



Kleines Handbuch Metadaten

Metadaten

Stefanie Rühle

1 Einleitung

Metadaten sind Daten, die dazu dienen, Objekte, Konzepte aber auch Daten strukturiert zu beschreiben. Das Prinzip das dahinter steht ist uralte und bekannt, seit Menschen begonnen haben, ihr "Hab und Gut" aufzulisten – sei es auf Tontafeln und Pergament, später in Katalogen und Findbüchern und schließlich in Datenbanken. Im Kontext der Datenbanken entstand dann auch der Begriff Metadaten – sind doch alle Objekte in einer Datenbank letztlich Daten und Metadaten die Daten, die eine Aussage über diese Daten machen. Seit den 90er Jahren wird dieser Begriff daher immer dann verwendet, wenn es um maschinenlesbare Beschreibungen von Ressourcen geht. Diese strukturierten Beschreibungen unterstützten das Wiederauffinden von relevanten Ressourcen in den Datenbanken, aber vor allem auch im "Informationsdschungel" des World Wide Web. Dabei muss alles, was dort von Interesse sein kann, mit Metadaten beschrieben werden: von Waren, die zum Verkauf angeboten werden, über Bücher, die in Bibliotheken zur Verfügung stehen, bis zu den Spielergebnissen der lokalen Fußballvereine. Diese sehr unterschiedlichen Arten von Ressourcen erfordern zum Teil sehr unterschiedliche Beschreibungen und somit auch unterschiedliche Formate und Standards. Der vorliegende Beitrag gibt eine Einführung in das Thema Metadaten und beschreibt einige Standards, die vor allem im Bereich des Kulturerbes Verwendung finden. Dabei gehen wir v.a. auf die Bedeutung von Metadatenstandards für die Interoperabilität ein.

2 Metadaten

2.1 Was sind Metadaten?

*"Metadata, then, can be thought of as data about other data. It is the Internet-age term for information that librarians traditionally have put into catalogs, and it most commonly refers to descriptive information about Web resources."*¹

Metadaten sind strukturierte Daten zur einheitlichen Beschreibung von Ressourcen jeglicher Art (z. B. Daten, Dokumente, Personen, Gemälde, Orte, Gebäude, Konzepte). Sie sorgen dafür, dass den Beschreibungen von Instanzen ein- und derselben Ressourcenklasse eine einheitliche Struktur zugrunde liegt und erleichtern so das Suchen, Finden und Selektieren relevanter Ressourcen aus einer Vielzahl möglicher Ressourcen. Dabei werden verschiedene Arten von Metadaten unterschieden:

- **Identifizierende/beschreibende Metadaten:** Beschreibende Metadaten sind Informationen, die nötig sind, um relevante Ressourcen zu suchen, zu finden und zu identifizieren. Hierzu gehören u.a. der Name oder Titel einer Ressource, der Urheber oder das Entstehungsdatum.
- **Administrative Metadaten:** Die administrativen Metadaten liefern Informationen, die der Verwaltung der Ressourcen dienen. Hierzu gehören Informationen zur Herkunft und Archivierung, Prozesse, die die Ressource durchlaufen hat, Technik, die bei der Entstehung Bearbeitung und den Zugriff auf die Ressource eine Rolle spielen usw.
- **Content Ratings Metadaten:** die Content Ratings Metadaten geben Auskunft über mögliche Nutzer (z. B. die Zielgruppe) einer Ressource.
- **Linkage/Relationship Metadaten:** Linkage/Relationship Metadaten beschreiben die Beziehungen zwischen Ressourcen (z. B. ist-Teil-von oder ist-Vorgänger-von Beziehungen).
- **Meta-Metadaten:** Meta-Metadaten informieren darüber, welche Modelle, Syntax und Formate den Metadaten zugrunde liegen, wer die Metadaten wann erstellt und modifiziert hat. Meta-Metadaten sind entweder identifizierende, administrative oder Relationship Metadaten sein.

Für die Verwendung von Metadaten hat die Dublin Core Metadata Initiative (DCMI) drei Prinzipien aufgestellt:²

1. Das One-to-One Prinzip, das besagt, dass ein Metadatensatz immer nur eine Manifestation einer Ressource beschreiben darf. Am Beispiel der Mona Lisa wird dies deutlich gemacht: Ein Foto der Mona Lisa hat einige Gemeinsamkeiten mit dem Gemälde, ist aber eine andere Ressource. Aus diesem Grund sollte das Foto für sich beschrieben werden unter Angabe des Fotografen als Urheber der Ressource und nicht des Malers, der das Gemälde erstellt hat. Die Beziehung zwischen dem Foto und dem Gemälde sollte durch die Verwendung von Relationship Metadaten hergestellt werden.

