



Metadaten an der BBAW

Karoline Lemke

Markus Schnöpf

- 20 Minuten Informationen über Metadaten (Karoline Lemke)
- 20 Minuten Übung
- 10 Minuten Auswertung
- 20 Minuten Informationen (Markus Schnöpf)



Metadaten

Workshop zur Einführung

Teil 1

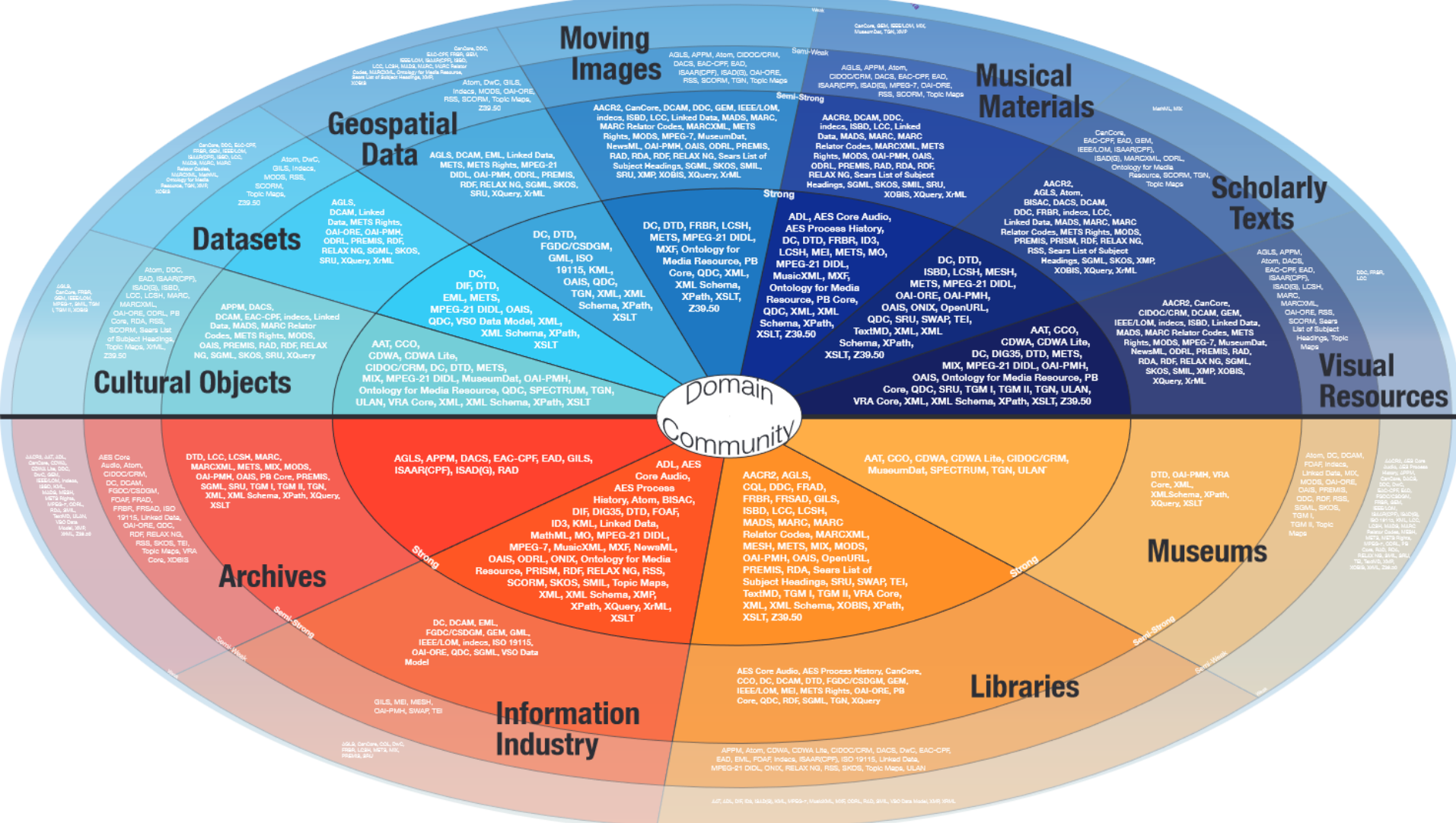
Karoline Lemke (Text+)

9. Juli 2024



1. Von Metadaten und wo sie zu finden sind
2. Metadatenformate und Standards
3. Metadatenformate der BBAW
4. Eine kurze Anleitung zur Beschreibung eines Metadatensets

Das Metadatenuniversum – ein Ausschnitt



Definition von der National Information Standards Organization (NISO) 2007

Metadaten sind strukturierte Informationen, die beschreiben, erklären, verorten oder in anderer Form etwas anderes repräsentieren.

