leichter zu finden.

### **Die Rolle der Datenproduzierenden**

Machen Sie sich mit den Zitierrichtlinien für Daten vertraut, die für Ihre Institution oder Ihr Publikationsorgan gelten, und zitieren Sie Forschungsdaten entsprechend. Dies gilt auch für publizierte und über PIDs zitierbare Ressourcen (z. B. digitalisierte Sammlungsbestände), Normdaten und Vokabulare, die in Ihren Daten gegenständlich sind.

### **Die Rolle des Betreibenden der Datenplattform**

Informieren Sie Ihre Nutzer:innen in geeigneter, gut auffindbarer Form über bewährte Praktiken beim Zitieren von Daten. Erleichtern Sie ihnen die Ausführung, z. B. durch

---

<sup>35</sup> <https://project-thor.readme.io/>

<sup>36</sup> <https://hdl.handle.net/11245.1/93e75978-419d-431b-81e7-289974e8e0a6>

<sup>37</sup> <https://doi.org/10.11588/ip.2016.2.33483>

<sup>38</sup> <https://tanc-ahrc.github.io/HeritagePIDs/>

<sup>39</sup>

<https://www.ukri.org/what-we-offer/browse-our-areas-of-investment-and-support/towards-a-national-collection-opening-uk-heritage-to-the-world/>

eine standardisierte Schaltfläche auf der Webseite mit der Aufschrift „Diesen Datensatz zitieren“, die das Zitat in verbreiteten Zitierstilen vorformatiert anbietet.

Empfehlungen für eine gute Praxis rund um zitierfähige Datenpublikationen

Data Citation Synthesis Group: Joint Declaration of Data Citation Principles, Martone, M. (Hg.), FORCE11, 2014<sup>40</sup>

Murdoch University Library: Chicago Referencing Guide, 2022<sup>41</sup>

Dataverse Project: Best Practices. Data Citation<sup>42</sup>

## Persistente Identifikatoren für Personen und Institutionen verwenden

Nutzen Sie dauerhafte Identifikatoren für Personen und Forschungsorganisationen und -institutionen, z. B. Open Researcher and Contributor ID (ORCID)<sup>43</sup>, International Standard Name Identifier (ISNI)<sup>44</sup>, Gemeinsame Normdatei (GND)<sup>45</sup> oder Virtual International Authority File (VIAF)<sup>46</sup>, für Institutionen auch Research Organization Registry (ROR)<sup>47</sup>. Kennzeichnen Sie die Beiträge aller Projekt-Teammitglieder, die verantwortlich benannt werden sollen, sei es als Autor:innen, als Beitragende oder zur Angabe ihrer Institutionszugehörigkeit (Affiliation). Wissenschaftler:innen unterstützen damit die Darstellung ihrer eigenen Forschungsleistung. Die Autor:innenidentifikation ermöglicht die Wiedererkennung und Auffindbarkeit von Personen und Institutionen. Sie hilft auch, Verknüpfungen zwischen Datensätzen, Forschungsaktivitäten, Veröffentlichungen und Forschenden herzustellen.

---

<sup>40</sup> <https://doi.org/10.25490/a97f-egyk>

<sup>41</sup> <https://libguides.murdoch.edu.au/Chicago/dataset>

<sup>42</sup> <https://dataverse.org/best-practices/data-citation>

<sup>43</sup> <https://orcid.org/>

<sup>44</sup> <https://isni.org/>

<sup>45</sup> [https://www.dnb.de/DE/Professionell/Standardisierung/GND/gnd\\_node.html](https://www.dnb.de/DE/Professionell/Standardisierung/GND/gnd_node.html)

<sup>46</sup> <http://viaf.org/>

<sup>47</sup> <https://ror.org/>

## Die Rolle der Datenproduzierenden

Unterscheiden Sie sich von gleichnamigen Forschenden oder Forschungsgruppen. Beantragen Sie eine Autor:innenkennung, wenn Sie noch keine besitzen. Registrieren Sie sich z. B. bei ORCID und verweisen Sie auf Ihre ORCID-ID in Ihrem Datensatz und an anderen Stellen, die Ihnen für die Zusammenführung forschungsrelevanter Informationen geeignet erscheinen.

Prüfen Sie auch, ob Ihre Institution bei ROR registriert ist und verwenden Sie in diesem Fall den entsprechenden Identifikator.

## Die Rolle des Betreibenden der Datenplattform

Zeigen Sie vorhandene Identifikatoren für Autor:innen und Institutionen mit den Metadaten an und erlauben Sie nach Möglichkeit eine Verlinkung zu den dahinter liegenden Profilen. Um komplexe Rollen rund um Autor:innenschaft, Mitarbeit und Verantwortlichkeiten transparenter auszuweisen, ermöglichen Sie die Zuweisung von Rollenbezeichnungen.

Weitere Informationen zu Identifikatoren für Personen und Rollenbezeichnungen

Hagemann-Wilholt, Stephanie / Burger, Marleen / Dreyer, Britta et al.: Autor:innenidentifikation mit der ORCID iD: Warum und für wen?, in: ORCID in Deutschland – Blog, 2022<sup>48</sup>

Beispiel für die Nutzung der ORCID in Nachweissystemen: GEPRIS<sup>49</sup>; Digital Humanities Lab (Leibniz-Institut für Europäische Geschichte)<sup>50</sup>

Beispiel für die Verlinkung zu den ORCID-Profilen der Autorinnen in einem Repository: Zenodo<sup>51</sup>

Rollenzuweisungen ermöglichen z. B. die Contributor Types im DataCite Metadata Schema v.4.4<sup>52</sup>, Appendix 1, S. 40

---

<sup>48</sup>

<https://www.orcid-de.org/support/blogbeitraege/autorinnenidentifikation-mit-der-orcid-id-war-um-und-fuer-wen>

<sup>49</sup> <https://gepris.dfg.de/gepris/person/229126145>

<sup>50</sup> <https://dhlabs.hypotheses.org/tag/orcid>

<sup>51</sup> <https://doi.org/10.5281/zenodo.5654634>

<sup>52</sup> [https://schema.datacite.org/meta/kernel-4.4/doc/DataCite-MetadataKernel\\_v4.4.pdf](https://schema.datacite.org/meta/kernel-4.4/doc/DataCite-MetadataKernel_v4.4.pdf)

## F2. Daten werden mit umfangreichen Metadaten (vgl. R1) beschrieben

Ausführliche Metadaten dokumentieren den Gegenstand der Forschungsdaten, aber auch, wie die Daten generiert wurden, wer sie erhoben, bearbeitet, oder publiziert hat und unter welchen Bedingungen (Lizenz) sie verwendet werden dürfen. Sie liefern somit den Kontext für die richtige Interpretation. Sie sind unerlässlich, um Daten auffindbar und zitierbar zu machen. Diese Informationen müssen maschineninterpretierbar sein.

Ein Metadatenschema ist eine Liste standardisierter Elemente zur Erfassung von Informationen über ein Objekt, z. B. ein Titel/Name, ein Identifikator, ein Name des Herstellers/Urhebers oder ein Datum. Durch die Verwendung bestehender Metadatenschemata wird sichergestellt, dass die internationalen Standards für den Datenaustausch eingehalten werden können.

Wählen Sie ein geeignetes Metadatenschema, das Ihrem Gegenstand angemessen ist und das in Ihrem Fachgebiet oder Ihrer Sparte verbreitet ist. Hier finden Sie [Empfehlungen für Metadatenstrukturen](#).

### **Die Rolle der Datenproduzierenden**

Um das Auffinden von Inhalten zu ermöglichen, sollten Forschungsdaten so konsistent und vollständig wie möglich beschrieben werden. Fügen Sie genügend Informationen hinzu, damit die Daten später auffindbar, zugänglich und verständlich sind. Oft ist das Metadatenschema bereits durch das Repositorium vorgegeben. Andere Repositorien erlauben auch die Verwendung eigener Metadatenschemata. Verwenden Sie dann nach Möglichkeit ein in Ihrer Fachcommunity etabliertes oder von der Datenplattform empfohlenes Metadatenschema, das gut für die Art des digitalen Objekts geeignet ist, das Sie beschreiben.

Konsultieren Sie die jeweiligen Empfehlungen zur Belegung von Elementen, die für bestimmte Metadatenschemata obligatorisch sind. Nutzen Sie auch optionale Elemente, die zusätzlich empfohlen werden – Sie erhöhen damit die Auffindbarkeit und Transparenz des Datensatzes. Gewährleisten Sie, dass Ihre Daten unter verschiedenen Fragestellungen, auch unter thematischen Gesichtspunkten, auffindbar sind.

## Die Rolle des Betreibenden der Datenplattform

Geben Sie klar und deutlich an, welches Metadatenschema Sie verwenden und den Forschenden empfehlen (z. B. DataCite für RADAR).

Unterstützen Sie Datengeber:innen bei der Eingabe der signifikanten Metadaten, indem Sie ein strukturiertes Formular oder ein dokumentiertes XML-Schema bereitstellen, in dem die für die Auffindung relevanten Metadaten erläutert werden. Unterstützen Sie dabei auch die einfache Verknüpfbarkeit zu Normdaten für die relevanten Metadaten-Entitäten. Ermöglichen Sie auch eine möglichst unkomplizierte Eingabe zusätzlicher optionaler Metadaten, z. B. über die Herkunft der Daten und den Projektkontext. Machen Sie es der oder dem Datengeber:in leicht, eine Normdatenreferenzierung für die thematische Verschlagwortung durchzuführen.

Bieten Sie neben der formularbasierten Eingabe alternative Ergänzungsoptionen für Metadaten an, z. B. über XML-Upload oder API. Ermöglichen Sie für komplexe, mehrteilige Forschungsdatensets auch die Angabe von Metadaten für Unterebenen oder die enthaltenen Einzeldateien.

Validieren Sie, ob die Kernfelder des Metadatensatzes vollständig belegt wurden.

Bestandsdaten-Plattformen nutzen meist technische Informationsseiten im Rahmen ihrer Angebote, um die verwendeten Metadatenschemata zu dokumentieren.

### Beispiele für die Schemadokumentation bei Datenplattformen

Dokumentation des Metadatenschemas bei RADAR<sup>53</sup>. Mit dem Metadatenschema v.9.1. ermöglicht RADAR nun eine umfassende GND-Referenzierung auch für die thematische Verschlagwortung von Forschungsdaten.

Metadatendokumentation bei Bestandsdaten-Plattformen: Europeana<sup>54</sup>; Rijksmuseum Amsterdam<sup>55</sup>; Getty Collections<sup>56</sup>

<sup>53</sup> <https://radar.products.fiz-karlsruhe.de/de/radarfeatures/radar-metadaten-schema>

<sup>54</sup> <https://pro.europeana.eu/page/edm-documentation>

<sup>55</sup> <https://data.rijksmuseum.nl/object-metadata/harvest/>

<sup>56</sup> <https://www.getty.edu/projects/open-data-apis/>

### F3. Metadaten enthalten eindeutig und explizit den Identifikator der Daten, die sie beschreiben

Werden Forschungsmetadaten und Daten separat gehalten, muss man sicherstellen, dass sie mit PIDs eindeutig aufeinander verweisen, um die Verbindung zwischen ihnen stets eindeutig zu gewährleisten.

#### **Die Rolle der Datenproduzierenden**

Nehmen Sie eine klare inhaltlich-strukturelle Gliederung der Datenbestände in Bezug auf Daten und Metadaten vor. Wo Bezüge zwischen Datensets absehbar benötigt werden, bereiten Sie die Referenzierung mit lokalen Identifikatoren vor.

#### **Die Rolle des Betreibenden der Datenplattform**

Unterstützen Sie die Vergabe von PIDs für Metadaten und Daten, die im besten Fall automatisiert erfolgt. Vermitteln Sie transparent, welche Datenbestände zusammengehören, so dass sie einfach gefunden und zusammen nachgenutzt werden können.

#### Beispiel einer mehrteiligen Forschungsdatenpublikation

Publikation von mehreren, in Beziehung gesetzten Datensätzen unterschiedlicher Formate: Arnoux, Mathilde: „OwnReality. Jedem seine eigene Realität [Forschungsdaten]“, heiDATA, 2020<sup>57</sup>

### F4. (Meta-)Daten werden in einer durchsuchbaren Ressource registriert oder indiziert

Zu den durchsuchbaren Ressourcen gehören nicht nur von Menschen nutzbare Angebote der Repositorien und Datenplattformen selbst, sondern auch auf erweiterte

---

<sup>57</sup> <https://doi.org/10.11588/data/ZMIHMY>

Suchräume orientierte Angebote, die die (Meta-)Daten einem für einen sehr viel größeren Kreis von Nutzer:innen auffindbar machen. Dazu zählt die Öffnung von Datenangeboten für bibliothekarische Discovery-Systeme und Internet-Suchmaschinen, aber auch die Bereitstellung der Daten auf Abfrage- oder Programmierschnittstellen (APIs) zum Harvesting durch Dritte.

Die Veröffentlichung strukturierter Metadaten in Webseiten kann ein einfaches und effizientes Mittel sein, um die FAIRness von Forschungsressourcen zu erhöhen. Forschungsdatenrepositorien und Bestandsdatenbanken im Web profitieren daher von einer Suchmaschinenoptimierung durch Integration von [Schema.org](https://schema.org/)<sup>58</sup>-Metadaten, die als Markup in die HTML-Webseiten eingebettet werden. Der Bezug auf das Schema.org-Vokabular liefert der Suchmaschine zusätzlich Kontext, um den semantischen Zusammenhang zwischen den Begriffen auf einer Webseite zu interpretieren.

### **Die Rolle der Datenproduzierenden**

Die Voraussetzung für gute Such- und Abfrageergebnisse liegt in der Vollständigkeit und Exaktheit der Metadaten. Nutzen Sie auch die nicht-obligatorischen Elemente des Schemas für die Verschlagwortung. Referenzieren Sie die in den Metadaten erwähnten Entitäten mit Normdaten-PIDs, um sie maschinell semantisch auswertbar zu machen.

### **Die Rolle des Betreibenden der Datenplattform**

Dokumentieren Sie, über welche Ressourcen und Schnittstellen die Daten Ihres Angebots erreichbar sind, und welche Metadaten in welchen Formaten zur Verfügung stehen. So werden Daten beispielsweise zum Download bereitgestellt, Metadaten stehen zum Harvesting über eine OAI-Schnittstelle zur Verfügung oder werden an weiteren Stellen, etwa bei DataCite oder B2Find, indiziert.

Öffnen Sie den Inhalt Ihres Angebots für Webcrawler, die ihn für Suchmaschinen sichten und indexieren. Integrieren Sie ein Markup mit Schema.org-Vokabular in die Webseiten Ihrer Plattform. Die RDA Research Metadata Schemas Working Group bietet dazu [Guidelines, einen Metadata Crosswalk und Toolempfehlungen](#)<sup>59</sup> an. Die

---

<sup>58</sup> <https://schema.org/>

<sup>59</sup> <https://www.rd-alliance.org/node/62494/outputs>

Integration von Schema.org wird von [DataCite](#)<sup>60</sup> und [Dataverse](#)<sup>61</sup> empfohlen und unterstützt.

## 6. Accessible – Zugänglich

Menschen und Maschinen erhalten Zugang zu Ihren Daten, gegebenenfalls unter bestimmten Bedingungen oder Einschränkungen.

Stellen Sie sicher, dass die Datenplattform, die Sie für die Bereitstellung Ihrer Daten verwenden, persistente Identifikatoren vergibt, über die die Daten abgerufen werden können. Gegebenenfalls erfordert das Zugriffsverfahren Authentifizierungs- und Autorisierungsschritte. Auch wenn die Daten nicht (mehr) zugänglich sind, bleiben die Metadaten zugänglich, wo immer möglich.

### A1. (Meta-)Daten sind über ihren Identifikator mithilfe eines standardisierten Kommunikationsprotokolls abrufbar

Grundsatz A1 besagt, dass eine FAIRe Datenabfrage ohne spezielle oder proprietäre Werkzeuge oder Kommunikationsmethoden erfolgen sollte. Hier liegt der Hauptakzent darauf, wie Daten und Metadaten anhand ihrer Identifikatoren abgerufen werden können.

#### A1.1 Das Protokoll ist offen, kostenlos und universell implementierbar

Durch die Verwendung standardisierter Kommunikationsprotokolle machen Repositorien (Meta-)Daten im Web öffentlich zugänglich und ermöglichen z. B. die Auswertung durch Suchmaschinen oder schnittstellenbasierte Abfragen. Jede:r Nutzer:in mit einem Computer und einer Internetverbindung kann zumindest auf die Metadaten zugreifen.

---

<sup>60</sup> <https://support.datacite.org/docs/schemaorg>

<sup>61</sup> <https://dataverse.org/blog/latest-dataverse-update-adds-support-schemaorg>

Gut dokumentierte und maschinenadressierbare Schnittstellen ermöglichen die automatische Indizierung, Abfrage und Kombination von (Meta-)Daten aus verschiedenen Datenbeständen. Die Zugänglichkeit wird so noch einmal erheblich verbessert.

### **Die Rolle des Betreibenden der Datenplattform**

Verwenden Sie standardisierte Protokolle und Schnittstellen wie OAI-PMH, REST, SPARQL Endpoints, ResourceSync oder SWORD. Nutzen Sie keine proprietären oder kommerziellen Kommunikationslösungen.

Konvertieren Sie Metadatenschemata in verbreitete Formate wie XML, JSON oder RDF und ermöglichen Sie ihre Übermittlung. Dokumentieren Sie Schnittstellen gut, führen Sie einen Verzeichnisdienst (Registry) für Schnittstellen-Endpunkte, über die die Forschungsdaten abgerufen werden können, und veröffentlichen Sie diese.

Zeigen Sie anhand von Beispielen, wie man erfolgreich Daten von verschiedenen Endpunkten abrufen und zu neuen Datensätzen kombinieren kann, die für neue Forschungsarbeiten genutzt werden können.

Ein Beispiel aus dem Bereich der Filmwissenschaft ist das Linked-Open-Data-Angebot von Cinema Context<sup>62</sup> mit gut dokumentierten "basic queries"<sup>63</sup>.

## **A1.2 Das Protokoll unterstützt bei Bedarf Verfahren zur Authentifizierung und Rechteverwaltung**

Daten sollten „so offen wie möglich, so geschlossen wie nötig“ publiziert werden. Nicht immer ist es möglich, alle Daten eines Projekts offen anzubieten, weil bestimmte Vertragsbedingungen oder rechtliche Einschränkungen berücksichtigt werden müssen.

Bei kommerziell sensiblen Daten kann es notwendig sein, dass diese nach der Erhebung und Auswertung über einen längeren Zeitraum unter Verschluss gehalten werden müssen. Dies kann bei Projekten der Fall sein, in denen mit Rechten belegte

---

<sup>62</sup> <https://uvacreate.gitlab.io/cinema-context/cinema-context-rdf/>

<sup>63</sup> <https://uvacreate.gitlab.io/cinema-context/cinema-context-rdf/queries/basic/>

Materialien digitalisiert werden und/oder die mit kommerziellen Partnern oder Verlagen kooperieren, so dass eine verzögerte Open Access-Publikation (moving wall) nach Projektende in Frage kommt.

In solchen Fällen kann ein Datenembargo eingesetzt werden. Während dieses Zeitraums wird nur die Beschreibung des Datensets veröffentlicht, die Daten selbst sind nicht zugänglich. Die vollständigen Daten werden nach einer bestimmten Zeit verfügbar gemacht.

Legen Sie die Zugänglichkeit der Daten und der Metadaten immer deutlich dar.

### **Die Rolle der Datenproduzierenden**

Berücksichtigen Sie bei der Wahl der Zugriffsoption rechtliche Anforderungen, fachspezifische Policies und gegebenenfalls ethische Richtlinien, die für Ihre Daten gelten. Wählen Sie Open Access, wenn das möglich ist. Wenn Sie personenbezogene Daten erheben, überprüfen Sie, ob dies durch die Nutzungsbedingungen der Datenplattform abgedeckt ist, ob die Personen ihre Zustimmung zu einer Nutzung gegeben haben und welche Maßnahmen Sie zum Schutz dieser Daten getroffen haben.

Achten Sie darauf, dass die gewählte Lizenzierung in maschinenauswertbarer Form darstellbar ist. Nähere Hinweise finden Sie dazu in der Dokumentation des gewählten Metadatenschemas.