¹ <http://dublincore.org/documents/usageguide/#whatismetadatas>

² <http://dublincore.org/documents/usageguide/>

2. Das Dumb-Down Prinzip, das die Möglichkeit bietet, Metadaten, die der granularen Beschreibung von Ressourcen dienen, in einer allgemeineren Umgebung nutzen zu können. Aus diesem Grund werden die granularen Metadatenelemente generischen Elementen zugeordnet (z. B. ist das Element `dcterms:issued` eine Subproperty von `dc:date`, das Element `dcterms:alternative` eine Subproperty von `dc:title`).
3. Appropriate Values: Die Informationen über eine Ressource müssen sowohl maschinen- als auch menschenlesbar sein. Die Verwendung eines Identifiers zur Beschreibung kann ggf. in einem bestimmten System aufgelöst werden, in anderen Systemen aber u. U. nicht. Aus diesem Grund ist es häufig sinnvoll, neben maschinenlesbaren Identifiern auch menschenlesbare Informationen (wie z. B. Namen von Personen oder Titelangaben) mit zu liefern.

2.2 Verwendung von Metadaten

Metadaten können in eine Ressource eingebettet sein oder separat von dieser gehalten werden. Typische Beispiele für Metadaten in Ressourcen sind beschreibende HTML-Tags in einer Webseite oder Metadaten Tags in digitalen Fotos. Handelt es sich bei den beschriebenen Objekten um nichtdigitale Ressourcen, müssen die Metadaten auf jeden Fall separat verwaltet werden. Aber auch die Beschreibung digitaler Ressourcen mit Metadaten erfolgt häufig von der Ressource selbst separiert, da dies das Datenmanagement erleichtert. Für die Verwaltung der separat gehaltenen Metadaten gibt es unterschiedliche Anwendungen. Aus dem bibliothekarischen Kontext kennen wir

- digitale Bibliothekskataloge (OPAC),
- Repositories wie OPUS oder Dspace, die sowohl digitale Ressourcen als auch die dazugehörigen Metadaten verwalten,
- Suchmaschinenindizes wie Base oder OAister, die Metadaten verschiedener Datenbanken zusammenführen,
- Metasuchen wie den KVK, der die Metadaten unterschiedlicher Datenbanken über eine gemeinsame Suchoberfläche abfragt.

Eine Metadatenanwendung kann verschiedene Module haben:

- Der Speicher, in dem die Metadaten vorgehalten werden,
- Erfassungsmodule, die die Erstellung und/oder Bearbeitung der Metadaten ermöglichen,
 - zwischen Erfassungsmodul und Speicher braucht es eine Import-Schnittstelle.
- Präsentationsebenen, über die die Metadaten den Nutzern zur Verfügung gestellt werden,
 - Zwischen Präsentationsebene und Speicher braucht es eine Export-Schnittstelle.
- Eine Importschnittstelle, über die die Metadaten aus anderen Anwendungen in den Speicher importiert werden.
- Eine Export-Schnittstelle, über die die Metadaten außerhalb der eigenen Anwendung zur Verfügung gestellt werden können.