Metadaten sind keine statische Einheit. Sie haben verschiedene Funktionen und stehen immer im Kontext anderer „Einheiten“. (Király/Brase 2021, S. 360)

Metadattentypen variieren je nach ihrer Funktion:

- » bibliografische Metadaten
- » administrative Metadaten
- » Prozessmetadaten
- » inhaltsbeschreibende bzw. deskriptive Metadaten

Metadattentypen werden in verschiedenen Standards ausgedrückt, z. B. Metadata Encoding and Transmission Standard (METS) (Containerstandard)

- » Kopfteil
- » Erschließungsangaben
- » Verwaltungsangaben
- » Dateiabschnitt
- » Strukturbeschreibung
- » Strukturverknüpfung
- » Verhalten

- » Katalog einer Forschungs- oder Bibliothekseinrichtung, die häufig auf Domänen spezialisiert sind
- » Verbundkataloge
- » Katalog mehrerer Verbünde, z. B. [Europeana](#), [WorldCat](#)
- » in Text+ z.B. die Registry für digitale Editionen, lexikalische Ressourcen, Sammlungen und Tools
- » NFDI-weit z.B. DALIA, ein Metadatenkatalog für Lernressourcen

Nach der Definition der Deutschen Digitalen Bibliothek (Glossar) 2015

Formale Beschreibung der Metadatenelemente, die für einen bestimmten Zweck entwickelt wurden.

Das Metadatenformat bestimmt die Struktur, in der die Metadatenelemente organisiert sind und legt fest, wie sie in maschinenlesbarer Form codiert werden.

Metadatenformate können zu Standards etabliert werden.

Werden Metadatenformate für eine bestimmte Nutzung angepasst, handelt es sich um ein Anwendungsprofil.

- » BibTex
- » Categories for the Description of Works of Art (CDWA)
- » Data Documentation Initiative (DDI)
- » Dublin Core
- » Lightweight Information Describing Objects (LIDO)
- » Extensible Metadata Platform (XMP)

usw.

Kataloge für Standards

- » [Metadata Standards Catalog der RDA](#)
- » [RDA Metadata Standards Directory Working Group](#)
- » [DDC \(Digital Curation Centre\)](#)

Beispiele für Strukturstandards

- » Dublin Core, DataCite, PREMIS, MARC
- » TEI, LIDO, MEI, CEI
- » METS, MODS, MADS
- » Exif, XMP, MIX, DDI

USW.

= Correspondence Metadata Interchange Format

- » Grundlage für die Erstellung eines digitalen Briefverzeichnisses mit dem Ziel digitale und gedruckte Briefeditionen nach Absender, Empfänger, Schreibort und -datum zu durchsuchen
- » baut auf TEI-Richtlinien auf mit dem Schwerpunkt auf Briefmetadaten
- » Aggregation und Auswertung von TEI-XML-Dateien
- » grundlegend: correspDesc-Element, das aus 11 Elementen besteht ([Dokumentation](#)), die Ort, Datum, Sender und Empfänger beschreiben
- » Ort dieser Metadaten: TEI/teiHeader/profileDesc
- » Datenlieferant muss Informationen über die CMIF-Datei und den digitalen Index der Briefe im teiHeader/fileDesc-Element aufnehmen
- » Webservice [correspSearch](#) (TELOTA)
- » Ausgabeformat: TEI-XML/CMIF, JSON, CSV