Machen Sie die Metadaten so bald wie möglich öffentlich zugänglich und versehen Sie sie mit einer möglichst offenen Lizenz, auch wenn die zugehörigen Daten nicht im Open Access veröffentlicht werden können. So werden Ihre zugangsbeschränkten Daten zumindest auffindbar.

Falls Sie mit einem Datenembargo arbeiten, geben Sie klar an, warum und für welchen Zeitraum es erforderlich ist.

### **Die Rolle des Betreibenden der Datenplattform**

Unterstützen Sie die Veröffentlichung von (Meta-)Daten im Open Access. Teilen Sie mit, welche Einschränkungen der Zugriffsmöglichkeiten für (Meta-)Daten Ihr

Repositorium bei Bedarf unterstützen kann. Geben Sie an, ob ein Embargo für eingelieferte Daten zulässig ist und welche Bedingungen dafür gelten.

Stellen Sie den Nutzenden deutlich dar, wie der beschränkte Zugang zu sensiblen (Meta-)Daten, die nicht Teil der öffentlich zugänglichen (Meta-)Daten sein können, erfolgen kann. In diesem Fall sollten Sie sich bemühen, die (Meta-)Daten über ein kontrolliertes und dokumentiertes Zugriffsverfahren verfügbar zu machen, das auch für maschinenbasierte Zugriffe geeignet ist. Die geltenden Kontakt- und Zuständigkeitsangaben müssen klar vermittelt werden. Manchmal ist die Einrichtung von Benutzerkonten notwendig, um die Autorisierung von Zugriffen und die Einstellung spezifischer Rechte zu regeln.

Ein Beispiel für die Dokumentation der vom Repositorium unterstützten Zugriffseinschränkungen findet sich bei RADAR<sup>64</sup>.

## A2. Metadaten bleiben verfügbar, auch wenn die zugehörigen Daten nicht (mehr) verfügbar sind

Die Gewährleistung der Verfügbarkeit von Datenbeständen über einen unbegrenzten Zeitraum hinweg bedeutet für Repositorien einen großen Aufwand, der durch die Anpassung an neue Formatstandards, Weiterentwicklung der verwendeten Forschungssoftware oder der Betriebsumgebung des Repositoriums notwendig wird. Deshalb wird das vollständige Forschungsdatenset häufig nur für einen bestimmten Zeitraum bereitgestellt. Die Metadaten sollten aber unbeschränkt verfügbar gehalten werden. So kann das Repositorium mit deutlich geringerem Aufwand umfassende Nachweissysteme zu allen aktuellen und nicht mehr verfügbaren Datenbeständen vorhalten.

Manchmal müssen Forschungsdaten auch aufgrund von Einschränkungen zurückgezogen werden, die erst nach der Publikation bekannt werden - auch dann sollten die Metadaten möglichst verfügbar bleiben.

---

<sup>64</sup> <https://radar.products.fiz-karlsruhe.de/de/radarsupport/faq#rollen+und+zugriffsrechte>

## **Die Rolle der Datenproduzierenden**

Strukturieren Sie Ihre Daten so, dass auch nach der mit der Datenplattform vereinbarten Vorhaltungsfrist für den Gesamtbestand der Projektdaten die Metadaten noch einen aussagekräftigen Nachweis über das gesamte Datenset liefern können. Auch wenn die Daten selbst nicht mehr zugänglich sind, erlauben die Metadaten anderen Forschenden die Einordnung im Rahmen ihrer eigenen Fragestellungen.

Halten Sie daher die Metadaten separat und betten Sie sie nicht ausschließlich in Komponenten ein, die sich in derselben Datei wie die Bitsequenz befinden. Ein Beispiel sind Metadaten zu Digitalbildern: Sie sollten nicht ausschließlich im Header der Bilddatei gespeichert werden. Halten Sie diese in einem externen Metadatenformat und ermöglichen Sie so eine gute Nachnutzbarkeit in Kontexten, die mit diesen Formaten arbeiten.

Für Anwendungskontexte, in denen direkt mit den Digitalbildern gearbeitet wird, ist es aber durchaus sinnvoll, bestimmte Informationen redundant in Bild-Headern zu speichern. Hier beugen Sie Informationsverlust vor, indem Sie etwa Angaben zur Lizenzierung, Urheber:in, Kurzbeschreibung und Verschlagwortung in den IPTC-Daten und bestimmte technische Informationen in den EXIF-Daten der Bilddateien unterbringen.

## **Die Rolle des Betreibenden der Datenplattform**

Stellen Sie sicher, dass Metadaten und Daten in separaten Dateien gehalten werden.

Sehen Sie eine Sperrmöglichkeit für Datensätze vor. Archivieren Sie Metadaten zeitlich unbegrenzt und stellen Sie sicher, dass sie über ihren PID erreichbar bleiben, auch wenn die zugehörigen Daten nicht mehr erreichbar sind.

Um die langfristige Zugänglichkeit von Metadaten über die Existenzdauer des Repositoriums hinaus zu gewährleisten, sollten Sie eine Exit-Strategie entwickeln, die vorsieht, die Metadaten ggf. in ein anderes Repository zu übergeben.

## 7. Interoperable – Interoperabel

Um Forschungsdaten effektiv in neue Auswertungs- und Nutzungskontexte integrieren zu können, sollten Daten so vorliegen, dass sie einfach ausgetauscht, interpretiert und in einer automatisierten oder semiautomatisierten Weise mit anderen Datensätzen kombiniert werden können – sowohl von Menschen als auch von Computersystemen.

Machen Sie Ihre Daten interoperabel, indem Sie möglichst anerkannte und vorzugsweise offene Formate und Software verwenden. Arbeiten Sie nach den einschlägigen Standards. Das sind von der Community anerkannte Metadatenschemata, kontrollierte Vokabulare, Thesauri und Ontologien.

Interoperabilität wird heute über semantische Technologien hergestellt, die auf Linked (Open) Data basieren.

### I1. (Meta-)Daten nutzen eine formale, zugängliche, gemeinsam genutzte und breit anwendbare Sprache für die Wissensrepräsentation

#### Zukunftssichere Dateiformate festlegen und verwenden

Alle in einem Repositorium gespeicherten Daten sollten in offenen, international anerkannten, standardisierten Dateiformaten vorliegen, um eine langfristige Interoperabilität in Bezug auf Nutzbarkeit, Zugänglichkeit und Nachhaltigkeit zu gewährleisten.

Hier finden Sie [Empfehlungen für Dateiformate](#) im Bereich der Kulturwissenschaften und des Kulturerbes.

#### **Die Rolle der Datenproduzierenden**

Bestimmen Sie schon zu Beginn Ihres Forschungsprojekts, welche zukunftssicheren Dateiformate für Ihr Projekt in Frage kommen. Verwenden Sie bevorzugte Formate, die von Ihrem Repositorium oder Ihrer Datenplattform empfohlen werden und die unabhängig von spezifischer Software, Entwicklern oder Anbietern genutzt werden können. Wenn Sie nicht von Anfang an mit diesen Datenformaten arbeiten, stellen Sie

sicher, dass eine inhaltlich verlustfreie Konvertierung der Daten in empfohlene Formate erfolgen kann. Publizieren Sie die (Meta-)Daten ggf. in mehreren Formaten.

### **Die Rolle des Betreibenden der Datenplattform**

Fördern Sie die Verwendung von Formaten, die für die langfristige Aufbewahrung geeignet sind. Stellen Sie einen leicht verständlichen und detaillierten Überblick über die akzeptierten bzw. empfohlenen Dateiformate bereit. Wenn Sie die Konvertierung von Dateiformaten vornehmen, stellen Sie dies als Teil Ihres Dienstleistungsangebots dar. Falls funktionale Langzeitarchivierung nicht durch Ihr Dienstleistungsspektrum abgedeckt ist, sollten Sie dies ebenfalls klarstellen (z. B. im Fall ausschließlicher Bitstream Preservation).

Beispiel: Das Rijksmuseum Amsterdam<sup>65</sup> publiziert seine Metadaten auf einer OAI-PMH-Schnittstelle in den Formaten LIDO, EDM und Dublin Core.

### **Interoperable Datenstandards festlegen und verwenden**

Die Verwendung eines Datenstandards, der von einer großen Community unterstützt wird, erhöht die Möglichkeit der gemeinsamen Nutzung, Wiederverwendung und Kombination von Datensammlungen. Nutzen Sie nach Möglichkeit die in Ihrer Community verwendeten oder übergreifende Standards. In diesem Abschnitt der Handreichung wird die Bedeutung von Datenstandards ausführlicher erläutert.

Sowohl die Syntax als auch die Semantik von Datenmodellen und -formaten, die für (Meta-)Daten in digitalen Objekten verwendet werden, sollten für Maschinen leicht zu identifizieren, zu analysieren, weiterzuverarbeiten oder zu übersetzen sein. Wie im Fall von Identifizierungsschemata und kontrollierten Vokabularen kann eine Vielzahl von Datenformaten prinzipiell FAIR sein.

Es liegt auf der Hand, dass jedes softwarebasierte Synchronisationsverfahren zur Analyse und Konvertierung von Daten fehleranfällig ist, und es wäre ideal, die Veröffentlichung von FAIRen Daten auf bestimmte von der Community akzeptierte Formate und Standards zu beschränken. Wenn ein Forscher jedoch nachweisen kann, dass ein alternatives Datenmodell oder -format eindeutig wie eines der

---

<sup>65</sup> <https://data.rijksmuseum.nl/object-metadata/harvest/>

akzeptierten FAIRen Formate verarbeitet werden kann, gibt es keinen Grund, ein solches Format nicht ebenfalls als FAIR anzusehen.

Hier finden Sie Empfehlungen für Rahmenwerke und Referenzmodelle und für Metadatenstrukturen (Elementsets, Schemata) im Bereich der Kulturwissenschaften und des Kulturerbes.

### **Die Rolle der Datenproduzierenden**

Erkundigen Sie sich frühzeitig bei dem Repository, in dem Sie Ihre Daten hinterlegen wollen, welche Datenstandards dort unterstützt werden. Prüfen Sie, ob diese Ihrem Forschungsgegenstand angemessen sind. Strukturieren Sie Ihre Daten von Beginn Ihres Forschungsprojekts entlang dieses Formats.

Wenn Ihr Forschungsprojekt besondere Herausforderungen mit sich bringt, die mit den vorhandenen Datenstandards nicht angemessen abgebildet werden können, versuchen Sie, eigene Erweiterungen auf ein Minimum zu begrenzen. Suchen Sie Beratung bei Fachleuten für den in Frage kommenden Standard. Sie können Ihnen oft angemessene Wege zur Modellierung der Daten im Rahmen bereits vorhandener Möglichkeiten aufzeigen. Wenn Sie dennoch lokale Erweiterungen vornehmen möchten, dokumentieren Sie diese so, dass sie ebenfalls FAIR nachnutzbar sind und erkundigen Sie sich beim Repository Ihrer Wahl, ob und in welcher Form diese Erweiterungen aufgenommen und verarbeitet werden können.

### **Die Rolle des Betreibenden der Datenplattform**

Stellen Sie, so weit wie möglich, maschinenlesbare Daten und Metadaten mit einem gut etablierten Rahmenwerk bereit. Legen Sie klar fest, welche Datenstandards Ihre Einrichtung unterstützt und publizieren Sie Hinweise, wie die Daten für eine bestmögliche Nachnutzung ausgestattet sein sollten, z. B. durch Belegung bestimmter Kernfelder.

Unterstützen Sie Fachinitiativen rund um die Pflege, den Ausbau und die Vermittlung dieser Standards, insbesondere mit Blick auf die Interoperabilität.

Beispiel: Hinweise der Deutschen Digitalen Bibliothek zu den Kernelementen der Lieferformate in den Teilnahmekriterien, Abschnitt Metadaten<sup>66</sup>

## Metadatenschemata dokumentieren

Die klare Dokumentation von Metadatenschemata hilft Entwickler:innen, Metadaten zu vergleichen und einander in einem Mapping zuzuordnen. In diesem Vorgang werden Datenelemente aus zwei verschiedenen Datenmodellen zum Zweck der Informationsintegration aufeinander abgebildet.

### **Die Rolle der Datenproduzierenden**

Nutzen Sie nach Möglichkeit existierende Metadatenschemata, die in Ihrer Fachcommunity häufig Verwendung finden oder die Ihr Repository vorgibt. Referenzieren Sie eindeutig, möglichst mit einem PID, auf die verwendeten Modelle und ihre Dokumentation, unter Angabe der verwendeten Version.

Lokale Erweiterungen oder Eigenentwicklungen sollten vermieden werden. Werden sie dennoch erstellt, sollten sie sorgfältig dokumentiert, mit freien Lizenzen versehen und mit den übrigen Forschungsdaten publiziert werden.

### **Die Rolle des Betreibenden der Datenplattform**

Veröffentlichen Sie die Metadatenschemata, die von Ihrer Forschungsinfrastruktur unterstützt werden. Dokumentieren Sie technische Spezifikationen und definieren Sie Klassen und Eigenschaften. Führen Sie auf, welche Komponenten obligatorisch und welche optional sind bzw. empfohlen werden.

Gewährleisten Sie, dass alle verwendeten Datenelemente definiert und eindeutig sind, um Fehlinterpretationen zu verhindern. Ein Element mit dem Label „Ort“ kann z. B. ein Entstehungsort, aktueller oder früherer Aufbewahrungsort oder Geburtsort einer Person sein.

Falls die Metadaten ursprünglich in einem anderen Datenschema erstellt wurden, erleichtern diese Informationen ein Mapping seitens des Datenproduzierenden. Liegen bereits sorgfältig erstellte Mappings von Metadatenschemata vor, sollten diese

---

<sup>66</sup> <https://pro.deutsche-digitale-bibliothek.de/daten-liefern/teilnahmekriterien>

publiziert werden, um eine möglichst große Konvergenz weiterer Datenbestände zu gewährleisten, die entsprechend gemappt werden müssen.

## Prozesse zur Verbesserung der Metadatenqualität einrichten

Um die Qualität der (Meta-)Daten und damit die Interoperabilität zu verbessern, sollten (automatische) Prozesse zur Bereinigung, Generierung und Anreicherung von (Meta-)Daten eingeführt werden. Sie können sowohl die Datenqualität in einem aktuellen Projekt verbessern als auch dazu dienen, bereits vorhandene Daten deutlich besser nachnutzbar zu machen.

### **Die Rolle der Datenproduzierenden**

Führen Sie Verfahren ein, die das Risiko von Fehlern bei der Datenerfassung minimieren. Nutzen Sie Funktionen der Datenbanksoftware, die Sie beim Identifizieren der korrekten Begriffe und Entitäten aus einem kontrollierten Vokabular unterstützen und die deren PIDs automatisiert in Ihre Daten übertragen. Wählen Sie z. B. Bezeichnungen und PIDs über ein in die Software integriertes Modul aus, oder übertragen Sie ein Datum aus einem Kalender, anstatt es von Hand einzutragen. Nutzen Sie konfigurierbare Obligatorikprüfungen Ihrer Software zur Unterstützung einer Dateneingabe gemäß definierter Pflichtvorgaben, so dass Sie bei abweichenden Eingaben gewarnt werden oder nicht speichern können. Machen Sie nachgängige Konsistenz- und Vollständigkeitsprüfungen für Daten-Kernfelder, indem Sie Daten gegen ein passendes Metadatenschema validieren oder redaktionelle Softwaretools einsetzen.

Für die nachgängige Qualitätsverbesserung von Datenbeständen hat sich das Tool OpenRefine<sup>67</sup> bewährt, mit dem Daten geordnet, strukturiert und transformiert und auf Normdaten referenziert werden können. Die nachgängige Referenzierung auf Normdaten und Vokabulare unterstützen auch Cocoda<sup>68</sup> und das Europeana Archaeology Vocabulary Tool<sup>69</sup>.

Wenn Sie strukturierte Metadaten zu den Entitäten erstellen, die im Rahmen Ihrer Forschungsarbeit behandelt werden (z. B. Werke, Objekte, Personen, Orte, Ereignisse) oder wenn es um die Bestandserschließung von Sammlungen in Bibliotheken, Museen

---

<sup>67</sup> <https://openrefine.org/>

<sup>68</sup> <https://coli-conc.gbv.de/cocoda/>

<sup>69</sup> <https://pro.carare.eu/en/services-guidelines/europeana-archaeology/>

oder Archiven geht, muss auch geplant werden, mit welcher Detailliertheit die Objekte erschlossen werden sollen. Berücksichtigen Sie hier die Empfehlungen der jeweiligen Metadatenschemata, die in Ihrer Domäne verwendet werden. Elemente mit der Kennzeichnung „Kernfeld“, „obligatorisch“ oder „empfohlen wenn vorhanden“ sollten mit FAIRem kontrollierten Vokabular (vgl. L2) besetzt werden, wenn es inhaltlich geboten ist und eine Information zu dem jeweiligen Sachverhalt vorliegt. Auf diese Weise wird auch unabhängig von bestimmten Dokumentationsschwerpunkten gewährleistet, dass das Objekt unter den Kerninformationen aufgefunden werden kann, mit denen Nutzer:innen meist suchen. Dies stärkt die Konsistenz der Datenbestände und damit die Interoperabilität in übergeordneten Kontexten.

Beschreiben Sie auch den Sammlungsbestand, bzw. bestimmte Teilbestände als solche und erschließen Sie, unter welchen inhaltlichen Gesichtspunkten sie relevant sind. So können Sammlungsbestände auch in Suchsystemen aufgefunden werden, die auf allgemeinere Ebenen von Bestandsnachweisen zielen.

### **Die Rolle des Betreibenden der Datenplattform**

Minimieren Sie die Wahrscheinlichkeit von Fehlern bei der Metadateneingabe durch automatisierte Validierungsregeln oder Prüfung der Angaben auf Vollständigkeit.

Richten Sie feste und transparente Geschäftsgänge ein, in denen eine Qualitätskontrolle nach Möglichkeit bereits verankert ist. Dies ist z. B. durch ein Rollenkonzept möglich, das verschiedenen Nutzenden unterschiedliche Rechte zuweist, um interne Verantwortlichkeiten abzubilden und Kurationsworkflows zu unterstützen und Reviewoptionen durch Dritte (externe Gutachter) zulässt. Beide Optionen sind bei RADAR realisiert.

Bündeln Sie Bestrebungen zur Entwicklung von Workflows und Softwarelösungen für solche automatischen Prozesse, z. B. durch den Einsatz von Werkzeugen für maschinelles Lernen. Wenn es die Auffindbarkeit und Nachnutzbarkeit der Daten stark verbessert, investieren Sie in Tools zur Bereinigung von (Meta-)Daten und zur Konvertierung von Daten in standardisierte und interoperable Datenformate.

Falls die Produktion und Kuratierung der Daten ein kontinuierlicher Prozess auf Seiten des Datenproduzierenden ist und Daten periodisch in modifizierter Form neu in das

Repositoryum eingebracht werden, sollten Sie anbieten, die von der Datenplattform erzielten qualitativen Verbesserungen in die Originaldaten re-integrierbar zu machen.

#### Weitere Informationen zu OpenRefine

Die OpenRefine-Website bietet eine Vielzahl von Empfehlungen zur Anwendung des Tools<sup>70</sup> an, die von Nutzer:innen erstellt wurden. Darunter sind auch deutschsprachige Tutorials<sup>71</sup> für Einsteiger:innen.

Über die OpenRefine Service API abgleichbare Datenquellen (darunter GND, VIAF, Getty Vokabulare, Wikidata, GeoNames) sind hier<sup>72</sup> aufgelistet.

#### Weitere Informationen zur Sammlungsbeschreibung

Die 2022 überarbeitete Version der DFG-Praxisregeln "Digitalisierung"<sup>73</sup> empfiehlt, die Sammlungsbeschreibung mit dem selben Metadatenschema anzulegen, das für die Bestandsobjekte genutzt wird (z. B. MODS, TEI-Header, EAD, LIDO) oder DCAT zu nutzen.

Weitere Beispiele: DARIAH Collection Registry<sup>74</sup>; Teilbestandsverzeichnung der Library of Congress mit EAD<sup>75</sup>

---

<sup>70</sup> <https://github.com/OpenRefine/OpenRefine/wiki/External-Resources>

<sup>71</sup> <https://github.com/OpenRefine/OpenRefine/wiki/External-Resources#german>

<sup>72</sup> <https://github.com/OpenRefine/OpenRefine/wiki/Reconcilable-Data-Sources>

<sup>73</sup> <https://doi.org/10.5281/zenodo.7435724>

<sup>74</sup> <https://colreg.de.dariah.eu/colreg-ui/collections/>

<sup>75</sup> <https://findingaids.loc.gov/browse/collections/>

## 12. (Meta-)Daten enthalten Vokabulare, welche den FAIR-Prinzipien folgen

### Offene, gut definierte Vokabulare verwenden

Grundlegend für Wissensorganisationssysteme sind Begriffe. Sie sind als „Einheiten des Denkens“<sup>76</sup> aufzufassen – Ideen, Bedeutungen oder (Kategorien von) Objekten und Ereignissen. Als solche existieren Begriffe im Bewusstsein als abstrakte Einheiten, die unabhängig von den Bezeichnungen sind, mit denen sie benannt werden.