2.3 Metadatenelemente

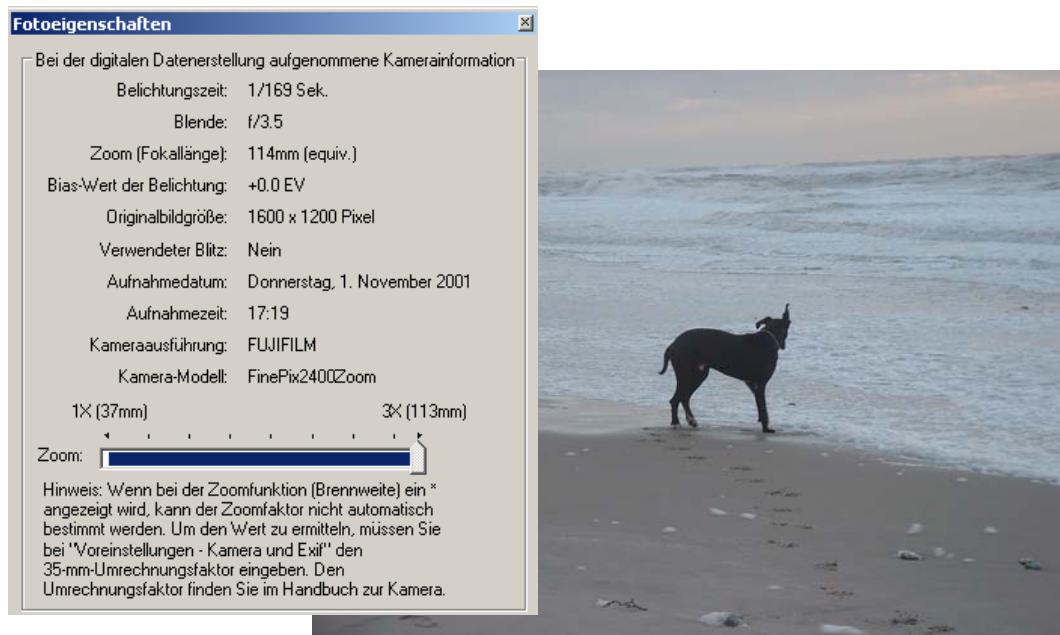
Eine Metadatenbeschreibung besteht aus den Metadatenelementen (terms), die verwendet werden, um eine bestimmte Art von Ressourcen zu beschreiben und den Werten (values), die verwendet werden, um eine bestimmte Ressource zu beschreiben.

Metadatenelemente sind:

- Classes – definieren die Art der Ressourcen, die beschrieben werden.
- Properties – definieren, was für Eigenschaften eine bestimmte Art von Ressourcen hat
- Encoding Scheme – definieren, wie der Eintrag bzw. Wert aussieht, der zu einer bestimmten Property gehört.

Bei der Beschreibung mit Metadaten werden einer bestimmten Klasse (class) von Ressourcen bestimmte Properties zugewiesen. (z. B. können der Klasse Foto die Properties Belichtungszeit, Blende, Aufnahmezeit, Aufnahmezeit, Kameramodell usw. zugewiesen werden). Diese Properties haben Werte (values), die für jede Ressource dieser Klasse individuell sind. Z. B. ist das Entstehungsdatum eines Fotos Donnerstag, 1. November 2001, das Entstehungsdatum eines anderen Fotos Freitag, 2. November 2001.

Abbildung 1: Properties eines digitalen Fotos



Die verwendeten Werte können einem Encoding Scheme entsprechen, das entweder die Syntax des verwendeten Werts bestimmt (z. B. entspricht die Angabe 2001-11-01 dem Standard ISO 8601) oder den Wert selbst (z. B. können Werte aus einem kontrollierten Vokabular wie der GND verwendet werden).

Abbildung 2: Identifizierende Metadaten am Beispiel von MODS-Daten

```
<mods:mods>
  <mods:titleInfo>
    <mods:title>Attila at the Beach</mods:title>
  </mods:titleInfo>
  <mods:originInfo>
    <mods:dateIssued encoding="iso8601">2001-11-01</mods:dateIssued>
  </mods:originInfo>
  <mods:subject>
    <mods:topic valueURI="http://d-nb.info/gnd/4113330-4">Dogge</mods:topic>
    <mods:topic valueURI="http://d-nb.info/gnd/4379510-9">Dänemark West</mods:topic>
  </mods:subject>
  <mods:name type="personal" valueURI="http://d-nb.info/gnd/131724126">
    <mods:namePart type="given">Stefanie</mods:namePart>
    <mods:namePart type="family">Rühle</mods:namePart>
  </mods:name>
</mods:mods>
```

3 Interoperabilität

Um Metadaten anwendungsübergreifend nutzen und verlinken zu können (z. B. in Metasuchen oder Semantic Web), müssen sie interoperabel sein. Interoperabilität ist die Fähigkeit, Daten zwischen verschiedenen

Systemen ohne allzu großen Informationsverlust auszutauschen. Im Bereich der Metadaten unterscheiden wir drei Arten der Interoperabilität³:

- Strukturelle Interoperabilität bedeutet, dass verschiedenen Metadatenformaten ein gemeinsames Datenmodell zugrunde liegt (z. B. das RDF-Modell oder das Dublin Core Abstract Model)
- Syntaktische Interoperabilität ist dann gegeben, wenn die Daten in einer gemeinsamen Syntax (XML oder auch Turtle) vorliegen.
- Semantische Interoperabilität erfordert die Verwendung derselben Metadaterme bzw. die gemeinsame Nutzung eines Metadatenstandards.

