

DEUTSCHE INITIATIVE FÜR NETZWERKINFORMATION E. V.



Einführung in Metadaten und Metadatenformate



Einführungen



Übungen



Lösungen

Zielgruppe:

Beschäftigte an Wissenschaftlichen Bibliotheken mit geringen Vorkenntnissen im Bereich Metadaten und Interoperabilität

Inhalt:

Metadaten dienen dazu, Nutzer bei der Suche, dem Finden, Identifizieren und Selektieren relevanter Informationen zu unterstützen. Dies gilt sowohl für lokale Informationssysteme als auch für das Internet.

In diesem Tutorial geht es um eine Einführung in Metadaten, eine Vorstellung der wichtigsten Metadatenstandards und einen Einblick in das Thema Metadatenprofile.

Inhalt

Thema 1

Einführung in Metadaten

Was sind Metadaten?

Übung: Metadaten

Übung: Metadaterme

Metadatentypen und Metadatenanwendungen

Übung: Typen von Metadaten

Thema 2

Metadatenstandarts

Metadatenstandarts

Schnittstellen

Dublin Core

Übung: Dublin Core Properties

Übung: Dublin Core Properties 2

Weitere Metadatenformate

Weitere Standards

Kontrollierte Vokabulare und Normdaten

Weitere kontrollierte Vokabulare und Crosskonkordanzen

Übung: Klassifikation, Thesaurus und Crosskonkordanz

Regelwerke und Modelle

Linked Data

Thema 3

Interoperabilität

Grundlagen der Interoperabilität

Mapping und Metadatenprofile

Thema 4

Zusammenfassung

Zusammenfassung

Thema 1



Einführung in Metadaten

Dieses Kapitel bietet eine kurze Einführung in Metadaten. Hier wird die Frage geklärt, was Metadaten überhaupt sind. Außerdem wird auf Metadatentypen und Metadatenanwendungen eingegangen.

Was sind Metadaten?

Wenn wir an Metadaten denken...

Metadaten ist ein Begriff, unter den man viele Aspekte packen kann. Wenn wir an Metadaten denken, dann denken wir nicht alle an das Gleiche. Wir denken beispielsweise an:

- Standards (Dublin Core, METS, MARC 21, ...)
- Regelwerke (RDA, RAK, ...)
- Metamodelle (FRBR, ...)
- Schnittstellen (OAI, Z 39.50, ...)
- Anwendungsprofile (SWAP, DCLib AP, ...)
- Markup Languages (HTML, XML, ...)
- Kontrollierte Vokabulare (DDC, SWD, Getty Thesaurus, ...)
- Anwendungen (OPAC, Metasuche, Repository, ...)



Doch was genau versteht man unter Metadaten? Hierzu findet man verschiedene Definitionen.

"Metadata is machine understandable information about web resources or other things."

(Quelle: Berners-Lee, Tim: Design Issues. Architectural and philosophical points, 6. Januar 1997. URL: <http://www.w3.org/Designissues/Metadata.html>)

"The World Wide Web affords unprecedented access to globally distributed information. Metadata, or structured data about data, improves discovery of and access to such information."

(Quelle: D-Lib Magazine, (May) 1998. URL: <http://www.dlib.org/dlib/may98/miller/05miller.html>)

"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."

(Quelle: Using Dublin Core, issued 07.11.2005. URL: <http://dublincore.org/documents/usageguide/#whatismetadata>)

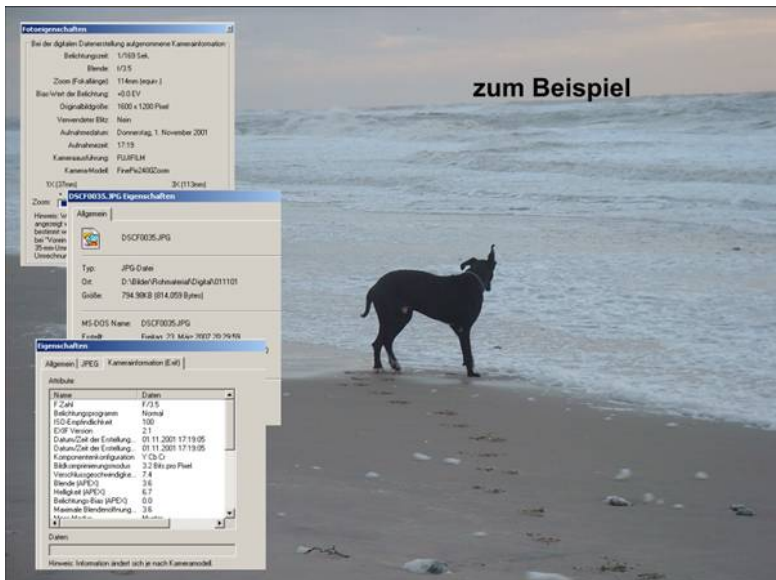
"Metadaten sind strukturierte Daten, die Objekte (wie Daten, Dokumente, Personen, Gemälde, Orte, Gebäude usw.) beschreiben. Sie bewirken, dass den Beschreibungen unterschiedlicher Objekte eine einheitliche Struktur zugrunde liegt und erleichtern so das Suchen, Finden und Selektieren relevanter Objekte aus der Vielzahl möglicher Objekte. Sie können in Objekte eingebettet sein oder separat (im eigenen Datensatz) zur Verfügung gestellt werden."

(Quelle: http://www.kim-forum.org/Subsites/kim/DE/Materialien/Glossar/glossar_node.html)



Metadaten sind strukturierte Daten zur einheitlichen Beschreibung von Objekten (resources) aller Art.

So kann man ein Foto mit Metadaten beschreiben. In den Metadaten stehen dann beispielsweise Angaben zur Uhrzeit und dem Datum, an dem das Foto gemacht worden ist, zu den Zoomeinstellungen oder zur Belichtungszeit.





Metadaten sind strukturierte Beschreibungen

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

Metadaterme sind:

Art des Objekts (classes) = beschreibt, was für eine Art von Ressource beschrieben wird

Elemente (properties) = beschreibt, was für Eigenschaften diese Art von Ressource hat

Kodierungen (encoding scheme) = Terme, die bestimmen, wie der Eintrag bzw. Wert aussieht, der zu einem bestimmten Term gehört.

Zum Beispiel:

Die Art des Objektes, also die "class", ist in diesem Fall ein Foto. Die Eigenschaften der Ressource ("properties") sind für Fotos typische Angaben wie Belichtungszeit, Blende, Aufnahmedatum, Aufnahmezeit und Kamera-Modell. Die "values" sind die Werte, die diesen Elementen zugeordnet werden können, z.B. "Donnerstag, 1. November 2001" als Wert für das Aufnahmedatum. Die Kodierung ("encoding") für die Uhrzeit ist in diesem Fall in Stunden und Minuten angegeben, die durch einen Doppelpunkt getrennt werden.



class **Fotoeigenschaften**

properties

Bei der digitalen Datenerstellung aufgenommene Kamerainformation

Belichtungszeit:	1/169 Sek.	values
Blende:	f/3.5	
Zoom (Fokallänge):	114mm (equiv.)	
Bias-Wert der Belichtung:	+0.0 EV	
Originalbildgröße:	1600 x 1200 Pixel	
Verwendeter Blitz:	Nein	
Aufnahmedatum:	Donnerstag, 1. November 2001	
Aufnahmezeit:	17:19	encoding
Kameraausführung:	FUJIFILM	
Kamera-Modell:	FinePix2400Zoom	

1X (37mm) 3X (113mm)

Zoom:

Hinweis: Wenn bei der Zoomfunktion (Brennweite) ein * angezeigt wird, kann der Zoomfaktor nicht automatisch bestimmt werden. Um den Wert zu ermitteln, müssen Sie bei "Voreinstellungen - Kamera und Exif" den 35-mm-Umrechnungsfaktor eingeben. Den Umrechnungsfaktor finden Sie im Handbuch zur Kamera.



Metadaten sind einheitliche Beschreibungen

Einer bestimmten Klasse (class) von Objekten (resources) sind bestimmte Elemente (properties) zugewiesen. So sind z.B. der Klasse Fotos die Elemente "Belichtungszeit", "Blende", "Zoom", "Aufnahmedatum" usw. zugewiesen. Bestimmte Elemente haben eine Kodierung (encoding scheme), die die Struktur eines Eintrags (syntax encoding) bestimmt oder die Terme, die für einen Eintrag verwendet werden dürfen (vocabulary encoding). Hierbei muss z.B. der Wert für das Element "Aufnahmedatum" die folgende Struktur haben: "Tag, Tagesdatum. Monat Jahr"

Alle Objekte, die zu dieser Klasse gehören, werden mit diesen und nur diesen Elementen (properties) beschrieben, unter Berücksichtigung der den Elementen zugewiesenen Kodierungen.

Zum Beispiel:

Diese beiden Fotos können miteinander verglichen werden, da sie der gleichen Klasse angehören, über die gleichen Elemente verfügen und eine einheitliche Struktur der Einträge haben. So kann beispielsweise ohne vorherige Konvertierung oder Anpassung das Aufnahmedatum miteinander verglichen werden.

The image displays two side-by-side screenshots of a photo metadata viewer. Both windows have a red box around the metadata fields, with an arrow labeled "Encoding" pointing to it. The metadata fields include: Belichtungszeit, Blende, Zoom (Fokallänge), Bias-Wert der Belichtung, Originalbildgröße, Verwendeter Blitz, Aufnahmedatum, Aufnahmezeit, Kameraausführung, and Kamera-Modell.

Property	Value (Left Photo)	Value (Right Photo)
Belichtungszeit	1/169 Sek.	1/274 Sek.
Blende	f/3.5	f/8.7
Zoom (Fokallänge)	114mm (equiv.)	91mm (equiv.)
Bias-Wert der Belichtung	+0.0 EV	+0.0 EV
Originalbildgröße	1600 x 1200 Pixel	1600 x 1200 Pixel
Verwendeter Blitz	Nein	Nein
Aufnahmedatum	Donnerstag, 1. November 2001	Mittwoch, 31. August 2005
Aufnahmezeit	17:19	03:30
Kameraausführung	FUJIFILM	FUJIFILM
Kamera-Modell	FinePix2400Zoom	FinePix2400Zoom

Übungen zu "Was sind Metadaten?"



Übung 1: Metadaten

Bitte füllen Sie die Lücken im Text, die Lösung finden Sie auf der nächsten Seite.

Metadaten sind strukturierte Daten, die (wie Daten, Dokumente, Personen, Gemälde, Orte, Gebäude usw.) beschreiben. Sie bewirken, dass den Beschreibungen unterschiedlicher Objekte eine einheitliche zugrunde liegt und erleichtern so das, und relevanter Objekte aus der Vielzahl möglicher Objekte. Sie können in Objekte eingebettet sein oder separat (im eigenen Datensatz) zur Verfügung gestellt werden.

Übung 2: Metadaterme

Bitte ordnen Sie die korrekten Beschreibungen zu, die Lösung finden Sie auf der nächsten Seite:

Art des Objektes (Classes)

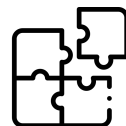
Elemente (Properties)

Kodierungen (Encoding Schemes)

= Terme, die bestimmen, wie der Eintrag bzw. Wert aussieht, der zu einem bestimmten Term gehört.

= beschreibt, was für eine Art Ressource beschrieben wird.

= beschreibt, was für Eigenschaften diese Art von Ressource hat.



Lösung 1: Metadaten

Metadaten sind strukturierte Daten, die **Objekte** (wie Daten, Dokumente, Personen, Gemälde, Orte, Gebäude usw.) beschreiben. Sie bewirken, dass den Beschreibungen unterschiedlicher Objekte eine einheitliche **Struktur** zugrunde liegt und erleichtern so das **Suchen**, **Finden** und **Selektieren** relevanter Objekte aus der Vielzahl möglicher Objekte. Sie können in Objekte eingebettet sein oder separat (im eigenen Datensatz) zur Verfügung gestellt werden.

Lösung 2: Metadaterme

Art des Objektes (Classes) = **beschreibt, was für eine Art Ressource beschrieben wird.**

Elemente (Properties) = **beschreiben, was für Eigenschaften diese Art von Ressource hat.**

Kodierungen (Encoding Schemes) = **Terme, die bestimmen, wie der Eintrag bzw. Wert aussieht, der zu einem bestimmten Term gehört.**

Metadattentypen und Metadatenanwendungen

Metadaten - wozu und wo?



Wozu werden Metadaten überhaupt gebraucht? Was kann man mit Metadaten machen?

Man benötigt sie für die strukturierte und einheitliche Beschreibung von Objekten.

Sie ermöglichen das Finden, das Identifizieren, das Selektieren und die Nutzung von Objekten, die für den Nutzer relevant sind. Metadaten verbessern den Austausch von Informationen.

Die andere Frage die sich stellt ist: Wo werden Metadaten benötigt? Wo kann man Metadaten nutzen und einsetzen?

Sie werden in Organisationen benötigt um Informationsobjekte in den Bestand aufzunehmen und zu organisieren.

Man benötigt sie zwischen Organisationen um den Austausch und die Nachnutzbarkeit von Informationen zu erleichtern.

Außerdem braucht man sie in Datenportalen um Informationen zu suchen, zu finden und zu nutzen

Metadaten kann man in fünf verschiedenen Typen unterscheiden.

Hierbei werden die Metadaten danach kategorisiert, welche Art von Daten sie beschreiben.

Es gibt:

- Identifizierende/beschreibende Metadaten
- Administrative Metadaten
- Content ratings Metadaten
- Linkage/relationship Metadaten
- Meta-Metadaten

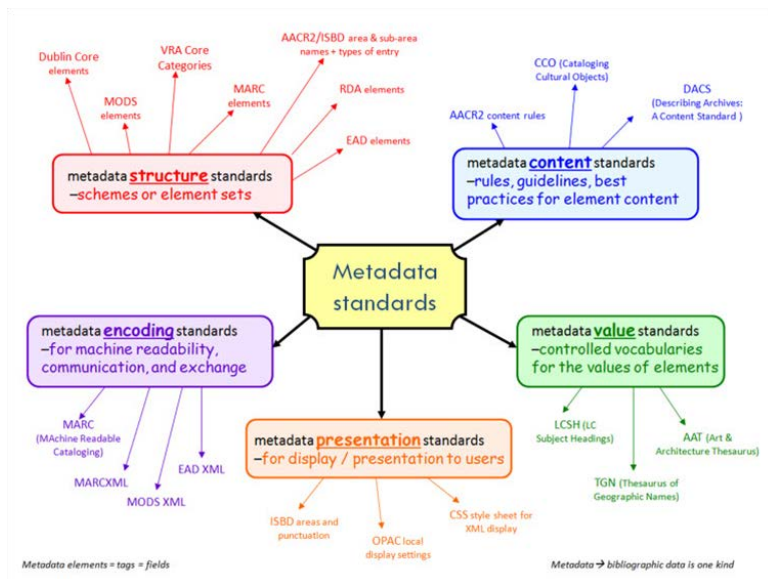
Die *beschreibenden Metadaten* sind Informationen, die nötig sind, um relevante Objekte zu suchen, zu finden und zu identifizieren. Hierzu gehören unter anderem der Titel, der "Schöpfer" einer Ressource oder deren Veröffentlichungsdatum (title, creator, issued).



Die *administrativen Metadaten* bieten Informationen zur Herkunft, Archivierung, Prozessen, Technik, Zugriff auf Objekte usw., die der Verwaltung der Objekte dienen (rights, owner, available, condition).

Informationen über mögliche Nutzer von Objekten (audience) stehen in den *Contentings Metadaten*. Die *Linkage / Relationship Metadaten* sind Informationen über die Beziehungen, die zwischen den verschiedenen Objekten bestehen (isPartOf, hasVersion, source).

Welche Modelle, Syntax und Formate zugrundeliegen, wer die Metadaten wann erstellt hat und weitere Informationen in dieser Richtung werden in den *Meta-Metadaten* festgehalten.



(Quelle: <https://pantherfile.uwm.edu/mlf/www/resource.html>)



Unterschiedliche Level - verschiedene Typen

In dieser Aufstellung werden Metadaten sogar in sieben Typen aufgeteilt. Bemerkenswert bei dieser Aufstellung ist die Tatsache, dass auf den ersten Blick gleiche Informationen (creator und source) an verschiedenen Stellen auftauchen können.

Einmal taucht der creator im Sinne des Autors eines Werkes bei den beschreibenden Metadaten auf. Auch die Herkunftsmetadaten (provenance Metadaten) haben einen Autor. Diese Daten haben auch eine Quelle, aus der sie stammen. Doch auch bei der Verlinkung verschiedener Ressourcen gibt es eine Quelle.

Typology of 7 types of metadata	Metadata Functions <i>"This type of metadata"</i>	Element examples*
Identification / description metadata	RESOURCE DISCOVERY / INFORMATION RETRIEVAL	Creator (Author), Title, Subject
Administrative metadata	RESOURCE MANAGEMENT	Price, Condition
Terms and conditions metadata	RESOURCE USAGE	Rights, Reproduction restrictions
Content ratings metadata	RESOURCE USE BY APPROPRIATE AUDIENCES	Audience
Provenance metadata	RESOURCE AUTHENTICATION AND OTHER PROVENANCE-RELATED ACTIVITIES	Creator , Source
Linkage / relationship metadata	RESOURCE LINKING WITH RELATED RESOURCES	Relation, Source
Structural metadata	RESOURCE HARDWARE AND SOFTWARE NEEDS	Compression ratio

(Quelle: <http://www.slideshare.net/mzeng/dublin-core-in-practice-1129092>)

Beispiel für Meta-Metadaten



Hier ein Beispiel für Meta-Metadaten in RDF:

```
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-  
rdf-syntax-ns#"  
  xmlns:dcterms="http://purl.org/dc/terms/"  
  xmlns:ex="http://example.org/taxonomy/"  
  xmlns:dcam="http://purl.org/dc/dcam/">  
  <rdf:Description rdf:about="http://example.org/123">  
    <dcterms:subject>  
      <rdf:Description>  
        <dcam:memberOf rdf:resource="http://example.org/  
taxonomy/MyVocab"/>  
        <rdf:value>Ornitology</rdf:value>  
      </rdf:Description>  
    </dcterms:subject>  
  </rdf:Description>  
</rdf:RDF>
```



Metadatenanwendung

Es gibt unterschiedlichste Metadatenanwendungen. Aus dem bibliothekarischen Kontext kennt man:

- OPAC
- Repository (z. B. OPUS, Dspace u.a.)
- Suchmaschinenindex (Base, OAister , u.a.)
- Metasuche (KVK u.a.) – ...