- = Corpus Vitrearum Medii Aevi, ein Projekt der Akademien Berlin und Mainz, das sich dem Erhalt, der Dokumentation und dem Studium mittelalterlicher Glasmalerei widmet
- » Entwicklung eines gemeinsamen Bildmetadatenstandards für digitale Bildressourcen
 - » Projekteigene XMP-Metadatenspezifikation, die auf Standards wie XMP, IPTC und Dublin Core aufbaut
 - » erfasst werden Informationen zu Personen, Standorten, Position im Fenster, Ikonographie und Art der Aufnahme
 - » eigener Namensraum für kunsthistorisch erforderliche Metadaten, die bisher noch nicht durch publizierte und kuratierte Standards erfasst sind, z. B. Aufnahmekontext, Himmelsrichtung, Fensternummer, Zeile, Bahn, ursprünglicher Standort, Angaben zu den Maßen der Scheibe (H, B, D), Normdatum (Anfang, Ende), Restaurierungsgeschichte, Person/Körperschaft (Name, Identifier, Rolle), Ikonographie (IconClass-Beschreibung und -Notation)
 - » Ort dieser Metadaten: JSON-Dateien zu jedem Bild und das Bild selbst
 - » Webservice [CVMA Digitaler Ressourcen Manager](#) (DRM)
 - » Ausgabeformat: JSON

```

MigrationStatusHelper: false
MigrationStatusTemporaryCopy: "M1"
PublishToArchive: false
PartOfBuilding: "Chor"
AgeDeterminationEnd: "1405-12-31"
AgeDeterminationStart: "1395-01-01"
PublishingStatus: true
Identifier: 119527
ICCProfileName: "ProPhoto RGB"
ColorMode: "RGB"
Marked: true
WebStatement: "https://creativecommons.org/licenses/by-sa/4.0/"
CreatorTool: "Fotoware FotoStation"
HasCrop: false
DigitalSourceType: "Originaldigitalaufnahme"
Description: "Bad Wilsnack, Nikolaikirche, Retabel, Detail"
▼ Rights: "CVMA Deutschland Potsdam/Berlin-Brandenburgische Akademie der Wissenschaften/ Foto: Holger Kupfer"
  Type: "Skulptur"
  Title: "Hochaltarretabel"
  Creator: "Holger Kupfer"
  Publisher: "CVMA Potsdam - http://corpusvitrearum.de"
  UsageTerms: "CC BY-SA 4.0"
  Owner: "CVMA Potsdam"
  XMPFileStamps: "2019:09:16 16:43:10+02:00"
  LocationCreatedCity: "Bad Wilsnack"
  LocationCreatedCountryName: "Deutschland"
  LocationCreatedProvinceState: "Brandenburg"
  LocationCreatedSublocation: "Stadtpfarrkirche St. Nikolaus"
  LocationCreatedWorldRegion: "Europa"
  LocationId: "http://sws.geonames.org/11203829"
  ArtworkCircaDateCreated: "um 1400"
▼ Copyright: "CVMA Deutschland Potsdam/Berlin-Brandenburgische Akademie der Wissenschaften/ Foto: Holger Kupfer"

```

Beispiel einer XMP-Metadaten im JSON-Format (JavaScript Object Notation) CVMA Potsdam (Screenshot, Anja Gerber)

Im [*Handbuchs zur Erstellung diskriminierungsfreier Metadaten*](#) (2024) wurden 7 praktische Schritte für die Beschreibung eines Metadatensets formuliert.

1. Klären, welche Metadaten zu welchem Zweck und für welches Publikum gesammelt werden sollen.
2. Metadatenfelder festlegen.
3. Vorhandene Metadaten zusammenstellen.
4. Zeitmanagement beachten.
5. Metadatenpunkte festlegen.
6. Datenwertstandards wählen.
7. Checkliste

7. Checkliste des [*Handbuchs zur Erstellung diskriminierungsfreier Metadaten*](#)

- » **Genauigkeit:** Sind die erfassten Daten korrekt und sachlich?
- » **Vollständigkeit:** Wurden alle relevanten Daten vollständig erfasst?
- » **Konsistenz:** Wurden die Daten konsistent eingegeben? Wird derselbe Satz von Metadatenelementen verwendet, um alle Ressourcen in der Sammlung zu beschreiben? Werden die Daten in demselben Format eingegeben?
- » **Interoperabilität:** Sind die Daten maschinenlesbar? Können die Metadaten leicht in ein anderes System migriert und von diesem verstanden werden? Können sie mit anderen Metadatenansätzen oder Sammlungen zusammengeführt werden?
- » **Inklusivität:** Sind die Daten inklusiv, nicht abwertend und frei von Vorurteilen und schädlicher Sprache? Falls historische Begriffe verwendet werden: Werden sie korrekt kontextualisiert? Sind die verwendeten Begriffe für die beschriebene Ressource geeignet? Stimmen die Begriffe und beschreibenden Informationen mit der Art und Weise überein, wie die Ersteller*innen oder Nutzer*innen einer Ressource diese beschreiben könnten?
- » **Ethische Überlegungen:** Enthalten die Daten persönliche, identifizierende oder anderweitig sensible Informationen? Sind die Rechte vorhanden, die in den Daten enthaltenen Informationen aufzuzeichnen oder zu veröffentlichen? Werden die Mitwirkenden an den Daten und die darin zitierten Ressourcen genannt?

