

FAIR-Prinzipien, Metadaten und Standards

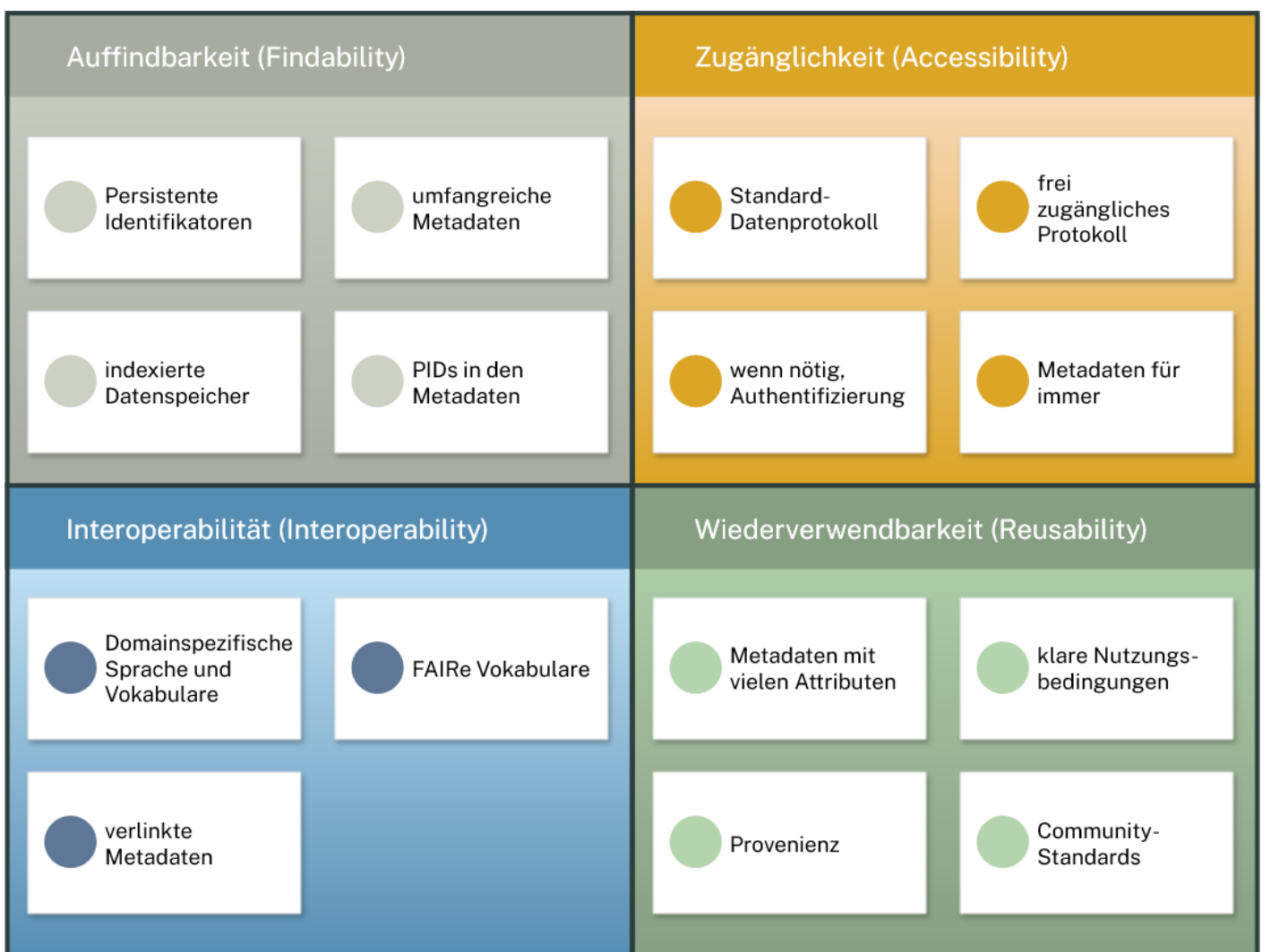
FAIR-Prinzipien, Empfehlungen zu Metadatenstandards wie EDM, LIDO und Klassifikationssystemen

- [FAIR-Prinzipien](#)
- [Metadatenstandards](#)
- [EDM, LIDO](#)
- [Für Bibliotheken: ISAD, MARC21, MODS, METS](#)
- [ISO-Standards: CIDOC CRM, Dublin Core](#)
- [TDWG-Standards: DwC, ABCD, AC](#)
- [Weitere Standards: EML, GGBN, Humboldt Core](#)
- [Zukünftige Standards: Lattimer Core, MIDS](#)
- [Standards für Digitalisate: IIF](#)

FAIR-Prinzipien

Die FAIR-Prinzipien garantieren die Nachhaltigkeit der Digitalisierungsvorhaben. Für die Erfüllung der FAIR-Prinzipien sind umfangreiche Metadaten, spezifische Sprache und Vokabulare sowie Lizenzangaben essenziell.