Hier ein Beispiel für Metadaten in einem OPAC. Man findet eine ganze Reihe von Metadaten über das Buch auf formaler und auf inhaltlicher Ebene. Aufgrund dieser Metadaten kann man entscheiden, ob man die Ressource bestellen möchte oder nicht.

	
Link zu diesem Datensatz	http://d-nb.info/986682225
Titel	Multimedia semantics - the role of metadata : with 20 tables / Michael Granitzer ... (eds.)
Person(en)	Granitzer, Michael [Hrsg.]
Verleger	Berlin ; Heidelberg : Springer
Erscheinungsjahr	2008
Umfang/Format	XII, 261 S. : Ill., graph. Darst. ; 24 cm
Gesamttitle	Studies in computational intelligence ; Vol. 101
Parallele Ausgabe(n)	Online-Ausg.: Multimedia semantics - the role of metadata
Anmerkungen	Literaturangaben
ISBN/Einband/Preis	978-3-540-77472-3 Pp. : EUR 106.95 (freier Pr.), ca. sfr 174.00 (freier Pr.)
Bestellnummer(n)	12047772
EAN	9783540774723
Schlagwörter	Multimedia ; Metadaten ; Semantisches Netz ; Aufsatzsammlung
DDC-Notation	006.7 [DDC22ger]
Sachgruppe(n)	004 Informatik
Links	Inhaltsverzeichnis

(Quelle: <http://d-nb.info/986682225>)

Weitere Beispiele für Metadatenanwendungen



OPUS ist eine Metasuche nach elektronischen Hochschulschriften. Man kann z.B. nach einem Autor in den Hochschulschriftendatenbanken mehrerer Hochschulen und Universitäten suchen.

(Quelle: http://elib.uni-stuttgart.de/opus/gemeinsame_suche.php)

Der Karlsruher Virtuelle Katalog (KVK) ermöglicht eine Metasuche in nationalen und internationalen Bibliothekskatalogen.

(Quelle: <http://www.ubka.uni-karlsruhe.de/kvk.html>)



Metadatenanwendung

Eine Metadatenanwendung kann verschiedene Ebenen haben:

Einen Speicher, in dem Metadaten vorgehalten werden.

Ein oder mehrere Erfassungsmodule, über die die Metadaten in den Speicher importiert werden.

Zwischen Erfassungsmodule und Speicher braucht es eine Import-Schnittstelle.

Eine oder mehrere Präsentationsebenen, über die die Metadaten den Nutzern zur Verfügung gestellt werden.

Zwischen Präsentationsebene und Speicher braucht es eine Export-Schnittstelle.

Zudem werden Metadaten aus anderen Anwendungen in den Speicher importiert.

Dazu braucht es ebenfalls eine Importschnittstelle.

Und es werden Metadaten aus der eigenen Anwendung nach außen exportiert.

Also noch eine Exportschnittstelle.

Weitere Informationen zu Schnittstellen auch im Kapitel **Metadatenstandards** in der Lektion "Kontrollierte Vokabulare und Schnittstellen".

Übung: Typen von Metadaten



Bitte ordnen Sie die korrekten Beschreibungen zu, die Lösung finden Sie auf der nächsten Seite:

Identifizierende / beschreibende Metadaten

Administrative Metadaten

Content ratings Metadaten

Linkage / relationship Metadaten

Meta-Metadaten

= Informationen, die nötig sind, um relevante Objekte zu suchen, zu finden und zu identifizieren.

= Informationen über mögliche Nutzer von Objekten

= Informationen zu Herkunft, Archivierung, Prozessen, Technik, Zugriff auf Objekte usw., die der Verwaltung der Objekte dienen.

= Informationen über die Beziehungen, die zwischen den verschiedenen Objekten bestehen

= Zugrunde liegende Modelle, Syntax und Formate



Lösung: Typen von Metadaten

Identifizierende / beschreibende Metadaten

Administrative Metadaten = Informationen, die nötig sind, um relevante Objekte zu suchen, zu finden und zu identifizieren.

Administrative Metadaten = Informationen zu Herkunft, Archivierung, Prozessen, Technik, Zugriff auf Objekte usw., die der Verwaltung der Objekte dienen.

Content ratings Metadaten = Informationen über mögliche Nutzer von Objekten

Linkage / relationship Metadaten = Informationen über die Beziehungen, die zwischen den verschiedenen Objekten bestehen

Meta-Metadaten = Zugrunde liegende Modelle, Syntax und Formate



Thema 2

Metadatenstandards

Das Kapitel führt in die Vielfalt der Metadatenstandards ein. Klassische Bibliotheksstandards finden hier ebenso ihre Beachtung wie neue Entwicklungen.

Was sind Metadatenstandards?

Ein Metadatenstandard ist eine Spezifikation, die anwendungsübergreifend die in einer bestimmten Community, aber auch communityübergreifend verbreiteten Metadaterme detailliert beschreibt - unter Berücksichtigung der Anforderungen, die diese Community hat. Sie dient der Festlegung von Elementen, der Struktur dieser Elemente und der Bedeutung dieser Elemente.

Ein Metadatenstandard ermöglicht die semantische Interoperabilität zwischen Anwendungen, die diesen Standard verwenden. Das Ziel ist eine einheitliche und maschinenverarbeitbare Beschreibung von Ressourcen und der Austausch dieser Beschreibungen.

(Quelle: http://www.kim-forum.org/Subsites/kim/DE/Materialien/Glossar/glossar_node.html#)



Warum brauchen wir Metadatenstandards?

Metadatenstandards könnte man als eine "gemeinsame Sprache" für den Austausch von Metadaten bezeichnen.

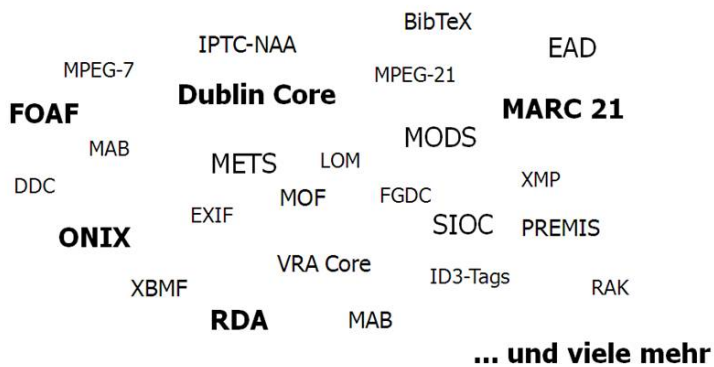
Mit Hilfe von Metadatenstandards sind Metadaten beim Austausch verständlich und interpretierbar. Sie sind nötig, um Metadaten effektiv einsetzen zu können. Denn sie bieten sowohl Datenformate für eine effektive Datenhaltung, als auch konsistente Daten für ein effektives Retrieval. Metadatenstandards dienen der Dokumentation, der Transparenz und der Verknüpfung von Objekten.

Somit sorgen Metadatenstandards für die Konsistenz der Metadaten und verbessern die Metadatenutzung. Außerdem erlauben sie den Austausch von Metadaten und die Verlinkung von Objekten.

Ein weiterer Vorteil ist, dass Metadaten anwendungsübergreifend sind und in verschiedenen Kontexten genutzt werden können.



Metadatenstandards - eine Auswahl





Hier einige Beispiele, welche Standards in welchen Institutionen oder Bereichen Anwendung finden:

Für Bibliotheken: AACR2 (Anglo-American Cataloguing Rules), BibTeX, DCMI und DCMI-Terms, MARC21, PICA (Project of Integrated Catalogue Automation), RAK (Regeln für die alphabetische Katalogisierung), RDA (Resource Description and Access) u.a.

Für Archive: EAD (Encoded Archival Description), SAFT (Standard-Austauschformat) u.a.

Für Museen: CDWA (Categories for the Description of Works of Art), LIDO (Leightweight Information Describing Objects), museumdat, CIDOC-CRM u.a.

Für Linked Data: Dublin Core, FOAF (Friend of a Friend), OAI-ORE (OAI-Object Reuse and Exchange), OWL (Web Ontology Language), SKOS (Simple Knowledge Organization System) u.a.

Typen von Metadatenstandards

Metadatenstandards kann man in folgende vier Typen einteilen:

Datenstruktur und Semantik	<ul style="list-style-type: none">• Formate (z. B. DC, PICA, MARC)• Syntax (z. B. XML)• Modelle (z. B. RDF)
Inhalte	<ul style="list-style-type: none">• Regelwerke (z. B. RAK-WB, AACR2, RDA)
Werte	<ul style="list-style-type: none">• Normdaten, kontrollierte Vokabulare (z. B. PND, LCSH, DDC)
Kommunikation	<ul style="list-style-type: none">• Schnittstellen, Protokolle (z. B. Z39.50, SRU, OAI)



Typen von Formaten

Bei den Formaten unterscheidet man zwischen dem Erfassungsformat, dem Internformat und dem Austauschformat.

Erfassungsformat

Das Erfassungsformat dient der Eingabe von Metadaten, z.B. zur Erfassung von Titel- oder Normdatensätzen. Das Format hierbei ist beispielsweise PICA3. Das Erfassungsformat sollte möglichst einfach für den Metadaten-Erfasser anwendbar sein.

Internformat

Das Internformat ist das Speicherformat, das zur Speicherung und Verwaltung von Metadaten in Datenbanken dient. Dieses Format ist beispielsweise PICA+.

Austauschformat

Das Austauschformat dient dem Austausch von Metadaten zwischen Partnern oder Datenbanken. D.h., dass das Format Absprachen zur Übertragung von Daten enthält. Dieses Format ist beispielsweise MARC 21.

Beispiel für Schlagwörter und DDC-Klassen in PICA3

```
5301 !150016492!54.72 / Künstliche Intelligenz
5302 !974195162!54.65 / Webentwicklung, Webanwendungen
5303 !150003889!06.74 / Informationssysteme
5400 [DDC22ger]004.67802856332
5401 004.678
5402 006.332
5403 -T1--0285
```

und Pica+

```
045F feDDC22gerfa004.67802856332
045F/01 fa004.678
045F/02 fa006.332
045F/03 ff0285
045Q/01 f9150016492f854.72 / Künstliche Intelligenz
045Q/02 f9974195162f854.65 / Webentwicklung, Webanwendungen
045Q/03 f9150003889f806.74 / Informationssysteme
```

Verwendung von Metadatenstandards



Wenn man einen Metadatenstandard verwenden möchte, sollte man sich zuerst folgende Frage stellen:

- Gibt es für meine Anwendung bereits einen Standard?
- Welcher Standard wird in meiner Community genutzt?
- Welcher Standard passt zu meiner Anwendung/in meinem Kontext?
- Welcher Standard entspricht am ehesten meinen Anforderungen?

Gleichzeitig sollte man noch daran denken, dass die Auswahlkriterien für einen Metadatenstandard die Anforderungen und die Nutzbarkeit der Anwendung sein sollten. Wichtig ist es außerdem, an die richtige Dokumentation zu denken.

Man sollte immer darauf achten, dass man die semantischen Grenzen eines existierenden Standards nicht verletzt.

Eine semantische Verfeinerung ist in Ordnung. Von einer semantischen Erweiterung sollte man jedoch absehen, da sonst die Interoperabilität nicht mehr gewährleistet werden kann.

Nehmen wir zum Beispiel eine statistische Tabelle, die die Jahre 1970 – 1975 behandelt. Diese Tabelle wurde 1976 erstellt und 1977 veröffentlicht.

Für die Beschreibung verwenden wir Elemente des Dublin Core Metadata Element Set.

Für die Datumsangaben wollen wir `dc:date` verwenden. Die Definition lautet:

"A point or period of time associated with an event in the lifecycle of the resource."

Dieses Element kann für verschiedene Datumsangaben verwendet werden. Hier für das Erscheinungsdatum (1977), aber auch für das Entstehungsdatum (1976). Beide Angaben fallen in den semantischen Kontext dieses Elements. Die Dublin Core Elemente `dcterms:created` und `dcterms:issued` sind dementsprechend Verfeinerungen des Elements `dc:date`.

Die Jahre, die die Tabelle behandelt (1970 – 1975) sind ebenfalls Datumsangaben, betreffen aber kein Ereignis im "Lebenszyklus" der Tabelle, sondern die Jahre, die die Ressource inhaltlich abdeckt. In diesem Fall darf dc:date daher nicht verwendet werden. An dieser Stelle nutzen wir das Element dc:coverage



"The spatial or temporal topic of the resource, the spatial applicability of the resource, or the jurisdiction under which the resource is relevant."

oder eine Verfeinerung dieses Elements, dcterms:temporal, definiert als:

"Temporal characteristics of the resource."

Schnittstellen



Um Metadaten mit anderen Institutionen austauschen zu können, braucht man Schnittstellen. Diese benötigen Interoperabilität. Interoperabilität wird gewährleistet durch definierte Metadaten-Schemata, durch Konkordanzen zwischen Schemata/Modellen und durch gemeinsame Transferprotokolle.

Die Schnittstellen dienen der Kommunikation. Es gibt eine ganze Reihe von verschiedenen Schnittstellen. So beispielsweise:

- Z39.50
- SRU/SRW
- OAI
- SOAP
- SOA
- SPARQL

Eine Auswahl dieser Schnittstellen wird auf den folgenden Seiten vorgestellt.



Z39.50

Z39.50 ist eine Schnittstelle bzw. ein Protokoll zur Kommunikation zwischen heterogenen Bibliothekssystemen. Es ist die technische Realisierung von Retrieval in verteilten Systemen. Mit Z39.50 kann man externe Datenquellen abfragen und gegebenenfalls strukturierte Daten übernehmen. Anwendung findet Z39.50 beispielsweise in Portalen.

Entwickelt und verwaltet wird die Schnittstelle von der Library of Congress (LOC). Inzwischen ist häufig der SRU-Standard im Einsatz, eine Weiterentwicklung des Z39.50-Protokolls.

SRU / SRW

SRU steht für "Search/Retrieve via URL", SRW für "Search/Retrieve Webservice". Es ist ein standardisiertes Webservice-Protokoll und stellt eine Weiterentwicklung von Z39.50 dar.

Anwendung findet es bei der Suchabfrage von Datenbanken und bei der gezielten Suche mittels Suchindex und Suchbegriffen (Boolesche Suche). Auch bei der Datenübernahme von Treffern in die eigene Anwendung kann es eingesetzt werden. Hier werden beispielsweise folgende Formate unterstützt (Beispiele aus dem Service der DNB): MARC-XML, MAB-XML und OAI-DC. Die Retrievalsprache ist CQL (Common/Contextual Query Language).

(Weitere Informationen unter: <http://www.loc.gov/standards/sru/>)

Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH)



Das OAI-Protokoll ermöglicht den Austausch von Metadaten zwischen Dokumentenservern und Internetdiensten. Dabei bietet ein Daten-Provider seine Metadaten über eine Internet-Schnittstelle an, die dort von einem Service-Provider geholt und für eigene/neue Dienste genutzt werden können. Mit dem OAI-PMH wurde ein anwendungsunabhängiges Framework entwickelt, dessen Austausch von Metadaten auf dem Prinzip von Metadata Harvesting beruht. Harvesting bedeutet in diesem Kontext das Zusammentragen von Metadaten aus verschiedenen Quellen in einen gemeinsamen Datenspeicher. Ein Harvesting-Protokoll besteht aus einem Set von Regeln, die die Kommunikation zwischen den Systemen definieren.

Quelle:

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

Offene Schnittstelle zum Austausch von XML-Daten

Abfragen eines OAI-Repository mit einem einfachen Web-Browser

DCMES 1.1 verpflichtend



[http://services.dnb.de/oai/repository?verb=ListRecords
&from=2008-06-09T08:11:11Z &until=2008-06-10T15:13:13Z
&set=ddb:publikationen:online
&metadataPrefix=onix](http://services.dnb.de/oai/repository?verb=ListRecords&from=2008-06-09T08:11:11Z&until=2008-06-10T15:13:13Z&set=ddb:publikationen:online&metadataPrefix=onix)

Quelle: <urn:nbn:de:101-2009120311>)

<http://www.openarchives.org/pmh/>

XML



XML (eXtensible Markup Language) ist eine Metasprache, auf deren Grundlage anwendungsspezifische Sprachen entwickelt werden können. Mit Hilfe von XML kann man eine maschinenlesbare Beschreibung von Metadaten erstellen.

Im konkreten Fall benötigt XML eine "Dokumenttypdefinition" (DTD), die das Vokabular festlegt, das für die Beschreibung gilt. Das Vokabular wird in der DTD oder im Schema spezifiziert.

XML-Sprachen sind: TEI (für Text), GML (für Geodaten), MPE7 (für Daten aus dem Multimedia-Bereich) u.v.m.

SPARQL

SPARQL ist das rekursive Akronym für "SPARQL Protocol and RDF Query Language". Es ist eine Anfragesprache auf RDF-Graphen und das Protokoll zur Übertragung von SPARQL-Anfragen und -Ergebnissen.

Mit der W3C-Spezifikation (seit 15. Januar 2008) können mehrere Datenquellen gleichzeitig abgefragt werden. Hierbei sind optionale Daten bei der Abfrage möglich oder auch Filtern und Sortieren.

Hier ein Beispiel für eine SPARQL-Anfrage. Diese Anfrage listet die Namen aller Mitarbeiter in alphabetischer Reihenfolge:

```
select ?name where {  
  ?emp rdf:type f:emp.  
  ?emp foaf:surname ?name.  
}  
ORDER BY ?name
```

(Quelle: http://en.wikibooks.org/wiki/XQuery/SPARQL_Tutorial)



(Quelle: <http://ganesha.fr/public/images/articles/dublinCoreMetaData.jpg>)

Dublin Core ist ein Standard zur Beschreibung aller Arten von Ressourcen. Er ist sowohl international, als auch domain-übergreifend. So wird der Dublin Core Standard beispielsweise in Museen, Bibliotheken, Archiven, aber auch Regierungsbehörden und in der Wirtschaft verwendet.

Entstanden ist der Dublin Core Standard auf dem ersten Dublin Core Metadata Workshop im März 1995 in Dublin, Ohio (USA). Die Ziele waren und sind die Einfachheit der Semantik und der Anwendung. Außerdem sollte die Basis für semantische Interoperabilität bereitgestellt werden. Es gibt viele Mappings von umfangreicheren Formaten zu DC.

Gepflegt und weiterentwickelt wird er durch die Dublin Core Metadata Initiative (DCMI).

Weitere Informationen unter <http://dublincore.org/>.