Die Checkliste für Metadaten sets ist eine Möglichkeit, selbstständig eine Qualitätskontrolle der Metadaten sets vorzunehmen.

Warum ist das so wichtig?

- » Metadaten von geringer Qualität wirken sich negativ aus auf die Wahrnehmung von Qualität und Nutzer:innenfreundlichkeit einer Sammlung.
- » Die Qualität auf Ebene der Entitäten wirkt sich direkt auf Linked Data Applications und Datensets aus.

- » Riley, Jenny: Seeing Standards. A Visualization of the Metadata Universe (2009-2010). <https://jennriley.com/metadatamap/seeingstandards.pdf>
- » Handbuch zur Erstellung diskriminierungsfreier Metadaten für historische Quellen und Forschungsdaten. Erfahrungen aus dem geschichtswissenschaftlichen Forschungsprojekt Stadt.Geschichte.Basel. <https://maehr.github.io/diskriminierungsfreie-metadaten/#grundlagen-zu-metadaten-f%C3%BCr-historische-quellen-und-forschungsdaten>
- » Király, Péter and Brase, Jan: "4.3 Qualitätsmanagement". Praxishandbuch Forschungsdatenmanagement, hrsg. von Markus Putnings, Heike Neuroth und Janna Neumann, Berlin/Boston 2021, S. 357-380. <https://doi.org/10.1515/9783110657807-020>
- » <https://forschungsdaten.info/themen/beschreiben-und-dokumentieren/metadaten-und-metadatenstandards/>
- » <https://pro.deutsche-digitale-bibliothek.de/glossar/metadatenformat>



Metadaten

Workshop zur Einführung
Teil 2

Markus Schnöpf (TELOTA)

9. Juli 2024

Übung (20 Minuten)



- In der Übung möchten wir gemeinsam über zwei Online-Metadaten-Editoren Metadaten für mehrere Objekte anlegen.
 - Online-Metadateneditoren
 - Dublin Core
 - DC simple und Advanced:
https://nsteffel.github.io/dublin_core_generator/generator_ng.html
 - DataCite
 - <https://dhvlab.gwi.uni-muenchen.de/datacite-generator/>
 - Beispiele:
 - Nehmen Sie sich ein Objekt aus Ihrer täglichen Forschungspraxis und versuchen, dieses mit Metadaten auszuzeichnen.
 - Wenn Sie kein Objekt zur Verfügung haben, dann versuchen Sie bitte, eine Projektwebseite von Ihnen mit Metadaten zu versehen.

Weitergehende Informationen zu Metadaten



- Minimaldatensatz der DDB
- TEI-Header
- Mapping
- Schnittstellen und welche Rolle sie für wen im Haus spielen (Entwickler vs. Wissenschaftler)
- Rückbindung an FDM

- Ablieferung von (Meta-)Daten für digitale Objekte, die in der DDB angeboten werden sollen in diesen Formaten:
- Jedes Profil behandelt unterschiedliche Objekte
- Basis ist das Europeana Data Model, Lightweight Information Describing Objects, MARCXML, EAD, METS/MODS
- Metadatenlieferung über FTP-Upload oder OAI-Schnittstelle.

Name des Profils
DDB-DC_RDF/XML
DDB-EDM
DDB-LIDO
DDB-MARC
EAD(DDB)
DFF-EDM
DDB-METS/MODS

- Identifikator für den Datensatz (Pflicht)
- Identifikator für den Datenpartner (Pflicht)
- Vorschaubild (Pflicht)
- Link zum Digitalen Objekt (Pflicht)
- Rechtsstatus für das Digitale Objekt (Pflicht)
- Objekttitel (Pflicht)
 - Max 200 Zeichen
- Objektyp (Pflicht)
 - Kontrolliertes Vokabular
- Medientyp (Pflicht)

- LIDO: Lightweight Information Describing Objects
- Einsatz vor allem im musealen Bereich.
- Leichter als CIDOC-CRM, orientiert sich aber daran.