FAIRe Daten folgen den Prinzipien der **Findability** (Auffindbarkeit), **Accessibility** (Zugänglichkeit), **Interoperability** (Interoperabilität) und der **Reusability** (Wiederverwendbarkeit). Es geht nicht nur darum, dass diese Eigenschaften für Menschen erfüllt werden, sondern auch für Computer. FAIR bietet eine nützliche Struktur, um Daten für eine vielfältige Nutzung und Nachnutzung bereitzustellen.



Diese Darstellung fasst die verschiedenen Prinzipien für die Metadaten von digitalen Objekten zusammen, damit diese sowohl für Menschen als auch für Maschinen auffindbar, zugänglich, interoperabel und wiederverwendbar sind. Grafik: Kulturpool, [CC0](#)

Die Einhaltung der FAIR-Prinzipien ist wichtig für langfristig gute Datenqualität und eine gute Anbindung an den Kulturpool. Ihre Daten sollten auffindbar sein, was eindeutige Identifikatoren und reichhaltige Metadaten erleichtern. Zugänglichkeit wird durch eine gute Benutzeroberfläche, dauerhafte Verfügbarkeit und klare Lizenzinformationen gewährleistet. Interoperabilität bedeutet, dass Ihre Daten und Metadaten in standardisierten Formaten vorliegen, die eine reibungslose Integration und Nutzung durch andere Systeme ermöglichen. Wiederverwendbarkeit sichert, dass Daten unter klaren Bedingungen für verschiedene Zwecke nutzbar sind. Gestützt wird dies durch umfassende Metadaten, die eine kontextuelle und technische Verständlichkeit gewährleisten, sowie klare Auszeichnung der Nutzungsrechte.

Die FAIR-Prinzipien lassen sich auf ein digitalisiertes Objekt wie folgt auslegen:

Auffindbarkeit

- Das Objekt hat einen [persistenten Identifikator \(PID\)](#).
- Das Objekt wird mit umfangreichen, semantischen Metadaten beschrieben.
- Die Metadaten enthalten den persistenten Identifikator.
- Das Objekt und die Metadaten sind in einer öffentlich-zugänglichen Ressource verzeichnet.

Zugänglichkeit

- Das Objekt kann mit dem Identifikator über ein Standardprotokoll (z. B. http) abgerufen werden.
- Dieses Protokoll ist offen, kostenlos und universell implementierbar.
- Wenn erforderlich, erlaubt das Protokoll eine Zugangsauthentifizierung.
- Die Metadaten bleiben erhalten, auch wenn das digitalisierte Objekt nicht mehr auffindbar ist.

Interoperabilität

- Es wird ein verständliches und formelles Schema für die Darstellung (z. B. RDF) verwendet.
- Es werden Vokabulare verwendet, die den FAIR-Prinzipien entsprechen.
- Es werden qualifizierte Referenzen zu anderen Objekten und Metadaten in der Darstellung des Objekts verwendet.

Wiederverwendbarkeit

- Das Objekt wird durch eine Vielzahl an fehlerfreien und relevanten Metadaten beschrieben.
- Das Objekt besitzt eine klare Angabe zu den Nutzungsbedingungen.
- Die umfangreiche Provenienz wird in den Metadaten angegeben. Diese soll auch enthalten, wie dieses Objekt zu zitieren ist.
- Das Objekt und dessen Beschreibung werden durch korrekte und relevante Standards im betreffenden Bereich dargestellt.

Linktipps

- [nature.com: The FAIR Guiding Principles for scientific data management and stewardship](#) – FAIR-Prinzipien und deren Rolle in der Datenverwaltung und des Datenmanagements
- [GoFAIR: FAIR Principles](#) – FAIR-Prinzipien und deren detaillierte Auslegung in den Metadaten

- [Angela Kailus: Handreichung für ein FAIRES Management kulturwissenschaftlicher Forschungsdaten](#)
- [Medium: How to make your data Findable, Accessible, Interoperable, and Resuable?](#) – kurze Einleitung zu FAIR sowie eine Anleitung, um Daten FAIR zu machen
- [Crosslateral Enterprises: FAIR data assessment tool](#) – Assessment-Tool für einen Wert zur FAIRness der Datensätze



Di

e

.

Metadatenstandards

Durch das Verwenden eines Metadatenstandards wie EDM, LIDO oder DCMI werden Metadaten als beschreibende Informationen zu einem digitalisierten Objekt untereinander logisch verbunden und strukturiert. Unterschiedliche Standards haben sich in verschiedenen Bereichen etabliert.

Informationen, die ein Objekt beschreiben, werden **Metadaten** genannt, etwa Titel, Autor, Erstellungsdatum, Dateigröße. Metadatenstandards sind strukturierte und vereinheitlichte Formate, die verwendet werden, um die beschreibenden Informationen zu einem Objekt zu speichern und zu verwalten.