Interoperabilität ist somit der gemeinsame Nenner, auf dessen Grundlage Daten ausgetauscht werden können. Interoperabilität herzustellen erscheint dabei ganz einfach: Alle verwenden

- denselben Metadatenstandard
- dasselbe Metadatenmodell
- dieselbe Syntax

und schon sind Metadaten interoperabel. Ein solches Vorgehen ist aber gar nicht möglich. Zum einen sind Metadatenformate und -modelle in verschiedenen Communities unabhängig voneinander historisch gewachsen, zum anderen erfordern die unterschiedlichen Anforderungen der Communities und die Beschreibung verschiedener Arten von Ressourcen i.d.R. unterschiedliche Modelle und Formate. Deutlich wird dies u. a. in den unterschiedlichen Datenmodellen, die den Metadatenstandards in Museen, Archiven und Bibliotheken zugrunde liegen.

Diese heterogenen Modelle und Standards, lassen sich mal mehr, mal weniger gut zusammenführen. Erforderlich sind dabei:

- eine gute Dokumentation der verwendeten Modelle und Metadatenelemente (z. B. in Form eines Metadatenprofils⁴),
- Ein Mapping zwischen den Modellen und den Elementen der unterschiedlichen Anwendungen.⁵

3.1 Strukturelle Interoperabilität

Strukturelle Interoperabilität entsteht, wenn verschiedenen Metadatenanwendungen ein gemeinsames Metadatenmodell zugrunde liegt. Im folgenden werden einige der im Kontext von Linked Data relevanten Modelle genannt.

3.1.1 Resource Description Framework (RDF)⁶

Im Kontext von Linked Data und Semantic Web hat diese Form der Interoperabilität in den letzten Jahren zunehmend an Bedeutung gewonnen. Sie macht es möglich, dass Metadaten aus unterschiedlichen heterogenen Anwendungen miteinander verknüpft werden. Die Verknüpfung erfolgt auf der Grundlage des RDF-Modells. Hier wird die traditionelle Beschreibung von Ressourcen in Form von Datensätzen aufgebrochen und eine Datenstruktur geschaffen, die es ermöglicht, Ressourcen flexibel miteinander zu verlinken. Dabei entspricht die Beschreibung der Ressource mehreren Tripeln aus Subjekt, Prädikat und Objekt, wobei gilt:

- Subjekt ist die Ressource, die beschrieben wird,
- Prädikat ist die Property, mit der diese Art von Ressourcen beschrieben wird,
- Objekt ist der Wert, der für genau diese Ressource gilt.

Handelt es sich bei dem Objekt um eine weitere Ressource, auf die verlinkt wird, dann wird aus dem Objekt des einen Triple das Subjekt eines anderen Triple. Voraussetzung dafür ist die Verwendung von URIs über die die Ressourcen eindeutig identifizierbar sind.

Mit Hilfe von RDF und URIs kann ich somit in maschinenlesbarer Form die Aussage machen, dass eine bestimmte Ressource (Subjekt) für eine bestimmte Eigenschaft (Prädikat) einen bestimmten Wert (Objekt) hat.

³ <http://www.kim->

[forum.org/Subsites/kim/DE/Materialien/Glossar/glossar_node.html;jsessionid=CF98C74A2D64B6354DA5D679A46AB9DD.prod-worker5](http://www.kim-forum.org/Subsites/kim/DE/Materialien/Glossar/glossar_node.html;jsessionid=CF98C74A2D64B6354DA5D679A46AB9DD.prod-worker5)