Kontrollierte Vokabulare sind Systeme zur Organisation von Wissen, die eine strukturierte Menge von Begriffen zur Organisation und Klassifizierung von Daten enthalten, um den späteren Zugriff auf sie und die Suche nach ihnen zu gewährleisten. Bei den Begriffen handelt es sich um Datendeskriptoren, die über explizite Beziehungen (hierarchisch oder assoziativ) miteinander verbunden sind. Diese Deskriptoren werden zur Unterscheidung und Definition der Merkmale von Wissensressourcen in einem bestimmten Bereich verwendet. Sie enthalten die Datenwerte für Allgemeinbegriffe, Individualnamen und andere Werte, die zur strukturierten Beschreibung von Daten notwendig sind.

Mit Hilfe von kontrollierten Vokabularen können die Ressourcen befragt, durchsucht, analysiert und mit anderen relevanten Informationsobjekten verknüpft werden.

Die gebräuchlichsten Arten kontrollierter Vokabulare sind:

- Thesaurus – eine geordnete Zusammenstellung von Begriffen und ihren (vorwiegend natürlichsprachigen) Bezeichnungen, die in einem Dokumentationsgebiet zum Erschließen, Speichern und Wiederauffinden eines Sachverhalts dient
- Klassifikation – ein System, das in erster Linie auf der Klassifizierung von Dingen oder Konzepten in Gruppen oder Klassen basiert, mit einer detaillierten Erläuterung dieser Klassifizierungsmethoden
- Anwendungsontologie – ein Klassifikationssystem für Daten, um diese maschineninterpretierbar einzuordnen und zu repräsentieren. Eine formal

---

<sup>76</sup> <https://www.w3.org/TR/skos-primer/#secconcept>

geordnete Darstellung einer Menge von Begriffen und der zwischen ihnen bestehenden Beziehungen in einem bestimmten Gegenstandsbereich, die von einem Referenzmodell abgeleitet sein kann.

- Schlagwortliste – eine Liste von Begriffen zur Beschreibung von Themen in einem Informationssystem
- Taxonomie – ein System, das Dinge und Begriffe auf der Grundlage ihrer gemeinsamen Merkmale und/oder Unterschiede in Gruppen organisiert
- Wortliste – eine Liste von Bezeichnungen, die zur Beschreibung von Begriffen in einem bestimmten Bereich verwendet werden
- Glossar – eine alphabetische Liste von Begriffen mit ihren Erklärungen, die in einem bestimmten Kontext verwendet werden
- Normdaten – Datenbestand, der einzelne Entitäten (Instanzen bestimmter Klassen), wie z. B. Personen, Körperschaften, Geografika und Werke soweit beschreibt, dass sie eindeutig identifizierbar und referenzierbar sind

Die ausgewählten Vokabulare sollen PIDs für ihre Entitäten aufweisen und zugänglich, interoperabel und sorgfältig dokumentiert, daher selbst FAIR sein. Verwenden Sie offene lizenzierte, gut ausgearbeitete und publizierte Vokabulare, die in der Fachcommunity allgemein anerkannt sind. Die Einbeziehung oder Hinterlegung solcher Vokabulare sichert die genauen Bedeutungen der Konzepte und Eigenschaften ab, die in den Daten vertreten sind.

Hier finden Sie [Empfehlungen für Vokabulare, Normdaten und Anwendungsentologien](#) im Bereich der Kulturwissenschaften und des Kulturerbes.

### **Die Rolle der Datenproduzierenden**

Verwenden Sie Vokabulare, die für Ihr Fachgebiet relevant sind, und ergänzen und strukturieren Sie Ihre Forschungsergebnisse von Beginn Ihres Projekts an entsprechend. Integrieren Sie neben den sprachlichen Bezeichnungen immer auch die PIDs der Begriffe in Ihre Daten, um Eindeutigkeit auch bei einer maschinellen Auswertung der Daten zu gewährleisten und die Nutzung im Kontext von Linked-Data-Anwendungen zu ermöglichen.

Nicht selten ist die Verwendung von in der Community verbreiteten und veröffentlichten Vokabularen nur teilweise möglich, weil Begriffe benötigt werden, die

noch nicht in einem der oder dem Forschenden bekannten publizierten Vokabular oder einer Ontologie abgedeckt sind. Beispiele sind unzureichende Angebote für Objekte oder Bildinhalte nicht-europäischer Kulturen oder das Fehlen von spezifischen Fachbegriffen, z. B. für historische Glasverarbeitungstechniken im Art and Architecture Thesaurus. Die GND und die Getty-Vokabulare bieten ihren Anwendergemeinschaften Möglichkeiten zur Ergänzung des Vokabulars gemäß ihrer redaktionellen Regeln an.

Man kann auch ein projektspezifisches Vokabular erstellen und es offen lizenziert veröffentlichen, vorzugsweise als Linked Open Data in maschinenauswertbarer Form. Nutzen Sie hierfür spezielle Werkzeuge zur Thesauruserstellung und zur Publikation, z. B. [VocBench](#)<sup>77</sup> (open source), [ACDH Vocabs Editor](#)<sup>78</sup> (MIT License), [xTree](#)<sup>79</sup> oder Vokabularmodule von Sammlungsmanagement-Software, oder auch den Editor [Protégé](#)<sup>80</sup> zur Modellierung von Ontologien. Die Publikation des Vokabulars kann z. B. mit [Skosmos](#)<sup>81</sup> erfolgen, ein offen lizenzierter webbasierter SKOS-Browser.

Das selbsterstellte Vokabular sollte sich in der Strukturierung so weit wie möglich an einem publizierten Vokabular orientieren und als dessen lokale Erweiterung konzipiert werden.

Das Vokabular oder die Ontologie, die für ein bestimmtes Datenfeld gilt, sollte eindeutig spezifiziert werden. Auch für Felder, auf die das nicht zutrifft, sollte der Wertetyp des Feldes mit Hilfe eines öffentlich zugänglichen Vokabulars oder einer Ontologie in den Metadaten des digitalen Objekts eindeutig angegeben werden.

### **Die Rolle des Betreibenden der Datenplattform**

Nennen Sie Beispiele für Vokabulare, die die von Ihnen vertretenen Fachcommunities verwenden können und die über die Schnittstellen der Plattform adressierbar sind.

Ermöglichen Sie, wo immer möglich, die Verwendung weit verbreiteter Normdaten bzw. Identifikationssysteme, z. B. Normdaten für Personen ([ORCID](#)<sup>82</sup>), Institutionen

---

<sup>77</sup>

<https://joinup.ec.europa.eu/collection/semantic-interoperability-community-semic/solution/vocbench/about>

<sup>78</sup> <https://github.com/acdh-oeaw/vocabeditor>

<sup>79</sup> <https://www.digicult-verbund.de/de/digicultxtree>

<sup>80</sup> <https://protege.stanford.edu/>

<sup>81</sup> <https://skosmos.org/>

<sup>82</sup> <https://orcid.org/>

([ROR](#)<sup>83</sup>), Förderorganisationen ([Crossref Funder Registry](#)<sup>84</sup>), [DFG-Fachsystematik](#)<sup>85</sup>, [GND](#)<sup>86</sup> etc.

Stellen Sie sicher, dass in den Metadaten die entsprechenden Attribute gespeichert werden, so dass Eindeutigkeit und maschinelle Auswertbarkeit auch hier gewährleistet sind.

Weitere Informationen zu FAIRen Vokabularen

[Hugo, Wim / Le Franc, Yann / Coen, Gerard / Parland-von Essen, Jessica / Bonino, Luiz: D2.5 FAIR Semantics Recommendations Second Iteration \(1.0\), 2020](#)<sup>87</sup>

Weitere Informationen zur Mitarbeit an Referenzvokabularen

[Gemeinsame Normdatei](#)<sup>88</sup>

[Getty-Vokabulare](#)<sup>89</sup>

Beispiel: Erstellung eines projektspezifischen Vokabulars mit LOD-Publikation

Projekt "Digitalisierung Gandharischer Artefakte (DiGA)": [Elwert, Frederik / Pons, Jessie: Brücken bauen für Buddha - Das Projekt "Digitalisierung Gandharischer Artefakte" \(DiGA\) und die Pelagios Working Group "Linked Data Methodologies in Gandharan Buddhist Art and Texts"](#), in: DHd 2022 Kulturen des digitalen

---

<sup>83</sup> <https://ror.org/>

<sup>84</sup> <https://www.crossref.org/services/funder-registry/>

<sup>85</sup> [https://www.dfg.de/dfg\\_profil/gremien/fachkollegien/faecher/](https://www.dfg.de/dfg_profil/gremien/fachkollegien/faecher/)

<sup>86</sup> <https://www.dnb.de/gnd>

<sup>87</sup> <https://doi.org/10.5281/zenodo.5362010>

<sup>88</sup> [https://gnd.network/Webs/gnd/DE/Mitmachen/mitmachen\\_node.html](https://gnd.network/Webs/gnd/DE/Mitmachen/mitmachen_node.html)

<sup>89</sup> <https://www.getty.edu/research/tools/vocabularies/contribute.html>

Gedächtnisses. 8. Tagung des Verbands "Digital Humanities im deutschsprachigen Raum" (DHd 2022), Potsdam, 7. März 2022<sup>90</sup>

Amato, Antonio / Elwert, Frederik / Pons, Jessie: Digitization of Gandharan Artefacts: A Project for the Preservation and the Study of the Buddhist Art of Pakistan. A Digitization Concept, 2022<sup>91</sup>

DiGA Thesaurus auf Github und bei Skosmos<sup>92</sup>

### 13. (Meta-)Daten enthalten qualifizierte Verweise auf andere (Meta-)Daten

Wenn ein Datensatz weitere Datensätze benötigt, um vollständige Informationen bereitzustellen, muss dies in den Daten angegeben werden. Die Referenzierung geschieht am besten über die persistenten Identifikatoren der Datensätze, ihre Verknüpfung wird über kontrolliertes, maschinenauswertbares Vokabular ausgewiesen.

#### **Die Rolle der Datenproduzierenden**

Planen Sie von Beginn Ihres Projektes an die Struktur der Datenveröffentlichung und sehen Sie die Vergabe bzw. Referenzierung von PIDs für alle relevanten Komponenten vor. Externe Ressourcen können ebenso über PIDs einbezogen werden.

Qualifizieren Sie die Referenzen nach Möglichkeit, indem Sie den Typ der Relation mit einem kontrollierten Vokabular angeben. Wählen Sie die richtige Bezeichnung aus der Perspektive der Ressource, die Sie gerade beschreiben.

Wenn Sie Sammlungsobjekte beschreiben, nutzen Sie Normdaten und kontrollierte Vokabulare, um Personen, Körperschaften, Geografika, beschreibende Sachschlagworte und in Beziehung stehende weitere Objekte (z. B. vorbereitende oder abgeleitete Werke), zu referenzieren. Die PIDs dieser Entitäten können zur

<sup>90</sup> <https://doi.org/10.5281/zenodo.6327971>

<sup>91</sup> <https://doi.org/10.46586/rub.204>

<sup>92</sup> <https://github.com/DiGArtefacts/thesaurus/>

automatisierten Ergänzung von Informationen dienen, die wertvolle weitere Such- und Filtermöglichkeiten ermöglichen. Bei Personen sind dies z. B. Lebensdaten und Namensvarianten, bei Sachschlagworten sind es Oberbegriffe, Synonyme, Bezeichnungen in mehreren Sprachen. Auch hier sollten Beziehungen zwischen Entitäten qualifiziert und mit kontrolliertem Vokabular ausgewiesen werden. In Frage kommen dafür die CIDOC-CRM-Properties<sup>93</sup>, die MARC Relators<sup>94</sup> oder Teile der LIDO-Terminologie<sup>95</sup>.

### **Die Rolle des Betreibenden der Datenplattform**

Ermöglichen Sie die qualifizierte Verlinkung von Metadaten und Daten im Repositorium. Ermöglichen Sie die umfassende und komfortable Anlage dieser Beziehungen über die zur Verfügung gestellten Formulare für Forschungsmetadaten. Stellen Sie diese Verweise und die zugehörigen PIDs so bereit, dass sie von Menschen und Maschinen einfach gefunden und nachgenutzt werden können.

Beispiel für die Nutzung kontrollierten Vokabulars für Bezüge zwischen Ressourcen bei DataCite: Bayer, Christiane / Frech, Andreas / Gabriel, Vanessa et. al.: DataCite Best Practice Guide (Version 2.0), 2022<sup>96</sup>, S. 17 ff.

Beispiel für eine durch Verweise verknüpfte komplexe Datenpublikation: Bender, K.: Thematic Research Collection of the Iconography of Venus from the Middle Ages to Modern Times, heiDATA, 2018<sup>97</sup>, bestehend aus acht Datensets mit 29 Dateien

---

<sup>93</sup> [https://www.cidoc-crm.org/html/cidoc\\_crm\\_v7.1.2.html](https://www.cidoc-crm.org/html/cidoc_crm_v7.1.2.html)

<sup>94</sup> <https://id.loc.gov/vocabulary/relators.html>

<sup>95</sup> <https://terminology.lido-schema.org/>

<sup>96</sup> <https://doi.org/10.5281/zenodo.7040047>

<sup>97</sup> <https://heidata.uni-heidelberg.de/dataverse/icovenus>

## 8. Reusable – Nachnutzbar

Ein wichtiges Ziel der FAIR-Prinzipien ist es, die Verlässlichkeit und Nachvollziehbarkeit der veröffentlichten Daten zu sichern.

Forschungsdaten sollen für künftige Forschung und Weiterverarbeitung so aufbereitet werden, dass es selbstverständlich ist, dass die Ergebnisse repliziert oder bestätigt werden können. So können neue Forschungen und Nutzungsszenarien effizient auf bereits gewonnenen, früheren Ergebnissen aufbauen.

Stellen Sie sicher, dass Ihre Daten gut genug dokumentiert sind, um Ihre forschungsbegleitende Qualitätssicherung darzustellen. Metadaten enthalten daher Herkunftsinformationen, die deutlich machen, wie, warum und von wem die Daten erstellt und verarbeitet wurden. Vergeben Sie klare und maschinell auswertbare Datennutzungslizenzen, damit andere wissen, welche Arten der Wiederverwendung erlaubt sind. Für die sachgerechte Nachnutzbarkeit ist es essentiell, dass Daten und Metadaten den einschlägigen Fachstandards entsprechen.

### R1. (Meta-)Daten sind detailliert beschrieben und enthalten präzise, relevante Attribute

Die Metadaten eines digitalen Objekts sollten so reichhaltig sein, dass eine Maschine oder ein:e menschliche:r Nutzer:in bei seiner Auffindung eine fundierte Entscheidung treffen kann, ob dieses digitale Objekt im Rahmen des angefragten Kontexts verwendet werden soll oder nicht. Die enthaltenen Metadaten sollten den oder die Anwender:in über die Lizenz und die daraus abzuleitenden Nutzungsbedingungen der Bitsequenz oder der Metadaten selbst informieren. Diese Metadaten sollten maschinenlesbar sein, um das automatisierte Daten-Harvesting zu erleichtern und gleichzeitig die korrekte Zuordnung von Herkunftsinformationen zu gewährleisten, um ein ordnungsgemäßes Zitieren zu ermöglichen. Die Metadaten sollten über jegliche Form der Zugriffskontrolle und gegebenenfalls über das für den Zugriff notwendige Authentifizierungsprotokoll informieren, so dass die Anwender:innen ermitteln können, auf welche Komponenten der Daten sie zugreifen dürfen.

## R1.1. (Meta-)Daten enthalten eine eindeutige, zugreifbare Angabe einer Nutzungslizenz

### Lizenzieren Sie Daten für die Nachnutzung

Um eine möglichst umfassende Weiterverwendung von Daten und Metadaten zu ermöglichen, sollte klar sein, wer Inhaber:in der Rechte an den jeweiligen Daten ist und welche Lizenz gilt.

Soweit möglich, sollen offene Lizenzmodelle wie das der Creative Commons<sup>98</sup> verwendet werden. Ist das nicht möglich, können auch die ebenfalls maschinenauswertbaren Rechteinweise von Rightsstatements.org<sup>99</sup> genutzt werden. Sie wurden speziell für Kulturerbe-Einrichtungen und -Plattformbetreiber entwickelt und enthalten auch maschinenauswertbare Rechteinweise für Objekte mit Urheberrechtsschutz und unklarem Urheberrechtsstatus.

Berücksichtigen Sie, dass die Urheberrechtsreform von 2021 Bibliotheken, Archiven und Museen neue rechtliche Möglichkeiten in der Zugänglichmachung ihrer Materialien eröffnet. Mehr Informationen finden Sie in der Handreichung Urheberrechtsreform 2021 – Neue Chancen für das kulturelle Erbe<sup>100</sup> von Paul Klimpel.

### **Die Rolle der Datenproduzierenden**

Stellen Sie sicher, dass Sie vor der Veröffentlichung Ihrer Forschungsdaten wissen, wer der oder die Inhaber:in der Rechte an Daten und Metadaten ist und welche rechtlichen Nachnutzungsbedingungen gelten. Möglicherweise müssen Urheber-, Persönlichkeits-, Leistungsschutz- und Archivrechte berücksichtigt werden, die für verschiedene Datentypen in unterschiedlicher Weise relevant sind.

In Digital Humanities-Projekten wird sehr häufig mit dem Datenmaterial Dritter (z. B. Text-, Bilddateien) gearbeitet. Beziehen Sie die Thematik der Rechtklärung daher von Anfang an in die Planung Ihres Projekts ein. Planen Sie genügend Ressourcen für das Feststellen der Rechteinhaber:innen und das Einholen von Nutzungsgenehmigungen

---

<sup>98</sup> <https://creativecommons.org/about/cclicenses/>

<sup>99</sup> <https://rightsstatements.org/de/>

<sup>100</sup>

<https://irights.info/artikel/neue-handreichung-urheberrechtsreform-2021-neue-chancen-fuer-das-kulturelle-erbe/31174>

ein. Eine Rechteklärung kann auf unterschiedlichen Ebenen notwendig sein, z. B. in Bezug auf die Künstler:innen-Urheberrechte eines abgebildeten Werkes, auf die Abbildung selbst, in Bezug auf verwendete Texte Dritter und in Bezug auf die Daten und Metadaten.

Sorgen Sie dafür, dass die Lizenz bzw. der Rechteinweis mit einem passenden URI in den Daten und Metadaten versehen werden kann, um die Angabe maschinenauswertbar zu machen.

Vermeiden Sie es, Daten ohne Lizenzangaben zu veröffentlichen, auch wenn die Nachnutzung Ihrer Meinung nach unproblematisch ist. Wegen der rechtlichen Risiken und der erhöhten Aufwände zur Feststellung der Nutzungsbedingungen seitens des Nachnutzenden wird die Verwendung nicht lizenzierter Daten häufig vermieden. Auch wenn bestimmte Nutzungsarten ausgeschlossen bleiben müssen, sollten diese bzw. die zulässigen Arten klar benannt werden, statt lediglich eine Angabe „Rechte vorbehalten“ zu machen.

### **Die Rolle des Betreibenden der Datenplattform**

Vermitteln Sie transparent, unter welcher Lizenz Daten und Metadaten stehen und unter welchen Bedingungen sie nachgenutzt werden dürfen. Stellen Sie diese Angaben auch in einem maschinenauswertbaren Format zur Verfügung. Um die Interoperabilität zu verbessern, versuchen Sie, Ihre Lizenzen auf bereits weit verbreitete Rahmenwerke wie Creative Commons abzubilden. Ermöglichen Sie die Angabe der Rechteinhaberin oder des -inhabers einschließlich PID. Setzen Sie ein effektives Rechtemanagement ein, um die Nutzung der Daten zu kontrollieren bzw. deren Verwendung nur in dem von den jeweiligen Rechteinhaber:innen gesetzten Rahmen zu ermöglichen. Der Zugriff auf den Datensatz könnte daran geknüpft sein, dass Interessent:innen die jeweiligen Nachnutzungsbedingungen online akzeptieren und bestätigen.