Dublin Core Metadata Element Set (DCMES)

Dublin Core besteht aus einem einfachen Element Set - auch als DC Simple bezeichnet. Dies sind 15 Elemente, die alle optional und alle wiederholbar sind.

Homepage: <http://dublincore.org/documents/dces/>

Alle Elemente des DC Simple - Sets:

(Quelle: <http://dublincore.org/documents/dcq-rdf-xml/>)

Die 15 Elemente geordnet nach formalen Kriterien, geistigem Eigentum und Inhalt:



Formale Kriterien	Geistiges Eigentum	Inhalt
Coverage	Contributor	Date
Description	Creator	Format
Language	Publisher	Identifier
Relation	Rights	Type
Source		
Subject		
Title		

DCMES wurde sowohl als ISO-Standard veröffentlicht (ISO 15836) als auch als NISO-Standard (Z39.85). Für OAI-PMH Schnittstellen ist die Verwendung des DCMES verpflichtend

"The OAI-PMH supports items with multiple manifestations (formats) of metadata. At a minimum, repositories must be able to return records with metadata expressed in the Dublin Core format, without any qualification."

<http://www.openarchives.org/OAI/openarchivesprotocol.html#Record>

U.a. aus diesem Grund ist DCMES das für die Erschließung von digitalen Ressourcen am weitesten verbreitete Metadatenformat.

Elemente bei Dublin Core

Beispiele für Elemente:

Term Name: title	
URI:	http://purl.org/dc/elements/1.1/title
Label:	Title
Definition:	A name given to the resource.
Comment:	Typically, a Title will be a name by which the resource is formally known.

Term Name: rights	
URI:	http://purl.org/dc/elements/1.1/rights
Label:	Rights
Definition:	Information about rights held in and over the resource.
Comment:	Typically, rights information includes a statement about various property rights associated with the resource, including intellectual property rights.

(Quelle: <http://dublincore.org/documents/dces/>)

DCMI Metadata Terms (DCMI-Terms)



Die DCMI-Terms sind eine Erweiterung des Dublin Core Metadata Element Set. Sie ermöglichen eine granularere Beschreibung von Ressourcen und definieren stellenweise auch die Werte, die bei der Beschreibung von Ressourcen verwendet werden dürfen. Die DCMI-Terms unterscheiden verschiedene Arten von Elementen. Dies sind:

Classes definieren die Art der Ressourcen, die beschrieben werden können aber auch die Art der Ressourcen, die bei der Beschreibung als Wert verwendet werden sollten.

Beispiel für eine class s. <http://dublincore.org/documents/dcmi-terms/#classes-Agent>

Properties definieren die Eigenschaften, die die Ressourcen haben und dienen der strukturierten Beschreibung dieser Ressourcen. Die DCMI-Properties umfassen zunächst die in den DCMES genannten Elemente, schränken die Verwendung derselben jedoch ein, indem sie die Werte, die verwendet werden dürfen, über das Attribut HasRange in vielen Fällen festlegen.

Beispiel für eine Property <http://dublincore.org/documents/dcmi-terms/#terms-creator>

Zudem umfassen die **DCMI-Terms** weitere Properties, die dazu dienen, Objekte granularer zu beschreiben. Diese Properties werden auch als Refinements bezeichnet. Zahlreiche Refinements gibt es z. B. für das Element dc:date (dcterms:issued, dcterms:created, dcterms:modified usw.) aber auch für das Element dc:relation (dcterms:hasFormat, dcterms:hasVersion, dcterms:isPartOf usw.). Diese Zuweisung erleichtert es, granulare Daten aus verschiedenen Anwendungen unter den weniger granularen DCMES Elementen zusammenzuführen, was u.a. für die Verwendung der Daten in OAI-Schnittstellen notwendig ist.

Beispiel für eine Property <http://dublincore.org/documents/dcmi-terms/#terms-isPartOf>

Die Verwendung der Dublin Core Properties wird im "Dublin Core Leitfaden" detailliert beschrieben.

Vocabulary Encoding Schemes (VES) definieren die kontrollierten Vokabulare und Normdateien, die verwendet werden sollten, um Ressourcen zu beschreiben.



Beispiel für VES: <http://dublincore.org/documents/dcmi-terms/#ves-TGN>

Syntax Encoding Schemes (SES) verweisen auf Standards, die für die Darstellung von Werten genutzt werden sollten.

Beispiel für SES <http://dublincore.org/documents/dcmi-terms/#ses-ISO639-2>

Term Name: alternative	
URI:	http://purl.org/dc/terms/alternative
Label:	Alternative Title
Definition:	An alternative name for the resource.
Comment:	The distinction between titles and alternative titles is application-specific.
Type of Term:	<i>Property</i>
Refines:	http://purl.org/dc/elements/1.1/title
Refines:	http://purl.org/dc/terms/title
Version:	http://dublincore.org/usage/terms/history/#alternative-003
Has Range:	http://www.w3.org/2000/01/rdf-schema#Literal

(Quelle: <http://dublincore.org/documents/dcmi-type-vocabulary/>)

DC Kernel

DC Kernel ist ein sehr generisches Format für die Beschreibung von Ressourcen jedweder Art.

Es besteht aus 4 Tag-Sets:

Elemente des ersten Sets: who , what , when , where , how

Elemente des zweiten Sets: about-who , about-what , about-when , about-where , about-how

Elemente des dritten Sets: meta- who , meta- what , meta-when , meta- where

Elemente des vierten Sets: support-who , support-what , support-when , support-where

Kernel Wiki: http://dublincore.org/kernelwiki/FrontPage?action=AttachFile&do=get&target=Kernel1_1.html

Links zum Thema Dublin Core



Metadaten Terme

Dublin Core Metadata Element Set (<http://dublincore.org/documents/dces/>)

DCMI Metadata Terms (<http://dublincore.org/documents/dcmi-terms/>)

DCMI Type Vocabulary (<http://dublincore.org/documents/dcmi-type-vocabulary/>)

Metadatenprofile

The Singapore Framework for Dublin Core Application Profiles (<http://dublincore.org/documents/singapore-framework/>)

Description Set Profiles - draft (<http://dublincore.org/documents/abstract-model/>)

Dublin Core Application Profile Guidelines – CWA 14855 (<ftp://ftp.cenorm.be/PUBLIC/CWAs/e-Europe/MMI-DC/cwa14855-00-2003-Nov.pdf>)

Weitere Empfehlungen

DCMI Abstract Model (<http://dublincore.org/documents/abstract-model/>)

Expressing Dublin Core Description Sets using XML (DC-DS-XML) (<http://dublincore.org/documents/2008/09/01/dc-ds-xml/>)

Expressing Dublin Core metadata using the Resource Description Framework (RDF) (<http://dublincore.org/documents/dc-rdf/>)



Übung: Dublin Core Properties 1

Die Lösungen finden Sie auf der nächsten Seite

1. Handelt es sich bei `dc:creator` um eine Dublin Core Property?
2. Handelt es sich bei `dc:author` um eine Dublin Core Property?
3. Handelt es sich bei `dcterms:contributor` um eine Dublin Core Property?
4. Handelt es sich bei `dcterms:agent` um eine Dublin Core Property?

Lösung: Dublin Core Properties 1



1. Handelt es sich bei dc:creator um eine Dublin Core Property? **JA**
2. Handelt es sich bei dc:author um eine Dublin Core Property? **NEIN**
3. Handelt es sich bei dcterms:contributor um eine Dublin Core Property? **JA**
4. Handelt es sich bei dcterms:agent um eine Dublin Core Property? **NEIN**

Übung: Dublin Core Properties 2



Die Lösungen finden Sie am Ende der Übung

1. Worum handelt es sich bei `dcterms:accessRights`?
 - A. Class
 - B. Property
 - C. Syntax Encoding Scheme
 - D. Vocabulary Encoding Scheme

2. Worum handelt es sich bei `dcterms:RightsStatement`?
 - A. Class
 - B. Property
 - C. Syntax Encoding Scheme
 - D. Vocabulary Encoding Scheme

3. Worum handelt es sich bei `dc:relation`?
 - A. Class
 - B. Property
 - C. Syntax Encoding Scheme
 - D. Vocabulary Encoding Scheme

4. Worum handelt es sich bei `dcterms:Period`?
 - A. Class
 - B. Property
 - C. Syntax Encoding Scheme
 - D. Vocabulary Encoding Scheme

5. Worum handelt es sich bei `dcterms:Frequency`?
 - A. Class
 - B. Property
 - C. Syntax Encoding Scheme
 - D. Vocabulary Encoding Scheme

6. Worum handelt es sich bei `dcterms:contributor`?
 - A. Class
 - B. Property
 - C. Syntax Encoding Scheme
 - D. Vocabulary Encoding Scheme



7. Worum handelt es sich bei `dcterms:BibliographicResource`?
 - A. Class
 - B. Property
 - C. Syntax Encoding Scheme
 - D. Vocabulary Encoding Scheme

8. Worum handelt es sich bei `dcterms:spatial`?
 - A. Class
 - B. Property
 - C. Syntax Encoding Scheme
 - D. Vocabulary Encoding Scheme

9. Worum handelt es sich bei `dcterms:URI`?
 - A. Class
 - B. Property
 - C. Syntax Encoding Scheme
 - D. Vocabulary Encoding Scheme

10. Worum handelt es sich bei `dc:description`?
 - A. Class
 - B. Property
 - C. Syntax Encoding Scheme
 - D. Vocabulary Encoding Scheme

11. Worum handelt es sich bei `dcterms:UDC`?
 - A. Class
 - B. Property
 - C. Syntax Encoding Scheme
 - D. Vocabulary Encoding Scheme



12. Worum handelt es sich bei dcterms:ISO639-3?
- A. Class
 - B. Property
 - C. Syntax Encoding Scheme
 - D. Vocabulary Encoding Scheme
13. Worum handelt es sich bei dcterms:SizeOrDuration?
- A. Class
 - B. Property
 - C. Syntax Encoding Scheme
 - D. Vocabulary Encoding Scheme
14. Worum handelt es sich bei dcterms:type?
- A. Class
 - B. Property
 - C. Syntax Encoding Scheme
 - D. Vocabulary Encoding Scheme
15. Worum handelt es sich bei dcterms:DCMIType?
- A. Class
 - B. Property
 - C. Syntax Encoding Scheme
 - D. Vocabulary Encoding Scheme
16. Worum handelt es sich bei dcmitype:Software?
- A. Class
 - B. Property
 - C. Syntax Encoding Scheme
 - D. Vocabulary Encoding Scheme

Lösungen: Dublin Core Properties 2



1. Worum handelt es sich bei `dcterms:accessRights`? **Property**

2. Worum handelt es sich bei `dcterms:RightsStatement`? **Class**

3. Worum handelt es sich bei `dc:relation`?

Property

4. Worum handelt es sich bei `dcterms:Period`?

Syntax Encoding Scheme

5. Worum handelt es sich bei `dcterms:Frequency`?

Class

6. Worum handelt es sich bei `dcterms:contributor`?

Property

7. Worum handelt es sich bei `dcterms:BibliographicResource`?

Class

8. Worum handelt es sich bei `dcterms:spatial`?

Property

9. Worum handelt es sich bei `dcterms:URI`?

Syntax Encoding Scheme



10. Worum handelt es sich bei dc:description?

Property

11. Worum handelt es sich bei dcterms:UDC?

Vocabulary Encoding Scheme

12. Worum handelt es sich bei dcterms:ISO639-3?

Syntax Encoding Scheme

13. Worum handelt es sich bei dcterms:SizeOrDuration

Class

14. Worum handelt es sich bei dcterms:type?

Property

15. Worum handelt es sich bei dcterms:DCMIType?

Vocabulary Encoding Scheme

16. Worum handelt es sich bei dcmitype:Software?

Class



Weitere Metadatenformate

PICA

PICA ist ein Format, das als Katalogisierungs- bzw. Erfassungsformat oder auch als Internformat Verwendung findet. PICA3 wird als Erfassungsformat genutzt, wohingegen PICA+ ein Internformat ist.

Die Software von OCLC-PICA gibt es für zentrale und lokale Bibliotheksdatenbanken. Anwender sind die Bibliotheksverbände (z.B. der GBV, hebis und der SWB) und die Deutsche Nationalbibliothek. Es ist ein sehr granulares Format.

Die Beispiele auf den folgenden Seiten haben einen Farbencode.

Rot sind Angaben zum Format, schwarz sind Angaben, die durch das Regelwerk bestimmt sind.

Die Angaben in der Farbe lila sind nur bei der Anzeige eingeblendet, aber nicht gespeichert. Die Nummern, z.B. !106405799!, sind IdNummern von Normsätzen, und nur in diesen Normsätzen stehen die Klartexte. Diese werden mit angezeigt und exportiert, damit ein empfangendes System auch damit, anstatt mit den Nummern arbeiten kann.



Pica3

0100 027252086

0500 Aau

1100 1980

1500 /1de

1700 /1de

2000 3-489-60534-9

2065 801020409

2100 80A080107

2105 79N390243

2190 1116/23P

3000 Rupert@Riedl

3010 Robert@Kaspar [Mitarb.]

4000@Biologie der Erkenntnis : die stammesgeschichtlichen
Grundlagen der Vernunft / Rupert Riedl. Unter Mitarb. von
Robert Kaspar

4030 Berlin [u.a.] : Parey

4060 230 S

4061 III

4201 Literaturverz. S. 215 - 223

5050 03a ; 18a

5301 !106405799!**42.88 Physische Anthropologie**

5302 !106420259!**42.65 Tiergeographie**

5550 !104591862!**Evolutionäre Erkenntnistheorie**

5550 !104755482!**Humanbiologie**

5550 !106098691!**Evolution**

(Quelle: <http://www.allegro-c.de/formate/examp/examp1.htm>)

Maschinelles Austauschformat für Bibliotheken (MAB)



Das Maschinelle Austauschformat für Bibliotheken (MAB) ist ein nationales Austauschformat für Titeldaten, Normdaten und Lokaldaten. Es wurde 1973 erstellt, 1995 folgte dann eine neue Version: MAB2.

MAB gibt es auch für XML: MAB-XML.

Bis zum Bibliografiejahrgang 2009 wurde MAB2 von der Deutschen Nationalbibliothek für ihre Datendienste eingesetzt, danach fand der Umstieg auf MARC 21 statt.

MAB2 (BVB)

MAB hat je 3 Felder fuer MARC 245 und 260. Was bei MARC in den "fixed fields" steht (008) ist bei MAB und Pica auf mehrere Felder verteilt.

```
### 00889nM2.01000024 h
001 00012645125
002a19890315
003 20001120
004 20020110152442.0
030 azzzc||z||17
037bger
050 a
051 m
100 Riedl, Rupert
331 Biologie der Erkenntnis
335 die stammesgeschichtlichen Grundlagen der Vernunft
359 von Rupert Riedl
403 2. Aufl.
410 Berlin [u.a.]
412 Parey
425 1980
433 230 S. : graph. Darst.
540aISBN 3-489-60734-1
900 00003313204
902 |Biologie
```

902 | Erkenntnistheorie
903 21904a024|008|013|004|079
905 00007729502
907 | Evolutionäre Erkenntnistheorie
909a056|017|027
910 00001427539
912 | Humanbiologie
914a069|079
915 00013336479
917 | Evolution
917 | Erkenntnistheorie
919a079

(Quelle: <http://www.allegro-c.de/formate/examp/examp1.htm>)





MARC 21 ist ein internationales Austauschformat für bibliografische Daten, Normdaten und Besitznachweise. Es findet aber auch Einsatz als Erfassungs- und Internformat.

Entstanden ist MARC 21 in den 60er Jahren aus UKMARC und CANMARC. Gepflegt und weiterentwickelt wird das Format vom Network Development and MARC Standards Office (NDMSO). Es entstammt der Formatfamilie UNIMARC, UKMARK, MARC 21 und MARCXML.

In Deutschland und Österreich fand ein Umstieg von MAB auf MARC 21 statt. In der Deutschen Nationalbibliothek werden Titeldaten seit Dezember 2008 und Normdaten sowie Titel- und Bestandsdaten der ZDB seit Februar 2009 in MARC 21 ausgetauscht.

Weitere Informationen unter <http://www.loc.gov/marc/>.

MARC21



Man sieht hier Abkürzungen im Titelzusatz, unten bei MAB dagegen nicht. AACR hat dafür jedoch nie Abkürzungen vorgeschrieben, RAK hat sie irgendwann abgeschafft. Vermutlich hat hier die LC den CIP-Satz der DB benutzt, der diese Abkürzungen enthielt. Die ISBN wird ohne Bindestriche eingegeben (Feld 020).

001 3252471
005 19850520000000.0
008 800326s1980 wb a b 001 0 ger
035 \$9(DLC)80461766
906 \$a7\$bcbc\$corignew\$d4\$encip\$f19\$gy-gencatlg
010 \$a 80461766
015 \$aGFR80-A
020 \$a3489607341
050 00\$aQP398\$b.R53 (\$b = CutterNr., hier fehlt noch das Jahr am Ende | \$b should have year)
100 1 \$aRiedl, Rupert.
245 10\$aBiologie der Erkenntnis :\$bd.
stammesgeschichtl.
Grundlagen d. Vernunft /\$c von Rupert Riedl, unter
Mitarbeit von
Robert Kaspar.
260 \$aBerlin ;\$aHamburg :\$bParey,\$c1980.
300 \$a230 p. :\$b60 ill. ;\$c24 cm.
504 \$aBibliography: p. 215-223.
500 \$aIncludes indexes.
650 0 \$aIntellect.
650 0 \$aPerception.
700 1 \$aKaspar, Robert,\$ejoint author.

(\$e wird selten verwendet | \$e not used in most libraries)

(Quelle: <http://www.allegro-c.de/formate/examp/examp1.htm>)



Der Metadata Encoding and Transmission Standard (METS) ist ein Format zur Beschreibung von komplexen digitalen Bibliotheksobjekten. Es werden sowohl die Struktur der Objekte, als auch die deskriptiven und administrativen Metadaten erfasst.

Eingesetzt wird METS für die Verwaltung von digitalen Objekten und den Austausch dieser Objekte.

METS ist ein XML-Dokumentformat. Zudem ist METS ein Containerformat, das die Struktureinheiten von digitalisierten Objekten beschreibt und für den Nutzer verfügbar macht. Für die bibliographische Beschreibung dieser Struktureinheiten können die unterschiedlichsten Metadatenformate verwendet werden. So erfordert die Anzeige von Digitalisaten im DFG-Viewer (<http://dfg-viewer.de/>) die Kombination von METS- und MODS-Daten.