Unterschiedliche Domänen verwenden unterschiedliche **Metadatenstandards**, um die Metadaten auszutauschen. Metadatenstandards definieren, welche Arten von Informationen erfasst werden sollen und wie sie formatiert und organisiert werden. Sie helfen bei der korrekten Verwendung, Interpretation und dem Austausch der beschreibenden Objektdaten.

Vernetzungsinitiativen wie der Kulturpool oder Europeana machen die beschreibenden Informationen zu den Digitalisaten von unterschiedlichen Institutionen über ein zentrales Suchsystem auffindbar. Auf die digitalen Objekte wird dabei per Link in den Metadaten verwiesen. Die Originalobjekte und Digitalisate verbleiben somit auf den Servern der vernetzten Organisationen.

Der erste Standard, der sich hier international etablierte, war **Dublin Core**. Für das kulturelle Erbe hat Europeana den Standard **EDM (Europeana Data Model)** entwickelt. Dieser Standard stellt ein offenes, bereichsübergreifendes, Semantic-Web-basiertes Framework bereit und hat sich international für die Aggregation und den Austausch kultureller Daten bewährt.

Der Kulturpool orientiert sich am EDM-Standard. Da die Institutionen selbst den Metadatenstandard zur internen Verwendung auswählen und somit mehrere Metadatenstandards parallel möglich sind, werden die Daten vom Kulturpool entweder bereits im EDM-Standard an Europeana weitergegeben oder zuvor in den EDM-Standard übersetzt. Dieser Prozess der „Übersetzung“ nennt sich [? Mapping](#).

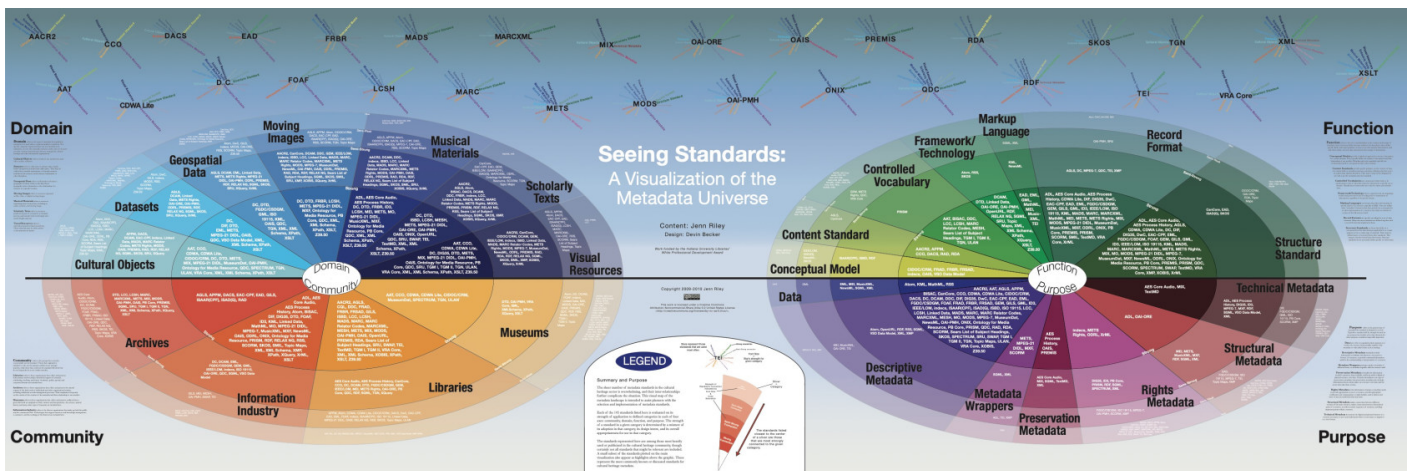
Bekannte Metadatenstandards

Es gibt viele verschiedene Metadatenstandards, welche für die Anforderungen in den unterschiedlichen Bereichen geschaffen wurden. Bekannte Standards sind zum Beispiel:

- Der [Dublin Core \(DC\)](#) ist ein Metadatenstandard zur Beschreibung von Dokumenten und Objekten im Internet. Er war der erste Metadatenstandard, der sich international etablieren konnte.
- Das [Europeana Data Model \(EDM\)](#) wurde entwickelt, um möglichst viele Objekte aus unterschiedlichen Bereichen mit einem Standard beschreibbar zu machen. Dabei bedient es sich

verschiedener Metadatenfelder aus anderen Metadatenstandards. Der Kulturpool verwendet diesen Metadatenstandard, um eine Weitergabe der Daten an Europeana zu gewährleisten.

- Für Museums- und Sammlungsobjekte wird oftmals der **LIDO**-Standard (Lightweight Information Describing Objects) verwendet.



Seeing Standards – A Visualization of the Metadata Universe: Eine (visuelle) Auflistung von 105 Metadatenstandards und der Bereiche, in denen sie verwendet werden.

Linktipp

- [Julian Palacz: Einführung in die Metadatenstandards \(Kulturpool Webinar\)](#)



u
nt
er

EDM, LIDO

Der Kulturpool orientiert sich am EDM-Standard von Europeana. Die Pflichtfelder im EDM-Standard müssen bei der Übermittlung von Daten in den Kulturpool in jedem Fall ausgefüllt sein, darüberhinaus können noch weitere optionale Felder damit abgedeckt werden.