Weitere Materialien zu Rechtklärung und Lizenzierung

E. Böker: Rechtssichere Nachnutzung von Forschungsdaten(-banken),  
Forschungsdaten.info<sup>101</sup>

iRights.info – Urheberrecht und kreatives Schaffen in der digitalen Welt<sup>102</sup>

Creative Commons FAQ<sup>103</sup>

Muß-Merholz, Jöran: Die CC-Lizenzen im Überblick – Welche Lizenz für welche  
Zwecke?, wb-web, 2021<sup>104</sup>

Klimpel, Paul: Freies Wissen dank Creative-Commons-Lizenzen: Folgen, Risiken und  
Nebenwirkungen der Bedingung „nicht-kommerziell – NC“, Hg. iRights.info, Creative  
Commons Deutschland, Wikimedia Deutschland, 2012<sup>105</sup>

Fischer, Veronika / Petri, Grischka: Bildrechte in der kunsthistorischen Praxis – ein  
Leitfaden, 2. Aufl., 2022<sup>106</sup>

Klimpel, Paul: In Bewegung. Die Rechtsfibel für Digitalisierungsprojekte in  
Kulturerbe-Einrichtungen, Hg. Digitales Deutsches Frauenarchiv (i.d.a.-Dachverband  
e.V.), digiS, Forschungs- und Kompetenzzentrum Digitalisierung Berlin, 2022<sup>107</sup>

Altenhöner, Reinhard / Berger, Andreas / Bracht, Christian et al.: DFG-Praxisregeln  
„Digitalisierung“ – Aktualisierte Fassung, 2022<sup>108</sup>, Abschnitt 6

---

<sup>101</sup>

<https://forschungsdaten.info/themen/rechte-und-pflichten/rechtssichere-nachnutzung-von-forschungsdaten-banken/>

<sup>102</sup> <https://irights.info/>

<sup>103</sup> <https://de.creativecommons.net/faqs/>

<sup>104</sup>

<https://wb-web.de/material/medien/die-cc-lizenzen-im-ueberblick-welche-lizenz-fur-welche-zwecke-1.html>

<sup>105</sup> [https://irights.info/wp-content/uploads/userfiles/CC-NC\\_Leitfaden\\_web.pdf](https://irights.info/wp-content/uploads/userfiles/CC-NC_Leitfaden_web.pdf)

<sup>106</sup> <https://doi.org/10.11588/artdok.00007769>

<sup>107</sup> <https://nbn-resolving.org/urn:nbn:de:0297-zib-86485>

<sup>108</sup> <https://doi.org/10.5281/zenodo.7435724>

## R1.2. (Meta-)Daten enthalten detaillierte Provenienz-Informationen

„Provenienz“ bezieht sich hier nicht auf die Besitz- und Herkunftsgeschichte der in den Daten beschriebenen Objekte, sondern auf die Daten und Metadaten selbst. Provenienz-Informationen geben an, zu welchem Zweck und unter welchen Projektfragestellungen die Daten generiert wurden und woher das Datenmaterial stammt. Sie erklären außerdem, wie die Daten modifiziert wurden, in welchem Kontext sie wiederverwendet werden können und wie zuverlässig sie sind. Ergänzen Sie daher die Metadaten oder die publizierte Begleitdokumentation Ihres Projekts um folgende Aspekte:

- Benennung aller an der Bearbeitung der Daten mitwirkenden Personen mit ihren Rollen und Anteilen, soweit datenschutzkonform möglich
- Beschreibung der Motivation, die zur Erstellung der Daten geführt hat (Projektziele und Schwerpunkte)
- Beschreibung der Methoden, die angewendet wurden
- Beschreibung des Arbeitsablaufs, der zur Erzeugung der Daten geführt hat: Wurden die Daten maschinell generiert oder händisch-intellektuell erzeugt? Wie wurden sie weiterverarbeitet?
- Wenn Daten aus anderen Quellen verwendet wurden: Beschreibung der Herkunft, ihrer Verwendung und Modifikation
- Für die Datenerzeugung, -bearbeitung oder -betrachtung der Daten verwendete Software: Skripte, Programmcode, DTDs oder XML-Schema-Dateien unter Angabe der Version
- verwendete Datenmodelle, Formate, Ontologien, Schreib- und Redaktionsanweisungen in maschineninterpretierbarer Form, z. B. als XML, DTDs oder XML-Schema-Dateien unter Angabe der Version
- Erstellungs- und Modifikationszeitpunkte der Daten und Metadaten
- Quellenbelege für Aussagen und Informationen, die Sie übernommen haben

Bei dynamischen Daten ist erstrebenswert, frühere Versionen verfügbar und über einen PID adressierbar zu halten, um die Zitierfestigkeit der Datensatzinhalte zu gewährleisten. Publizieren Sie auch im Projektkontext erstellte Software unter einer möglichst offenen Lizenz, z. B. über Github.

Weitere Informationen zur Dokumentation von Daten

E. Böker: Warum dokumentieren?, in: Forschungsdaten.info, 14.11.2022<sup>109</sup>

## Benennungskonventionen einhalten

Die Einhaltung von präzisen und konsistenten Benennungskonventionen – allgemein anerkannter Muster für die Benennung von Daten – erleichtert künftigen Generationen von Forschenden das Auffinden, den Zugriff und das Verständnis von digitalen Objekten und Datensätzen erheblich. Hinweise zum Benennen und Organisieren von Dateien<sup>110</sup> gibt der Verbund Forschungsdaten Bildung.

### **Die Rolle der Datenproduzierenden**

Konsultieren Sie die Richtlinien der bewährten Verfahren für Ihre Forschungsdisziplin oder Ihren Fachbereich, um die am besten geeignete Namenskonvention zu finden. Beginnen Sie mit der Umsetzung bereits zu Beginn Ihres Projekts und halten Sie die Konventionen von Anfang an ein.

### **Die Rolle des Betreibenden der Datenplattform**

Nennen Sie die Best Practices zur Erstellung und Anwendung spezifischer Namenskonventionen.

## Daten systematisch dokumentieren

Um klarzustellen, was in einem Datenbestand erwartet werden kann und was nicht, sollten die Daten systematisch dokumentiert werden. Die diesbezügliche Transparenz fördert das Vertrauen und damit die Wiederverwendung der Daten.

### **Die Rolle der Datenproduzierenden**

Integrieren Sie detaillierte Provenienz-Informationen in Ihre Forschungsdaten und publizieren Sie diese mit.

---

<sup>109</sup>

<https://forschungsdaten.info/themen/beschreiben-und-dokumentieren/datendokumentation/>

<sup>110</sup> <https://www.forschungsdaten-bildung.de/dateien-benennen>

Bittet Ihr Repository Sie um die aussagekräftige Beschreibung Ihres Datensets mit Metadaten, füllen Sie das bereitgestellte Formular sorgfältig und ausführlich aus. Tun Sie dies auch für Abschnitte, die als nicht-obligatorisch und optional bezeichnet werden.

### **Die Rolle des Betreibenden der Datenplattform**

Stellen Sie Dokumentationsvorlagen zur Verfügung, die eine Beschreibung des Layouts, der Struktur und der Inhalte von Datensammlungen enthalten und die Erstellung der zu Beginn dieses Abschnitts genannten Begleitdokumentation erleichtert. Ebenso kann die Methodik, eine Liste der Abkürzungen, eine Beschreibung der Datenlücken, der Aufbau der Datenbank usw. bereitgestellt werden. Sie erleichtern Datengeber:innen eine systematische Dokumentation der Datenprovenienz.

### **Datenintegrität bewahren**

Erhobene Forschungsdaten sollten mit den später zugänglichen Forschungsdaten identisch sein. Um die Zuverlässigkeit der Daten zu gewährleisten, sollten Prüfungen zur Datenintegrität durchgeführt werden.

### **Die Rolle der Datenproduzierenden**

Implementieren Sie eine Methode zur Versionskontrolle. Hierbei können Sie häufig bestimmte Funktionen Ihrer Software nutzen. Die Garantie, dass jede Änderung in einer überarbeiteten Version eines Datensatzes korrekt dokumentiert wird, ist von wesentlicher Bedeutung für die Authentizität eines jeden Datensatzes.

### **Die Rolle des Betreibenden der Datenplattform**

Um festzustellen, ob eine Datei geändert wurde, ist es unerlässlich, im Rahmen einer Versionskontrolle die Provenienz aufzuzeichnen – die Herkunft der Daten und alle Änderungen, die im Laufe der Zeit vorgenommen wurden – und jede Kopie mit dem Original zu vergleichen. Eine Prüfung der Datenintegrität kann mit Hilfe eines Prüfzeichens, z. B. einer Prüfsumme, oder durch einen direkten Vergleich zweier Dateien erfolgen. Es sollte ein Mechanismus für den Umgang mit unterschiedlichen Versionen vorgesehen werden, z. B. durch Hinzufügen einer Versionskomponente zum Identifikator als Suchparameter.

Beispiel: Verfahren zur DOI-Versionierung bei [Zenodo](#)<sup>111</sup>, Abschnitt DOI versioning

## R1.3. (Meta-)Daten entsprechen den fachgebietsrelevanten Community-Standards

Es ist einfacher, Datensätze nachzunutzen, die ähnlich strukturiert sind: Daher sollten vergleichbare Daten auf standardisierte Art organisiert vorliegen, unter Nutzung etablierter und nachhaltiger Dateiformate, begleitet von Metadaten und einer Dokumentation auf einer einheitlichen Grundlage, unter Verwendung eines gängigen Vokabulars.

Wenn Community-Standards oder Best Practices für die Datenarchivierung und gemeinsame Nutzung von Daten existieren, sollten diese genutzt werden.

Möglicherweise werden in einer Community auch mehrere Standards nebeneinander angewendet. Weitere Hinweise finden Sie im Abschnitt [Die Rolle von Datenstandards](#).

Hier finden Sie Empfehlungen für [Dateiformate](#), für [Rahmenwerke und Referenzmodelle](#), für [Metadatenstrukturen](#) und für [Vokabulare, Normdaten und Anwendungsentologien](#) im Bereich der Kulturwissenschaften und des Kulturerbes.

### **Die Rolle der Datenproduzierenden**

Erstellen Sie Ihre Daten und Metadaten gemäß den Community-Standards sowie den bewährten Verfahren für die Archivierung und den Austausch von Daten in Ihrem Forschungsgebiet. Wählen Sie auch unter diesem Aspekt Referenzmodelle (Ontologien), Inhaltsstandards, Datenstrukturstandards (Metadatenschemata) und Vokabulare, die in Ihrer Community geläufig sind.

Kulturerbe-Institutionen orientieren sich bei der Auswahl an den Metadatenschemata, die für die jeweilige Sparte (Bibliotheken, Archive, Museen) empfohlen werden. Hilfreich zur Orientierung sind auch die [DFG-Praxisregeln "Digitalisierung" \(aktualisierte Version 2022\)](#)<sup>112</sup> sowie die Empfehlungen der für die Publikation in Frage kommenden großen Datenplattformen [Deutsche Digitale Bibliothek](#)<sup>113</sup> und

<sup>111</sup> <https://help.zenodo.org/>

<sup>112</sup> <https://doi.org/10.5281/zenodo.7435724>

<sup>113</sup> <https://pro.deutsche-digitale-bibliothek.de/daten-liefern/teilnahmekriterien>

Europeana<sup>114</sup>. Auch wenn Linked-Data-Komponenten wie z. B. Referenzierung kontrollierter Vokabulare in diesen Materialien lediglich als optional empfohlen werden, sollten Sie sie implementieren, um Ihre Daten zukunftssicherer zu machen.

Forscher:innen sind gut beraten, die Erstellung eines Datenmodells zu den Gegenständen ihres Projekts entlang des CIDOC-CRM<sup>115</sup> vorzunehmen. Die formale Semantik dieses ISO-Standards für die Datenmodellierung im Kulturerbe-Bereich ist die Grundlage für die Interoperabilität der Daten.

Es kann Situationen geben, in denen Best Practices für Ihre Forschungsdaten existieren, es aber triftige Gründe gibt, davon abzuweichen. Bauen Sie dennoch so weit wie möglich auf der Standardpraxis auf, um die Interoperabilität und Wiederverwendbarkeit in diesem Rahmen weitestgehend zu gewährleisten. Begründen und dokumentieren Sie in den Metadaten die Abweichungen umfassend und publizieren Sie diese Materialien, um dennoch eine maximale Nachnutzbarkeit der Daten zu ermöglichen.

### **Die Rolle des Betreibenden der Datenplattform**

Wenn Ihr Repository auf ein bestimmtes Forschungsfeld spezialisiert ist, kann es bestimmte Standards in Bezug auf die einzubringenden Daten und Metadaten implementieren. Überprüfen und überarbeiten Sie Ihr Metadatenschema regelmäßig, um zu gewährleisten, dass es dem aktuellen Entwicklungsstand der Standards entspricht.

---

<sup>114</sup> <https://pro.europeana.eu/page/edm-documentation>

<sup>115</sup> <https://www.cidoc-crm.org/Version/version-7.1.2>

## 9. FAIR-Prinzipien umsetzen – Hinweise zur Projektplanung

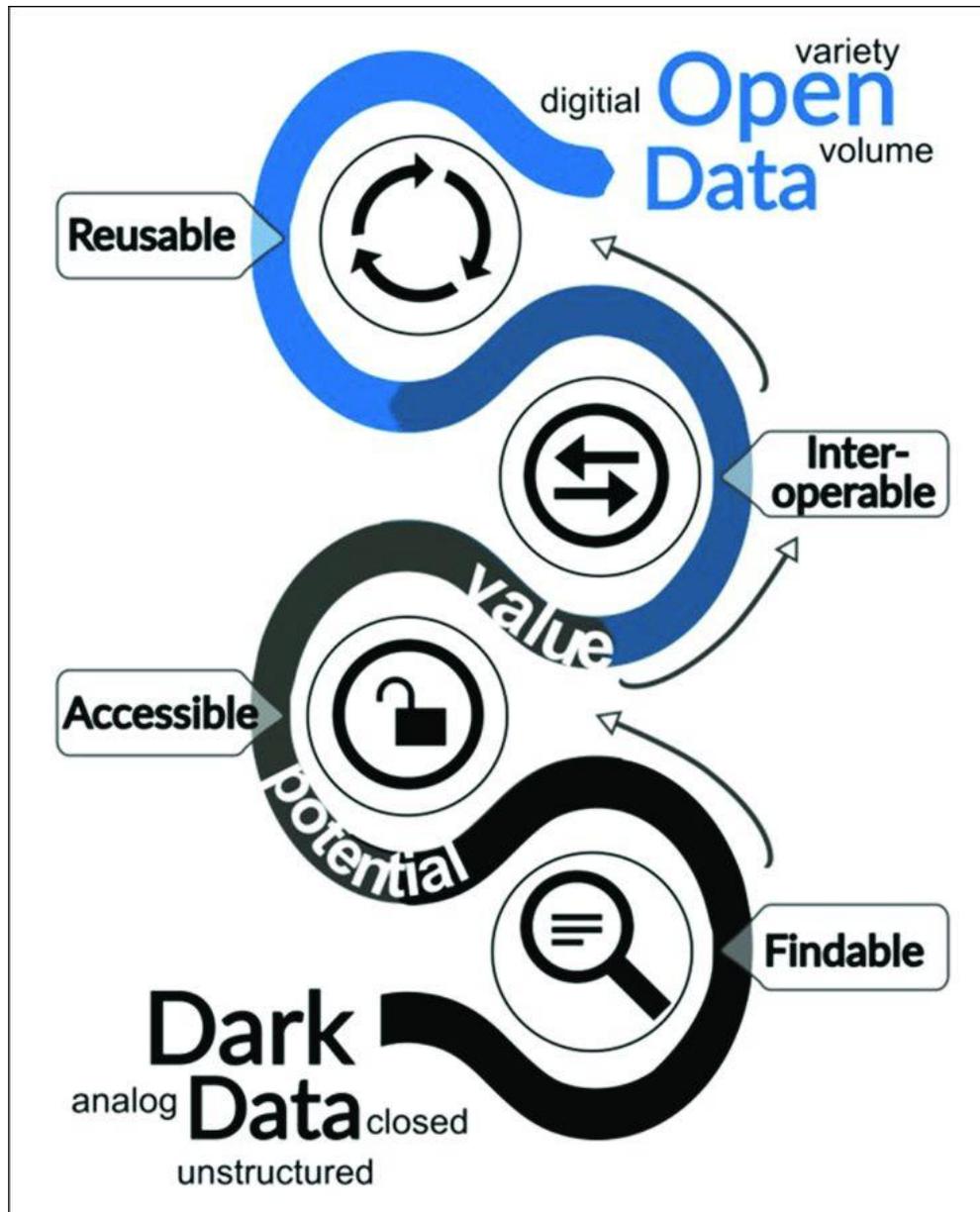


Abb. 3: Konzeptdiagramm des FAIR-Fahrplans für „dunkle“ (nicht-digitale, nicht genutzte, nicht ausgewertete) Daten: Jede Kurve steht für einen Schritt zur Steigerung des Werts und des Potenzials „dunkler“ Daten für die Wissenschaft. Easterday, K. / Paulson, T. / DasMohapatra, P. et al.: From the Field to the Cloud: A Review of Three Approaches to Sharing Historical Data From Field Stations Using Principles From Data Science, in: *Frontiers in Environmental Science*, 2018. CC-BY 4.0

Wissenschaftler:innen sehen sich immer häufiger der Anforderung gegenüber, die FAIR-Prinzipien in mit digitalen Materialien und Methoden durchgeführten Forschungsprojekten umzusetzen. Da dieses Ziel innerhalb einer begrenzten Projektlaufzeit für eine definierte Fragestellung erreicht werden muss, sollte die Umsetzung von FAIR von Anfang an integraler Bestandteil der Planung und Durchführung von Arbeitspaketen sein.

Komplexer ist diese Herausforderung für Museen, Archive und andere sammelnde Gedächtnisinstitutionen: Sie verfügen in der Regel über Bestandsdaten, die als dynamische Daten über einen langen Zeitraum ohne definierten Endpunkt kontinuierlich fortentwickelt werden. Solche Daten stellen besondere Anforderungen: Sie müssen aktuell, vollständig und korrekt gehalten werden, auch wenn die Kenntnisstände über die Objekte sich ändern und andere Fragen an sie gestellt werden. Oft entsteht der Wunsch, kritischer und differenzierter mit der Dokumentation selbst zu verfahren. Man stellt neue Anforderungen an die Verfügbarkeit früherer Kenntnisstände durch Versionierung der Daten, die umfassendere Einbringung von Quellenbelegen, den Umgang mit unsicherem Wissen. Die Ansprüche an den Leistungsumfang der Datenbasis wandeln sich: Die Bestandsdatenbank, die zuerst nur Inventarfunktion hatte, wird zum Sammlungsmanagement-System, das eine Vielzahl von Kommunikations- und Arbeitsprozessen der Institution und ihrer Partner unterstützen soll. Forschungsergebnisse, die im Rahmen der Provenienzforschung oder für ein Ausstellungsprojekt gewonnen wurden, müssen integriert werden. Immer wichtiger wird die Publikation der Bestandsdaten in der eigenen Online-Datenbank, die Weitergabe an Kulturerbe-Portale, die Bereitstellung von Daten auf Schnittstellen. Da den Daten ein höherer Wert beigemessen wird, sucht man neue professionelle Lösungen für die langfristige Datensicherung.

Die erforderlichen Weichenstellungen zu einer Neuausrichtung müssen häufig während des laufenden Betriebs unter gesetzten Bedingungen vorgenommen werden. Sie betreffen überwiegend die Daueraufgaben, und meist sind die Möglichkeiten begrenzt, dafür zusätzliche Ressourcen einzuwerben.

Möchte man eine nachhaltigere und breitere Nutzbarkeit der Daten im Sinne der FAIR-Prinzipien erreichen, sollte man bei der Überprüfung und Anpassung der institutionellen digitalen Strategie beginnen. Die Leitungsebene einer Institution sollte die Neuausrichtung mittragen, um dem Vorhaben den notwendigen Stellenwert einzuräumen und die notwendigen Personalressourcen und die finanziellen Mittel zu beschaffen und abzusichern. Soll es nicht bei Absichtserklärungen bleiben, ist ein Maßnahmenplan notwendig, der Ziele und Prioritäten im Verhältnis zu den verfügbaren Ressourcen benennt und der in einer konkreten Neuausrichtung der Aktivitäten in der Sammlungserschließung und im Datenmanagement mündet.