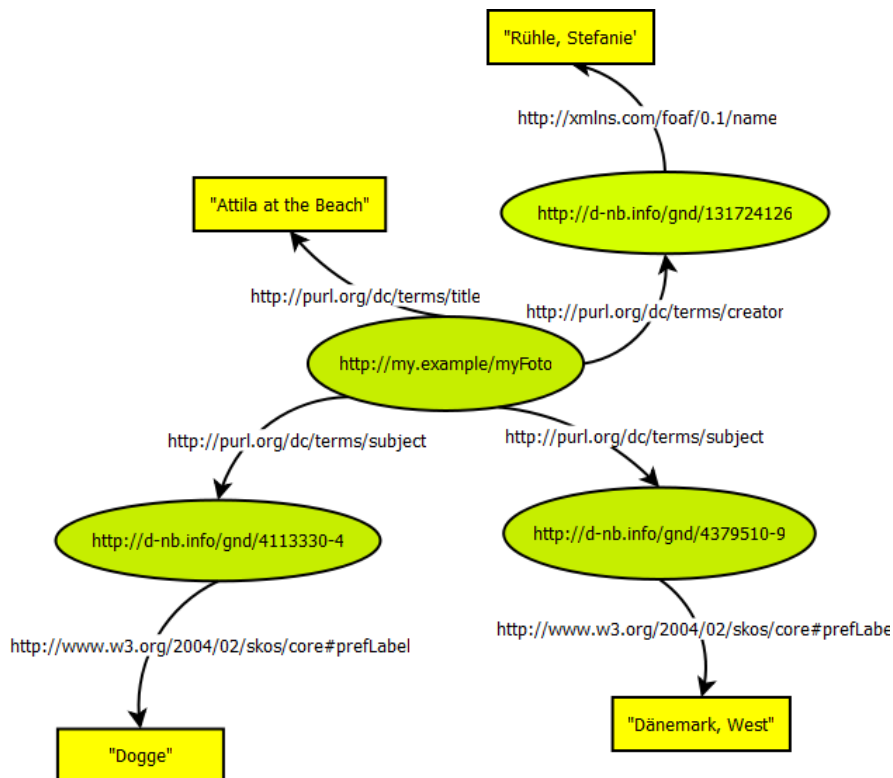
⁴ S. den Beitrag Metadatenprofile in diesem Handbuch

⁵ S. den Beitrag Mapping und Crosskonkordanzen in diesem Handbuch

⁶ <http://www.w3.org/RDF/>

Bei dem Wert kann es sich um eine einfache Zeichenkette (literal value) handeln, die keinerlei Link zu einem weiteren Objekt enthält, oder um eine weitere Ressource (non-literal value), die wiederum für eine bestimmte Eigenschaft (Prädikat) einen bestimmten Wert (Objekt) hat usw.

Abbildung 3: RDF Graph



3.2 DCMI Abstract Model (DCAM)⁷

Das DCMI Abstract Model (DCAM) ist ein syntaxunabhängiges Meta-Modell, das auf RDF basiert und die syntaxübergreifende Konvertierung von Metadaten ermöglicht. DCAM definiert Elemente, die benötigt werden, um Informationsstrukturen und insbesondere die Beziehungen innerhalb dieser Strukturen auf der Metaebene zu beschreiben. Ebenso wie RDF unterscheidet auch DCAM verschiedene Typen von Metadaten, nämlich Eigenschaften (`rdf:property`), Klassen (`rdfs:class`), syntaktische Kodierungsschema (`rdfs:Datatype`) und semantische Kodierungsschema (`dcm:VocabularyEncodingScheme`). Ein Metadaterm ist dann DCAM konform, wenn er explizit einem dieser Typen zugewiesen wurde und über einen den Term identifizierenden URI verfügt.

3.3 Functional Requirements of Bibliographic Records (FRBR)

Die Functional Requirements of Bibliographic Records (FRBR)⁸ ist ein Modell zur Strukturierung bibliographischer Daten. Das Modell geht von verschiedenen Gruppen von Ressourcen aus, die bei der Beschreibung von Informationsressourcen eine Rolle spielen und miteinander verlinkt werden.

Gruppe 1: WEMI (Work – Expression – Manifestation – Item)

Diese Gruppe macht es möglich, verschiedene Ausprägungen eines Werks zusammenhängend anzuzeigen indem Work, Expression, Manifestation und Item miteinander in Beziehung gesetzt werden. Der Nutzer erhält

⁷ <http://www.dublincore.org/documents/abstract-model/>

⁸ <http://www.ifla.org/en/publications/functional-requirements-for-bibliographic-records>

damit einen Überblick, welche Versionen eines Werks vorhanden sind und kann die ihm relevant erscheinende auswählen. Die WEMI-Entitäten und ihre Eigenschaften sind in den FRBR definiert.

Gruppe 2: Personen und Körperschaften

Diese Gruppe beschreibt die Beziehungen zwischen den WEMI-Entitäten und Personen und Körperschaften, die für diese Entitäten relevant sind. Die Entitäten der Gruppe 2 und die Eigenschaften derselben werden in den Functional Requirements for Authority Data (FRAD)⁹ beschrieben.

Gruppe 3: Konzept, Objekt, Ereignis, Ort

Diese Gruppe beschreibt die Beziehungen zwischen den WEMI-Entitäten und Konzepten, Objekten, Ereignissen und Orten, die Thema der WEMI-Entitäten sind. Diese Entitäten und ihre Eigenschaften werden in den Functional Requirements for Subject Authority Data (FRSAD)¹⁰ beschrieben.