Europeana Data Model (EDM)

Das [Europeana Data Model \(EDM\)](#) zielt darauf ab, die Integration, Zugänglichkeit und gemeinsame Nutzung von kulturellen und wissenschaftlichen Sammlungen in Europa zu verbessern. EDM ermöglicht die Beschreibung von Objekten, deren digitalen Repräsentationen und den Kontexten, in denen sie stehen. Durch die Anwendung von [?Linked Open Data](#) Prinzipien unterstützt das Modell die Verknüpfung von Sammlungen aus verschiedenen Quellen und fördert so die Entstehung eines vernetzten digitalen Kulturraums.

EDM unterstützt die Darstellung komplexer Beziehungen zwischen Objekten, Sammlungen, Ereignissen und Orten und ermöglicht so ein tiefgehendes Verständnis der digital erfassten Kulturgüter. Mit seiner Fähigkeit, vielfältige Datenquellen in einem einheitlichen Format zu aggregieren, spielt das Europeana Data Model eine Schlüsselrolle in der digitalen Kulturerbe-Landschaft Europas und trägt zur Erhaltung und Verbreitung des kulturellen Erbes bei.

[? Europeana Data Model \(EDM\)](#)

Lightweight Information Describing Objects (LIDO)

[LIDO](#) ist das Ergebnis einer 2008 begonnenen Zusammenarbeit zwischen internationalen Akteuren des Museumssektors, um eine gemeinsame Lösung für die Bereitstellung von Inhalten des Kulturerbes für Portale und andere aggregierte Repositorien zu schaffen. Als Anwendung des CIDOC Conceptual Reference Model (CIDOC CRM) bietet es ein explizites Format für die standardisierte Bereitstellung von Informationen über Museumsobjekte.

LIDO unterstützt das gesamte Spektrum an beschreibenden Informationen zu Sammlungsobjekten und kann für alle Arten von Objekten, wie Kunst, Architektur, Kulturgeschichte oder sogar Naturgeschichte, verwendet werden. Das Schema unterstützt mehrsprachige Portale durch die Verwendung eines eigenen Sprachattributs und gewährleistet so internationale Interoperabilität.

LIDO ermöglicht die Veröffentlichung von Sammlungsdaten in Multi-Source-(Web-)Datenbanken, unabhängig von der in der Einrichtung verwendeten Katalogisierungssoftware. Durch den Einsatz von LIDO können die Daten in einem einheitlichen, standardisierten Format zur Verfügung gestellt und von mehreren Aggregatoren und externen Repositorien genutzt werden. Ohne weitere Transformation oder Anpassung ermöglicht es die Interoperabilität auf der Ebene der Datenstruktur aus verschiedenen Quellen innerhalb eines Webportals.

Linktipp

- [Bildarchiv Foto Marburg: Serie LIDO-Handbuch](#)



ell

t.

Für Bibliotheken: ISAD, MARC21, MODS, METS

Internationaler Standard für die archivische Beschreibung ISAD(G)

[ISAD\(G\)](#) steht für International Standard Archival Description und stellt einen internationalen Anwendungsstandard für die Beschreibung von Archivgut dar, der im Jahr 2000 vom International Council of Archives (ICA) verabschiedet wurde. Der Standard ermöglicht die Verwaltung und Erschließung von Informationen über die logischen Einheiten von Archivgut durch Beschreibungselemente, die nach Informationsbereichen geordnet sind. Die Beschreibung folgt damit einheitlichen, sachgerechten und unmittelbar verständlichen Kriterien, erleichtert die Ermittlung und den Austausch von Informationen über Archivgut und ermöglicht den Austausch von übergreifenden, allgemeinen Informationen. Der ISAD(G) dient als Leitfaden für die archivische Beschreibung unter Verwendung bereits bestehender nationaler Standards, als Instrument für den internationalen Austausch von Beschreibungsinformationen und als ein Standard, der unabhängig von Form und Medium angewendet werden kann. In Kombination mit dem ebenfalls vom Internationalen Archivrat (ICA) verabschiedeten [International Standard Archival Authority Record for Corporate Bodies, Persons, and Families \(ISAAR\(CPF\)\)](#) kann dem Benutzer die Zuverlässigkeit und Authentizität der Dokumente garantiert werden. Basierend auf diesem Standard folgt eine Empfehlung des Verbands Österreichischer Archivare (VÖA) dem Prinzip der mehrstufigen Beschreibung. Die mehrstufige Beschreibung bezieht sich auf ein hierarchisches Modell, dessen Ebenen vom Archiv über die Sammlungsgruppen (Abteilungen), Sammlung, Untersammlung, Serien, Unterserien und Akt/Konvolut bis zum Einzelstück reichen.