## 9.1 FAIRification benötigt Ressourcen und Expertise

Viel stärker als traditionelle Publikationen verlangen Bestandsdaten eine kontinuierliche Kuratierung, um über lange Zeit ein hohes Qualitätsniveau bei sich entwickelnden Anforderungen der Nutzungsszenarien halten zu können. Verankern Sie Datenqualitätsmanagement als definierte Aufgabe in Ihren Projektaktivitäten oder im Aufgabenprofil Ihrer Institution. Auch an Bibliotheken, Archiven und Museen ist die Kuratierung der Bestandsdaten eine wichtige Aufgabe, für die ausreichend Ressourcen bereitstehen müssen.

Legen Sie ein Zeit- und Kostenbudget für Investitionen in technische Infrastruktur und Personal fest.

Investieren Sie in Dateninfrastrukturen oder suchen Sie sich kompetente Partnerinstitutionen oder -netzwerke. Stellen Sie Datenfachleute ein, oder bilden Sie sich und Ihr Team weiter. Machen Sie sich mit digitalen Methoden und den Praktiken des Forschungsdatenmanagements vertraut.

Hier unterstützen Sie der [NFDI4Culture-Helpdesk](https://nfdi4culture.de/de/aufgaben/services.html)<sup>116</sup> oder die bei NFDI4Culture gelisteten [Repositorien](https://nfdi4culture.de/de/ressourcen/repositorien.html)<sup>117</sup>. Weitere Anlaufstellen: [Forschungsdienstleistungen der AG Datenzentren des Verbands Digital Humanities im deutschsprachigen Raum e. V.](https://dhd-ag-datenzentren-dienstekatalog.github.io/allgemeine-dienste/)<sup>118</sup>, die

---

<sup>116</sup> <https://nfdi4culture.de/de/aufgaben/services.html>

<sup>117</sup> <https://nfdi4culture.de/de/ressourcen/repositorien.html>

<sup>118</sup> <https://dhd-ag-datenzentren-dienstekatalog.github.io/allgemeine-dienste/>

Forschungsdatenmanagement-Landesinitiativen<sup>119</sup>, die Koordinierungsstelle für wissenschaftliche Universitäts-sammlungen in Deutschland<sup>120</sup>, digiS<sup>121</sup> für Kulturerbe-Einrichtungen in Berlin, digiCULT<sup>122</sup> für Museen, MusIS<sup>123</sup> für Museen in Baden-Württemberg und die Museums-<sup>124</sup> und Archivberatungen der Länder.

Weitere Materialien zu Projektplanung und Forschungsdatenmanagement

Minn, Gisela / Lemaire, Marina: Forschungsdatenmanagement in den Geisteswissenschaften. Eine Planungshilfe für die Erarbeitung eines digitalen Forschungskonzepts und die Erstellung eines Datenmanagementplans, 2017<sup>125</sup>

PARTHENOS Training, 2016/2019<sup>126</sup>

RWTH Aachen: Lehrvideos "Datenmanagement nach Plan" und "Inhalte eines Datenmanagementplans (DMP)", 2020<sup>127</sup>

Alge, Barbara: Forschungsdatenmanagement in der Musikethnologie, 2019<sup>128</sup>

Deutscher Museumsbund: Handreichung Digitale Grunderfassung. 10 Grundsätze, 2022<sup>129</sup>

Knaus, Gudrun: Leitfaden für digitales Sammlungsmanagement an Kunstmuseen, 2021<sup>130</sup>

---

<sup>119</sup> <https://forschungsdaten.info/fdm-im-deutschsprachigen-raum/deutschland/>

<sup>120</sup> <https://wissenschaftliche-sammlungen.de/de/>

<sup>121</sup> <https://www.digis-berlin.de/>

<sup>122</sup> <https://www.digicult-verbund.de/de>

<sup>123</sup> <https://www.bsz-bw.de/MusIS.html>

<sup>124</sup>

<https://kmbi.de/de/museumsberatungsstellen/oeffentliche-museumsberatungsstellen-in-deutschland>

<sup>125</sup> <https://nbn-resolving.org/urn:nbn:de:hbz:385-10715>

<sup>126</sup> <https://training.parthenos-project.eu/>

<sup>127</sup>

<https://www.rwth-aachen.de/cms/root/Forschung/Forschungsdatenmanagement/Weiterbildung/sangebote/~udzt/Lehrvideos/>

<sup>128</sup> <https://doi.org/10.18442/031>

<sup>129</sup> <https://www.museumsbund.de/digitale-grunderfassung/>

<sup>130</sup> <https://doi.org/10.11588/arthistoricum.775>

## 9.2 Projektziel Nachnutzung: bei der Planung die Weichen stellen

Die Planungsphase eines Projekts wird meist von der inhaltlichen Konzeption und Herausforderungen bestimmt, die zur Erfüllung des Projektziels bewältigt werden müssen. Doch es ist wichtig, dass Datenproduzentinnen und -produzenten sich bereits in dieser initialen Projektphase mit den gewünschten Nachnutzungsszenarien für Menschen und Maschinen befassen. Sie erhöhen die Reichweite und Bedeutung der produzierten Forschungsdaten und dienen damit auch der wissenschaftlichen Reputation der Datenproduzierenden.

Die Anwendung der FAIR-Prinzipien hat Auswirkungen auf die Methoden, wie Daten zusammengestellt, modelliert und dokumentiert werden. Es hat Konsequenzen für Entscheidungen, in welchen Formaten die Speicherung erfolgt, wie die Daten aufbewahrt und weitergegeben und wie sie für die Weiterverwendung lizenziert werden. Von ihnen hängen zahlreiche weitere Schritte der Projekt- und Ressourcenplanung ab.

Doch die FAIR-Kriterien sollen nicht nur den Mehrwert von regulär publizierten Forschungsergebnissen maximieren. Die Anforderung der FAIRness richtet sich auch auf die Algorithmen, Werkzeuge und Arbeitsabläufe, die zur Erzeugung der Forschungsdaten geführt haben. Alle wissenschaftlichen digitalen Forschungsobjekte – von Daten bis hin zu Analysepipelines – sollen in den Prozess der FAIRification einbezogen werden, da alle Komponenten des Forschungsprozesses verfügbar sein müssen, um Transparenz, Reproduzierbarkeit und Wiederverwendbarkeit zu gewährleisten.

Der erste Schritt zur Nachhaltigkeit besteht darin, Ihr Projekt aus der Perspektive bestmöglicher Nachnutzbarkeit zu analysieren. Planen Sie dabei nicht zu eng nur auf konkrete und in nächster Zeit anstehende Nutzungsszenarien hin, sondern auch bereits auf mittelfristig oder langfristig angestrebte. Erwägen Sie, für welche Fragestellungen verwandter Fachrichtungen Ihre Daten auch interessant sein könnten. Priorisieren Sie ggf. bestimmte Nutzungsarten, um Ihre Aktivitäten skalieren zu können, wenn nicht alles mit den absehbar zur Verfügung stehenden Ressourcen

umsetzbar ist. Diese mittel- und langfristigen Ausbaustufen brauchen Sie anfangs nur grob zu skizzieren, die Umsetzung dieser Ziele kann stufenweise erfolgen. Analysieren Sie aber sorgfältig, welche davon bereits Auswirkungen auf die Weichenstellungen am Anfang haben, und welche Konsequenzen für die anderen Stationen des Datenlebenszyklus aus Ihrer Schwerpunktsetzung folgen. Sie sehen dann klarer, welche Kriterien oder Aufgaben aufeinander aufbauen oder einander bedingen und können die Grundlagen für einen zukünftigen Ausbau und neue Nutzungen richtig legen.

Besonders in Gedächtnisinstitutionen, die die FAIRification ihrer Datenhaltung als internes Projekt planen, ist es sinnvoll, eine "FAIR Roadmap" aufzustellen, die Ziele und die zugehörigen Maßnahmen mit Prioritäten versieht. Besonderes Augenmerk sollte man auf die folgenden Themenbereiche legen, denn aus ihnen resultieren häufig eine Reihe von weiteren Aufgabenpaketen, die konkret geplant sein wollen:

- die Vergabe von (möglichst offenen) maschinell auswertbaren Lizenzen und Nutzungshinweisen für die physischen Objekte, das zugehörige digitale Bildmaterial und die Metadaten
- die Gewährleistung der dauerhaften Zugänglichkeit der Daten und Metadaten
- der Einsatz von kontrolliertem Vokabular
- die Referenzierung auf Normdaten
- die Vergabe von PIDs für publizierte Datenbestände
- die Konversion der Daten in ein oder mehrere gängige maschinenauswertbare Bereitstellungsformate
- die Publikation der Daten über Schnittstellen oder über eine oder mehrere Datenplattformen in passenden Bereitstellungsformaten

Manche FAIR-Aktivitäten können von zeitlich und in der Aufgabenstellung klar begrenzten Forschungsprojekten von Anfang an relativ einfach in die Projektplanung einbezogen werden, sind aber nur mit sehr großem Aufwand in die kontinuierliche Bestandserschließung zu integrieren. Dies gilt zum Beispiel für die sehr aufwändige retrospektive Aufarbeitung und Integration der Quellenmaterialien, die einer Bestandsdokumentation zugrunde liegen. Sie könnten zunächst eine nachgeordnete Priorität erhalten.

## 9.3 Aktives Projektmanagement mit dem Datenmanagementplan

Datenmanagementpläne (DMP) sind ein Schlüsselement eines guten Datenmanagements. Ein DMP beschreibt den Lebenszyklus der Daten. Er macht Aussagen über den Umgang mit Forschungsdaten während und nach dem Ende des Projekts. Dazu gehört, welche Daten gesammelt, verarbeitet und/oder erzeugt werden, welche Methoden und Standards angewandt werden, zu welchen Konditionen Daten genutzt werden können und wie die Daten kuratiert und aufbewahrt werden, auch nach Projektende. Er unterstützt Sie bei der Identifikation aller an einem Projekt Beteiligten und bei der realistischen Kalkulation der notwendigen Aufwände in Bezug auf Personal, Zeit, Fachexpertise und finanzieller Mittel.

### Tipps zum Umgang mit dem Datenmanagementplan

Schreiben Sie bereits in der Planungsphase einen Datenmanagementplan. Er ist Ihre wichtigste Planungsunterlage und die Richtschnur für die tatsächliche Handhabung der Daten. Tun Sie das nicht nur dann, wenn Ihre Daten in einem Drittmittelprojekt mit einer definierten Zielsetzung und einer zeitlichen Begrenzung entstehen und kuratiert werden, sondern auch, wenn Sie mit dynamischen Daten befasst sind, deren Aktualisierung und Fortschreibung eine Daueraufgabe ist. Dies ist z. B. bei Bestandsdaten in Kulturerbe-Institutionen der Fall.

- Beginnen Sie bei der konkreten Planung mit der Datenlebenszyklus-Stufe der Nachnutzung, da von ihr viele Prioritätensetzungen für die anderen Stufen abhängen.
- Beginnen Sie mit einem Grobkonzept und verfeinern Sie Ihren DMP schrittweise. Konzipieren Sie alle Schritte so, dass sie im Projektalltag realistisch umsetzbar sind.
- Planen Sie genau, welche Personen und Institutionen Sie zur Umsetzung bestimmter Ziele konsultieren oder beteiligen müssen. Diese gehören evtl. nicht zu Ihrem Projektteam oder Ihrer Institution. Planen Sie auch Ressourcen für diese Kommunikationsprozesse ein.
- Planen Sie auch beigeordnete Arbeitspakete innerhalb des Projekts sorgfältig, da sie aufwändiger sein können als zunächst angenommen. Dies gilt z. B. für

das Ausheben eines Bestands vor der Digitalisierung oder für die Rechtklärung.

- Halten Sie den DMP während des Projekts aktuell. Passen Sie ihn an die tatsächlichen Gegebenheiten und Herausforderungen an, sodass er auch im Projektverlauf eine realistische Planungsgrundlage bleibt.
- Auch wenn ein DMP von Ihrer Förderinstitution oder Ihrem Vorgesetzten nicht verlangt wird, sollten Sie einen anfertigen und für Ihr Projekt als verbindlich ansehen.
- Machen Sie den DMP für alle Projektmitarbeitenden und ggf. weitere relevante Projektbeteiligte offen zugänglich und erinnern Sie Ihr Team regelmäßig daran, diesen zu konsultieren und ggf. zu aktualisieren.
- Ermutigen Sie Ihr Team, sich abzeichnende Schwierigkeiten in der Umsetzung des DMP frühzeitig zur Sprache zu bringen. So kann man zeitnah eine Lösungsmöglichkeit suchen, bevor das Problem gravierender wird.
- Nutzen Sie ein webbasiertes Softwaretool zur Erstellung und Pflege Ihres DMP. Es unterstützt sie durch seine Fragebogenstruktur und passende Begleitinformationen.

Beispiel: Datenmanagementplan mit FAIR-Schwerpunkt im EU-Programm Horizon 2020<sup>131</sup>

Mehr Informationen zu Datenmanagementplänen

[Forschungsdaten.info](https://forschungsdaten.info)<sup>132</sup>

Webbasiertes Tool: An vielen deutschen Hochschulen ist der Research Data Management Organiser (RDMO)<sup>133</sup> für Institutionsangehörige verfügbar. Je nach Institution sind unterschiedliche Templates für Datenmanagementpläne hinterlegt, die man für das eigene Vorhaben anpassen kann.

---

<sup>131</sup>

[https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management\\_en.htm](https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management_en.htm)

<sup>132</sup> <https://forschungsdaten.info/themen/informieren-und-planen/datenmanagementplan/>

<sup>133</sup> <https://rdmorganiser.github.io/>

## Tipps zur konkreten Projektplanung

- Identifizieren Sie die Zielgruppen, die Ihnen besonders wichtig sind.
- Identifizieren Sie die Objektbestände, die Gegenstand des Projekts sind.
- Identifizieren Sie die Informationen, die für das Auffinden Ihrer Daten besonders relevant sind.
- Machen Sie eine Bestandsaufnahme: Erstellen Sie einen Überblick über
  - alle Sammlungen der Einrichtung sowie über alle Datenbanken, Repositorien und Systeme, in denen die im Projektkontext relevanten Sammlungen beschrieben und präsentiert werden,
  - die existierenden Rechtsgrundlagen, Vereinbarungen und Verträge, die in Bezug auf das Rechtemanagement der physischen Objekte, ihrer Abbildungen und der Metadaten relevant sind
  - den aktuellen Umsetzungsstand der FAIR-Prinzipien
- Welche digitalen Objekte sind in Ihrem Datenbestand identifizierbar? Legen Sie fest, welche Schichten „Daten“ sind, welche „Metadaten“ und welche Einheiten durch einen gemeinsamen Identifikator adressiert werden müssen.
- Legen Sie anhand der geplanten Nachnutzungskontexte fest, bis zu welcher Granularitätsstufe identifizierbare digitale Objekte adressierbar sein sollen. Im Archiv könnte dies ein Konvolut mit Hunderten zugehöriger Schriftstücke sein. Sind diese allerdings einzeln verzeichnet und vorhandene Digitalisate entsprechend zugeordnet, wäre es möglich, jedes als digitales Objekt anzusprechen.
- Legen Sie fest, welche Metadaten-Elemente Ihrer Daten Sie benötigen, um Ihre Daten gegen die Kriterien der FAIR-Prinzipien prüfen zu können. Eine sehr konkret gehaltene Checkliste für die FAIRness von Daten, die gut in einen Anforderungskatalog für einen FAIR-Entwicklungsplan integrierbar ist, bietet das FAIR Data Maturity Model<sup>134</sup>.
- Nicht alle notwendigen Informationen befinden sich in den Projektdaten selbst. Klären Sie, woher Sie ggf. die Informationen beziehen können, die Sie zur Bereitstellung aussagekräftiger Metadaten benötigen, z. B. in Bezug auf die Datenprovenienz oder auf die begleitende Dokumentation von Metadatenschemata oder Software.

---

<sup>134</sup> <https://doi.org/10.5281/zenodo.5834115>

- Legen Sie im DMP die zu Ihrem Projekt passenden Kriterien für die Umsetzung der FAIR-Prinzipien fest, notieren Sie Klärungsbedarf und anstehende Aufgaben.
- Wahrscheinlich wird es nicht möglich sein, alle Kriterien für FAIRness gleich schnell umzusetzen. Skalieren Sie, setzen Sie Prioritäten.
- Erstellen Sie für alle Datenlebenszyklus-Stufen Kriterienkataloge. Hier sammeln Sie die Kriterien, die Ihre Daten aufweisen oder die beteiligten Akteure (Personen und Infrastrukturen) erfüllen müssen, um die verschiedenen FAIR-Aspekte umsetzbar zu machen. Listen Sie die Konsequenzen auf, die bestimmte Maßnahmen in späteren Phasen des Datenlebenszyklus haben. Dadurch werden wechselseitige Bedingtheiten klarer, die notwendigen Arbeiten lassen sich besser planen.

#### Beispiel: Integration von Normdatenreferenzen in einen Datenbestand

Ein Datenbestand einer Sammlung soll mit GND-Normdatenreferenzen für die erwähnten Personen, für Sachschlagworte und Geografika ausgestattet werden. Große zu bearbeitende Mengen bewältigt man effektiv mit Hilfe von Tools wie OpenRefine in separaten Arbeitsgängen im Rahmen der Datenkuratierung. Sollen zahlreiche Referenzen während der fortlaufenden Erschließung ermittelt werden, lohnt sich die Integration von Software-Funktionen, die den Katalogisierenden die schnelle Identifizierung und Verknüpfung des richtigen Normdatensatzes im Alltagsbetrieb erleichtern. Bei kleineren Mengen kann ein händisch-intellektuelles Verfahren das effektivste sein, da es einen vergleichsweise geringen Änderungsbedarf an anderen Stellen erzeugt. Die involvierten Mitarbeiter:innen müssen die einschlägigen Kompetenzen für das durchzuführende Datenmatching aufweisen oder erwerben. Änderungen an der Konfiguration der Datenbanksoftware oder des Metadatenschemas sind erforderlich. Korrektheit, Konsistenz und Abdeckung der eingetragenen Referenzen sind zu prüfen. Vor der Bereitstellung der Daten für die Publikation ist zu berücksichtigen, dass die Einträge korrekt und vollständig in das Bereitstellungsformat (z. B. MARC, LIDO, EAD) übernommen werden, die Datentransformationsskripte müssen ggf. angepasst werden.

Mit dem Betreibenden der Publikationsplattform ist ggf. abzustimmen, ob und wie die neuen Normdatenreferenzen die Funktionalitäten der Plattform unterstützen oder ob sie sogar verbesserte Linked-Data-Funktionen ermöglichen. Unkritisch ist es, falls die von Ihnen bereitgestellten Daten bereits Normdatenreferenzen enthalten, die die Plattform aktuell noch nicht verarbeitet. Die Plattformen arbeiten kontinuierlich am Qualitätsausbau ihrer Angebote und analysieren dazu die eingehenden Daten auf neue Nutzungspotentiale.

## Wählen Sie eine vertrauenswürdige Datenplattform

Machen Sie Ihre Daten über ein vertrauenswürdiges Repositorium zugänglich. Prüfen Sie schon früh verschiedene Repositorien, ob und in welcher Weise diese Sie bei der Umsetzung von FAIR unterstützen. Ein zertifiziertes Repositorium bietet einen vertrauenswürdigen Speicherort für Datensätze. Die Zertifizierung ist eine Garantie dafür, dass die Daten sicher gespeichert werden und langfristig verfügbar, auffindbar und zugänglich sind. Beispiele für Zertifizierungsstandards sind CoreTrustSeal<sup>135</sup>, das nestor-Siegel für vertrauenswürdige digitale Langzeitarchive<sup>136</sup> und ISO 16363-Zertifizierung<sup>137</sup>. Der Zertifizierungsgrad sollte auf der Webseite des Repositoriums deutlich angegeben sein. Wenn es (noch) nicht zertifiziert ist, sollte es klare Aussagen bereitstellen, wie Verfügbarkeit, Zugänglichkeit und Wiederverwendbarkeit der (Meta-)Daten über einen definierten Zeitraum sichergestellt werden. Fragen Sie nach, wenn Ihnen Informationen fehlen oder die Umsetzung der beschriebenen Konditionen nicht klar ist. Erkundigen Sie sich auch nach eventuell anfallenden Kosten für die Speicherung und die damit verbundenen Bedingungen.