3.4 Syntaktische Interoperabilität

Von syntaktischer Interoperabilität spricht man dann, wenn Daten, die zusammengeführt werden sollen, in ein- und derselben Syntax vorliegen. Im Folgenden werden kurz zwei Sprachen vorgestellt, die beim Austausch von Daten zurzeit weit verbreitet sind.

3.5 Extensible Markup Language (XML)

XML¹¹ ist eine von SGML abgeleitete relativ einfache und flexible Auszeichnungssprache. Sie wird für die Datenhaltung aber vor allem auch für den Austausch von Daten (u.a. via OAI-PMH) verwendet. Als Metasprache bietet XML keine Tags für die Beschreibung von Ressourcen, diese Tags können vielmehr anwendungsspezifisch definiert werden bzw. können bekannte Standards verwendet werden, die XML konform sind (z. B. DCMI Metadata Terms, MODS, EAD, LIDO usw).

Die Verwendung dieser Standards unter Berücksichtigung anwendungsspezifischer Einschränkungen lässt sich in einer XML Document Type Definition (DTD) oder einem XML-Schema¹² abbilden. DTD oder XML-Schema beschreiben die verwendeten Metadatenelemente und die geltenden Einschränkungen in maschinenlesbarer Form und ermöglichen so eine nicht vollständige aber doch recht weitreichende automatische Validierung von XML-Datensätzen, die diese Elemente verwenden. Geprüft werden kann allerdings nur das Vorhandensein von verpflichtenden Elementen, Einschränkungen hinsichtlich der Wiederholbarkeit von Elementen sowie die Korrektheit von Encodings.

3.6 Terse RDF Triple Language (Turtle)¹³

Turtle ist eine einfache und leicht verständliche Syntax für die Darstellung von RDF-Tripeln in Textform. Während bei der Verwendung von RDF in XML einige Einschränkungen gelten, die sich aus XML selbst sowie den Regeln für XML-Namespace ergeben, ist Turtle ein Subset von Notation 3 (N3)¹⁴ und kann in allen Anwendungen verwendet werden, die N3 unterstützen; außerdem kann es auch in SPARQL, der Abfragesprache für RDF, verwendet werden. Zudem wird die Abbildung von RDF-Tripel in XML schnell sehr umfangreich und unübersichtlich, im Vergleich dazu ist die Abbildung in Turtle einfach strukturiert und übersichtlich.

Abbildung 4: Abbildung eines RDF-Graphs mit Turtle

⁹ <http://www.ifla.org/publications/functional-requirements-for-authority-data>

¹⁰ <http://www.ifla.org/files/classification-and-indexing/functional-requirements-for-subject-authority-data/frsad-final-report.pdf>

¹¹ <http://www.w3.org/XML/>

¹² <http://www.w3.org/XML/Schema>

¹³ <http://www.w3.org/TeamSubmission/turtle/>

¹⁴ <http://www.w3.org/DesignIssues/Notation3>

```

<http://my.example/myFoto> <http://purl.org/dc/terms/title> "Attila at the Beach" ;
<http://purl.org/dc/terms/creator> <http://d-nb.info/gnd/131724126> ;
<http://purl.org/dc/terms/subject> <http://d-nb.info/gnd/4113330-4> , <http://d-nb.info/gnd/4379510-9> .

<http://d-nb.info/gnd/131724126> <http://xmlns.com/foaf/0.1/name> "Rühle, Stefanie" .

<http://d-nb.info/gnd/4113330-4> <http://www.w3.org/2004/02/skos/core#prefLabel> "Dogge" .

<http://d-nb.info/gnd/4379510-9> <http://www.w3.org/2004/02/skos/core#prefLabel> "Dänemark West" .

```

4 Semantische Interoperabilität

Von semantischer Interoperabilität sprechen wir dann, wenn dieselben Metadatenelemente in verschiedenen Metadatenformaten bzw. -schemata verwendet werden. In der Regel geschieht dies unter Verwendung von Metadatenstandards wie z. B. Dublin Core, MARC 21 u.a.