Bsp. für einen METS-MODS-Datensatz:

```
http://gdz.sub.uni-goettingen.de/oai2/?  
verb=GetRecord&metadataPrefix=mets&identifier=gdz.sub.uni-  
goettingen.de:PPN62386889X
```

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

Metadata Object Description Schema (MODS)



Das Metadata Object Description Schema (MODS) ist ein Standard für deskriptive Metadaten. Es ist ein Derivat von MARC 21. Der Fokus liegt auf der Beschreibung von digitalen Objekten. Auch einzelne Komponenten von Objekten können beschrieben werden.

Auch MODS ist XML-basiert. Die Aktuelle Version ist 3.4 (2010). Gepflegt und weiterentwickelt wird das Schema vom Network Development and MARC Standards Office (NDMSO). Ein verwandtes Format zu MODS ist MADS, der Metadata Authority Description Standard.

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

Link zu MADS: <http://www.loc.gov/standards/mads/>



EDITEUR

ONIX ist ein internationaler Metadatenstandard für Produkte des Buchhandels. Er wird als Austauschformat für Verlage und Buchhandel genutzt und zum Teil auch von Bibliotheken eingesetzt. Beschrieben werden bibliografische und vertriebsrelevante Metadaten.

Der Standard hat ein XML-basiertes Format. Die Pflege und Weiterentwicklung liegt in den Händen des ONIX International Steering Committee von EDItEUR. Es gibt eine ganze Reihe von Formatfamilien: ONIX for Books, ONIX for Serials, ONIX for Licensing Terms, ONIX for DOI Registration Format und ONIX ISTC Registration Formats.

Homepage: <http://www.editeur.org/8/ONIX/>

Weitere Standards



DOI metadata specification

Der Digital Object Identifier (DOI ist ein persistenter und eindeutiger Identifier für digitale Objekte, der vor allem für Zeitschriftenaufsätze verwendet wird.

Die DOI metadata specification dient der Beschreibung von Objekten, die mit einem DOI identifiziert werden.

Die Spezifikation besteht aus einigen wenigen „deskriptiven“ und administrativen properties.

Spezifikation: http://www.doi.org/handbook_2000/metadata.html

ContextObjects in Span (COinS)



ContextObjects in Span (COinS) dient der Einbindung von Metadaten in HTML-Seiten auf der Grundlage des OpenURL - Standards.

Der zugrundeliegende Metadatenstandard ist das KEV (Key encoded values) Format, das unterscheidet zwischen

KEV-Format to represent a book

[http://alcme.oclc.org/openurl/servlet/OAIHandler/extension?](http://alcme.oclc.org/openurl/servlet/OAIHandler/extension?verb=GetMetadata&metadataPrefix=mtx&identifier=info:ofi/fmt:kev:mtx:book)

[verb=GetMetadata&metadataPrefix=mtx&identifier=info:ofi/fmt:kev:mtx:book](http://alcme.oclc.org/openurl/servlet/OAIHandler/extension?verb=GetMetadata&metadataPrefix=mtx&identifier=info:ofi/fmt:kev:mtx:book)

KEV-Format to represent a journal publication

[http://alcme.oclc.org/openurl/servlet/OAIHandler/extension?](http://alcme.oclc.org/openurl/servlet/OAIHandler/extension?verb=GetMetadata&metadataPrefix=mtx&identifier=info:ofi/fmt:kev:mtx:journal)

[verb=GetMetadata&metadataPrefix=mtx&identifier=info:ofi/fmt:kev:mtx:journal](http://alcme.oclc.org/openurl/servlet/OAIHandler/extension?verb=GetMetadata&metadataPrefix=mtx&identifier=info:ofi/fmt:kev:mtx:journal)

KEV-Format to represent the Dublin Core Element Set

[http://alcme.oclc.org/openurl/servlet/OAIHandler/extension?](http://alcme.oclc.org/openurl/servlet/OAIHandler/extension?verb=GetMetadata&metadataPrefix=mtx&identifier=info:ofi/fmt:kev:mtx:dc)

[verb=GetMetadata&metadataPrefix=mtx&identifier=info:ofi/fmt:kev:mtx:dc](http://alcme.oclc.org/openurl/servlet/OAIHandler/extension?verb=GetMetadata&metadataPrefix=mtx&identifier=info:ofi/fmt:kev:mtx:dc)

Spezielle Browser Plugins (z.B. Zotero oder Citavi) können die Metadaten auslesen und für weitere Anwendungen zur Verfügung stellen.

Standards für Textobjekte



Text Encoding Initiative (TEI) P5 ist ein XML-basierter Standard für die differenzierte Auszeichnung von Text.

Anwendung vor allem in den Geisteswissenschaften im Bereich der Editionswissenschaften und der Linguistik.

Es gibt zahlreiche, teilweise sehr granulare Element Sets, sog. Tag Sets, die sich in zwei Arten einteilen lassen:

Base Tag Sets = TEI.prosa , TEI.drama , TEI.verse , TEI.spoken usw.

Additional Tag Sets = TEI.linking , TEI.analysis , TEI.certainty , TEI.name.dates usw.

TEI-Homepage: <http://www.tei-c.org/index.xml>

Für Einsteiger: <http://computerphilologie.uni-muenchen.de/praxis/tei/prax.html> TEI

P5: <http://www.tei-c.org/Guidelines/P5/>

Standards aus dem Archivwesen

Encoded Archival Description (EAD) ist ein internationaler XML-basierter Standard für die Beschreibung von Archivmaterial und den Austausch von Metadaten aus diesem Bereich.

Besteht aus ca. 150 Elementen/Properties

EAD Homepage: <http://www.loc.gov/ead/>

Tag Library: <http://www.loc.gov/ead/tglib/index.html>

Standard-Austauschformat (SAFT)

Standard-Austauschformat (SAFT) ist ein deutscher XML-basierter Standard für die Beschreibung von Archivmaterial.

<http://internet.archivschule.uni-marburg.de/SAFT/doku.php>

Standards in Museen



Categories for the Description of Works of Art (CDWA)

http://www.getty.edu/research/conducting_research/standards/cdwa/

relationale Datenstruktur

Dient der Beschreibung von Museumsbeständen

CDWA lite

http://www.getty.edu/research/conducting_research/standards/cdwa/cdwalite.html

xml -Schema

Kern-Set aus CDWA

museumdat

museumdat

<http://www.museumdat.org/>

Deutsches XML -basiertes Austauschformat

dient der Beschreibung von Museumsbeständen

Lightweight Information Describing Objects (LIDO)

Lightweight Information Describing Objects (LIDO)

<http://www.athenaeurope.org/getFile.php?id=535>

Basiert auf verschiedenen Standards, u. a. CDWA lite und

museumdat

CIDOC-CRM

Kontrollierte Vokabulare und Normdateien



Kontrollierte Vokabulare

Ein kontrolliertes Vokabular verzeichnet Namen bzw. Bezeichnungen von Objekten, Konzepten, Personen, Organisationen usw. Ziel eines kontrollierten Vokabulars ist, dass die in einem bestimmten Kontext (z. B. bei der Beschreibung von Büchern) verwendeten Werte möglichst einheitlich und eindeutig sind, um so das Finden von Objekten zu erleichtern.

Dabei reichen kontrollierte Vokabulare von einfachen Termlisten über umfangreiche Glossare hin zu komplexen Normdateien.

Beispiele für kontrollierte Vokabulare sind:

- Einfache Verzeichnisse wie die MARC Relator Codes oder das DCMI Type Vokabular.
- Komplexe Normdateien wie die GND, LCSH
- Klassifikationen wie DDC oder UDC
- Thesauri wie STW oder TGN usw.

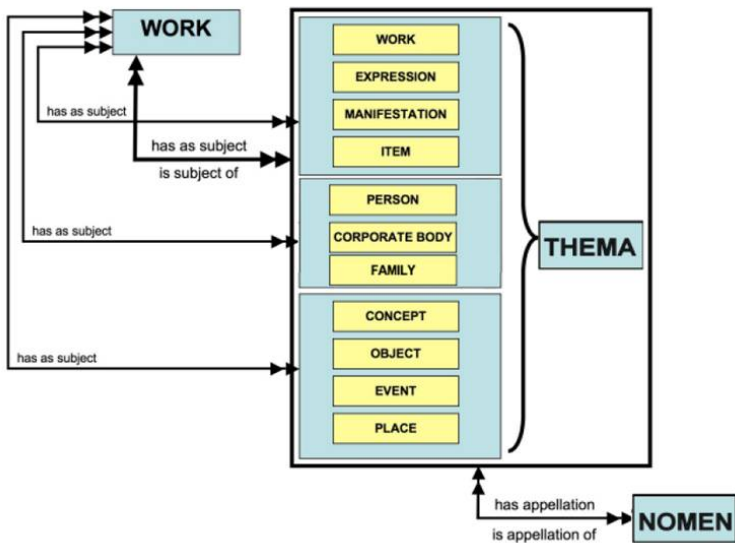
Normdaten



Eine Normdatei ist ein komplexes kontrolliertes Vokabular abgelegt in Normdatensätzen, die zusammen genommen die Normdatei bilden.

Nach den Functional Requirements for Subject Authority Data (FRSAD) gibt es folgende Typen von Normdaten:

- Personen
 - Körperschaften
 - Familien
 - WEMI (Work, Expression, Manifestation, Item)
- Objekt
 - Geografikum
 - Sachbegriff
 - Zeit



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

Datenelemente sind:

- Identifizier
- Namensformen (Ansetzungsform, alternative Benennungen, Synonyme, Äquivalente)
- Attribute
- Relationen



Die Praxis bei der Beschreibung von Normdaten weicht stellenweise stark voneinander ab, doch gibt es vier Attribute, die fast allen Normdaten gemein sind:

Identifizier – Normdaten sind i.d.R. durch eine eindeutige Zeichenkette identifiziert, die der beschriebenen Ressource (einem Objekt, Konzept, einer Person oder Organisation) innerhalb der Normdatei zugewiesen wird.

Namensformen – Normdaten enthalten die Namen bzw. Bezeichnungen, die für diese Ressource verwendet werden.

Beschreibung – Normdaten enthalten eine Definition oder eine andere Art der Beschreibung der Ressource, für die der Name bzw. die Bezeichnung steht.

Relationen – I.d.R. werden Normdaten zueinander in Beziehung gesetzt (z. B. Unterbegriff – Oberbegriff-Relationen)

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

Normdateien



Bei den Normdateien unterscheidet man zwischen fachspezifischen und fachübergreifenden Normdateien sowie Mappings und Crosskonkordanzen.

Fachübergreifende Normdateien sind beispielsweise die Personennamendatei (PND), die Gemeinsame Körperschaftsdatei (GKD), die Schlagwortnormdatei (SWD) und auch die Zeitschriftendatenbank (ZDB). PND, GKD und SWD werden im März 2012 zur Gemeinsamen Normdatei (GND) zusammengeführt.

Ebenso die Schlagwortdatei der Library of Congress LCSH (Library of Congress Subject Headings) und die französische Schlagwortdatei RAMEAU (Répertoire d'autorité-matière encyclopédique et alphabétique unifié).

Fachspezifische Normdateien sind beispielsweise der STW und der agrovoc-Thesaurus.

Beispiele für Mappings und Crosskonkordanzen sind die Projekte VIAF, MACS und CrissCross.

Fachübergreifende Normdateien



Zu den in Deutschland verwendeten fachübergreifenden Normdateien zählen:

- die Personennamendatei (PND)
- die Schlagwortnormdatei (SWD)
- die Gemeinsame Körperschaftsdatei (GKD)

Ab März 2012 werden diese drei Dateien zur Gemeinsamen Normdatei (GND) zusammengeführt.

Zeitschriftendatenbank (ZDB)

Auf die nationalen Normdateien kann auf unterschiedliche Weise zugegriffen werden. Beispielsweise gibt es einen Datenbankzugriff zum PICA-Iltis-System, eine SRU- und OAI-Schnittstelle, das DNB-Portal sowie einen Linked Data Service.

Personennamendatei (PND)



In der PND werden alle für Formal- und Sacherschließung sowie nationale Katalogisierungsunternehmungen wesentlichen Namen zusammengeführt. Sie stellt vor allem eine Dienstleistung zur Vermeidung von Mehrfacharbeit bei der Ansetzung von Personennamen sowie zur Vereinheitlichung der Ansetzungsformen dar. Darüber hinaus ist sie die Bezugsdatei bei der internationalen Kooperation im Bereich der Ansetzung der Personennamen. Sie ist somit eine wichtige überregionale Normdatei für Personennamen aus allen Jahrhunderten, aus allen Fach- und Wissensgebieten und aus allen Staaten. Sie wird für die formale und inhaltliche Erschließung von Ressourcen verwendet. Die PND wurde von 1995 bis 1998 in einem von der Deutschen Forschungsgemeinschaft (DFG) geförderten Projekt aufgebaut. Sie wird kooperativ von allen deutschen und österreichischen Bibliotheksverbänden sowie der Schweizerischen Nationalbibliothek, verschiedenen Projekten, Archiven und Museen gepflegt.

Die PND umfasst zur Zeit rund 3,6 Mio. Namen (davon ca. 1,8 Mio. Personen mit individualisierten Datensätzen) aus:

- der Katalogisierung der PND-Kooperationspartner und der Deutschen Nationalbibliothek
- Projekten der Altbestandserschließung und Namen der Bayerischen Staatsbibliothek
- aus dem osteuropäischen und islamischen Kulturkreis
- aus den Bereichen Musik und Karten
- aus den Personennamen der Zentralkartei der Autographen der Staatsbibliothek zu Berlin und den Personennamen der SWD

Die Bibliotheken nutzen die PND als integrierte Normdatei bei der Katalogisierung und verknüpfen die Normdatensätze mit den Titelsätzen.

(Quelle: http://www.dnb.de/DE/Standardisierung/Normdaten/PND/pnd_node.html)



Gemeinsame Körperschaftsdatei (GKD)

Die GKD ist die im deutschen Sprachraum relevante Normdatei für deutsche und ausländische Körperschaftsnamen. Sie ist in den 70er-Jahren aus den Körperschaftsaufnahmen der Bayerischen Staatsbibliothek München, der Staatsbibliothek Preußischer Kulturbesitz Berlin und der Deutschen Bibliothek Frankfurt am Main aufgebaut worden. Heute wird sie von der Bayerischen Staatsbibliothek, der Staatsbibliothek zu Berlin, der Deutschen Nationalbibliothek und von der Österreichischen Nationalbibliothek aktualisiert und redaktionell betreut, sowie laufend um die Körperschaftsaufnahmen aus der Zeitschriften- und Monografiensammlung der deutschen und österreichischen Bibliotheksverbände erweitert.

Die GKD-Datensätze enthalten neben Ansetzungs- und Verweisungsformen von Körperschaftsnamen, die nach den "Regeln für die alphabetische Katalogisierung in wissenschaftlichen Bibliotheken" (RAK-WB) erstellt worden sind, auch chronologische und hierarchische Relationen sowie weitere Angaben, die die repräsentierten Körperschaften charakterisieren. Zurzeit enthält die GKD mehr als 1 Mio. Datensätze mit wachsender Tendenz. Sie ist eine überregionale Normdatei für deutsche und ausländische Körperschaftsnamen.

Die GKD ist ein wesentliches Instrument zur Erschließung von gedruckten und digitalen Ressourcen und findet breite Anwendung im bibliothekarischen Bereich sowie beim Aufbau von Wissens- und Informationssystemen im außerbibliothekarischen Bereich. Die Deutsche Nationalbibliothek nutzt die GKD bei der Erstellung der Deutschen Nationalbibliografie (DNB). Die in der DNB-Datenbank enthaltenen Titelsätze sind mit Normdatensätzen, darunter GKD-Sätzen, verknüpft.

In der GKD sind folgende Typen von Körperschaften enthalten: religiöse Körperschaften, musikalische Körperschaften, Kongresse, Firmen, Gebietskörperschaften und Organe von Gebietskörperschaften.

Schlagwortnormdatei (SWD)



Die SWD bietet einen normierten, terminologisch kontrollierten Wortschatz an. Die dafür benutzten Quellen sind in der "Liste der fachlichen Nachschlagewerke zu den Normdateien" verzeichnet. Sie enthält Ansetzungs- und Verweisungsformen von Schlagwörtern, die nach den "Regeln für den Schlagwortkatalog (RSWK)" und den "Praxisregeln zu den RSWK und der SWD" festgelegt werden. Die Schlagwörter umfassen alle Fachgebiete und Schlagwortkategorien. Sie werden mittels ISO-Ländercode, ISO-Sprachencode und den Notationen der SWD-Sachgruppen klassifiziert.

Die SWD ist eine überregionale Normdatei für Schlagwörter. Sie hat folgende Inhalte:

- alle Fachgebiete
- Allgemein- und Individualbegriffe (Produkte, Werktitel, Körperschaften, Kongresse, Bauwerke)
- Relationen (Oberbegriffe, verwandte Begriffe, Äquivalenzbegriffe)
- Einordnung in SWD-Systematik und DDC-Notationen

Die SWD wird kooperativ von den deutschen und österreichischen Bibliotheksverbänden, der Schweizer Nationalbibliothek und der Deutschen Nationalbibliothek erstellt, gepflegt und weiterentwickelt. Bibliotheken, die die RSWK anwenden, aber keinem Bibliotheksverbund angehören, können sich mit Änderungs- und Ergänzungswünschen direkt an die SWD-Redaktion sowie die Clearingstelle für Öffentliche Bibliotheken bzw. die Clearingstelle für theologische Bibliotheken an der Deutschen Nationalbibliothek in Frankfurt am Main wenden.

Gemeinsame Normdatei (GND)



Im Projekt "Gemeinsame Normdatei (GND)" sollen die bestehenden Normdateien PND, SWD und GKD sowie die Einheitssachtitel-Datei des Deutschen Musikarchivs (EST-DMA) in einer gemeinsamen Normdatei (GND) zusammengeführt werden. Bestehende Formatunterschiede, die parallele Haltung von Datensätzen für Körperschaften und Geografika in den beiden Normdateien GKD und SWD sowie unterschiedliche Ansetzungsregeln für Formal- und Sacherschließung sollen für die GND der Vergangenheit angehören. Ziel ist der Aufbau einer gemeinsamen Normdatei als gemeinsames, eindeutiges Bezugssystem für die bibliografischen Daten der Bibliotheken sowie für die Erschließungsdaten anderer Normdatenanwender wie Archive, Museen, Projekte und Wissenschafts- und Kultureinrichtungen.