Richten Sie Ihre Datenhaltung auf die Konditionen des Repositoriums aus. Wenn Sie darüber hinaus die Standards des Repositoriums befolgen (bevorzugte Dateiformate, Metadatenschemata usw.), können Sie sicherstellen, dass alle Anforderungen für die Bereitstellung der Daten erfüllt sind.

Zur Publikation von Daten aus Bibliotheken, Archiven und Museen kommen neben den bibliothekarischen Online Public Access Catalogues (OPACs), dem

---

<sup>135</sup> <https://www.coretrustseal.org/>

<sup>136</sup> <http://nbn-resolving.de/urn:nbn:de:0008-2019062507>

<sup>137</sup> <https://www.iso.org/standard/56510.html>

Archivinformationssystem Arcinsys<sup>138</sup> und von einzelnen Museen oder Verbänden getragene Webdatenbanken auch Fachportale oder Kulturerbe-Portale wie die Deutsche Digitale Bibliothek (DDB)<sup>139</sup> infrage, die als Aggregatoren für Europeana<sup>140</sup> fungieren. DDB und Europeana unterstützen beratend das lokale Qualitätsmanagement, bereiten Daten auf und präsentieren diese nicht nur in den Portalen, sondern auch über Schnittstellen zur weiteren Nachnutzung.

Beispielhafte Darstellungen der FAIR-Policy von Repositorien

Crosas, Mercè: The FAIR Guiding Principles: Implementation in Dataverse, 2019<sup>141</sup>

Zenodo Best Effort Principles<sup>142</sup>: Abschnitte "FAIR Principles", "Self-assessment against the Plan S requirements for Open Access Repositories", "Strongly recommended additional criteria"

## 10. Empfehlungen für Dateiformate

Verwenden Sie möglichst weit verbreitete und mit einem ISO-Standard verbundene Dateiformate. Sie sollten nicht proprietär, also nicht an eine Software oder einen Hersteller gebunden, mit unterschiedlichen Programmen verwendbar und mit einer offenen Lizenz versehen sein und über eine frei zugängliche Dokumentation einschließlich der technischen Spezifikationen verfügen. Sie erlauben eine verlustfreie Speicherung ohne Kompression und sie sind einfach dekodierbar oder unmittelbar lesbar.

---

<sup>138</sup> <https://www.arcinsys.de/>

<sup>139</sup> <https://pro.deutsche-digitale-bibliothek.de/daten-liefere>

<sup>140</sup> <https://pro.europeana.eu/page/aggregators>

<sup>141</sup>

<https://scholar.harvard.edu/merceecrosas/presentations/fair-guiding-principles-implementation-dataverse>

<sup>142</sup> <https://about.zenodo.org/principles/>

Konsultieren Sie Ihre Datenplattform, welche Dateiformate dort entgegengenommen bzw. empfohlen werden.

Für die Kulturwissenschaften und für Kulturerbe-Sammlungen empfehlen wir die Verwendung folgender Standard-Dateiformate:

## Text

- [Extensible Markup Language \(XML\) 1.1](#)<sup>143</sup>, mit XML Schema Definition (XSD)
- [Resource Description Framework in Attributes \(RFDa\)](#)<sup>144</sup> für das Einbetten von RDF-Statements in XML
- [JavaScript Object Notation \(JSON\)](#)<sup>145</sup>
- [JavaScript Object Notation for Linked Data \(JSON LD\)](#)<sup>146</sup>
- [Comma-Separated Values \(CSV\)](#)<sup>147</sup>
- Textdatei (TXT) (Codierung UTF-8)
- Präsentationsformat: [Portable Document Format A \(PDF-A\)](#)<sup>148</sup>

## Musik

- [Music Encoding Initiative \(MEI\)](#)<sup>149</sup>
- [MusicXML](#)<sup>150</sup>
- [Parsons Code](#)<sup>151</sup>
- Präsentationsformat: [Portable Document Format A \(PDF-A\)](#)<sup>152</sup>

## Bild (Rastergrafiken)

- Rohdaten: [Digital Negative \(DNG\)](#)<sup>153</sup>

---

<sup>143</sup> <https://www.w3.org/TR/xml11/>

<sup>144</sup> <https://www.w3.org/TR/xhtml-rdfa-primer/>

<sup>145</sup> <https://www.ecma-international.org/publications-and-standards/standards/ecma-404/>

<sup>146</sup> <https://www.w3.org/TR/json-ld11/>

<sup>147</sup> <https://www.rfc-editor.org/rfc/rfc4180>

<sup>148</sup> <https://www.slub-dresden.de/veroeffentlichen/open-access-publizieren/pdfa-erstellung>

<sup>149</sup> <https://music-encoding.org/>

<sup>150</sup> <https://www.musicxml.com/>

<sup>151</sup> <https://www.dcode.fr/parsons-code>

<sup>152</sup> <https://www.slub-dresden.de/veroeffentlichen/open-access-publizieren/pdfa-erstellung>

<sup>153</sup> <https://helpx.adobe.com/de/camera-raw/digital-negative.html>

- Master: baseline Tagged Image File Format (TIFF)<sup>154</sup>, unkomprimiert; TIFF mit Lempel-Ziv-Welch-Komprimierung (TIFF-LZW)<sup>155</sup>
- Joint Photographic Experts Group (JPEG 1<sup>156</sup> und JPEG 2000<sup>157</sup>), verlustfrei komprimiert, lizenzfreie Bereiche
- Präsentationsformate (Derivate): JPEG, JPEG 2000, Portable Network Graphics (PNG)<sup>158</sup>

## Bild (Vektorgrafiken) und CAD

- Scalable Vector Graphics (SVG)<sup>159</sup>

## Audio

- Archivformat: Waveform Audio File-Format (WAV)<sup>160</sup> in Verbindung mit Pulse Code Modulation (PCM); Free Lossless Audio Codec (FLAC)
- Präsentationsformat: MPEG-2 Audiolayer III (MP3)<sup>161</sup>

## Video/Film

- Archivformate:
  - Moving Picture Experts Group Motion JPEG 2000 (MJPEG2000)<sup>162</sup>
  - Moving Picture Experts Group, Standard MPEG-4<sup>163</sup>
  - Digital Picture Exchange (DPX)<sup>164</sup>, SMPTE 268M-2003, v 2.0
  - Material Exchange Format (MXF)<sup>165</sup>, SMPTE 377M
  - Codec FFV1<sup>166</sup> / Container MKV
  - TIFF mit FFV1 in Matroschka codiert

<sup>154</sup> <https://www.loc.gov/preservation/digital/formats/fdd/fdd000022.shtml>

<sup>155</sup> <https://www.loc.gov/preservation/digital/formats/fdd/fdd000074.shtml>

<sup>156</sup> <https://jpeg.org/jpeg/index.html>

<sup>157</sup> <https://jpeg.org/jpeg2000/>

<sup>158</sup> <https://www.w3.org/TR/2003/REC-PNG-20031110/>

<sup>159</sup> <https://www.w3.org/Graphics/SVG/>

<sup>160</sup> [https://www.tactilemedia.com/info/MCI\\_Control\\_Info.html](https://www.tactilemedia.com/info/MCI_Control_Info.html)

<sup>161</sup> <https://en.wikipedia.org/wiki/MP3>

<sup>162</sup> <https://www.loc.gov/preservation/digital/formats/fdd/fdd000127.shtml>

<sup>163</sup> <https://mpeg.chiariglione.org/standards/mpeg-4>

<sup>164</sup> <https://doi.org/10.5594/SMPTE.ST268-1.2014>

<sup>165</sup> <https://www.loc.gov/preservation/digital/formats/fdd/fdd000013.shtml>

<sup>166</sup> <https://github.com/FFmpeg/FFV1/blob/master/ffv1.md>

- Präsentationsformat: MP4 (MPEG-4, part 14)<sup>167</sup>
- weitere Empfehlungen der nestor AG Media<sup>168</sup>

## 3D

- Graphics Language Transmission Format (glTF)<sup>169</sup>
- Wavefront Object (OBJ)<sup>170</sup>
- COLLABorative Design Activity (COLLADA)<sup>171</sup>
- Extensible 3D (X3D)<sup>172</sup>
- Präsentationsformat: glTF, X3D, OBJ

## Tabellen

- Extensible Markup Language (XML) 1.1<sup>173</sup>, mit XML Schema Definition (XSD)
- Comma-Separated Values (CSV)<sup>174</sup>

## Datenbanken

- Software Independent Archival of Relational Databases (SIARD)<sup>175</sup>
- Structured Query Language (SQL)<sup>176</sup>
- Export und externe Speicherung der Datenbankinhalte empfohlen, z. B. als XML, CSV, JSON mit Entitäten-Relationen-Diagramm (ERD)

## Geoinformationssysteme (GIS)

- Vektordaten: Environmental Systems Research Institute ESRI Shapefile<sup>177</sup>

<sup>167</sup> <https://mpeg.chiariglione.org/standards/mpeg-4/mp4-file-format>

<sup>168</sup> <https://wiki.dnb.de/display/NESTOR/Digitalisierungsempfehlungen>

<sup>169</sup> <https://www.khronos.org/glTF/>

<sup>170</sup> <https://www.fileformat.info/format/wavefrontobj/egff.htm>

<sup>171</sup> <https://www.khronos.org/collada/>

<sup>172</sup> <https://www.web3d.org/>

<sup>173</sup> <https://www.w3.org/TR/xml11/>

<sup>174</sup> <https://www.rfc-editor.org/rfc/rfc4180>

<sup>175</sup> <https://www.bar.admin.ch/bar/de/home/archivierung/tools---hilfsmittel/siard-suite.html>

<sup>176</sup> <https://www.torsten-horn.de/techdocs/sql.htm>

<sup>177</sup>

<https://desktop.arcgis.com/de/arcmap/latest/manage-data/shapefiles/geoprocessing-considerations-for-shapefile-output.htm>

- Rasterdaten: Open Geospatial Consortium GeoTIFF<sup>178</sup>

## Webseiten

- Hypertext Markup Language (HTML5)<sup>179</sup>
- Extensible Hypertext Markup Language (XHTML)<sup>180</sup>
- Resource Description Framework in Attributes (RFDa)<sup>181</sup> für das Einbetten von RDF-Statements in HTML und XHTML

Weitere Erläuterungen zu den Dateiformaten

Altenhöner, Reinhard / Berger, Andreas / Bracht, Christian et al.: DFG-Praxisregeln „Digitalisierung“ – aktualisierte Fassung, 2022<sup>182</sup>

nestor AG Media: Leitfaden für die digitale Langzeitarchivierung audiovisueller Medien, nestor Materialien 19, 2016<sup>183</sup>

IANUS IT-Empfehlungen für den nachhaltigen Umgang mit digitalen Daten in den Altertumswissenschaften, 2016<sup>184</sup>

Neuroth, Heike / Oßwald, Achim / Scheffel, Regine et al.: nestor-Handbuch: Eine kleine Enzyklopädie der digitalen Langzeitarchivierung, Version 2.3, 2010<sup>185</sup>

Sustainability of Digital Formats: Planning for Library of Congress Collections, 2004–2022<sup>186</sup>

<sup>178</sup> <http://www.opengis.net/doc/IS/GeoTIFF/1.1>

<sup>179</sup> <https://www.w3.org/TR/html5/>

<sup>180</sup> <https://www.w3.org/TR/xhtml2/>

<sup>181</sup> <https://www.w3.org/TR/xhtml-rdfa-primer/>

<sup>182</sup> <https://doi.org/10.5281/zenodo.7435724>

<sup>183</sup> <http://nbn-resolving.de/urn:nbn:de:0008-2016102107>

<sup>184</sup> <https://ianus-fdz.de/it-empfehlungen/dateiformate>

<sup>185</sup> <http://www.nestor.sub.uni-goettingen.de/handbuch/index.php>

<sup>186</sup> <https://www.loc.gov/preservation/digital/formats/index.html>

SLUB Archiv. Technische Standards für die Ablieferung von digitalen Dokumenten<sup>187</sup>

lzw.nrw Landesinitiative Langzeitverfügbarkeit: Interaktive Tafel gängiger Dateiformate<sup>188</sup>

File Format Developer's Guide<sup>189</sup>

## 11. Empfehlungen für Rahmenwerke und Referenzmodelle

### Grundlegende Semantic-Web-Technologien

Diese W3C-Standards stellen die grundlegenden Rahmenwerke für Semantic-Web-Technologien dar:

Resource Description Framework (RDF)<sup>190</sup>, Einführung: RDF 1.1 Primer<sup>191</sup>

RDF Schema (RDFS)<sup>192</sup>

Web Ontology Language (OWL)<sup>193</sup>, Einführung: OWL 2 Web Ontology Language Primer (Second Edition)<sup>194</sup>

Simple Knowledge Organization System (SKOS)<sup>195</sup>, Einführung: SKOS Simple Knowledge Organization System Primer<sup>196</sup>

---

<sup>187</sup>

<https://slubarchiv.slub-dresden.de/technische-standards-fuer-die-ablieferung-von-digitalen-dokumenten>

<sup>188</sup> [https://www.lzv.nrw/dateiformate/?mtm\\_campaign=tafellaunch&mtm\\_medium=social](https://www.lzv.nrw/dateiformate/?mtm_campaign=tafellaunch&mtm_medium=social)

<sup>189</sup> <https://www.fileformat.com/>

<sup>190</sup> <https://www.w3.org/RDF/>

<sup>191</sup> <https://www.w3.org/TR/rdf11-primer/>

<sup>192</sup> <https://www.w3.org/TR/rdf-schema/>

<sup>193</sup> <https://www.w3.org/OWL/>

<sup>194</sup> <https://www.w3.org/TR/owl2-primer/>

<sup>195</sup> <https://www.w3.org/2004/02/skos/>

<sup>196</sup> <https://www.w3.org/TR/skos-primer/>

## Referenzmodelle

Für die Kulturwissenschaften und für Kulturerbe-Sammlungen empfehlen wir die Orientierung an folgenden Referenzmodellen (top-level ontologies).

Anwendungsontologien finden Sie im Abschnitt Empfehlungen für Vokabulare, Normdaten und Anwendungsontologien.

**Conceptual Reference Model (CIDOC-CRM)**<sup>197</sup> des International Council of Museums (ICOM), International Council for Documentation (CIDOC)

CIDOC-CRM ist ein Instrument für die Informationsmodellierung und -integration im Bereich des kulturellen Erbes. Es stellt Definitionen und eine formale Ontologie zur Beschreibung der impliziten und expliziten Konzepte und Beziehungen bereit, die in der Dokumentation des kulturellen Erbes verwendet werden und die für die Abfrage und Auswertung solcher Daten von Interesse sind. Seit 2006 ist es offizieller Standard ISO 21127, der Status wurde 2014 erneuert.

Publikation der Ontologie als Erlangen CRM / OWL<sup>198</sup>

Für bestimmte Fragestellungen sind auch die CRM-Erweiterungen<sup>199</sup> relevant.

Besonders sei auf folgende hingewiesen:

- CRM Model for Provenance Metadata (CRMdig)<sup>200</sup>  
ist eine Ontologie und ein RDF-Schema zur Kodierung von Metadaten über die Schritte und Methoden der Herstellung ("Provenienz") von Digitalisierungsprodukten und digitalen Darstellungen wie 2D-, 3D- oder sogar animierten Modellen, die mit verschiedenen Verfahren erstellt wurden. Das Modell ist darüber hinaus für alle Arten von wissenschaftlichen Daten und anderen digitalen Objekten anwendbar und verfügt über ein großes Spektrum an Möglichkeiten in Bezug auf Abdeckung, Universalität, Ausdruckskraft und Detaillierungsgrad der Informationsmodellierung.
- CRM Model for Archaeological Buildings (CRMba)<sup>201</sup>  
ist eine Ontologie und ein RDF-Schema zur Kodierung von Metadaten über die

---

<sup>197</sup> <https://www.cidoc-crm.org/>

<sup>198</sup> <http://erlangen-crm.org/about>

<sup>199</sup> <https://cidoc-crm.org/collaborations>

<sup>200</sup> <https://cidoc-crm.org/ModelVersion/version-3.2.1>

<sup>201</sup> <https://cidoc-crm.org/crmba/ModelVersion/version-1.4>

Dokumentation von Gebäuden gemäß archäologischer Fragestellungen. Das Modell wurde entwickelt, um den Prozess der Aufzeichnung von Zeugnissen und Diskontinuitäten der Materie an archäologisch relevanten Gebäuden zu unterstützen, um die Entwicklung der Anlage im Laufe der Jahrhunderte zu bestimmen und die Beziehungen zwischen den einzelnen Gebäudekomponenten untereinander und mit dem Gebäude als Ganzes zu erfassen.

- CRM Argumentation Model (CRMinf)<sup>202</sup>  
ist eine Ontologie, die für die Integration von Metadaten über Argumentationsprozesse, Prämissen, Beweisführung und Schlussfolgerungen in deskriptiven und empirischen Wissenschaften verwendet werden soll. Es wird empfohlen, sie mit dem CRM Scientific Observation Model (CRMsci) zusammen anzuwenden. Das CRMinf-Dokument finden Sie hier<sup>203</sup>.
- CRM Scientific Observation Model (CRMsci)<sup>204</sup>  
ist eine Ontologie, die als globales Schema für die Integration von Metadaten über wissenschaftliche Beobachtungen, Messungen und verarbeitete Daten in beschreibenden und empirischen Wissenschaften verwendet werden soll. Ihr Hauptzweck ist die Erleichterung der Verwaltung, Integration und Vermittlung von Forschungsdaten, ihres Austausches und ihrer Nachnutzung.

**Library Reference Model (LRM)**<sup>205</sup> der International Federation of Library Associations and Institutions (IFLA)

Das Library Reference Model wurde für den Bibliotheksbereich entwickelt. Es konsolidiert die Vorgängermodelle Functional Requirements for Bibliographical Records (FRBR), Functional Requirements for Authority Data (FRAD) und Functional Requirements for Subject Authority Data (FRSAD) und deckt alle Aspekte bibliografischer Daten ab.

Publikation der Ontologie<sup>206</sup>

---

<sup>202</sup> <https://cidoc-crm.org/crminf/>  
<sup>203</sup>

<http://www.cidoc-crm.org/crminf/sites/default/files/CRM%20Inference%20extension%20CRM-INF%200.9.pdf>

<sup>204</sup> <https://cidoc-crm.org/crmsci/ModelVersion/version-1.2.3>

<sup>205</sup> <https://repository.ifla.org/handle/123456789/40>

<sup>206</sup> <https://www.iflstandards.info/lrm/lrmer.html>

**Definition of FRBROO: A Conceptual Model for Bibliographic Information in Object-Oriented Formalism, v.2.4**<sup>207</sup> der International Federation of Library Associations and Institutions (IFLA)

FRBROO ist eine formale Ontologie, die die zugrundeliegende Semantik bibliographischer Informationen erfasst und durch die Harmonisierung mit dem CIDOC-CRM die Integration, Vermittlung und den Austausch von Informationen aus Bibliotheken und Museen und ähnlichen Institutionen erleichtert.

Publikation der Ontologie<sup>208</sup>

**Records in Contexts Conceptual Model (RiC-CM)**<sup>209</sup> des International Council on Archives (ICA)

Ziel des neuen Standards RiC-CM für den Archivbereich ist es, die vier vorhandenen Standards für den Archivbereich General International Standard Archival Description (ISAD(G)), International Standard Archival Authority Records—Corporate Bodies, Persons, and Families (ISAAR(CPF)), International Standard Description of Functions (ISDF), und International Standard Description of Institutions with Archival Holdings (ISDIAH) aufeinander abzustimmen und sie in einer integrierten weiterentwickelten Form nutzbar zu machen.

Publikation der Ontologie<sup>210</sup>

## 12. Empfehlungen zu Metadatenstandards

Metadatenbezogene Standards sind formale Dokumente, die einheitliche Kriterien, Methoden, Prozesse und Praktiken für die Dokumentation festlegen. Sie sollen als Leitfaden für den Entwurf, die Erstellung und die Implementierung von Datenstrukturen, Datenwerten, Dateninhalten dienen und die Nutzung der Daten auf effiziente und einheitliche Weise gewährleisten. Basierend auf dem Zweck, für den ein Metadatenstandard entwickelt wurde, können Metadatenstandards im Kontext von

---

<sup>207</sup> <https://repository.ifla.org/handle/123456789/659>

<sup>208</sup> <https://www.iflastandards.info/fr/frbr/frbroo>

<sup>209</sup> <https://www.ica.org/en/records-in-contexts-conceptual-model>

<sup>210</sup> <https://www.ica.org/standards/RiC/ontology>

Forschungsdaten und Kulturerbe-Dokumentation in drei Kategorien unterteilt werden: Standards für Dateninhalte, für Datenstrukturen (Elementsets und Schemata) und für Datenwerte.