Die ISAD(G) in der vorliegenden Form ist die Grundlage für die „[Encoded Archival Description](#)“ (EAD), die als technisches Werkzeug für die Online-Darstellung von Beschreibungsergebnissen dienen kann – dies unterstützt die Verwendung von EAD als internationales Austauschformat für Online-Findmittel.

Die durch den ISAD(G)-Standard definierten Datenstrukturen und -elemente werden durch verschiedene ISO-zertifizierte Normen ergänzt und runden den präzisen und bedarfsgerechten Informationsbedarf ab.

[ISO 15489](#): Allgemeine Internationale Norm für ordnungsgemäße aktenbasierte Geschäftsverwaltung. Garantiert dem Urheber/Dateiersteller selbst Zuverlässigkeit und Authentizität (Geschäftsverwaltung)

[ISO 999](#): Enthält Richtlinien für den Inhalt, die Anordnung und die Darstellung von Verzeichnissen von Büchern, Zeitschriften, Berichten, Patentdokumenten und anderen schriftlichen Dokumenten, einschließlich nicht gedruckter Materialien wie elektronische Dokumente, Filme, Ton- und Videoaufnahmen

[ISO 5963](#): Methoden zur Prüfung von Dokumenten, zur Bestimmung ihres Gegenstands und zur Auswahl von Indexbegriffen

[ISO 25964](#): Richtlinien für die Erstellung und Entwicklung einsprachiger Thesauri

- [ISDF: International Standard for Describing Functions \(Allgemeine Internationale Standard-Archivbeschreibung\)](#)
- [Umsetzungsempfehlung der VÖA \(PDF\)](#)

MARC21 – Machine Readable Cataloging

Die [MARC21-Formate](#) definieren Standards für die Darstellung und Übertragung von bibliographischen Informationen in maschinenlesbarer Form. Der Datenaustausch erfolgt entweder über die beiden Austauschformate [ISO 2709](#) und [ANSI / NISO Z39.2](#) oder über XML. MARC ist ein „bibliografisches Format“ für Bücher, Tonträger, bewegte Bilder und Archivbestände und ist der am weitesten verbreitete Standard in Bibliotheken und Archiven.

Der MARC-Datensatz besteht aus drei Elementen: der Datensatzstruktur, dem Inhaltsetikett und dem Dateninhalt des Datensatzes. Die Datensatzstruktur ist eine Implementierung des internationalen Standards Format for Information Exchange (ISO 2709) und seines amerikanischen Gegenstücks Bibliographic Information Interchange (ANSI/NISO Z39.2). Der Content Identifier – die Codes und Konventionen, die ausdrücklich zur Identifizierung und weiteren Charakterisierung der Datenelemente in einem Datensatz und zur Unterstützung der Bearbeitung dieser Daten festgelegt wurden – wird von jedem der MARC-Formate definiert. Der Inhalt der Datenelemente, aus denen ein MARC-Datensatz besteht, wird normalerweise durch Standards außerhalb der Formate definiert. Beispiele hierfür sind die [International Standard Bibliographic Description \(ISBD\)](#), anglo-amerikanische Katalogisierungsregeln, [Library of Congress Subject Headings \(LCSH\)](#) oder andere Katalogisierungsregeln, Thesauri und Klassifizierungspläne, die von der Organisation verwendet werden, die einen Datensatz erstellt. Der Inhalt bestimmter kodierter Datenelemente ist in den MARC-Formaten definiert (z. B. Führer, Feld 007, Feld 008).

Metadata Object Description Standard (MODS)

Das [Metadata Object Description Schema \(MODS\)](#) ist ein XML-Format für bibliografische Metadaten. Es wird als XML-Schema vom Network Development and MARC Standards Office der [Library of Congress](#) entwickelt und gepflegt und ist seit Januar 2018 in der Version 3.7 verfügbar.

[MODS](#) wurde als Kompromiss zwischen der Komplexität von [MARC](#) (einem sehr detaillierten bibliografischen Austauschformat) und Dublin Core entwickelt, das für viele Anwendungen zu einfach ist. Daher gibt es für viele – aber nicht für alle – MARC-Felder in MODS entsprechende XML-Elemente. Ebenso gibt es in MODS Elemente, die nicht mit MARC kompatibel sind. Im Gegensatz zu den MARC-Nummerncodes wurden die Elemente in Englisch benannt. Die Library of Congress stellt eine Reihe von XSLT-Skripten für die Konvertierung zwischen Dublin Core, MARC und MODS zur Verfügung.

Metadata Encoding and Transmission Standard (METS)

Das [METS-Format](#) ist ein XML-Format für die strukturierte Beschreibung von digitalen Objekten. Mit METS können Informationen von der Herkunft des Objekts bis zur internen Struktur abgebildet werden. Für allgemeine Metadaten wird METS mit Formaten wie [MODS](#) kombiniert.

In einem digitalen Archiv müssen neben den digitalen Objekten auch die Metadaten zu diesen Objekten verwaltet werden. Diese Metadaten können neben der Objektbeschreibung (z. B. Titel, Erstellungsdatum) auch Informationen über das Dokument im Archiv enthalten (z. B. welcher Archivar es wann ins Archiv gestellt hat). METS bietet eine Struktur für diese Informationen.