Die Gemeinsame Körperschaftsdatei (GKD), die Personennamendatei (PND) und die Schlagwortnormdatei (SWD) sowie die Einheitssachtitel-Datei des Deutschen Musikarchivs (DMA-EST-Datei), die in Zukunft ebenfalls als überregionale Normdatei geführt werden soll, sind "historisch gewachsen", und sind, in voneinander divergierenden Daten-modellen und in jeweils unterschiedlichen Formaten weiter-entwickelt worden, die den heutigen Kommunikations-anforderungen nicht mehr entsprechen.

Im GND-Projekt sollen sie nun zusammengeführt und zu einer modernen, web-fähigen Normdatei umgebaut werden, geeignet, die vielfältigen Ressourcen und Informationsangebote der Bibliotheken und anderer Kultureinrichtungen im deutschsprachigen Raum zu vernetzen und zugreifbar zu machen. Die "funktionale" Differenzierung in Normdaten für die Formalerschließung und Normdaten für die Sacherschließung wird dabei zu Gunsten einer objektorientierten Sicht aufgegeben. Die Gemeinsame Normdatei (GND) soll in MARC 21 Authority ausgetauscht werden und soweit absehbar die künftigen Regelungen der RDA (Resource Description and Access) berücksichtigen.

Die GND soll eine gemeinsame Normdatei sein, um gemeinsam und eindeutig von bibliografischen Daten und der Daten anderer Normdatenanbieter (z.B. Archive und Museen) zu verweisen.



Auch soll das neue Format für das Semantic Web geeignet sein.

Wie die jetzigen Normdateien wird auch die Gemeinsame Normdatei von den Normdatenteilnehmern kooperativ geführt und an der Deutschen Nationalbibliothek gehalten werden. Am GND-Projekt sind dementsprechend neben der Deutschen Nationalbibliothek alle Bibliotheksverbände im deutschsprachigen Raum beteiligt.

(Quelle: http://www.dnb.de/DE/Standardisierung/Normdaten/GND/gnd_node.html)

Zeitschriftendatenbank (ZDB)



Die Zeitschriftendatenbank ist die umfassendste und weltweit führende Datenbank der in deutschen Bibliotheken vorhandenen periodischen Veröffentlichungen (z.B. Zeitschriften, Zeitungen und Schriftenreihen). Sie umfasst gegenwärtig etwa 1,5 Mio. Titel sowie 10,3 Mio. dazugehörige Bestandsnachweise von rund 4.300 Bibliotheken. Außerdem weist die ZDB etwa 1,3 Mio. Körperschaften der Gemeinsamen Körperschaftsdatei (GKD) nach.

Verzeichnet werden periodische Veröffentlichungen aus allen Ländern, in allen Sprachen. Es werden sowohl gedruckte, als auch elektronische Formen verzeichnet. Eine zeitliche Einschränkung findet nicht statt. Die Verantwortung für den Betrieb und die Weiterentwicklung der ZDB tragen die Staatsbibliothek zu Berlin, Preußischer Kulturbesitz und die Deutsche Nationalbibliothek als gleichberechtigte Partner gemeinsam. Insgesamt arbeiten ca. 4.300 Bibliotheken an der Pflege der ZDB mit.

Homepage: <http://www.zeitschriftendatenbank.de>



Sie sind hier: Startseite

- **Startseite**
- Aktuelles
- Suche
- Erschließung
- Services
- Über uns
- Kontakt

Schnelleinstieg:

ZETA

Schriftgröße:

Barrierefreiheit?

Willkommen bei der Zeitschriftendatenbank (ZDB)



Die ZDB weist den Weg zu Zeitschriften, Zeitungen und Datenbanken in deutschen und österreichischen Bibliotheken. Sie wird in Kooperation von 4.300 Bibliotheken erstellt.

Suche im ZDB-Katalog

Was ist der ZDB-Katalog ▶?

Teilnahme

- ZDB für Katalogisierende ▶
- ZDB für Neueinsteigende ▶

Aktuelles

- Katalogisierung mit dem WebCat startet ▶ – Katalogisierung in der ZDB mit dem webbasierten Katalogisierungsclient WebCat
- AGDBT ▶ – 39. Sitzung der AGDBT am 8./9. November 2011
- ZDB-Website ▶ – Suchtipps für Bibliothekare ist überarbeitet worden.

▲ Seitenanfang

Letzte Änderung: 09.12.2011

Ein Angebot der



STAATSBIBLIOTHEK ZU BERLIN
PREUSSISCHER KULTURBESITZ

(Quelle: <http://www.dnb.de/wir/kooperation/zdb.htm#header>)

Internationale fachübergreifende Normdateien



Zu den internationalen fachübergreifenden Normdateien zählen:

- Library of Congress Authorities
- Virtual International Authority File
- CrissCross

Beispiel: Library of Congress Name Authority File

Die LC Name Authority File ist die amerikanische Normdatei für Personen- und Körperschaftsnamen. Sie liegen im Format MARC 21 for Authority Data vor und sind nach dem Regelwerk AACR2 erstellt. Die Library of Congress Subject Headings werden für die inhaltliche Erschließung verwendet. Sie werden beispielsweise in der Library of Congress vergeben. Die LCSH beinhalten Sachschlagwörter und Geografika.

Insgesamt besteht die Datei aus 265.000 Schlagwörtern, 5,3 Mio. Personen-, Körperschafts-, Kongress- und geografischen Namen, 350.000 Serien- und Einheitssachtiteln und 340.000 Namen/ Titel Normdaten.

Homepage: <http://authorities.loc.gov/>

Weitere kontrollierte Vokabulare und Cross-Konkordanzen



Klassifikationen

Definition: Klassifizieren

"Klassifizieren bedeutet, gleiche Gegenstände in Gruppen zusammenzufassen. Alle Mitglieder einer durch Klassifizieren entstandenen Gruppe - oder Klasse - haben mindestens ein gemeinsames Merkmal, das die Mitglieder anderer Klassen nicht besitzen."

(Quelle: Brian Buchanan: Bibliothekarische Klassifikationstheorie. München u.a.: Saur, 1989. S. 9)

Klassifikationen dienen dazu, Themen oder Objekte systematisch zu ordnen. Dabei werden meistens hierarchische Systeme verwendet, deren Ebenen die Themen oder Objekte unterschiedlich detailliert unterscheiden. Strenge hierarchische Systeme lassen sich als Bäume darstellen: In der Wurzel werden alle Themen oder Objekte zusammengefasst; die Blätter sind einzelne Themen, die nicht mehr weiter unterschieden werden. Die Bezeichnungen der detailliertesten Klassen können als Pfad von der Wurzel bis zu einem Blatt dargestellt werden.

Klassifikationen eignen sich insbesondere dazu, physische Objekte systematisch anzuordnen. So kann ein Buch in einer Bibliothek nur an einem Platz stehen, es sollte dort aber thematisch auffindbar sein. Dazu werden die Bücher zunächst in Abteilungen wie Naturwissenschaften oder Geisteswissenschaften eingeteilt. Innerhalb dieser Abteilungen gibt es dann weitere Unterabteilungen wie Physik, Biologie oder Philosophie und Sprachwissenschaften. Diese werden dann weiter in Fachgebiete aufgegliedert, bis schließlich in den einzelnen Regalen Bücher zu sehr spezifischen Themen zusammengefasst sind.

(Quelle: http://information-retrieval.de/irb/ir.part_1.chapter_3.section_3.html)

Dewey Dezimalklassifikation (DDC)



Die Dewey Dezimalklassifikation (DDC) ist die international am weitesten verbreitete Universalklassifikation. Sie wurde 1873 von Melvil Dewey entwickelt und 1876 erstmals veröffentlicht. Einsatz findet sie in Bibliotheken in mehr als 135 Ländern. Übersetzt wurde sie in über 30 Sprachen

(<http://www.oclc.org/dewey/>).

Geordnet ist die Klassifikation nach Fachgebieten. Es gibt 10 Hauptklassen. Jede Hauptklasse ist in 10 Klassen der Ebene 2 (divisions) aufgeteilt und jede Klasse der Ebene 2 in 10 Klassen der Ebene 3 (subdivisions).

DDC-Übersicht	
000	<u>Informatik, Informationswissenschaft, allgemeine Werke</u>
100	<u>Philosophie & Psychologie</u>
200	<u>Religion</u>
300	<u>Sozialwissenschaften</u>
400	<u>Sprache</u>
500	<u>Naturwissenschaften</u>
600	<u>Technik, Medizin, angewandte Wissenschaften</u>
700	<u>Künste und Unterhaltung</u>
800	<u>Literatur</u>
900	<u>Geschichte und Geografie</u>

(Quelle: <http://melvil.dnb.de/melvilsearch?bs=dnb-portal>)

Pflege und Lizenz der Klassifikation liegen bei OCLC. Die DDC Summaries sind als Linked Open Data unter <http://dewey.info/> verfügbar.

Seit Oktober 2005 gibt es eine deutsche Übersetzung der DDC. Es handelt sich um die 22. Ausgabe der DDC (<http://ddc-deutsch.de>).

Diese Übersetzung findet beispielsweise in der Deutschen Nationalbibliothek Anwendung. Hier werden die Reihen A, B und H der Deutschen Nationalbibliografie mit DDC-Notationen erschlossen.

Thesauri



In gewisser Weise bilden Thesauri, Singular Thesaurus, das sprachliche oder terminologische Gegenstück zu hierarchischen Klassifikationssystemen. Während Klassifikationen Themen oder Objekte inhaltlich ordnen, erfassen Thesauri Wörter, Terme und Ausdrücke eines Sachgebiets und beschreiben die Beziehungen zwischen ihnen. Thesauri haben vor allem zwei Funktionen: Sie definieren ein kontrolliertes Vokabular.

Sie stellen Beziehungen (Relationen) zwischen den Termen dieses Vokabulars her.

Dabei werden nicht nur hierarchische Systeme definiert wie bei der Klassifikation, sondern es können eine Vielzahl von weiteren Beziehungen zwischen Wörtern dargestellt werden. Nach DIN 1463 ist ein Thesaurus eine geordnete Zusammenstellung von Begriffen mit ihren (natürlich-sprachlichen) Beziehungen.

Man sollte allerdings zwischen allgemeinen Thesauri und den in einem IR-System verwendeten Thesauri unterscheiden. Ein allgemeiner Thesaurus listet zu jedem Wort, neben einer kurzen Definition des Begriffs, verschiedene Relationen zu Gruppen von anderen Wörtern auf:

Synonyme (Wörter mit der gleichen Bedeutung)

Antonyme (Wörter mit gegensätzlicher Bedeutung)

verwandte Wörter

Oberbegriffe

speziellere Begriffe

(Quelle: http://information-retrieval.de/irb/ir.part_1.chapter_3.section_4.html)

Standard-Thesaurus-Wirtschaft (STW)



Der Standard-Thesaurus-Wirtschaft (STW) ein Wirtschafts-thesaurus. Er enthält Vokabular zu allen ökonomischen Themenstellungen und verwandten Themen, wie beispielsweise die Sozialwissenschaften, Recht oder Politik. Von der Terminologie her sind Fach-, Allgemein-, und geografische Begriffe zu finden.

Der STW ist Mitte der 90er Jahre entstanden. Er wird durch die Deutsche Zentralbibliothek der Wirtschaftswissenschaften (ZBW) gepflegt und kontinuierlich weiterentwickelt. Enthalten sind ca. 600 Deskriptoren und ca. 20.000 Nicht-Deskriptoren in deutsch und englisch.

Der Thesaurus steht unter der Creative Commons-Lizenz. Eine SKOS-Repräsentation ist unter <http://zbw.eu/stw> zu finden.

(vgl.: <http://www.zbw.eu/stw/versions/8.08/about.de.html>)



RDF/XML RDF/Turtle English

Liste der Deskriptoren
Download
Andere Versionen
Wiki
STW Web Services



- ▶ A Allgemeinwörter
- ▶ B Betriebswirtschaft
- ▶ G Geographische Begriffe
- ▶ N Nachbarwissenschaften
- ▶ P Produkte
- ▶ V Volkswirtschaft
- ▶ W Wirtschaftssektoren und spezielle Wirtschaftslehren

Standard-Thesaurus Wirtschaft

Version 8.08 (Änderungen)

Subthesauri

- A Allgemeinwörter
- B Betriebswirtschaft
- G Geographische Begriffe
- N Nachbarwissenschaften
- P Produkte
- V Volkswirtschaft
- W Wirtschaftssektoren und spezielle Wirtschaftslehren

Der Standard-Thesaurus Wirtschaft soll mit seinen genormten Schlagwörtern die thematische Suche im Wirtschaftsportal EconBiz und im ZBW-Katalog ECONIS erleichtern (einfach per  bzw.  Icon).

Der Thesaurus enthält Vokabular zu allen ökonomischen Themenstellungen: mehr als 6.000 Schlagwörter und zusätzlich 10.000 Begriffe als Synonyme, um individuelle

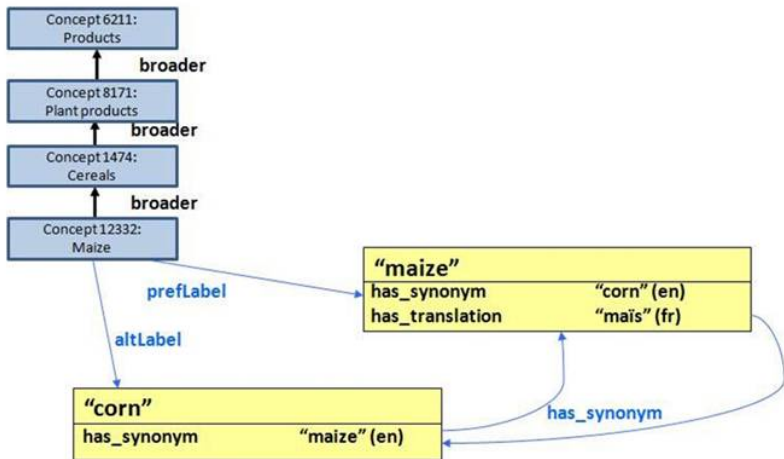
AGROVOC (Agricultural Information Management Standard)



Der AGROVOC Thesaurus enthält mehr als 40 000 Begriffe in bis zu 21 Sprachen zu Themen, die im Zusammenhang mit Lebensmitteln, Ernährung, Landwirtschaft, Fischerei und Forstwirtschaft, Umwelt und anderen verwandten Bereichen stehen.

Eine weltweite Gemeinschaft von Editoren (Bibliothekare, Terminologen, Information Manager und Software-Entwickler) halten AGROVOC mit Hilfe eines web-basierten Open-Source Editors aktuell. AGROVOC wurde als Linked Data veröffentlicht.

Weitere Informationen: <http://aims.fao.org/website/AGROVOC-Thesaurus/sub>



(Quelle: <http://aims.fao.org/standards/agrovoc/concept-scheme>)

TEMA (Gebiete der Technik, Wirtschaft)



Der Thesaurus Technik und Management umfaßt die Fachgebiete Betriebsführung und Betriebsorganisation, Medizinische Technik, Maschinenbau und Fertigungstechnik, industrielle Informationstechnik, Textil, Werkstoffe, Elektrotechnik und Elektronik. Er ist ein strukturiertes Schlagwortverzeichnis und ein wichtiges Werkzeug zur Verbesserung der Suche.

Der Thesaurus wird zur Indexierung (Verschlagwortung) der Datenbanken TEMA® (Technik und Management), BEFO (Betriebsführung und -organisation), BERG (Bergbau), DOMA® (Maschinen- und Anlagenbau), ENTEC (Energietechnik), ITEC (Industrielle Informationstechnik), MEDITEC (Medizinische Technik), TOGA® (Textil), WEMA® (Werkstoffe) und ZDE (Elektrotechnik und Elektronik) eingesetzt.

(Quelle: <http://www.hsb.hs-anhalt.de/ivs/DOWNLOAD/FIZT-WEB-Intranet-Thesaurus.pdf>)

Beispiel für einen Deskriptoreintrag:

	Hauptbenennung		Drehmaschine
W	Schreibvariante	Dieser Begriff ist eine Schreibvariante und dient nur als Einsprungsadresse. Er verweist auf den eigentlichen Begriff (USE).	
B	Deutsches Synonym		Drehbank
E	Englisches Synonym	Bei englischen Synonymen wird auf die deutsche Vorzugsbenennung (USE) verwiesen.	lathe machines lathes turning lathes
O	Oberbegriff der nächsten Ebene	Allgemeiner Begriff	Werkzeugmaschine
U	Unterbegriff	Spezifischerer Begriff	Drehautomat Feindrehmaschine Hinterdrehmaschine Karusselldrehmaschine ...
V	Verwandter Begriff	Verweis auf andere Begriffe, die nicht in der Hierarchie enthalten sind, aber eventuell auch zutreffend sind.	Drehbohrmaschine Drehen (Bearbeiten) Drehmaschinenbett ...

(Quelle: <http://www.hsb.hs-anhalt.de/ivs/DOWNLOAD/FIZT-WEB-Intranet-Thesaurus.pdf>)

Getty Thesaurus of Geographic Names (TGN)



Der Getty Thesaurus of Geographic Names (TGN) enthält Geografika (politische und physische Entitäten) und hat außerdem einen Fokus auf Orte, die im Bereich Kunst und Architektur von Bedeutung sind.

Er wurde vom J. Paul Getty Trust entwickelt und wird durch Getty gepflegt und weiterentwickelt.

Der TGN umfasst ca. 895.000 Deskriptoren (multilingual) und insgesamt über 1 Million Terme. Dies sind Namen, Koordinaten, Datumsangaben, Ortstypen, Erläuterungen und Quellenangaben.

Homepage: <http://www.getty.edu/>

The screenshot shows the homepage of the Getty Research Institute. The main navigation bar includes links for Home, Museum, Research Institute, Conservation Institute, Foundation, Trust, Connect with Us, Visit, and Shop. A search bar is located in the top right corner. Below the navigation bar, there are several tabs: Exhibitions & Events, Special Collections, Library, Search Tools & Databases (selected), Scholars & Projects, Publications, and About the GRI. The main content area is titled "Search Tools & Databases" and features a section for "Getty Thesaurus of Geographic Names® Online". This section includes a search form with fields for "Find Name or ID:", "Place Type:", "Lookup:", and "Nation:", each with a corresponding "Lookup" button. There are also "Pop-up Search" and "Browse the TGN hierarchies" links. A "Search Tips" section provides instructions on using AND and OR operators. On the left side, there is a sidebar menu with various search tools and databases. On the right side, there is a "Have a Question?" section with a link to "Contact the Vocabulary Program".