## 12.1 Empfehlungen für Metadateninhalte

Inhaltsstandards legen fest, wie bestimmte Ressourcen und Objekte beschrieben und dokumentiert werden sollen. Sie werden auch als Katalogisierungsregeln bezeichnet. Sie können Vorgaben für die Erfassung und Formatierung von Informationen sowie die zu verwendenden Vokabulare (Datenwerte) enthalten. Ihre Verwendung ist nicht an ein bestimmtes Metadatenschema gebunden.

Für die Kulturwissenschaften und für Kulturerbe-Sammlungen empfehlen wir die Verwendung folgender Inhaltsstandards:

### Bibliotheken

Resource Description and Access (RDA)<sup>211</sup>

### Archivgut

General International Standard Archival Description (ISAD(G))<sup>212</sup>

Nachlässe: Ressourcenerschließung mit Normdaten in Archiven und Bibliotheken (RNAB)<sup>213</sup>

### Materielles und visuelles Kulturgut in Museen, Bauwerke, Monumente, archäologische Stätten

Cataloging Cultural Objects: A Guide to Describing Cultural Works and Their Images (CCO)<sup>214</sup>

Categories for the Description of Works of Art (CDWA)<sup>215</sup>

---

<sup>211</sup> <https://www.dnb.de/rda>

<sup>212</sup>

<https://www.ica.org/en/isadg-general-international-standard-archival-description-second-edition>

<sup>213</sup> <https://www.dnb.de/rnab>

<sup>214</sup> <https://vraweb.org/resources/cataloging-cultural-objects/>

<sup>215</sup> [https://www.getty.edu/research/publications/electronic\\_publications/cdwa/](https://www.getty.edu/research/publications/electronic_publications/cdwa/)

MIDAS Heritage<sup>216</sup>

International Core Data Standard for Archaeological and Architectural Heritage (CDS)<sup>217</sup>

Arbeitskreis Provenienzforschung: Leitfaden zur Standardisierung von Provenienzangaben, 2018<sup>218</sup>

Annotationen

Web Annotation Data Model<sup>219</sup> und Web Annotation Protocol<sup>220</sup>

## 12.2 Empfehlungen für Metadatenstrukturen

Daten-Inhaltsstandards kommen meist in Form von spezifischeren, auf bestimmte Dokumentations- oder Anwendungskontexte bezogenen Metadatenmodellen mit begleitenden Regelwerken zur Anwendung. Diese Regelwerke bestehen aus Metadaten-Elementsets (Feldkatalogen). Die Bedeutung der Elemente und ihre Beziehung zueinander wird definiert (Semantik). Sie werden durch Anweisungen ergänzt, zu welchen Aussagen und auf welche Weise den Elementen in einer Anwendung Datenwerte (Vokabular) zugewiesen werden sollten.

Datenmodelle erfahren oft eine Ausprägung als technische Formate (Metadatenschemata) für die Publikation, Austausch und Bereitstellung von Daten. Diese Formatdefinitionen ermöglichen eine Kodierung oder Auszeichnung der Daten für die maschinelle Verarbeitung. Sie können in verschiedenen Serialisierungen vorliegen, z. B. als XML, JSON oder RDF.

Für die Kulturwissenschaften und für Kulturerbe-Sammlungen empfehlen wir die Verwendung folgender Metadaten-Elementsets und -schemata:

---

<sup>216</sup> <http://www.heritage-standards.org.uk/midas-heritage/>

<sup>217</sup>

<https://cidoc-data.org/aswg-international-core-data-standard-for-archaeological-and-architectural-heritage>

<sup>218</sup> <https://www.arbeitskreis-provenienzforschung.org/arbeitsgruppen/ag-standardisierung/>

<sup>219</sup> <https://www.w3.org/TR/annotation-model/>

<sup>220</sup> <https://www.w3.org/TR/annotation-protocol/>

## fachübergreifende Verbreitung

Dublin Core Metadata Element Set (DCMES)<sup>221</sup>

DCMI Metadata Terms (DC Terms)<sup>222</sup>

Schema.org<sup>223</sup> Types

Europeana Data Model (EDM)<sup>224</sup>, Definition of the Europeana Data Model v5.2.8<sup>225</sup>,  
Einführung: EDM Primer<sup>226</sup>

International Image Interoperability Framework (IIIF) Presentation API<sup>227</sup>

## Bibliotheken

Machine-Readable Cataloging (MARC)<sup>228</sup>, aktuell als MARC21

Metadata Encoding and Transmission Standard (METS)<sup>229</sup>

Metadata Object Description Schema (MODS)<sup>230</sup>

Bibliographic Framework Initiative (BIBFRAME)<sup>231</sup>

---

<sup>221</sup> <https://www.dublincore.org/specifications/dublin-core/dces/>

<sup>222</sup> <https://www.dublincore.org/specifications/dublin-core/dcmi-terms/>

<sup>223</sup> <https://schema.org/docs/schemas.html>

<sup>224</sup> <https://pro.europeana.eu/page/edm-documentation>

<sup>225</sup>

[https://pro.europeana.eu/files/Europeana\\_Professional/Share\\_your\\_data/Technical\\_requirements/EDM\\_Documentation/EDM\\_Definition\\_v5.2.8\\_102017.pdf](https://pro.europeana.eu/files/Europeana_Professional/Share_your_data/Technical_requirements/EDM_Documentation/EDM_Definition_v5.2.8_102017.pdf)

<sup>226</sup>

[https://pro.europeana.eu/files/Europeana\\_Professional/Share\\_your\\_data/Technical\\_requirements/EDM\\_Documentation/EDM\\_Primer\\_130714.pdf](https://pro.europeana.eu/files/Europeana_Professional/Share_your_data/Technical_requirements/EDM_Documentation/EDM_Primer_130714.pdf)

<sup>227</sup> <https://iiif.io/api/presentation/3.0/>

<sup>228</sup>

[https://www.dnb.de/DE/Professionell/Metatadendienste/Exportformate/MARC21/marc21\\_node.html](https://www.dnb.de/DE/Professionell/Metatadendienste/Exportformate/MARC21/marc21_node.html)

<sup>229</sup> <https://www.loc.gov/standards/mets/>

<sup>230</sup> <https://www.loc.gov/standards/mods/>

<sup>231</sup> <https://www.loc.gov/bibframe/>

## Archivgut

Encoded Archival Description (EAD)<sup>232</sup>

Metadata Encoding and Transmission Standard/Metadata Object Description Schema (METS/MODS)<sup>233</sup>

Materielles und visuelles Kulturgut in Museen, Bauwerke,  
Monumente, archäologische Stätten

Lightweight Information Describing Objects (LIDO)<sup>234</sup>, Einführung: LIDO Primer<sup>235</sup>

Visual Resources Association Core (VRA Core)<sup>236</sup>

Linked Art<sup>237</sup>

CARARE 2.0<sup>238</sup>

## Textuelle Quellen

Text Encoding Initiative (TEI)<sup>239</sup>

gedruckte Texte: Metadata Encoding and Transmission Standard / Metadata Object Description Schema (METS/MODS)<sup>240</sup>

Handschriften: TEI-Anwendungsprofil für digitalisierte Handschriften<sup>241</sup>

---

<sup>232</sup> <https://www.loc.gov/ead/>

<sup>233</sup> <https://wiki.deutsche-digitale-bibliothek.de/pages/viewpage.action?pageId=69124677>

<sup>234</sup> <https://www.lido-schema.org/>

<sup>235</sup> <https://lido-schema.org/documents/primer/latest/lido-primer.html>

<sup>236</sup> <https://www.loc.gov/standards/vracore/>

<sup>237</sup> <https://linked.art/>

<sup>238</sup>

<https://pro.carare.eu/en/introduction-carare-aggregation-services/carare-metadata-schema/>

<sup>239</sup> <https://tei-c.org/>

<sup>240</sup> <https://wiki.deutsche-digitale-bibliothek.de/pages/viewpage.action?pageId=69124677>

<sup>241</sup> <https://dfg-viewer.de/metadaten>

Layoutinformationen: PAGE XML<sup>242</sup>, Analyzed Layout and Text Object (ALTO)<sup>243</sup>

## Quellen der Musik

Music Encoding Initiative (MEI)<sup>244</sup>

## Metadaten in Bilddateien

Extensible Metadata Platform (XMP)<sup>245</sup>

International Press Telecommunications Council Information Interchange Model (IPTC-IIM)<sup>246</sup>

## Annotationen

Web Annotation Data Model<sup>247</sup>

Texte: Text Encoding Initiative (TEI)<sup>248</sup>

Musik: Music Encoding Initiative (MEI)<sup>249</sup>, MusicXML<sup>250</sup>

Artikel in wissenschaftlichen Zeitschriften: Journal Article Tag Suite (JATS)<sup>251</sup>

Layoutinformationen: PAGE XML<sup>252</sup>, Analyzed Layout and Text Object ALTO<sup>253</sup>

---

<sup>242</sup> <https://ocr-d.de/de/gt-guidelines/trans/trPage.html>

<sup>243</sup> <https://www.loc.gov/standards/alto/>

<sup>244</sup> <https://music-encoding.org/>

<sup>245</sup> <https://www.adobe.com/devnet/xmp.html>

<sup>246</sup> <http://www.iptc.org/standards/iim/>

<sup>247</sup> <https://www.w3.org/TR/annotation-model/>

<sup>248</sup> <https://tei-c.org/>

<sup>249</sup> <https://music-encoding.org/>

<sup>250</sup> <https://www.musicxml.com/>

<sup>251</sup> <https://www.niso.org/standards-committees/jats>

<sup>252</sup> <https://ocr-d.de/de/gt-guidelines/trans/trPage.html>

<sup>253</sup> <https://www.loc.gov/standards/alto/>

## 3D

Historic Building Information Modelling (HBIM)<sup>254</sup>

CARARE 2.0<sup>255</sup>

## Geodaten

INSPIRE Metadata Implementing Rules<sup>256</sup>

## Forschungsmetadaten in Repositorien

Datacite Best Practice Guide<sup>257</sup>

Data Catalog Vocabulary (DCAT 2<sup>258</sup>, DCAT 3<sup>259</sup>)

## Langzeitsicherung

PREservation Metadata: Implementation Strategies (PREMIS)<sup>260</sup>, dt. Einführung<sup>261</sup>

Nachweis weiterer Metadatenschemata

FAIRsharing.org – Standards, Databases, Policies<sup>262</sup>

Metadata Standards Catalog<sup>263</sup>

<sup>254</sup> <https://doi.org/10.5194/isprs-archives-XLII-2-W9-581-2019>

<sup>255</sup>

<https://pro.carare.eu/en/introduction-carare-aggregation-services/carare-metadata-schema/>

<sup>256</sup>

<https://inspire.ec.europa.eu/documents/inspire-metadata-implementing-rules-technical-guidelines-based-en-iso-19115-and-en-iso-1>

<sup>257</sup> <https://doi.org/10.5281/zenodo.7040047>

<sup>258</sup> <https://www.w3.org/TR/vocab-dcat-2/>

<sup>259</sup> <https://www.w3.org/TR/vocab-dcat-3/>

<sup>260</sup> <http://www.loc.gov/standards/premis/>

<sup>261</sup> [https://www.loc.gov/standards/premis/understanding\\_premis\\_german.pdf](https://www.loc.gov/standards/premis/understanding_premis_german.pdf)

<sup>262</sup> <https://fairsharing.org/>

<sup>263</sup> <https://rdamsc.bath.ac.uk/>

Empfehlungen für die Erstellung oder Anpassung von Metadatenprofilen

Kompetenzzentrum Interoperable Metadaten (KIM) / Rühle, Stefanie: Kleines Handbuch Metadaten, o. J.<sup>264</sup>

Guidelines for Dublin Core™ Application Profiles<sup>265</sup>

## 12.3 Empfehlungen für Vokabulare, Normdaten und Anwendungsontologien

Kontrollierte Vokabulare sind Systeme zur Organisation von Wissen, die eine strukturierte Menge von Begriffen zur Organisation und Klassifizierung von Daten enthalten, um den späteren Zugriff auf sie und die Suche nach ihnen zu gewährleisten. Zu ihnen gehören unter anderem Anwendungsontologien, Thesauri, Klassifikationen, Schlagwortlisten und Normdaten.

Während Ontologien die Ordnung der Informationskategorien mit ihren Eigenschaften und Relationen leisten, enthalten Thesauri, Klassifikationen und Normdaten die Datenwerte, die zur Identifikation und Benennung der Instanzen dienen, die eine konkrete Ausprägung einer Klasse darstellen. Hamburg ist z. B. eine Instanz der Klasse Ort. Für Allgemeinbegriffe, Individualnamen und andere Werte werden sie zur Belegung der innerhalb der Metadatenschemata definierten Datenelemente verwendet. Die zugehörigen PIDs sollten immer mit übernommen werden.

Alle empfohlenen Vokabulare sind selbst FAIR, denn sie verfügen über folgende Merkmale:

- Sie bieten PIDs für die nachgewiesenen Entitäten.

---

<sup>264</sup> [https://wiki.dnb.de/download/attachments/43523047/201209\\_metadaten.pdf](https://wiki.dnb.de/download/attachments/43523047/201209_metadaten.pdf)

<sup>265</sup> <https://dublincore.org/specifications/dublin-core/profile-guidelines/>

- Sie sind offen lizenziert und damit frei nachnutzbar.
- Die Datenangebot werden über Standardschnittstellen angeboten.
- Das Vokabular ist gut dokumentiert und mit einer Ontologie hinterlegt.
- Das Vokabular wird kontinuierlich gepflegt.
- Pflege und Ausbau liegen in einer klaren institutionellen Verantwortlichkeit (außer Geonames, Wikidata).

Mehr Informationen zu kontrollierten Vokabularen

Zaytseva, Ksenia / Durčo, Matej: Controlled Vocabularies and SKOS. Version 1.1.0. Edited by Matej Ďurčo and Tanja Wissik. DARIAH-Campus, [Training module], 2020<sup>266</sup>

## Kontrollierte Vokabulare / Normdaten

Für die in den Daten und Metadaten der Kulturwissenschaften und Kulturerbe-Sammlungen beschriebenen Klassen empfehlen wir die Verwendung folgender Standards für Datenwerte:

### Personen

Gemeinsame Normdatei (GND)<sup>267</sup>, Personen

GND-Explorer<sup>268</sup> der Deutschen Nationalbibliothek

Lobid-GND<sup>269</sup> des Hochschulbibliothekszentrums des Landes Nordrhein-Westfalen

Online-GND<sup>270</sup> des Bibliotheksservice-Zentrums Baden-Württemberg

Virtual International Authority File (VIAF)<sup>271</sup>

Wikidata<sup>272</sup>

Getty Union List of Artist Names (ULAN)<sup>273</sup>

<sup>266</sup> <https://campus.dariah.eu/id/D8d6OrLdpLIGRqBSQDVNO>

<sup>267</sup> <https://www.dnb.de/gnd>

<sup>268</sup> <https://explore.gnd.network/>

<sup>269</sup> <https://lobid.org/gnd>

<sup>270</sup> <https://www.bsz-bw.de/ognd.html>

<sup>271</sup> <https://viaf.org/>

<sup>272</sup> <https://www.wikidata.org/>

<sup>273</sup> <https://www.getty.edu/research/tools/vocabularies/ulan/>

CERL Thesaurus<sup>274</sup>

International Name Standard Identifier (ISNI)<sup>275</sup>

Forschende: Open Researcher and Contributor ID (ORCID)<sup>276</sup>

Körperschaften

Gemeinsame Normdatei (GND), Körperschaften

Virtual International Authority File (VIAF)

Wikidata

Getty Union List of Artist Names (ULAN)

CERL Thesaurus

International Name Standard Identifier (ISNI)

International Standard Identifier for Libraries and Related Organizations (ISIL)<sup>277</sup>

Lobid-Organisations<sup>278</sup> des Hochschulbibliothekszentrums des Landes

Nordrhein-Westfalen

Forschungsinstitutionen: Research Organization Registry (ROR)<sup>279</sup>

Forschungsförderer: Crossref Funder Registry<sup>280</sup>

Geografika

Gemeinsame Normdatei (GND), Geografika (außer Bauwerke und Denkmäler)

---

<sup>274</sup> [https://data.cerl.org/thesaurus/\\_search](https://data.cerl.org/thesaurus/_search)

<sup>275</sup> <https://isni.org/>

<sup>276</sup> <https://orcid.org/>

<sup>277</sup> <https://sigel.staatsbibliothek-berlin.de/suche>

<sup>278</sup> <https://lobid.org/organisations>

<sup>279</sup> <https://ror.org/>

<sup>280</sup> <https://www.crossref.org/documentation/funder-registry/>

Virtual International Authority File (VIAF)

Wikidata

GeoNames<sup>281</sup>

Thesaurus of Geographic Names (TGN)<sup>282</sup>

CERL Thesaurus

MARC Code List for Countries<sup>283</sup>

Werke, materielle Kulturerbe-Objekte

Gemeinsame Normdatei (GND), Werke

Virtual International Authority File (VIAF)

Wikidata

Cultural Objects Name Authority (CONA)<sup>284</sup>

Bauwerke, Monumente

Gemeinsame Normdatei (GND), Geografika - Bauwerke und Denkmäler

Virtual International Authority File (VIAF)

Wikidata

GeoNames

Cultural Objects Name Authority (CONA)

---

<sup>281</sup> <https://www.geonames.org/>

<sup>282</sup> <https://www.getty.edu/research/tools/vocabularies/tgn/>

<sup>283</sup> <https://id.loc.gov/vocabulary/countries.html>

<sup>284</sup> <https://www.getty.edu/research/tools/vocabularies/cona/>

## Sachschlagworte

Gemeinsame Normdatei (GND), Schlagworte sensu stricto, Schlagworte - Sprachen, Schlagworte - Ethnografika

## Wikidata

Art & Architecture Thesaurus (AAT)<sup>285</sup>

Darstellungsinhalte, Themen: Iconclass<sup>286</sup>

Musical Instrument Museums Online (MIMO) Thesaurus<sup>287</sup>

Hornbostel-Sachs Systematik der Musikinstrumente<sup>288</sup>

MARC Instruments and Voices Code List<sup>289</sup>

UNIMARC: Medium of performance<sup>290</sup>

Sachthesaurus Wortnetz Kultur<sup>291</sup>

Unesco Thesaurus<sup>292</sup>

Taxonomy of Digital Research Activities in the Humanities (TaDiRAH)<sup>293</sup>

## Rollen, Beziehungen

Rollen von Akteuren: Art & Architecture Thesaurus (AAT), Agents Facet

---

<sup>285</sup> <https://www.getty.edu/research/tools/vocabularies/aat/>

<sup>286</sup> <https://iconclass.org/>

<sup>287</sup> <http://www.mimo-db.eu/>

<sup>288</sup> <http://www.oberlin.edu/faculty/rknight/Organology/H-S-1914-German.pdf>

<sup>289</sup> <https://www.loc.gov/standards/valuelist/marcmusperf.html>

<sup>290</sup> <https://www.iflstandards.info/unimarc/terms/mop>

<sup>291</sup>

[https://www.lvr.de/de/nav\\_main/kultur/kult\\_land\\_\\_testkopie/das\\_team\\_1/dokumentation/digicult\\_xtree/digicult\\_xtree.jsp](https://www.lvr.de/de/nav_main/kultur/kult_land__testkopie/das_team_1/dokumentation/digicult_xtree/digicult_xtree.jsp)

<sup>292</sup> <https://vocabularies.unesco.org/browser/thesaurus/en/>

<sup>293</sup> <https://de.dariah.eu/web/guest/tadirah>

Rollen von Akteuren im Forschungsprozess: CRediT Taxonomy<sup>294</sup>

Rollen von Akteuren im Forschungsprozess: Taxonomy of Digital Research Activities in the Humanities (TaDiRAH)

Bezüge zwischen Ressourcen und Akteuren: MARC Relators<sup>295</sup>

Bezüge zwischen Werken/Objekten: LIDO Terminology<sup>296</sup>, RelatedWorkRelType

Kodierungen

Kodierungen für Zeit- und Datumsangaben

Date and Time Format ISO 8601<sup>297</sup>

Extended Date/Time Format (EDTF)<sup>298</sup>

Das auf Initiative der Library of Congress entwickelte Extended Date/Time Format bietet weitere Möglichkeiten, unsichere oder vage Angaben auszudrücken und ist in die aktuellste Version des Standards ISO 8601:2019 integriert worden.