METS kann verwendet werden, um hierarchische und auf andere Art strukturierte Werke zu kodieren. Ein Werk kann z. B. eine Buchreihe, ein einzelnes Buch mit Kapitel- und Seitenstruktur oder ein Film mit mehreren Szenen sein.



er

ei

tg

ISO-Standards: CIDOC CRM, Dublin Core

CIDOC Conceptual Reference Model (CIDOC CRM)

[Das CIDOC Conceptual Reference Model \(CIDOC CRM\)](#) stellt einen ISO-zertifizierten Standard für die Informationsintegration im Bereich des kulturellen Erbes dar, der über 20 Jahre hinweg entwickelt wurde. Das Modell bietet einen gemeinsamen und erweiterbaren semantischen Rahmen für die evidenzbasierte Integration von Informationen und ermöglicht die Integration von Daten aus verschiedenen Quellen auf eine software- und schemaunabhängige Weise. Der Datenaggregationsdienst [ARIADNEplus](#) basiert auf dem CIDOC-CRM und nutzt diesen Standard, um über entsprechende CRM-Erweiterungen wie [CRMarchaeo](#) (Ausgrabungen), [CRMtex](#) (antike Texte) oder [CRMba](#) (archäologische Gebäude) eine Vielzahl von fachspezifischen Daten zu integrieren.

Dublin Core (DC)

[Dublin Core \(DC\)](#)[™] ist ein Metadatenschema zur Beschreibung von elektronischen Ressourcen. Es handelt sich dabei um eine Sammlung von einfachen und standardisierten Konventionen zur Beschreibung von Dokumenten und anderen Objekten im Internet, die mithilfe von Metadaten leichter auffindbar gemacht werden sollen. Das Schema wurde von der [Dublin Core Metadata Initiative \(DCMI\)](#) erstellt und gepflegt. Der Metadatenstandard enthält 15 Kernfelder, die als Dublin Core Metadata Element Set, Version 1.1, ISO-zertifiziert wurden.



si

n

d

TDWG-Standards: DwC, ABCD, AC

Die TDWG-Standards wurden von der Taxonomic Database Working Group erstellt und werden kontinuierlich evaluiert und an aktuelle Entwicklungen angepasst.

DarwinCore (DwC)

[DarwinCore \(DwC\)](#) ist eine Sammlung von Standards für den Austausch von Biodiversitätsdaten. Es handelt sich um eine Reihe von Begriffen mit klar definierter Semantik, die von Menschen verstanden und von Maschinen interpretiert werden können. DwC wurde von der TDWG (Taxonomic Database Working Group) entwickelt und gepflegt. Sie basiert in erster Linie auf Taxa und ihrem Vorkommen in der Natur, wie es durch Beobachtungen und Proben und damit verbundene Informationen dokumentiert ist.

ABCD – Zugang zu biologischen Sammlungsdaten

[Access to Biological Collection Data \(ABCD\)](#) ist ein sich entwickelnder, umfassender Standard für den Zugriff auf und die gemeinsame Nutzung von Primärdaten zur biologischen Vielfalt. Das Schema unterstützt eine Vielzahl von Datenbanken und ist mit mehreren bestehenden Datenstandards wie DarwinCore kompatibel. ABCD wird u. a. von Datenaggregator-Netzwerken wie GBIF (Global Biodiversity Information Facility) und BioCAsE (Biological Collection Access Service) verwendet. Es handelt sich um ein hochkomplexes XML-Datenschema mit etwa 1.000 Elementen, das nahezu alle Informationen einer naturkundlichen Probensammlung oder einer Beobachtungsdatenbank speichern kann und gleichermaßen für lebende wie konservierte Objekte verwendet wird.

ABCDEFG – Erweiterter Zugang zu biologischen Sammlungsdatenbanken für Geowissenschaften

[Access to Biological Collection Databases Extended for Geosciences \(ABCDEFG\)](#) ist ein XML-Schema, das für die Verarbeitung paläontologischer, mineralogischer und geologischer digitalisierter Sammlungsdaten entwickelt wurde. Es stellt eine Erweiterung des bestehenden ABCD-Schemas dar und ergänzt die bestehenden Elemente um Daten, die sich auf geologische Beobachtungen beziehen. Mithilfe der Software BioCAsE Provider können die Daten zwischen der lokalen Quelle und dem GeoCAsE-Portal in Form von XML-Dokumenten gesendet und empfangen werden. Der Standard wird auch von der TDWG-Gruppe (Taxonomic Databases Working Group) verwaltet und soll so bald wie möglich für die Verwendung in GBIF übernommen werden.