(Quelle: <http://www.getty.edu/research/tools/vocabularies/tgn/index.html>)

Crosskonkordanzen



Crosskonkordanzen sind intellektuell erstellte Verbindungen zwischen zwei Thesauri oder Klassifikationen. Sie setzen die Terme eines Ausgangsvokabulars mit denen des Zielvokabulars durch Relationen in Verbindung und machen damit Aussagen über die Ähnlichkeit semantischer Konzepte beider Vokabulare. Eingesetzt werden sie als Transfermodule bei der integrierten Suche über Datenbanken, die mit unterschiedlichen Vokabularen inhaltlich erschlossen wurden. Ohne Transformation der Terme zwischen den Vokabularen würden lediglich Ergebnisse geliefert, bei denen zur Indexierung eines semantischen Konzepts der gleiche Begriff in exakt der gleichen Ansetzung verwendet wurde. Variationen in der Ansetzung (z. B. Singular vs. Plural) oder die Verwendung unterschiedlicher Begriffe würden zum Nicht-Auffinden von potentiell relevanten Ergebnissen in einzelnen Datenbanken führen.

(Quelle: http://www.ib.hu-berlin.de/~mayr/arbeiten/mayr_etal_dgi05.pdf)

Typischerweise werden die Vokabulare bilateral verbunden, d. h. eine Konkordanz verbindet Terme eines Vokabulars A zu einem Vokabular B und eine weitere Konkordanz verbindet Terme von Vokabular B zurück zu A. Bilaterale Relationen sind dabei nicht notwendigerweise symmetrisch. Beispielsweise wird der Term ‚Computer‘ aus Vokabular A auf den Term ‚Information System‘ in Vokabular B abgebildet, aber der gleiche Term ‚Information System‘ in Vokabular B kann mit einem anderen Term, z. B. ‚Data base‘, in Vokabular A relationiert werden.

(Quelle: <http://www.ib.hu-berlin.de/~mayr/arbeiten/Oberhof2010.pdf>)



VIAF dient zum Abgleich und zur Verlinkung von Normdateien mit maschinellen Verfahren. Aktuell geschieht dies mit 19 Normdateien.

Die Normierungsergebnisse können zur Verbesserung der weltweiten Recherche genutzt werden. Sie ermöglichen die Suche von Namensformen in der eigenen Sprache und in der vom Benutzer bevorzugten Schrift. Nationale bzw. regionale Ansetzungsformen können nebeneinander beibehalten werden.

Die VIAF findet Anwendung in 13 Nationalbibliotheken. Ihr Fokus liegt auf Individualnamen, im ersten Schritt auf Personennamen.

Homepage: <http://www.viaf.org/>

Beispiel:

Webster, Max, 1864-1920

Webster, Max, sociologue

Webster, Max

1864-1920 **מקס, מאקס**

1864-1920 **מקס, מאקס**

VIAF ID: 100180950 (Personal)

Permalink: <http://viaf.org/viaf/100180950>

Vorzugsbezeichnungen

- 100 1 [_1a_Weber_Max_d_1864-1920](#)
- 100 1 [_1a_Weber_Max_d_1864-1920](#)
- 100 1 [_1a_Weber_Max_d_1864-1920](#)
- 100 1 0 [_1a_Weber_Max](#)
- 100 1 [_1a_Weber_Max_d_1864-1920](#)
- 100 1 [_1a_Weber_Max_d_1864-1920](#)
- 100 1 [_1a_מקס, מאקס_1864-1920](#)
- 100 1 [_1a_מקס, מאקס_d_1864-1920](#)
- 200 [_1_1a_Weber_zb_Max_zf_1864-1920](#)
- 200 [_1_1a_Weber_zb_Max_zf_1864-1920](#)
- 100 1 [_1a_מקס, מאקס_d_1864-1920](#)

Deutsche Nationalbibliothek
DNB|118629743

(Quelle: <http://viaf.org/viaf/100180950/>)

Folgende Institutionen sind an der VIAF beteiligt:

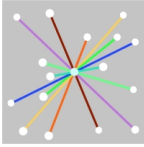


- National Library of Australi
- Library and Archives Canada -Test
- National Library of the Czech Republic
- Bibliotheca Alexandrina (Ägypten)
- Bibliothèque nationale de France
- SUDOC (France)
- Deutsche Nationalbibliothek
- Getty Research Institute
- National Széchényi Library (Hungary)
- National Library of Israel
- Istituto Centrale per il Catalogo Unico (Italien)
- NUKAT Center (Polen) -Test
- Biblioteca Nacional de Portugal
- Biblioteca Nacional de España
- National Library of Sweden
- Swiss National Library RERO (Switzerland)
- Library of Congress/NACO
- Vatican Library

Für Tests werden außerdem in VIAF geladen:

- Flemish Public Libraries-test
- Russian State Library-test

CrissCross



(Quelle: <http://linux2.fbi.fh-koeln.de/crisscross/index.html>)

Ziel des DFG-geförderten Projektes CrissCross war die Erstellung eines multilingualen, thesaurusbasierten und benutzer-gerechten Recherchevokabulars zu heterogen erschlossenen Dokumenten.

Hierfür wurden Mappings zwischen Sachschlagwörtern der Schlagwortnormdatei (SWD) und Notationen der Dewey-Dezimalklassifikation (DDC) erstellt. Die Verknüpfungsrichtung ist von der SWD ausgehend zur DDC. Außerdem gibt es Mappings zwischen SWD-Sachschlagwörtern mit ihren Äquivalenten in den Library of Congress Subject Headings (LCSH) und der französischen Schlagwortnormdatei RAMEAU. Diese Mappings stellten eine Fortführung der Arbeiten dar, die während des MACS-Projektes gemacht wurden.

Projektpartner bei CrissCross waren die FH Köln und die Deutsche Nationalbibliothek. Das Projekt begann 2006 und wurde 2010 mit einem Abschlussworkshop beendet

(<http://linux2.fbi.fh-koeln.de/cisko2010/index.html>).

Projekthomepage: <http://www.fbi.fh-koeln.de/institut/projekte/CrissCross/index.html>

Mapping zwischen STW und SWD



Es gibt ein Mapping vom Standard-Thesaurus Wirtschaft (STW) auf die Schlagwortnormdatei (SWD). Dieses Mapping wurde aufgebaut von den Expertinnen der ZBW (Leibniz-Informationszentrum Wirtschaft) und der Deutschen Nationalbibliothek. Herausgeber des Mappings ist die ZBW.

Das Mapping steht unter der CC0-Lizenz (<http://creativecommons.org/publicdomain/zero/1.0/>). Die CC0-Lizenzierung des Mappings soll eine möglichst breite und einfache Wiederverwendung ohne rechtliche Einschränkungen fördern.

Zwischen STW und SWD wurden folgende Relationen vergeben:

4.608 skos:exactMatch

5.528 skos:narrowMatch

63 skos:broadMatch

(Quelle: <http://zbw.eu/stw/versions/latest/mapping/gnd/about>)

Hier einige Beispiele:

"<" = skos:broadMatch

">" = skos:narrowMatch

"=" = skos:exactMatch

"0" = kein Mapping möglich

STW	Relation	SWD
Abwasserwirtschaft	<	Entsorgungswirtschaft
Aktienindex	=	Aktienindex
Aktienindex	>	Deutscher Aktienindex
Aktienindex	>	Dow-Jones-Index
Autovermietung	=	Mietwagenunternehmen
Bilanztheorie	=	Bilanztheorie
Bilanztheorie	>	Dynamische Bilanztheorie
Bilanztheorie	>	Statische Bilanztheorie
Dauerhafte Entwicklung	=	Nachhaltige Entwicklung
Nahrungsmittelpreispolitik	<	Staatliche Preispolitik
	0	Stärken-Schwächen-Analyse
Verschuldungsrestriktion	0	
Wettbewerb	>	Wettbewerbsverhalten
Wettbewerb	>	Wettbewerbsvorsprung
Zentralbank	=	Notenbank

(Quelle: http://www.dnb.de/wir/projekte/cross_beispiele.htm)

Übungen: Klassifikation, Thesaurus und Crosskonkordanz



Die Lösungen finden Sie auf der nächsten Seite

Worum handelt es sich bei VIAF?

- A. Klassifikation
- B. Thesaurus
- C. Crosskonkordanz

Worum handelt es sich bei STW?

- A. Klassifikation
- B. Thesaurus
- C. Crosskonkordanz

Worum handelt es sich bei DDC?

- A. Klassifikation
- B. Thesaurus
- C. Crosskonkordanz

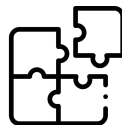
Worum handelt es sich bei CrissCross?

- A. Klassifikation
- B. Thesaurus
- C. Crosskonkordanz

Worum handelt es sich bei THEMA?

- A. Klassifikation
- B. Thesaurus
- C. Crosskonkordanz

Lösungen: Klassifikation, Thesaurus und Crosskonkordanz



Worum handelt es sich bei VIAF?

Crosskonkordanz

Worum handelt es sich bei STW?

Thesaurus

Worum handelt es sich bei DDC?

Klassifikation

Worum handelt es sich bei CrissCross?

Crosskonkordanz

Worum handelt es sich bei THEMA?

Thesaurus

Regelwerke und Modelle



ISBD - International Standard Bibliographic Description

Die International Standard Bibliographic Description (Internationaler Standard zur Beschreibung bibliographischer Daten, kurz ISBD) ist ein Regelwerk, das von der IFLA betreut wird. Mit ihm ist es möglich Bibliotheksmaterialien zu erfassen. Die letzte Version der ISBD wurden 2007 vorgestellt. Ursprüngliche Idee der ISBD war es, einen bibliographischen Standard für den internationalen Datenaustausch zu entwickeln.

Übersicht über die Elemente der ISBD(G) (General International Standard Bibliographic Description. - 1977 bzw. 1992)



Bereich der Sachtitel- und Verfasserangabe

- Hauptsachtitel
- Allgemeine Materialbenennung
- Parallelsachtitel
- Zusätze zum Sachtitel
- Verfasserangabe

Bereich des Ausgabevermerks

- Ausgabebezeichnung
- Parallele Ausgabebezeichnung
- Verfasserangabe zur Ausgabe ...

Bereich der materialspezifischen (oder die Veröffentlichungsart betreffenden) Angaben

Bereich des Erscheinungsvermerks

- Erscheinungsort, Vertriebsort usw. ...
- Name des Verlegers, Vertreibers usw. ...
- Erscheinungsjahr, Vertriebsjahr usw. ...

Bereich der physischen Beschreibung

- Spezifische Materialbezeichnung und Umfang der Vorlage
- Andere physische Details
- Ausmaße des Objekts
- Angabe des Begleitmaterials

Bereich der Gesamttitelangabe

- Hauptsachtitel der Reihe
- Paralleltitel der Reihe
- Zählung innerhalb der Reihe
- Angaben zur Unterreihe

Bereich der Fußnoten

Bereich Bereich der Standardnummer und ähnlicher Nummern, Angaben zur Beschaffung und Preis

(Quelle: <http://www.ib.hu-berlin.de/~pannier/isbd.html>)

Beispiel für ein gedrucktes Buch:



Area 1	Klassische Dramen [Gedruckter Text] : [Text und Kommentar] / Friedrich Schiller ; herausgegeben von Matthias Luserke-Jaqui
Area 2	
Area 3	
Area 4	. — Frankfurt, M. : Dt. Klassiker-Verl., 2008
Area 5	. — 888 S. ; 18 cm
Area 6	. — (Deutscher Klassiker-Verlag im Taschenbuch ; Bd. 26)
Area 7	. — Literaturverz. S. 879-885. — Frühere Aufl. als: Werke und Briefe : in zwölf Bänden / Friedrich Schiller ; Bd. 5. — Enthält u.a.: Maria Stuart. Die Jungfrau von Orleans
Area 8	. — ISBN 978-3-618-68026-0

(Quelle: http://www.ifla.org/files/cataloguing/isbd/isbd-examples_2009.pdf)

Weitere Informationen unter: <http://www.ifla.org/en/isbd-rg/superseded-isbd-s>

Regeln für die alphabetische Katalogisierung (RAK)



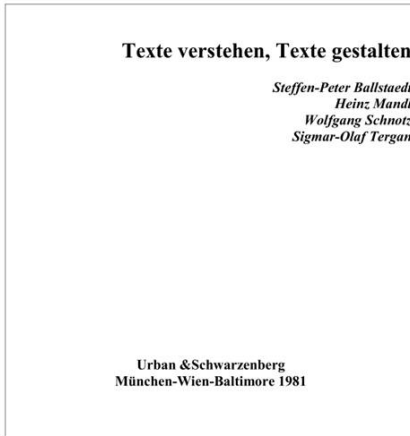
Die Regeln für die alphabetische Katalogisierung (RAK) sind ein Regelwerk für die formale Erschließung. Entstanden ist das erste Rahmenregelwerk für öffentliche Bibliotheken (ÖB) und wissenschaftliche Bibliotheken (WB) bereits 1977. 1983 erschienen dann RAK für weitere Bibliotheks- und Dokumenttypen.

Insgesamt gibt es folgende Teile: RAK-WB, RAK-ÖB, RAK-Musik, RAK-Karten, RAK-Körperschaften und RAK-NBM.

Die Basis, auf die die RAK aufbauen ist die International Standard Bibliographic Description (ISBD).

Hier ein Beispiel für eine Titelaufnahme:

<http://elib.tu-darmstadt.de/texte/RAK-Schulung.pdf>



Rückseite der Haupttitelseite: ISBN 3-541-09901-1

im Buch:

- Literatur S. 289 - 316
- das Werk wurde gemeinschaftlich von diesen 4 Autoren verfasst
- 328 S.
- wenige grafische Darstellungen

Das Katalogisierungsergebnis:

Texte verstehen, Texte gestalten / Steffen-Peter Ballstaedt ...
- München [u.a.] : Urban & Schwarzenberg, 1981. - 328 S. : graf. Darst.
Literaturverz. S. 289 - 316
ISBN 3-541-09901-1

Anglo-American Cataloguing Rules (AACR)



Die AACR sind ein Regelwerk für die formale Erschließung. Sie werden in über 30 Ländern weltweit verwendet, beispielsweise in Großbritannien, den USA und Kanada. Die Library of Congress verwendet die AACR. Die erste Ausgabe der AACR gab es bereits 1967. Eine überarbeitete Fassung, die AACR2, gab es 1978. Die letzte Revision gab es 2002, die AACR2r. Updates gibt es in regelmäßigen Abständen. Gepflegt werden die AACR vom Joint Steering Committee (for Revision of AACR).

Homepage: <http://www.aacr2.org/>



Resource Description and Access (RDA) ist ein neuer internationaler Erschließungsstandard zur Erschließung aller Objekttypen. Bei diesem Standard stehen die Nutzerbedürfnisse im Vordergrund .

RDA ist ein community-übergreifender Standard für Bibliotheken, Museen und Archive. Er basiert auf FRBR und FRAD.

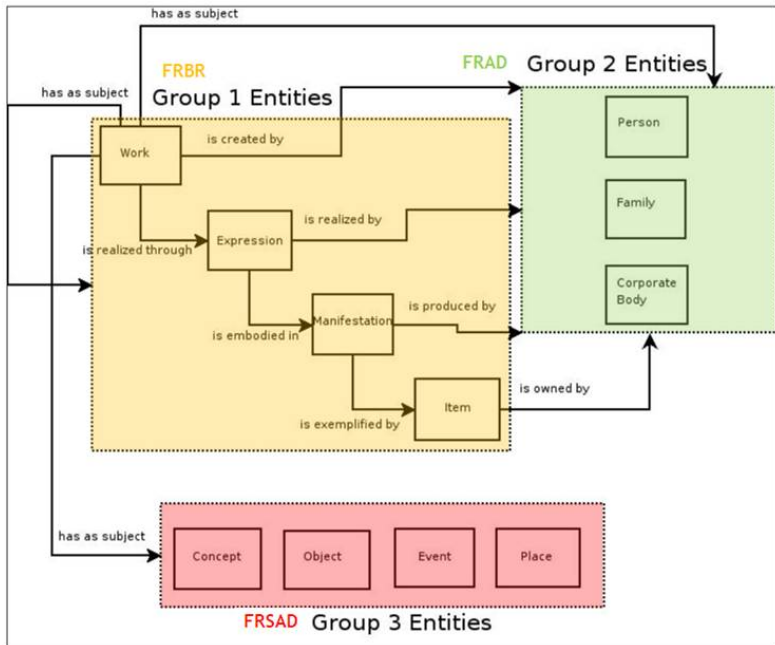
RDA ist unabhängig von Metadatenformaten. Der Fokus liegt auf Relationen, Rollen und kontrolliertem Vokabular. Eine erste Veröffentlichung des "RDA-Toolkit" gab es im Juni 2010. Es gibt Mappings von RDA zu Dublin Core, ISBD und MARC 21.

Weitere Informationen unter: <http://www.rdatoolkit.org/>.

FRBR, FRAD, FRSD



Die Functional Requirements for Bibliographic Records (FRBR) sind ein Modell zur Strukturierung bibliographischer Daten. Dabei unterscheidet FRBR drei Gruppen von Entitäten:



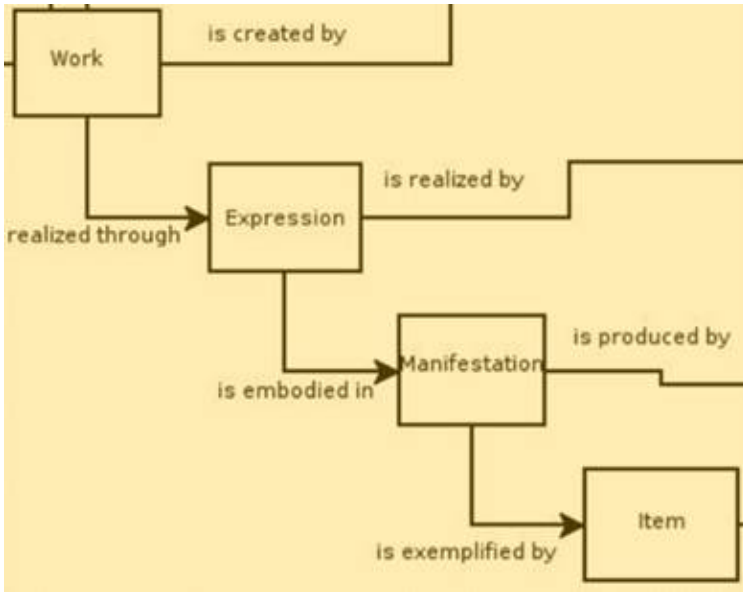
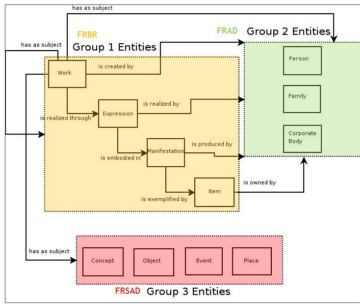
(Quelle: <http://www.fbr.org/files/entity-relationships.png>)

Drei Gruppen:

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

Gruppe 2: Personen und Körperschaften

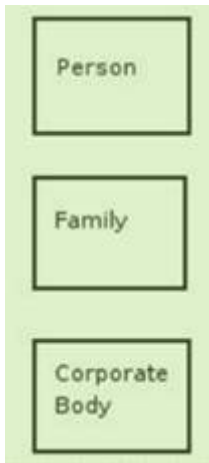
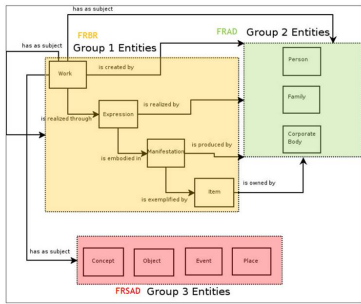
Gruppe 3: Begriff, Gegenstand, Ereignis, Ort



Gruppe 1:

WEMI (Work – Expression – Manifestation – Item) soll es ermöglichen, verschiedene Ausprägungen eines Werks zusammenhängend anzuzeigen, so dass der Nutzer einen schnellen und einfachen Überblick über relevante Objekte bekommt. Um dies zu erreichen, werden die WEMI-Entitäten über entsprechende Properties miteinander in Beziehung gesetzt.