4.1 Metadatenchema

Entsprechend den Anforderungen, die er an seine Anwendung hat, entwickelt ein Metadatenprovider ein mehr oder weniger umfangreiches und komplexes Metadatenchema, in dem alle Elemente, die er verwendet, genannt und spezifiziert werden (hinsichtlich ihres Verpflichtungsgrads, ihrer Wiederholbarkeit usw.) Dieses Schema ist konform zu der Datenbank, in der der Provider die Metadaten erfassen und speichern will und ist somit passgerecht für die eigenen Anwendungen. Probleme kann es allerdings geben, wenn die Daten für eine andere Anwendung zur Verfügung gestellt werden sollen. In der Regel sind daher im Kontext der Metadatenanwendungen drei Arten von Metdatenschemata bzw. -formaten zu betrachten:

Erfassungsformat

Das Erfassungsformat dient der Eingabe von Metdaten, z. B. zur Erfassung von Titel- oder Normdatensätzen. Das Erfassungsformat muss alle Elemente abbilden, die für die manuelle Beschreibung einer Ressource benötigt werden. Um welche Elemente es sich dabei handelt, kann sich von Ressource zu Ressource ändern. Es sollte zudem möglichst einfach und übersichtlich sein, um so den Katalogisierer bei seiner Arbeit zu unterstützen.

Speicherformat

Das Speicherformat ist das Format, in dem die Daten im Speicher abgelegt werden. Hier können neben den im Erfassungsformat erfassten Metadaten weitere Metadaten (z. B. Meta-Metadaten) vorhanden sein, die bei der Übernahme der Metadaten in den Speicher automatisch generiert wurden. Das Speicherformat kann mit dem Erfassungsformat identisch sein, kann von diesem aber auch abweichen wenn dies die maschinelle Verarbeitbarkeit der Daten erfordert.

Austauschformat

Das Austauschformat/Exportformat dient dem Austausch von Metadaten zwischen Partnern bzw. Datenbanken. Dies erfordert Absprachen zwischen den Partnern, die die Metadaten austauschen. Aus diesen Gründen werden vor allem im Bereich der Austauschformate i.d.R. Metadatenstandards verwendet. Bekannte Austauschformate sind MARC 21, LIDO, EAD.

4.2 Metadatenstandards

Ein Metadatenstandard spezifiziert anwendungsübergreifend die Elemente, die in einem bestimmten Kontext verwendet werden sollen, sowie die Struktur und Bedeutung dieser Elemente. Damit schafft er Transparenz hinsichtlich der Struktur und des Inhalts der Metadaten und ermöglicht eine effiziente Datenhaltung und – wenn der verwendete Metadatenstandard dem Zielformat entspricht – einen effektiven Datenaustausch. Indem Metadatenstandards als Austauschformat verwendet werden, unterstützen sie somit die semantische

Interoperabilität zwischen Anwendungen. Sie sind die „gemeinsame Sprache“ in der Information zur Verfügung gestellt wird und ermöglichen Retrieval und Anzeige von Metadaten aus unterschiedlichsten Datenbeständen. Ziel ist darum die einheitliche und maschinenverarbeitbare Beschreibung der Ressourcen und der möglichst reibungslose Austausch dieser Beschreibungen,¹⁵ denn aufgrund eines gemeinsamen Metadatenstandards sind die Metadaten auch in anderen Anwendungen verständlich und interpretierbar.

Im Kontext von Semantic Web und Linked Data sorgen Metadatenstandards aber nicht nur für die Konsistenz der Metadaten, sie ermöglichen außerdem die Verlinkung von Ressourcen mittels eindeutiger und verbreiteter Relationship Elemente. Dennoch ist zu bedenken, dass in den verschiedenen Communities sehr unterschiedliche Metadatenstandards verwendet werden, die häufig nicht nur ganz unterschiedliche Elemente verwenden, sondern denen zudem oft auch ein ganz anders geartetes Metadatenmodell zugrunde liegt. Hier einige Beispiele für Metadatenstandards aus dem Bereich des Kulturerbes:

MARC (Machine Readable Cataloging) 21¹⁶

MARC 21 ist ein Format für das Speichern und den Austausch von bibliographischen Datensätzen, sowie mit diesen verknüpften Informationen wie Normdaten, Besitznachweise u.a. Als Austauschformat ist MARC 21 international verbreitet, es wird stellenweise aber auch als Erfassungsformat verwendet. MARC 21 ist Teil einer ganzen Familien von MARC Standards – zu nennen seien hier UNIMARC, UKMARC, CANMARC aber auch MARCXML.¹⁷

Metadata Object Description Schema (MODS)¹⁸

MODS ist ein XML-basiertes Derivat von MARC 21 und dient der Beschreibung bibliographischer Daten. Im Gegensatz zu MARC 21 verwendet es jedoch für die Beschreibung seiner Elemente keine numerischen Kategorien sondern menschlesbare Begriffe. MODS wird überwiegend für die Beschreibung von digitalen Objekten – oft in Zusammenhang mit METS - eingesetzt. Neben MODS ist der Metadata Authority Description Standard (MADS)¹⁹ ein weiteres Derivat aus MARC 21.