AudubonCore (AC)

[Der Audubon Core \(AC\)](#) ist eine Sammlung von Begriffen zur Darstellung von Metadaten für Multimedia-Ressourcen und Sammlungen mit Bezug zur biologischen Vielfalt und wurde 2013 als TDWG-Standard ratifiziert. Ziel ist es, Informationen zu präsentieren, die dabei helfen, vor dem Erwerb einer bestimmten Ressource oder Sammlung festzustellen, ob diese für eine bestimmte Anwendung in der Biodiversitätswissenschaft geeignet ist. Die Vokabulare befassen sich unter anderem mit Fragen der Verwaltung der Medien und Sammlungen, Beschreibungen des Inhalts, ihrer taxonomischen, geografischen und zeitlichen Abdeckung und den geeigneten Möglichkeiten, sie abzurufen, zu klassifizieren und zu reproduzieren. Die verwendeten Begriffssammlungen werden aus verschiedensten Quellen wie dem DarwinCore-Vokabular zusammengestellt und bei Bedarf kontinuierlich erweitert.



st
ell
t.

Weitere Standards: EML, GGBN, Humboldt Core

Ecological Metadata Language (EML)

Die [Ecological Metadata Language \(EML\)](#) bietet ein umfassendes Vokabular und eine lesbare XML-Syntax für die Dokumentation von Forschungsdaten. Sie ist in den Erd- und Umweltwissenschaften weit verbreitet und wird zunehmend auch in anderen Forschungsdisziplinen eingesetzt. Der EML-Standard wird vom Knowledge Network for Biocomplexity (KNB) verwaltet und gepflegt. Der Standard umfasst Module zur Identifizierung und Zitierung von Datenpaketen, zur Beschreibung des räumlichen, zeitlichen, taxonomischen und thematischen Umfangs der Daten, zur Beschreibung von Forschungsmethoden und -protokollen sowie zur Annotation mit semantischen Vokabularen. Jedes Darwin Core-Archiv enthält eine EML-Datei als Komponente.

Global Genome Biodiversity Network Standard (GGBN)

Um den Austausch von Informationen über genomische Proben und die daraus gewonnenen Daten zu erleichtern, wurde auf der Grundlage des ABCDDNA-Standards ein neuer, umfassenderer Datenstandard, der [GGBN-Standard](#), für die Sammlung von DNA-, Gewebe- und Umweltproben entwickelt. Es handelt sich dabei nicht um einen eigenständigen Standard, da er nur in Verbindung mit DarwinCore oder ABCD verwendet werden kann und die dort vorhandenen Informationen um Details zu Definitionen, Beschreibungen, Qualität, unterstützenden Informationen und zugehörigen Metadaten für genomische Proben ergänzt.

Humboldt Core

Der [Humboldt Core](#) stellt einen Datenstandard für bestandsbezogene Informationen dar. Die Inventardaten enthalten Informationen wie z. B. die Erfassung mehrerer Arten für einen bestimmten Ort und eine bestimmte Zeit, was wiederum Aufschluss über die Charakterisierung der Biodiversität und deren Veränderung im Laufe der Zeit gibt. Die Möglichkeit, mehrere verschiedene Bestandsdatenbanken miteinander zu vergleichen, ermöglicht eine bessere und umfassendere Nutzung dieser sammlungsbezogenen Daten.

Das Projekt [Map of Life](#) diente als Pilotprojekt, um die Machbarkeit der Umsetzung des neu entwickelten Datenstandards zu testen. Die Aufnahme und Bereitstellung von Metadaten zu Inventarisierungsprozessen aus dem Map of Life-Datenaggregator ermöglicht die Aggregation von Inventarisierungsinformationen aus Hunderten von zuvor veröffentlichten Inventaren. Der Humboldt Core war ursprünglich als ein von der TDWG ratifizierter Standard geplant, wurde aber noch nicht umgesetzt. Es wird eine breitere Nutzung angestrebt, um mehr Input über die Nutzbarkeit des Standards und seine Kompatibilität mit bestehenden Datenstandards wie DarwinCore

zu erhalten. Zu diesem Zweck wurde mit der Taxonomic Database Working Group (heute [Biodiversity Information Standards](#)) eine Arbeitsgruppe eingerichtet, die sich mit der möglichen Integration des Humboldt Core als Erweiterung des DarwinCore-Standards befasst.



Di

e

.

Zukünftige Standards: Lattimer Core, MIDS

Aufgrund der rasch fortschreitenden Digitalisierung natur- und geowissenschaftlicher Sammlungen werden ständig neue und angepasste Standards entwickelt. Einige dieser Entwicklungen sind hier beispielhaft aufgeführt.