FRAD



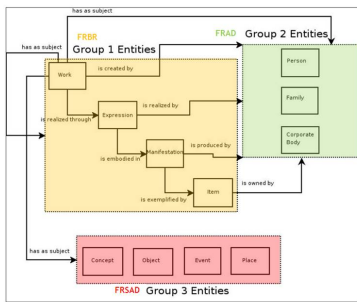
Gruppe 2:



Personen und Körperschaften

Diese Gruppe beschreibt, welche Relationen zwischen Personen und Körperschaften auf der einen Seite und den Entitäten der Gruppe 1 auf der anderen Seite bestehen. Die Beschreibung von Personen und Körperschaften im Rahmen von Normdateien und die Beziehungen zwischen diesen, werden in den Functional Requirements for Authority Data (FRAD) näher spezifiziert.

FRSAD



Gruppe 3:

Begriff, Gegenstand, Ereignis, Ort



Diese Gruppe besteht aus Entitäten, die vor allem als Thema von bibliographischen Ressourcen eine Rolle spielen. Das Modell, das der Beschreibung dieser Entitäten zugrundeliegt – die Functional Requirements for Subject Authority Data (FRSAD) - haben wir bereits im Abschnitt "Kontrollierte Vokabulare und Normdateien" kurz vorgestellt.

Weitere Informationen:

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

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

Linked Data



Linked Data ist eine Methode, um Daten im Web mittels referenzierbarer URIs miteinander zu verknüpfen.

Die Entwicklungen im Bereich Semantic Web zielen darauf ab, semantische Strukturen nachnutzbar und Daten besser zugreifbar zu machen. Des Weiteren erlauben die Konzepte des Semantic Web, Daten aus heterogenen Quellen miteinander zu verknüpfen, was eine Etablierung neuer Dienstleistungen mit sich führt. Im Zuge der Linked-Data-Bewegung bieten bereits eine Vielzahl von Anbietern (meist gemeinnützige Organisationen, Universitäten oder öffentliche Einrichtungen) ihre Daten in einer Semantic-Web-konformen Repräsentation an. Zu diesen Daten gehören insbesondere Daten, die von generellem Nutzen für die Allgemeinheit sind. Beispiele hierfür sind geografische Informationen, Thesauri, Lexika und auch bibliografische Daten und Normdaten.

(Quelle: http://files.dnb.de/pdf/linked_data.pdf)

Weitere Informationen:

<http://linkeddata.org/>

<http://dbpedia.org/About>

Metadaten - Vom Datensilo ins Semantic Web

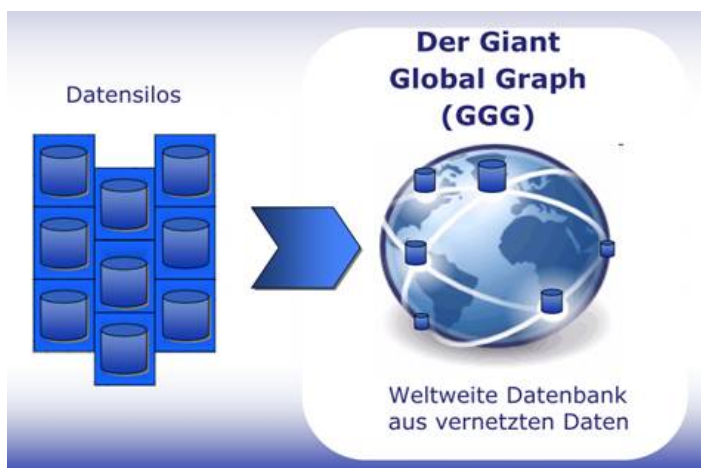


In den letzten Jahren hat sich die Nutzung und damit die Sicht auf und die Anforderungen an Metadaten entscheidend geändert.

Datensilos

Metadaten werden vor allem für lokale Anwendungen (z. B. den bibliothekseigenen Katalog) gebraucht.

Hauseigene Metadatenformate unterstützen das Finden von Objekten innerhalb dieser lokalen Anwendung und sind einerseits sehr granular, andererseits aber nur in geringem Maß interoperabel mit anderen Metadatenformaten.



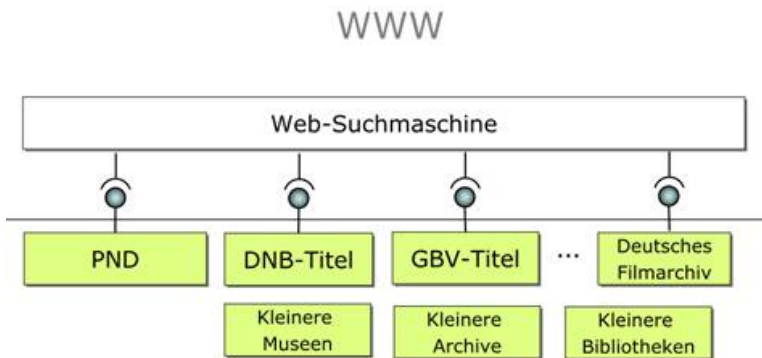
(Quelle: http://www.opus-bayern.de/bib-info/volltexte/2010/892/pdf/2010-03-15_Bibliothekskongress_kett.pdf)



Metasuchen und Suchmaschinen

Metadaten werden in anwendungsübergreifenden Suchen (Metasuche und Suchmaschinenindex) zur Verfügung gestellt.

Metadaten unterstützen den Austausch und das Zusammenführen von Daten in übergreifenden Suchportalen. Für die Nutzung der Metadaten werden Schnittstellen bereitgestellt, die die Metadaten in Standardformaten zur Verfügung stellen.



(Quelle: http://www.opus-bayern.de/bib-info/volltexte/2010/892/pdf/2010-03-15_Bibliothekskongress_kett.pdf)

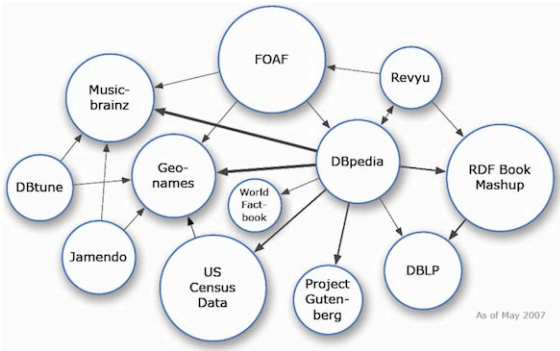
Linked Open Data



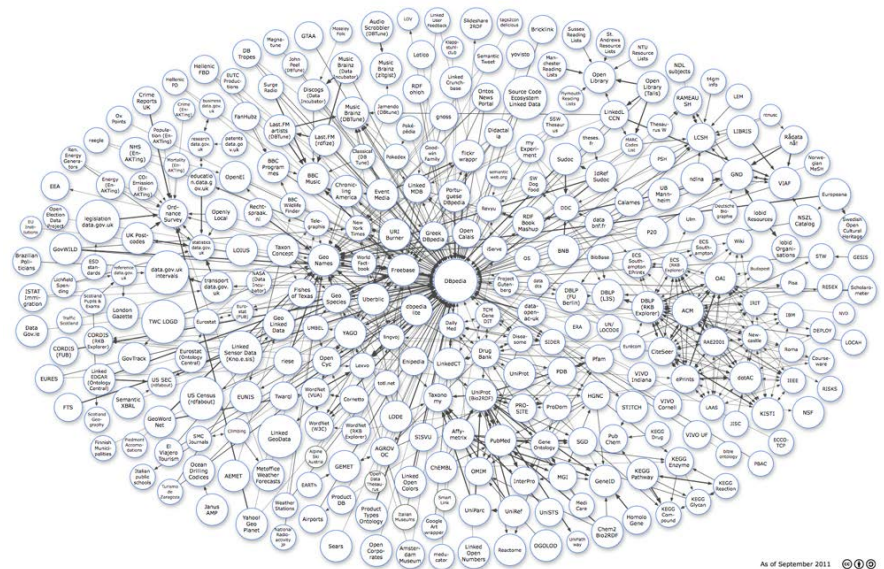
Metadaten werden frei im Internet zur Verfügung gestellt und mit anderen Metadaten verlinkt.

Um Daten verlinken zu können, müssen sie interoperabel und maschinenlesbar sein. Die Maschinenlesbarkeit wird durch die Nutzung des RDF-Modells realisiert, die Verlinkung mit Hilfe von Identifiern (HTTP URIs).

Die Linked Data Cloud hat im Mai 2007 sehr klein angefangen:



Die Linked Data Cloud einige Jahre später im September 2011:



As of September 2011

Semantic Web - Prinzipien und Schichten



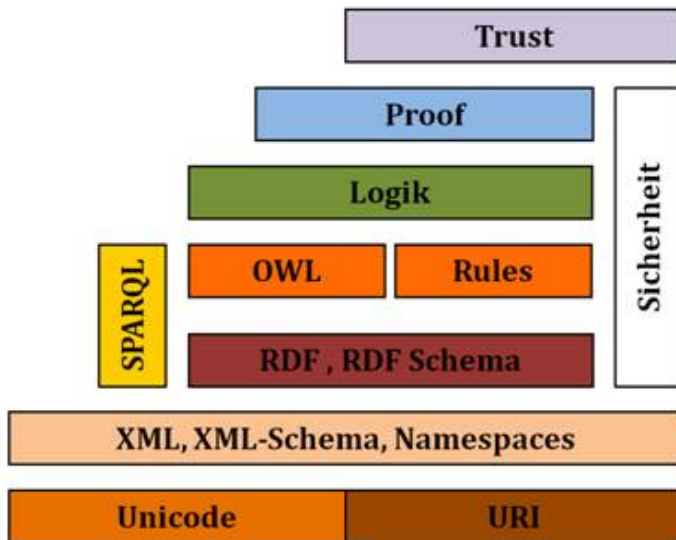
Es gibt vier Prinzipien, nach denen Daten für Linked Data aufgebaut sein sollten:

- Use URIs as names for things
- Use HTTP URIs so that people can look up those names.
- When someone looks up a URI, provide useful information, using the standards (RDF*, SPARQL)
- Include links to other URIs. so that they can discover more things.

(Quelle: <http://www.w3.org/DesignIssues/LinkedData.html>)

Diese Prinzipien spiegeln sich in den Schichten des Semantic Web wider. URIs bilden hier die Basis, auf denen alle anderen Schichten aufbauen.

Die Basis bilden Unicode, URI und XML bzw. XMLNS. Alle anderen Schichten bauen auf diesen Grundlagen auf. So beispielsweise RDF zur Beschreibung der Daten und OWL als Ontologie Vokabular.



(Quelle: <http://www.w3.org/2000/Talks/1206-xml2k-tbl/slide10-0.html> und <http://winfwiki.wi-fom.de/images/thumb/9/9f/Stack.png/400px-Stack.png>)

URI



"URIs are at the heart of the web"

(Quelle: <http://www.ariadne.ac.uk/issue65/thompson-hs>)

- Identifizieren Ressourcen / Objekte
- Ermöglichen den Zugriff auf Beschreibungen einer Ressource / eines Objektes (representation = Metadaten)
- Ermöglichen Zugriff auf die Ressource / das Objekt selbst (presentation)
- Im Semantic Web gilt: Alles ist eine Ressource = nicht nur Objekte werden mittels URI identifiziert, sondern alle Terme, die im Rahmen von Metadatenstandards und -profilen definiert und annotiert werden.

Eine URI steht eindeutig und dauerhaft für eine ganz bestimmte Ressource.

Dabei gilt:

"things are resources because someone created a URI to identify them"

(Quelle: <http://www.ariadne.ac.uk/issue65/thompson-hs>)

Eine Ressource kann in diesem Sinne alles sein: ein Buch, ein Gebäude, eine Idee aber auch ein Metadaterm.

Ein Metadaterm wird mittels einer URI identifiziert, die auf die Definition und Beschreibung des Terms verweist und so die Nutzung des Terms nachvollziehbar macht



RDF ist ein Modell, das die traditionelle Beschreibung von Ressourcen aufbricht, und eine Datenstruktur schafft, die es möglich macht, Ressourcen flexibel miteinander zu verlinken. Dabei entspricht die Beschreibung einer Ressource einem Triple aus Subjekt, Prädikat und Objekt

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

Voraussetzung dafür:

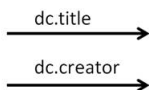
Die Ressource muss über einen URI eindeutig identifizierbar sein.

Ist dies der Fall, so kann ich mittels RDF in maschinenlesbarer Form die Aussage machen, dass eine bestimmte Ressource (Subjekt) für eine bestimmte Eigenschaft (Prädikat) einen bestimmten Wert (Objekt) hat. Bei dem Wert kann es sich um eine einfache Zeichenkette handeln, die keinerlei Link zu einem weiteren Objekt beinhaltet. Es kann sich aber auch um eine weitere Ressource handeln, die wiederum für eine bestimmte Eigenschaft (Prädikat) einen bestimmten Wert (Objekt) hat usw.

subject = resource



predicat = property



object = value

Hund am Strand

<http://d-nb.info/gnd/131724126>

Weitere Informationen unter <http://www.w3.org/TR/rdf-concepts/>

Linked Data Standards



Einige der Standards für Linked Data sind:

OAI-ORE
OWL
SKOS
FOAF

OAI-ORE

Ziel von OAI-ORE ist es, mittels URIs den Austausch und die Nachnutzung von digitalen Objekten und deren Aggregaten zu ermöglichen und diese eindeutig zu identifizieren. ORE schafft Interoperabilität auf Objektebene und basiert auf Prinzipien der W3C Web-Architektur, d.h. es gibt eine Resource, die durch verschiedene Aggregate repräsentiert wird und durch einen Identifier identifiziert wird.

ORE bildet die Binnenstruktur von einzelnen Aggregaten eines digitalen Objektes und deren Verknüpfungen ab. Verknüpfungen können interne Verknüpfungen logisch zusammenhängender Aggregate eines Objektes sein oder externe Verknüpfungen logisch zusammenhängender Objekte.

(Quelle: http://www.kim-forum.org/Subsites/kim/DE/Materialien/Glossar/glossar_node.html)

OAI- Object Reuse and Exchange (OAI-ORE) ist ein Standard, der Beziehungen zwischen Objekten beschreibt.

Es wird unterschieden zwischen
classes = die Objekte bzw. Entitäten, die im OAI-ORE-Kontext eine Rolle spielen
relationships = die Beziehungen, die zwischen diesen Entitäten oder zwischen einer Entität und einem „ literal value “ bestehen.

Neben eigenen Elementen empfiehlt OAI-ORE die Verwendung von Elementen aus folgenden Namespaces :



Dublin Core Elements

Dublin Core Terms

FOAF Elements

RDF Terms

OAI-ORE Spezifikation

<http://www.openarchives.org/ore/>

OAI-ORE Vocabulary

<http://www.openarchives.org/ore/1.0/vocabulary.html>

Web Ontology Language (OWL)



Die Web Ontology Language (OWL) ist ein RDF-basierter Standard zur Beschreibung von Ontologien.

Es gibt drei verschiedene Versionen:

OWL Full

OWL DL (diverse Einschränkungen, die in OWL Full nicht gelten)

OWL Lite

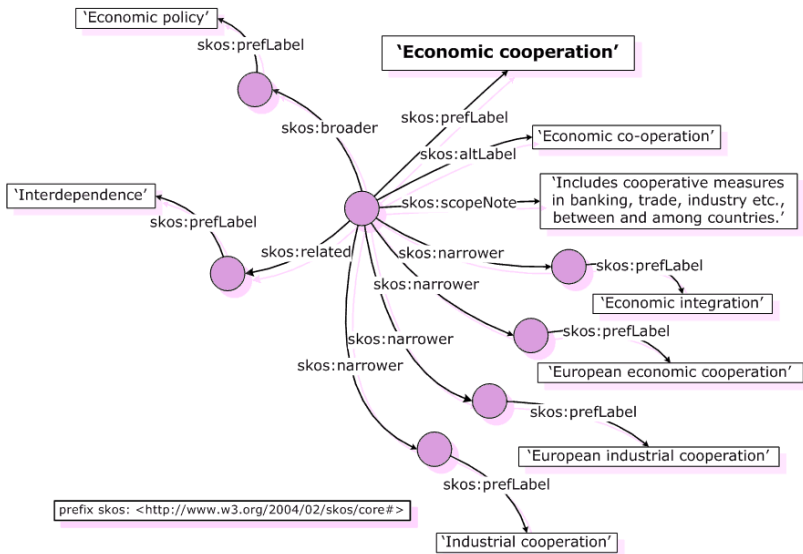
Einen Überblick gibt es unter <http://www.w3.org/TR/owl-features/>.

Simple Knowledge Organisation System (SKOS)



Das Simple Knowledge Organisation System (SKOS) ist ein RDF-basierter Standard zur Beschreibung von Dokumentationsprachen (Klassifikationen, Thesauri u.a. kontrollierte Vokabulare). So bietet SKOS alle notwendigen Ausdrücke um klassische Thesaurusstrukturen abbilden zu können.