Datenfelder (Erfassung)

Datenfelder, die üblicherweise bei der Erfassung befüllt werden:

- Objekttitle oder -benennung
- Objekttyp oder -bezeichnung
- Klassifikation
- Inventarnummer
- Objektbeschreibung
- Material
- Technik
- Maße
- Ereignis in der Objektgeschichte (Feldgruppe)
 - Ereignistyp
 - Person/Körperschaft
 - Datierung
 - Ort
- Inhaltsschlagwort
- Mediendatei (Feldgruppe)
 - Link zur Mediendatei
 - Nutzungsrechte Mediendatei
 - Rechtswahrnehmung Mediendatei
 - Alternativtext

Datenfelder (Export)

Datenfelder, die üblicherweise erst beim oder nach dem Export aus dem lokalen D:

- ID Datensatz
- Sprache des Datensatzes
- Datensatzart
- Verwahrende Einrichtung
- Datensatzerstellende Einrichtung
- Mediendatei: Medientyp
- Nutzungsrechte Metadaten
- Link zum veröffentlichten Metadatensatz
- Datierung des Datensatzes

- Metadaten können (und sollten) auch im TEI-Header bei textbasierten Projekten verzeichnet werden.

- Struktur:

- `<teiHeader>`

 - `<fileDesc>`

 - `<titleStmt>`

 - `<title>REGISTRES DE L'ACADEMIE ROYALE DES SCIENCES ET </title>`

 - `<author>Wolfgang Knobloch</author>`

 - `</titleStmt>`

 - `<editionStmt>`

 - `<edition><date>2011-03-20T21:19:00</date></edition>`

 - `</editionStmt>`

 - `<publicationStmt>`

 - `<p>no publication statement available</p>`

 - `</publicationStmt>`

 - `<sourceDesc>`

 - `<p>Written by OpenOffice</p>`

 - `</sourceDesc>`

 - `</fileDesc>`

 - `<profileDesc>`

 - `<textClass>`

 - `<keywords scheme="unknown">`

 - `<list>`

 - `<item/>`

 - `</list>`

 - `</keywords>`

 - `</textClass>`

 - `</profileDesc>`

 - `<revisionDesc>`

 - `<listChange>`

 - `<change><name>Knobi</name><date>2020-02-26T12:46:00</date></change>`

 - `</listChange>`

 - `</revisionDesc>`

- `</teiHeader>`

- In den Projekten werden detaillierte Metadaten zu den Forschungsobjekten erhoben. Dabei kommen eigene Erfassungssysteme auf XML- oder SQL-Basis zum Einsatz.
- Um diese Metadaten dann in andere Systeme einzuspielen, ist das Mapping von Metadaten notwendig.
- Dabei wird in einer Tabelle dokumentiert, welche Felder der Ausgangsdaten in welche Zielfelder geschrieben werden müssen.
- Beispiel: Mapping des Minimaldatensatzes der DDB zu lido:
 - <https://docs.nfdi4culture.de/lido-schulung/modul-lido-mapping>
 - Mapping-Tabelle LIDO-> DDB: [Externe Excel-Tabelle](#), bzw. Download hier: <https://pro.deutsche-digitale-bibliothek.de/downloads-links/lido-mappingtabelle>

- OAI-Schnittstellen
 - Bedeutung für Wissenschaftler:innen
 - Nutzende, z.B. Bielefeld Academic Search Engine (Aggregator)
 - Bedeutung für Developer
 - Arbeitsaufwand
 - Mapping von Daten
 - An der BBAW im Einsatz:
 - [Edoc-Server der BBAW](#)
 - [Edition-Humboldt](#)

- Metadaten ermöglichen die Auffindbarkeit von Daten
- FAIR-Prinzipien gelten im Mindesten für Metadaten
- Im Forschungsdatenmanagement muss mit einer Vielzahl an Metadaten umgegangen werden
- Unterscheidung von administrativen Metadaten, strukturellen Metadaten, juristischen Metadaten
- Metadaten, kontrollierte Vokabulare, Normdaten, Schnittstellen erhöhen und ermöglichen Findbarkeit, Zugang, Interoperabilität und Weiderverwendung von Forschungsdaten.