Lattimer Core – Collection Descriptions (CD)

Die Arbeitsgruppe [Collections Descriptions \(CD\)](#) der TDWG (Taxonomic Database Working Group) ist aus der Gruppe Natural Collection Descriptions (NCD) hervorgegangen, die im Laufe ihrer Arbeit einen erhöhten Bedarf an detaillierteren Sammlungsinformationen festgestellt hat. Die Mitglieder der Arbeitsgruppe entwickeln einen Datenstandard für die Beschreibung ganzer Sammlungen naturkundlicher Materialien. Dazu gehören Objektsammlungen, Beobachtungsdaten, visuelle Ressourcen, Fotos und andere Daten, die auf vielen Entdeckungsreisen gesammelt wurden. Ein CD-Datensatz sollte Informationen über die Sammlung, den Zugang und die Nutzung der Sammlung sowie einen Verweis auf detaillierte Informationsquellen enthalten. Derzeit werden die Daten auf Sammlungsebene über den EML-Standard (Ecological Metadata Language) verarbeitet. Da die über die EML-Datei verfügbaren Daten auf Sammlungsebene nicht ausreichen, um den Datenbedarf für institutionelle, regionale, nationale oder internationale Planungsaktivitäten zu decken, werden umfassendere Informationen über Sammlungsbestände und die sie beherbergenden Institutionen benötigt, um strategische Prioritäten für die Mobilisierung von Sammlungsdaten zu setzen. Alle Informationen auf Sammlungsebene müssen derzeit weltweit von Menschen aktualisiert werden und können daher schnell veralten und bei unterschiedlichen Ressourcen nicht mehr synchronisiert werden. Die CD-Arbeitsgruppe wurde initiiert, um der Weiterentwicklung der Datenaggregationsplattform GBIF (Global Biodiversity Information Facility) Rechnung zu tragen, die in Zukunft auch Daten auf Sammlungsebene zur Verfügung stellen möchte. Mit der Einführung des CD-Standards soll es in Zukunft möglich sein, aus den Daten beispielsweise abzulesen, welcher Teil einer Sammlung digitalisiert oder georeferenziert ist.

Minimum Information about a Digital Specimen (MIDS)

Die [Minimal Information about a Digital Specimen \(MIDS\)](#) wird von einer Untergruppe der Collections Descriptions (CD) Interest Group der TDWG (Taxonomic Database Working Group) erstellt und verwaltet. Sie definieren derzeit den Digitalisierungsgrad eines digitalisierten Sammlungsobjekts auf vier verschiedenen Ebenen und geben so Auskunft über den Detailgrad der verfügbaren Informationen. Der Standard kann dazu beitragen, den größtmöglichen Nutzen aus den verfügbaren Daten zu ziehen und eine einheitliche Messung des erreichten Digitalisierungsgrades und seiner Entwicklung aufzuzeigen. Der Standard befindet sich derzeit noch in der Entwicklung.

Im Rahmen des [CETAF Collections Digitisation Dashboard \(CDD\)](#), das durch das Projekt SYNTHESIS+ ins Leben gerufen wurde, wurden bereits Informationen auf MIDS-Ebene sowie zuverlässige und aktuelle

Informationen über den taxonomischen und geografischen Umfang sowie den Digitalisierungsgrad der Sammlungen erfasst und visualisiert.

Das CDD soll als dynamisches und visuelles Bewertungsinstrument für die Entscheidungsfindung dienen, um Prioritäten bei der Digitalisierung zu setzen und die Auffindbarkeit der europäischen naturwissenschaftlichen Sammlungen (NSCs) zu verbessern. Durch die konsequente Pflege des Dashboards und die daraus resultierende Vergleichsbasis für den Digitalisierungsfortschritt erfüllt das CDD die Anforderungen des DiSSCo-Projekts zur Messung und Priorisierung von Digitalisierungsinitiativen.



ell

t.

Standards für Digitalisate: IIF

International Image Interoperability Framework (IIF)

Das [International Image Interoperability Framework \(IIF\)](#) (Englisch ausgesprochen: *triple-eye-eff*) standardisiert die Bereitstellung von Bildern und audiovisuellen Daten von Servern in einer Vielzahl von Webumgebungen. Die IIF-Spezifikationen stehen im Einklang mit den allgemeinen Webstandards und funktionieren daher mit allen wichtigen Browsern. Sie ermöglichen umfangreiche Zusatzfunktionen wie das Zoomen/Vergleichen/Strukturieren und Kommentieren von Bildern oder das Hinzufügen von Untertiteln, Transkriptionen oder Übersetzungen sowie von Anmerkungen zu Audio-/Videoinhalten.

Vorteile von IIF

- **Interoperabilität:** IIF ermöglicht Austausch von Bildmaterial zwischen verschiedenen Plattformen ohne Kompatibilitätsprobleme.
- **Hohe Auflösung:** IIF unterstützt die Anzeige von hochauflösenden Bildern, was besonders für Details in Kunstwerken oder Manuskripten interessant ist.
- **Anpassbare Bildanzeige:** Im IIF-Viewer können spezifische Bereiche eines Bildes direkt für die Ansicht vergrößert und in Helligkeit und Kontrast angepasst werden.
- **Annotationen:** IIF unterstützt die Annotation von Bildern, was für Bildung und Forschung wertvoll sein kann.

Im Kulturpool ist IIF bereits implementiert und die digitalen Objekte können in der IIF-Ansicht betrachtet werden.

Linktipps

- [International Image Interoperability Framework \(IIF\): Why IIF?](#)



s

er

S