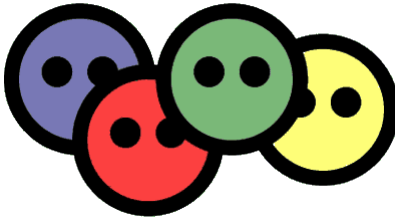
Hier ein Beispiel für eine SKOS-Modellierung eines Thesaurus-Eintrages. Es geht um den Begriff "Economic cooperation". Für diesen Begriff können Vorzugsbenennungen, alternative Benennungen, erklärende Notizen, Ober- und Unterbegriffe und verwandte Begriffe angegeben werden.



(Quelle: <http://www.w3.org/TR/2005/WD-swp-skos-core-guide-20051102/>)

Die SKOS-Spezifikation ist unter <http://www.w3.org/2004/02/skos/specs> zu finden. Hier wird das Simple Knowledge Organisation System detailliert beschrieben.

Friend of a Friend (FOAF)



(Quelle: <http://www.science3point0.com/wp-content/uploads/2010/08/foaf.gif>)

Friend of a Friend (FOAF) ist ein Projekt mit dem Ziel Informationen über Personen miteinander zu verlinken. Das FOAF Vocabulary ist ein RDF-basierter Standard, der der Beschreibung von Beziehungen zwischen Personen und zwischen Personen und anderen Ressourcen bzw. „literal values“ dient.

Hier ein Beispiel für eine ganz einfache Beschreibung einer Person in FOAF. Angegeben sind der Name (`foaf:name`), die persönliche Internet-Mailbox (`foaf:mbox_sha1sum`), die Homepage der Person (`foaf:homepage`) und ein Bild der Person (`foaf:img`).

```
<foaf:Person rdf:about="#me" xmlns:foaf="http://xmlns.com/foaf/0.1/">
  <foaf:name>Dan Brickley</foaf:name>
  <foaf:mbox_sha1sum>241021fb0e6289f92815fc210f9e9137262c252e</foaf:mbox_sha1sum>
  <foaf:homepage rdf:resource="http://danbri.org/" />
  <foaf:img rdf:resource="http://rdfweb.org/people/danbri/mugshot/danbri-small.jpeg" />
</foaf:Person>
```

(Quelle: <http://xmlns.com/foaf/0.1/>)

Weitere Beispiele zu FOAF unter <http://semantisches-web.net/beispiele/beispiele-fur-den-privaten-bereich/foaf/>.

Die FOAF-Spezifikation ist unter <http://xmlns.com/foaf/0.1/> zu finden.

Weitere Informationen zu FOAF auch auf <http://www.foaf-project.org/>.

FOAF-a-matic zum Erstellen eine FOAF-Beschreibung: <http://www.ldodds.com/foaf/foaf-a-matic.de.html>.

Thema 3

Interoperabilität



Das Kapitel behandelt den Punkt Interoperabilität. Es beschäftigt sich außerdem mit dem Singapore Framework und Mapping.

Grundlagen der Interoperabilität

Was ist Interoperabilität?

Was versteht man unter dem Begriff Interoperabilität? Interoperabilität ist die Fähigkeit unterschiedlicher Systeme miteinander zu kommunizieren und so Informationen möglichst effizient und sinnvoll auszutauschen.

"... the ability of two or more systems or components to exchange information and to use the information that has been changed."

(Quelle: <http://en.wikipedia.org/wiki/interoperability>)

Im Metadatenkontext bedeutet dies den systemübergreifenden Austausch von Metadaten unter möglichst geringem Aufwand.

Man unterscheidet zwischen struktureller, syntaktischer und semantischer Interoperabilität:

Strukturelle Interoperabilität

Strukturelle Interoperabilität ist dann gegeben, wenn unterschiedlichen Metadatenformaten ein gemeinsames Datenmodell zugrunde liegt (z. B. das DCMI Abstract Model).

(Quelle: http://www.kim-forum.org/Subsites/kim/DE/Materialien/Glossar/glossar_node.html)

Syntaktische Interoperabilität

Syntaktische Interoperabilität ist dann gegeben, wenn unterschiedliche Metadatenformate in einer gemeinsamen Kodierungssyntax (z. B. XML) wiedergegeben werden. (Quelle: http://www.kim-forum.org/Subsites/kim/DE/Materialien/Glossar/glossar_node.html)

"If a system is capable of communicating and exchanging data, it is syntactic interoperable. For communicating data, specified data formats, communication protocols, interfaces of descriptions and the like are fundamental. In general XML or SQL standards provide syntactic interoperability." (Quelle: <http://en.wikipedia.org/wiki/Interoperability>)



Semantische Interoperabilität

Semantische Interoperabilität ist dann gegeben, wenn unterschiedliche Metadatenformate ein gemeinsames Vokabular verwenden (z. B. Dublin Core Metadata Element Set) (Quelle: http://www.kim-forum.org/Subsites/kim/DE/Materialien/Glossar/glossar_node.html)

Voraussetzung für die Interoperabilität von Metadatenformaten ist

ein gemeinsamer Nenner = Metadatenstandards
standardisierte Dokumentation

(Quelle: http://www.kim-forum.org/Subsites/kim/DE/Materialien/Glossar/glossar_node.html)

Interoperabilität ist der gemeinsame Nenner, auf dessen Grundlage Daten ausgetauscht werden.

Vier Level der Interoperabilität



Dublin Core unterscheidet vier Level der Interoperabilität:

> 4: Description Set Profile Interoperability
• Shared formal vocabularies and constraints in records
> 3: Description Set syntactic interoperability
• Shared formal vocabularies in exchangeable records
> 2: Formal semantic interoperability
• Shared vocabularies based on formal semantics
> 1: Shared term definitions
• Shared vocabularies defined in natural language

"Shared term definition" = die Beschreibung der Ressourcen basiert auf einem gemeinsamen Vokabular, zur Beschreibung derselben classes werden dieselben properties und encoding schemes verwendet => Gemeinsamer Metadatenstandard (z.B. Dublin Core terms)

"Formal semantic interoperability" = Die Beschreibung der Ressourcen basiert auf einem gemeinsamen formalen Modell (z.B. dem RDF Modell)

"Description Set Syntactic Interoperability" = Der Beschreibung der Ressourcen liegt das DCMI Abstract Modell (DCMA) zugrunde.

"Description Set Profile Interoperability" = Für die Beschreibung werden nicht nur dieselben Vokabulare verwendet, sondern auch dieselbe Syntax (z.B. dieselbe XML-Auszeichnung).

(Quelle: <http://dublincore.org/documents/interoperability-levels/>)

Interoperabilität herstellen



Fasst man die vorangegangenen Aussagen zusammen, so scheint es ganz einfach, Interoperabilität herzustellen. Alle verwenden

dasselbe Vokabular

dasselbe Modell

dieselbe Syntax

und schon sind die Metadaten interoperabel.

Eine solches Vorgehen ist aber häufig gar nicht möglich. Neben der Tatsache, dass Vokabulare und Modelle in verschiedenen Communities unabhängig voneinander historisch gewachsen sind, muss beachtet werden, dass unterschiedliche Anforderungen der Community und die Beschreibung verschiedener Arten von Objekten häufig unterschiedliche Modelle und Vokabulare erfordern.

Deutlich wird dies u.a. in den sehr unterschiedlichen Datenmodellen, die den Metadatenstandards in Museen, Archiven und Bibliotheken zugrunde liegen.

Die heterogenen Modelle und Formate, die in diesem Zusammenhang entstehen, lassen sich aber dennoch zusammenführen. Erforderlich ist dafür eine gute Dokumentation der verwendeten Modelle und Terme in Form eines Metadatenprofils Mapping zwischen den Modellen und Metadatermen unterschiedlicher Anwendungen

Mapping



Liegen Datenbestände aus unabhängigen, heterogenen Systemen vor, kann nicht einfach damit weitergearbeitet werden. Die Daten müssen auf ein gemeinsames Datenformat gemappt werden. Erst dann ist eine sinnvolle Weiterverwertung und Anwendung möglich.

Grundlage des Mappings ist der kleinste gemeinsame Nenner der Daten. Denn nur so lassen sich verschiedene detaillierte Strukturen aufeinander abbilden. Das Mapping ist das Konvertieren der Feldinhalte eines Datenformats aus einer Struktur in eine andere. Dies funktioniert jedoch nur, wenn die Metadatenformate interoperabel sind.

Die Schwierigkeiten, die sich beim Mappen auftun, sind vielfältig. Sie reichen von leichter zu behebenden Problemen, bis hin zu Punkten, die Qualitätsverlust der Ursprungsdaten bedeuten. Die Ausgangsdaten können beispielsweise in unterschiedlichen Zeichencodes vorliegen wie Latin1 oder UTF. Bei mehrsprachigen Daten kann die Gefahr bestehen, dass jeweils unterschiedliche Codes für Sprachen oder Länder verwendet wurden. In dem einen Datenbestand wird Deutsch mit "de", im anderen Datenbestand vielleicht mit "ger" ausgezeichnet. Auch Interpunktionsregeln, die sich unterscheiden, können Schwierigkeiten hervorrufen. Liegen zwei Datenbestände in jeweils einer anderen Sprache vor, so kann es nötig sein ganze Textelemente von einer Sprache in eine andere Sprache zu übersetzen.

Mapping und Metadatenprofile



Beispiel Mapping von MAB zu Dublin Core

Diese Tabelle zeigt, wie man MAB-Elemente auf DC-Elemente abbilden kann.

MAB	Definition MAB	DC-Element	DC Definition
331 / 310	Hauptsachtitel in Vorlageform oder Mischform / Hauptsachtitel in Ansetzungsform	DC.Title	A name given to the resource.
335 / 370	Zusätze zum Hauptsachtitel / Weitere Sachtitel	DC.Title.Alternative	Any form of the title used as a substitute or alternative to the formal title of the resource.
304	Einheitssachtitel	DC.Title.Alternative	Any form of the title used as a substitute or alternative to the formal title of the resource.
376	Normierter Zeitschriftentitel	DC.Title.Alternative	Any form of the title used as a substitute or alternative to the formal title of the resource.
670	Sachtitel in abweichender Orthographie	DC.Title.Alternative	Any form of the title used as a substitute or alternative to the formal title of the resource.
104; 108-196	Name der Person(en)	DC.Creator	An entity primarily responsible for making the content of the resource.
204; 208-296	Name der Körperschaft(en)	DC.Creator	An entity primarily responsible for making the content of the resource.

(Quelle: http://www.kim-forum.org/material/pdf/kim_schulunggoettingen_teil3.pdf)

Mapping - verschiedene Arten



Das Prinzip der Dublin Core Community ist, dass sich granulare Strukturen auf allgemeinere Strukturen mappen lassen. Das nennt man "Dump Down".

dcterms.issued => dc.date

dcterms.alternative => dc.title

dcterms.extent => dc.fomat

Die granularen Terme werden im Metadatenprofil als "refinements" der allgemeineren Terme beschreiben. Sie stellen eine Verfeinerung bzw. eine genauere Angabe dar. Die allgemeineren Terme hingegen können mehrere feinere Terme mit dem gleichen Vokabular beschreiben.

Prinzipiell gibt es vier Arten, in die man das Mapping unterscheiden kann.

Vom Besonderen zum Allgemeinen

dcterms.alternative => dc.title

Vom Allgemeinen zum Besonderen

dc.date => dcterms.issued

dcterms.modified oder dcterms.created

Vom Besonderen zum Besonderen

MAB Kategorie 670 => dcterms.alternative

Vom Allgemeinen zum Allgemeinen

dc.title => MAB Kategorie 331 oder 310

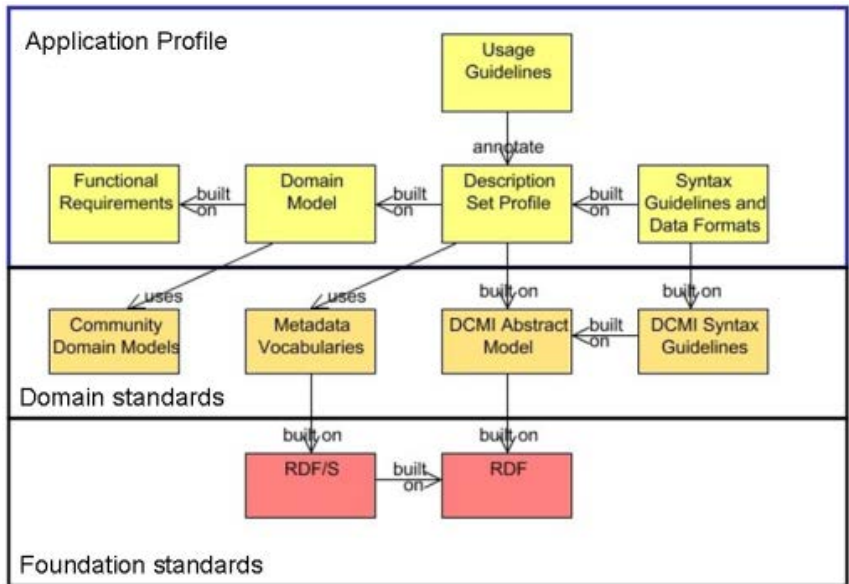
In allen Fällen gilt: Mapping ist aufwendig und führt zu Informationsverlust.

Metadatenprofil



Ein Metadatenprofil ist die formale Beschreibung der Metadaten-terme, die ich verwende, um Objekte in meiner Anwendung zu beschreiben. Dabei müssen die Anforderungen der Nutzer, der Objekte sowie des Systems an sich (hier ist neben der technischen Komponente auch die Frage nach Personalkapazitäten zu berücksichtigen) beschrieben werden. Was bei der Gestaltung eines Metadatenprofils beachtet werden muss, beschreibt das Singapore Framework.

Singapore Framework



Das Singapore Framework ist ein Standard für die Dokumentation von Metadatenprofilen. Es beschreibt, wie ein Metadatenprofil aussehen muss, damit es Dublin-Core-konform ist. Dabei erlaubt es die Verwendung von Metadaten-terminen aus verschiedenen Metadatenstandards sowie die Erstellung eigener Metadaten-terme, insoweit diese mit dem Dublin Core Abstract Model konform sind.

Verpflichtende Bestandteile eines Dublin-Core-konformen Metadatenprofils sind gemäß dem Singapore Framework die *funktionellen Anforderungen* und das *Domain-Modell* sowie das *Beschreibungssatzprofil*. Unter den *funktionellen Anforderungen* wird aufgelistet, welche Funktionalitäten durch die Metadaten erfüllt werden sollen, wobei die Bedürfnisse der Nutzer und die Möglichkeiten des Systems hier besondere Berücksichtigung finden sollten. Das *Domain-Modell* definiert, welche Objekte mit den Metadaten beschrieben werden und was für Beziehungen es zwischen diesen Objekten gibt. Schließlich werden im *Beschreibungssatzprofil* die Metadaterme beschrieben, die notwendig sind, um die funktionellen Anforderungen umsetzen zu können. Neben diesen verpflichtenden Bestandteilen, kann ein Metadatenprofil weitere Informationen enthalten, wie Regeln für die Verwendung der Terme (Regeln für die Erstellung von Metadaten, für die Präsentation der Terme in einem bestimmten Kontext usw.) oder für die syntaktische Kodierung der Terme (z. B. in XML) usw.



Wie man Metadatenprofile im Sinne des Singapore Framework erstellt, finden Sie in dem Tutorial "Einführung in die Gestaltung von Metadatenprofilen".

Homepage: <http://dublincore.org/documents/singapore-framework/>

Thema 4

Zusammenfassung



Die wichtigsten Informationen kompakt zusammengefasst.

Zusammenfassung I

Metadaten sorgen für eine strukturierte und einheitliche Beschreibung von Ressourcen jeder Art und in jeglichem Kontext.

Zusammenfassung II

Man unterscheidet folgende Arten von Metadatermenen:

- Classes
- Properties
- Encoding Schemes

Zusammenfassung III

Man unterscheidet verschiedene Arten von Metadaten, wobei es häufig zu Überschneidungen kommt:

- Administrative Metadaten
- "Descriptive" Metadaten
- Technische Metadaten ...

Zusammenfassung IV

Interoperabilität von Metadaten kann sehr unterschiedlich aussehen.

Zusammenfassung V



Metadatenstandards und -formate basieren auf unterschiedlichen
- teilweise aufeinander aufbauenden - Modellen

RDF und
DCAM

Zusammenfassung VI

Es gibt verschiedene syntaktische Auszeichnungsmöglichkeiten,
um Metadaten zu speichern und auszutauschen.

XML
RDF
...

Zusammenfassung VII

Es gibt unterschiedliche Anwendungen für Metadaten

OPAC
Repository
Suchmaschine
...

Zusammenfassung VIII

Metadatenprofile dokumentieren,

was
wozu
womit
für wen

mit Metadaten beschrieben wird

"menschenslesbar"
"maschinenslesbar"

Zusammenfassung IX



Metadatenprofile erleichtern

- die Nachnutzbarkeit
- den Austausch von Metadaten
- die Verlinkung zwischen Konzepten

Zusammenfassung X

Das Singapore Framework ist ein Standard, der beschreibt, wie ein Metadatenprofil aufgebaut ist. Komponenten eines Metadatenprofils sind:

- Functional requirements
- Domain model
- Description Set Profile (DSP)
- Usage Guidelines
- Encoding syntax guidelines

Zusammenfassung XI

Das DSP fordert für jede Art von Ressourcen, die in einer Anwendung beschrieben werden, ein eigenes Beschreibungstemplate.

Zusammenfassung XII

Dabei gilt: Im DSP können Terme aus verschiedenen Standards zusammengeführt werden, vorausgesetzt sie sind DCAM-kompatibel.

Zusammenfassung XIII

Um unterschiedliche Formate in einem gemeinsamen Format anzubieten braucht es ein Mapping.

Zusammenfassung XIV

Mapping hat fast immer Informationsverlust zur Folge

LEGAL NOTICE

The online version of this publication is available via the following link: DOI: <https://doi.org/10.18452/22242.2>

Publisher: Deutsche Initiative für Netzwerkinformation e. V.

Contact details

DINI – Deutsche Initiative für Netzwerkinformation e. V.

Geschäftsstelle | c /o Niedersächsische Staats- und

Universitätsbibliothek Göttingen | Platz der Göttinger Sieben 1 |

37070 Göttingen |

Tel.: 0551 39-33857 | E-Mail: gs@dini.de | <https://www.dini.de>

Last updated: December 2020

DEUTSCHE INITIATIVE FÜR NETZWERKINFORMATION E. V.

All of the texts used in this publication, excluding citations, are licensed under a Creative Commons “Attribution 4.0 International” (CC BY 4.0) agreement. See <https://creativecommons.org/licenses/by/4.0/>.