Metadata Encoding and Transmission Standard (METS)²⁰

METS ist ein XML-basiertes Standardformat für die Verwaltung und den Austausch von komplexen digitalisierten Ressourcen. METS dient vor allem dazu, die komplexe Struktur dieser Ressourcen wiederzugeben, um es so zu ermöglichen, dass die gescannten Bilder in der richtigen Reihenfolge den richtigen Kapiteln, Aufsätzen, Abschnitten zugewiesen werden. Neben den Strukturen werden auch administrative und bibliographische Metadaten in die METS-Daten aufgenommen, für letztere gibt es jedoch keine METS-spezifischen Tags, vielmehr empfiehlt die METS-Spezifikation die Einbindung von XML-konformen Metadatenstandards wie Dublin Core oder MODS. In Deutschland ist METS in Kombination mit MODS für die Darstellung von Digitalisaten im DFG-Viewer²¹ verpflichtend.

Encoded Archival Description (EAD)²²

EAD ist ein internationaler Standard für die Beschreibung von Archivmaterial und den Austausch von Metadaten aus diesem Bereich. EAD ist ein XML-basierter Standard und besteht aus ca. 150 Elementen bzw. Tags, die der Erstellung von Findbüchern dienen.

Lightweight Information Describing Objects (LIDO)²³

LIDO ist ein Standardformat für die Beschreibung von Museumsobjekten, wird aber auch für die Beschreibung von Dokumenten über Museumsobjekte verwendet. Grundlage von LIDO waren verschiedene Standards aus der Museumswelt – CDWA lite, museumdat und SPECTRUM – und das Ziel des Standards ist ein einfaches

¹⁵ S. auch http://www.kim-forum.org/Subsites/kim/DE/Materialien/Glossar/glossar_node.html

¹⁶ <http://www.loc.gov/marc/>

¹⁷ CANMARC ist einer der beiden Vorgänger von MARC 21, während Unimarc und UKMARC ähnliche Formate sind.

¹⁸ <http://www.loc.gov/standards/mods/>

¹⁹ <http://www.loc.gov/standards/mads/>

²⁰ <http://www.loc.gov/standards/mets/>

²¹ <http://dfg-viewer.de>

²² <http://www.loc.gov/ead/>

²³ <http://www.lido-schema.org/schema/v1.0/lido-v1.0-specification.pdf>

Metadatenformat für Museen, das dazu dient, Metadaten auszutauschen sowie im Web zu veröffentlichen und zu verlinken. I.d.R. werden LIDO-Daten in XML geliefert.

Dublin Core²⁴

Dublin Core ist ein Standard zur Beschreibung unterschiedlichster Arten von Ressourcen – damit wird er nicht nur international, sondern auch communityübergreifend genutzt (Museen, Bibliotheken, Archive, aber auch Regierungsbehörden und Wirtschaft verwenden Dublin Core Elemente zur Beschreibung ihrer Internetressourcen). Bekannt ist vor allem das Dublin Core Metadata Element Set²⁵, auch bekannt als ISO 15836: ein einfaches Element Set aus 15 Elementen, das für OAI-PMH Schnittstellen verpflichtend ist und deshalb häufig auch OAI DC genannt wird. Die DCMI Metadata Terms²⁶ sind eine Erweiterung dieses einfachen Sets. Sie beinhalten nicht nur granulare Properties (sogenannte Refinements zu den 15 Elementen), sondern definieren auch Dublin Core Klassen und Kodierungsschema.

Dublin Core Standards sind syntaxunabhängig und können daher sowohl mit XML als auch mit Turtle verwendet werden.

Weitere Standards

Ein umfassende Liste von Metadatenstandards wird im Open Metadata Handbook veröffentlicht.

Informationen dazu gibt es unter:

http://en.wikibooks.org/wiki/Open_Metadata_Handbook/Metadata_Standards

²⁴ <http://www.dublincore.org/>

²⁵ <http://www.dublincore.org/documents/dces/>

²⁶ <http://www.dublincore.org/documents/dcmi-terms/>