Kodierungen für Sprachen

Language Codes ISO 639-1, 639-2<sup>299</sup>, 639-3<sup>300</sup>

Vokabular- und Normdatenangebote: Linked Open Data, Datenabzüge, Schnittstellen

CERL Thesaurus<sup>301</sup>

Crossref Funder Registry<sup>302</sup>

---

<sup>294</sup> <https://credit.niso.org/>

<sup>295</sup> <http://www.loc.gov/loc.terms/relators/>

<sup>296</sup> <http://terminology.lido-schema.org/>

<sup>297</sup> [https://en.wikipedia.org/w/index.php?title=ISO\\_8601&oldid=1131560436](https://en.wikipedia.org/w/index.php?title=ISO_8601&oldid=1131560436)

<sup>298</sup> <https://www.loc.gov/standards/datetime/>

<sup>299</sup> [https://www.loc.gov/standards/iso639-2/php/code\\_list.php](https://www.loc.gov/standards/iso639-2/php/code_list.php)

<sup>300</sup> <https://iso639-3.sil.org/>

<sup>301</sup> [https://www.cerl.org/resources/cerl\\_thesaurus/linkeddata](https://www.cerl.org/resources/cerl_thesaurus/linkeddata)

<sup>302</sup> <https://www.crossref.org/documentation/funder-registry/funder-data-via-the-api/>

EuroVoc<sup>303</sup>

GeoNames Datenabzug, Webservice<sup>304</sup> und Ontologie<sup>305</sup>

Getty Vokabulare (AAT, CONA, TGN, ULAN) Linked Data-Dokumentation<sup>306</sup> und Ontologie<sup>307</sup>

GND bei der Deutschen Nationalbibliothek<sup>308</sup> und bei Lobid<sup>309</sup>

Iconclass<sup>310</sup>

Hornbostel-Sachs-Systematik<sup>311</sup>

ISIL bei der Staatsbibliothek Berlin<sup>312</sup> und bei Lobid<sup>313</sup>

ISNI<sup>314</sup>

Library of Congress Linked Data Service<sup>315</sup> (EDTF, ISO 639, MARC Instruments and Voices Code List, MARC Relators, MARC List for Countries)

MIMO Thesaurus<sup>316</sup>

ORCID<sup>317</sup>

ROR<sup>318</sup>

<sup>303</sup> <https://data.europa.eu/data/datasets/eurovoc>

<sup>304</sup> <https://www.geonames.org/export/>

<sup>305</sup> <https://www.geonames.org/ontology/documentation.html>

<sup>306</sup> <https://www.getty.edu/research/tools/vocabularies/lod/>

<sup>307</sup> <http://vocab.getty.edu/ontology>

<sup>308</sup> <https://www.dnb.de/lids>

<sup>309</sup> <https://lobid.org/gnd/api>

<sup>310</sup> <https://iconclass.org/help/lod>

<sup>311</sup> <https://uri.gbv.de/terminology/hornbostelsachs/>

<sup>312</sup> <https://sigel.staatsbibliothek-berlin.de/suche/linked-data-service>

<sup>313</sup> <https://lobid.org/organisations/api/de>

<sup>314</sup> <https://isni.org/page/linked-data/>

<sup>315</sup> <https://id.loc.gov/>

<sup>316</sup> <https://vocabulary.mimo-international.com/InstrumentsKeywords/en/>

<sup>317</sup> <https://info.orcid.org/documentation/>

<sup>318</sup> <https://ror.readme.io/docs/rest-api>

TaDiRAH<sup>319</sup>

UNIMARC Vokabulare<sup>320</sup> (Medium of Performance)

UNESCO Thesaurus<sup>321</sup>

VIAF: Datenabzug<sup>322</sup>, Schnittstellendokumentation<sup>323</sup>

Wikidata<sup>324</sup>, s. Dokumentation unter Data Access

## Anwendungsontologien

Die Grundlage für eine semantische Auswertbarkeit der Daten ist ihre Hinterlegung mit einer Ontologie. Per Definition ist eine Ontologie ein formales Modell, das die Darstellung von Wissen für einen bestimmten Bereich ermöglicht. Sie beschreibt (a) die Arten von Dingen, die es gibt (Klassen), (b) die Beziehungen zwischen ihnen (Eigenschaften) und (c) die logischen Möglichkeiten, wie diese Klassen und Eigenschaften zusammen verwendet werden können (Axiome).

Es gibt zwei allgemeine Kategorien von Ontologien. „Referenzontologien“ (top-level ontologies) sind umfassende, axiomatische Modelle, deren Schwerpunkt auf der Präzisierung der vorgesehenen Bedeutungen von Begriffen in bestimmten Bereichen liegt. Sie werden verwendet, um die semantische Integration von Domänen-Ontologien zu erleichtern. Zu diesem Zweck enthalten sie allgemeine Kategorien, die auf mehrere Domänen anwendbar sind. Die für den Bereich der Kulturwissenschaften relevanten Referenzontologien finden Sie unter Empfehlungen für Rahmenwerke und Referenzmodelle.

„Anwendungsontologien“ (lightweight ontologies) werden für einen bestimmten Zweck oder Anwendungsschwerpunkt entwickelt. Sie bieten eine minimale terminologische

---

<sup>319</sup> <https://vocabs.dariah.eu/tadirah/en/>

<sup>320</sup> <https://www.iflstandards.info/unimarc>

<sup>321</sup> <https://vocabularies.unesco.org/>

<sup>322</sup> <https://viaf.org/viaf/data/>

<sup>323</sup> <https://www.oclc.org/developer/api/oclc-apis/viaf.en.html>

<sup>324</sup> [https://www.wikidata.org/wiki/Wikidata:Data\\_access](https://www.wikidata.org/wiki/Wikidata:Data_access)

Struktur, die auf die Bedürfnisse einer bestimmten Community oder eines Anwendungskontextes zugeschnitten ist. Die Anwendungsontologie verwendet häufig kanonische Ontologien oder bezieht sich auf diese, um ontologische Klassen und Beziehungen zwischen Klassen zu konstruieren. Anwendungsontologien werden häufig zu den Metadatenvokabularen gezählt, daher werden sie in diesem Abschnitt aufgeführt. Einige Vokabulare mit ontologieähnlichen Strukturen werden auch als „Schemata“ und nicht als „Ontologien“ bezeichnet.

Für die Kulturwissenschaften und Kulturerbe-Sammlungen empfehlen wir die Verwendung folgender Anwendungsontologien:

[Erlangen Conceptual Reference Model \(ECRM\)](http://erlangen-crm.org/)<sup>325</sup>

[Doing Reusable Musical Data \(DOREMUS\)](https://www.doremus.org/)<sup>326</sup>

[Music Ontology](http://purl.org/ontology/mo/)<sup>327</sup>

[Online Music Recognition and Searching II \(OMRAS2\) Chord Ontology](http://purl.org/ontology/chord/)<sup>328</sup>

[Ontology for Scientific Documentation of source-based 3D reconstruction of architecture \(OntSciDoc3D\)](https://www.ontscidoc3d.hs-mainz.de/ontology/)<sup>329</sup>

[Audio-visual rhetorics of affect \(AdA\) Filmontology](https://projectada.github.io/ontology/)<sup>330</sup>

Numismatik: [NOMISMA](http://nomisma.org/)<sup>331</sup>

[Schema.org](https://schema.org/)<sup>332</sup>

[Records in Contexts \(RiC\) Ontology](https://www.ica.org/en/records-in-contexts-ontology)<sup>333</sup>

---

<sup>325</sup> <http://erlangen-crm.org/>

<sup>326</sup> <https://www.doremus.org/>

<sup>327</sup> <http://purl.org/ontology/mo/>

<sup>328</sup> <http://purl.org/ontology/chord/>

<sup>329</sup> <https://www.ontscidoc3d.hs-mainz.de/ontology/>

<sup>330</sup> <https://projectada.github.io/ontology/>

<sup>331</sup> <http://nomisma.org/>

<sup>332</sup> <https://schema.org/>

<sup>333</sup> <https://www.ica.org/en/records-in-contexts-ontology>

Provenance Ontology (PROV-O)<sup>334</sup>

Web Annotation Vocabulary<sup>335</sup>

Nachweis und Linked Open Data-Angebote weiterer Wissensorganisationssysteme

Linked Open Vocabularies (LOV)<sup>336</sup>

Library of Congress Linked Open Data Service (ID.LOC.GOV)<sup>337</sup>

Basic Register of Thesauri, Ontologies & Classifications (BARTOC)<sup>338</sup>

Wissensorganisationssysteme im Gemeinsamen Bibliotheksverbund (GBV)<sup>339</sup> mit  
Vokabularserver DANTE<sup>340</sup>

FAIRsharing. A curated, informative and educational resource on data and metadata standards, inter-related to databases and data policies<sup>341</sup>

Forum on Information Standards in Heritage (FISH) Vocabularies<sup>342</sup>

xTree.public<sup>343</sup>

---

<sup>334</sup> <https://www.w3.org/TR/prov-o/>

<sup>335</sup> <https://www.w3.org/TR/annotation-vocab/>

<sup>336</sup> <https://lov.linkeddata.es/dataset/lov/>

<sup>337</sup> <https://id.loc.gov/>

<sup>338</sup> <https://bartoc.org/>

<sup>339</sup> <https://uri.gbv.de/terminology/>

<sup>340</sup> <https://api.dante.gbv.de/>

<sup>341</sup> <https://fairsharing.org/>

<sup>342</sup> <http://www.heritage-standards.org.uk/fish-vocabularies/>

<sup>343</sup> <http://xtree-public.digicult-verbund.de/vocnet/>

Terminologieservices zu Geografika

[gazetteers.net](https://gazetteers.net)<sup>344</sup>

[World Historical Gazetteer](https://whgazetteer.org/)<sup>345</sup>

## 13. Verwendete Literatur

Altenhöner, Reinhard / Berger, Andreas / Bracht, Christian / Klimpel, Paul / Meyer, Sebastian / Neuburger, Andreas / Stäcker, Thomas / Stein, Regine: DFG-Praxisregeln „Digitalisierung“, aktualisierte Fassung 2022 (dt.), <https://doi.org/10.5281/zenodo.7435724> und dies.: DFG Practical Guidelines on Digitisation, Updated version 2022 (engl.), <https://doi.org/10.5281/zenodo.7561148>

Cremer, Fabian / Klaffki, Lisa / Steyer, Timo: Der Chimäre auf der Spur: Forschungsdaten in den Geisteswissenschaften, in: O-Bib. Das Offene Bibliotheksjournal, Bd. 5/2, 2018, S. 142–162. <https://doi.org/10.5282/o-bib/2018H2S142-162>

Deutsche Forschungsgemeinschaft: DFG-Praxisregeln „Digitalisierung“, 2016, [https://www.dfg.de/formulare/12\\_151/](https://www.dfg.de/formulare/12_151/). Letzter Zugriff 28.02.2023

Digital Repository of Ireland (DRI), Research Data Alliance Ireland (RDA): Publishing GLAM data as FAIR data, Europeana Research Webinar, 2020, <https://www.rd-alliance.org/publishing-glam-data-fair-data>. Letzter Zugriff 28.02.2023

Easterday, Kelly / Paulson, Tim / DasMohapatra, Proxima / Alagona, Peter / Feirer, Shane / Kelly, Maggi: From the Field to the Cloud: A Review of Three Approaches to Sharing Historical Data From Field Stations Using Principles From Data Science, in: Frontiers in Environmental Science, 6, 2018, <https://doi.org/10.3389/fenvs.2018.00088>

---

<sup>344</sup> <https://gazetteers.net/>

<sup>345</sup> <https://whgazetteer.org/>

FORCE11: Guiding Principles for Findable, Accessible, Interoperable and Re-usable Data Publishing version b1.0, 2011/2021, <https://force11.org/info/guiding-principles-for-findable-accessible-interoperable-and-re-usable-data-publishing-version-b1-0/>. Letzter Zugriff 28.02.2023

Go FAIR, <https://www.go-fair.org/>. Letzter Zugriff 28.02.2023

Higman, Rosie / Bangert, Daniel / Jones, Sarah: Three Camps, One Destination: The Intersections of Research Data Management, FAIR and Open, in: Insights 32 (1), 18, <http://doi.org/10.1629/uksg.468>

Kraft, Angelina: Die FAIR Data Prinzipien für Forschungsdaten, in: TIB Blog, 2017, <https://blogs.tib.eu/wp/tib/2017/09/12/die-fair-data-prinzipien-fuer-forschungsdaten/>. Letzter Zugriff 28.02.2023

Hollander, Hella / Morselli, Francesca / Uiterwaal, Frank / Admiraal, Femmy / Trippel, Thorsten / Di Giorgio, Sara: PARTHENOS Guidelines to FAIRify data management and make data reusable, 2019, <https://doi.org/10.5281/zenodo.3368858>

PARTHENOS – Pooling Activities, Resources and Tools for Heritage E-research Networking, Optimization and Synergies, Horizon 2020-EU.1.4.1.1. (2015–2019), <https://www.parthenos-project.eu/>. Letzter Zugriff 28.02.2023

RDA FAIR Data Maturity Model Working Group: FAIR Data Maturity Model. Specification and guidelines (engl.), 2020, <https://doi.org/10.15497/RDA00050> und dies.: Das FAIR Data Maturity Model. Spezifikation und Leitlinien (dt.), 2020, <https://doi.org/10.5281/zenodo.5834115>

Schöch, Christof: Big? Smart? Clean? Messy? Data in the Humanities, in: Journal of Digital Humanities Volume 2, Issue 3, 2–13, 2013, <https://nbn-resolving.org/urn:nbn:de:bvb:20-opus-129492>

Smedt, Koenraad / Koureas, Dimitrios / Wittenburg, Peter: FAIR Digital Objects for Science: From Data Pieces to Actionable Knowledge Units, in: Publications, 2020, 8(2), 21, <https://doi.org/10.3390/publications8020021>

Wilkinson, Mark D. / Dumontier, Michel / Aalbersberg, IJsbrand Jan et al.: The FAIR Guiding Principles for scientific data management and stewardship, in: Scientific Data 3, 160018 EP, 2016, <https://doi.org/10.1038/sdata.2016.18>

Zeng, Marcia Lei / Qin, Jian: Metadata, 3. Aufl., 2022

## 14. Weiterführende Literatur

Baca, Murtha (Hg.): Introduction to Metadata. 3. Aufl., 2016, <https://www.getty.edu/publications/intrometadata/>. Letzter Zugriff 28.02.2023

Böker, E.: Metadaten und Metadatenstandards, in: Forschungsdaten.info, 14.11.2022, <https://forschungsdaten.info/themen/beschreiben-und-dokumentieren/metadaten-und-metadatenstandards/>. Letzter Zugriff 28.02.2023

Chue Hong, Neil P. / Katz, Daniel S. / Barker, Michelle / Lamprecht, Anna-Lena / Martinez, Carlos / Psomopoulos, Fotis E. / Harrow, Jen / Castro, Leyla Jael / Gruenpeter, Morane / Martinez, Paula Andrea / Honeyman, Tom / Struck, Alexander / Lee, Allen / Loewe, Axel / van Werkhoven, Ben / Jones, Catherine / Garijo, Daniel / Plomp, Esther / Genova, Francoise et al., RDA FAIR4RS WG.: FAIR Principles for Research Software (FAIR4RS Principles) (1.0), 2022, <https://doi.org/10.15497/RDA00068>

Engelhardt, Claudia / Biernacka, Katarzyna / Coffey, Aoife / Cornet, Ronald / Danciu, Alina / Demchenko, Yuri / Downes, Stephen / Erdmann, Christopher / Garbuglia, Federica / Germer, Kerstin / Helbig, Kerstin / Hellström, Margareta / Hettne, Kristina / Hibbert, Dawn / Jetten, Mijke / Karimova, Yulia / Kryger Hansen, Karsten / Kuusniemi, Mari Elisa / Letizia, Viviana et al.: D7.4 How to be FAIR with your data. A teaching and training handbook for higher education institutions (V1.2.1) [Computer software], 2022, <https://doi.org/10.5281/zenodo.6674301>

FAIRsFAIR. Fostering FAIR Data Practices In Europe, Horizon 2020-INFRAEOSC-2018-2020 Grant Agreement 831558, <https://www.fairsfair.eu/>.  
Letzter Zugriff 28.02.2023

FAIRsFAIR: Clearing some of the highest FAIR hurdles: PIDs, Metadata, and Semantic Interoperability for Researchers, Webinar, 04.02.2021,  
<https://www.fairsfair.eu/events/clearing-some-highest-fair-hurdles-pids-metadata-and-semantic-interoperability-researchers>. Letzter Zugriff 28.02.2023

Harpring, Patricia: Introduction to Controlled Vocabularies: Terminology for Art, Architecture, and Other Cultural Works, 2010,  
<https://www.getty.edu/publications/resources/virtuallibrary/160606018X.pdf>. Letzter Zugriff 28.02.2023

Harrower, Natalie / Maryl, Maciej / Biro, Timea / Immenhauser, Beat / ALLEA Working Group E-Humanities: Sustainable and FAIR Data Sharing in the Humanities: Recommendations of the ALLEA Working Group E-Humanities. Digital Repository of Ireland, 2020, <https://doi.org/10.7486/DRI.tq582c863>

Hugo, Wim / Le Franc, Yann / Coen, Gerard / Parland-von Essen, Jessica / Bonino, Luiz: D2.5 FAIR Semantics Recommendations Second Iteration (1.0), 2020,  
<https://doi.org/10.5281/zenodo.5362010>

Jettka, Daniel / Henny-Kramer, Ulrike: Leitfaden für die nachhaltige Entwicklung und Nutzung von Forschungssoftware, 2022, <https://nfdi4culture.de/go/E3332>. Letzter Zugriff 28.02.2023

Koster, Lukas: Persistent identifiers for heritage objects. Code4lib Journal. Issue 47, 2020, <https://hdl.handle.net/11245.1/93e75978-419d-431b-81e7-289974e8e0a6>

Martinez, Paula Andrea / Erdmann, Christopher / Simons, Natasha / Otsuji, Reid / Labou, Stephanie / Johnson, Ryan / Castelao, Guilherme / Villas Boas, Bia / Lamprecht, Anna-Lena / Martinez Ortiz, Carlos / Garcia, Leyla / Kuzak, Mateusz / Stokes, Liz / Honeyman, Tom / Wise, Sharyn / Quan, Josh / Peterson, Scott / Neeser, Amy / Karvovskaya, Lena et al.: Top 10 FAIR Data & Software Things, 2019,

<https://doi.org/10.5281/zenodo.3409968> und  
<https://librarycarpentry.org/Top-10-FAIR/>. Letzter Zugriff 28.02.2023

Neuroth, Heike / Oßwald, Achim / Scheffel, Regine / Strathmann, Stefan / Huth, Karsten (Hg.): Nestor-Handbuch: Eine kleine Enzyklopädie der digitalen Langzeitarchivierung. Version 2.3, 2010,  
<http://www.nestor.sub.uni-goettingen.de/handbuch/index.php>. Letzter Zugriff 28.02.2023

Poveda-Villalón, María / Espinoza-Arias, Paola / Garijo, Daniel / Corcho, Oscar: Coming to Terms with FAIR Ontologies, 2020,  
[https://www.researchgate.net/publication/344042645\\_Coming\\_to\\_Terms\\_with\\_FAIR\\_Ontologies](https://www.researchgate.net/publication/344042645_Coming_to_Terms_with_FAIR_Ontologies). Letzter Zugriff 28.02.2023

THOR – Technical and Human Infrastructure for Open Research, H2020-EU.1.4. - EXCELLENT SCIENCE - Research Infrastructures Grant Agreement ID: 654039 (2015-2017), Persistent Identifier Platform, <https://project-thor.readme.io/>. Letzter Zugriff 28.02.2023

## Danksagung

Die Autorin dieser Handreichung spricht folgenden Personen ihren Dank aus – für die vielen hilfreichen inhaltlichen Anregungen und Ergänzungen, aber auch für das Korrekturlesen und die technische Unterstützung bei der Publikation: Martin Albrecht-Hohmaier, Christian Bracht, Christoph Eggersglüß, Melanie Gruß, Sandra Göller, Jörg Heseler, Dietmar Kammerer, Celia Krause, Desiree Mayer, Frodo Podschwadek, Kristina Richts-Matthaei, Kerstin Soltau, Martha Stellmacher und Barbara Wiermann.