

Metadaten-Schemata im Bibliothekswesen

Stand: 12. Oktober 2023

Dieses Dokument enthält eine alphabetische Liste von an österreichischen Bibliotheken verwendeten Metadaten-Schemata mit Kurzbeschreibungen und weiterführenden Links, die einen ersten Einstieg in die Materie bieten sollen. Ausgangspunkt sind die beiden Umfragen »Umgang mit Metadaten in Repositorien«, die 2016 bzw. 2018 im Rahmen der Projekte »e-Infrastructures Austria« und »e-Infrastructures Austria Plus« an wissenschaftlichen Bibliotheken durchgeführt wurden und unter anderem auch die Frage nach den im jeweiligen Bereich verwendeten Schemata enthielten. In die Liste nicht aufgenommen wurden eigene, nur lokal verwendete Standards, zu denen keine ausreichende Dokumentation vorliegt.

Die Ergebnisse der Umfragen im Detail:

- <https://phaidra.univie.ac.at/o:441216> (2016), Frage 2.3., Seite 9
- <https://phaidra.univie.ac.at/o:1032466> bzw. <https://doi.org/10.25651/1.2019.0018> (2018), Frage 2.4., Seite 21

1. A Resource Center for the Humanities (ARCHE)

ARCHE ist eine Plattform der Österreichischen Akademie der Wissenschaften, deren Zweck darin besteht, »stabiles und dauerhaftes Hosting sowie die Verbreitung digitaler Forschungsdaten und Ressourcen für die Geisteswissenschaften in Österreichisch anzubieten«. Für die Erfassung von Texten wird TEI-XML (s.d.), das Dokumentenformat der Text Encoding Initiative, verwendet.

URL: <https://arche.acdh.oeaw.ac.at>

Technische Beschreibung: <https://arche.acdh.oeaw.ac.at/browser/about-service>

2. Aleph Sequential Format (ASEQ)

ASEQ ist ein auf MAB2 und MARC 21 basierendes internes Datenformat, das vom Bibliothekssystem ALEPH für den Import und Export von Datensätzen verwendet wird. Da ALEPH im österreichischen Bibliothekswesen keine Rolle mehr spielt, dürfte die Beschäftigung damit nur noch von historischem Interesse sein.

Technische Beschreibung (Konkordanz mit MAB2 / MARC 21):

<https://wiki.obvsg.at/Katalogisierungshandbuch/KonKordanz>

3. Component Metadata Infrastructure (CMDI)

CMDI ist ein Software-Entwicklungs-Framework des europäischen Forschungsverbunds CLARIN, der sich mit der Archivierung und Verarbeitung von Sprachdaten in der geistes- und sozialwissenschaftlichen Forschung beschäftigt. CMDI ist kein Metadatenschema, sondern ein Meta-Modell, das den Anwendern die Möglichkeit bietet, eigene Schemata zu entwickeln und untereinander auszutauschen, wobei unterschiedliche Bestandteile («Components») miteinander kombiniert werden können.

URL: <https://www.clarin.eu/content/component-metadata>

Technische Beschreibung:

http://media.dwds.de/clarin/userguide/text/metadata_CMDI.xhtml

<https://www.clarin.eu/content/cmd-specification-version-12>

4. DataCite Metadata Schema

Das DataCite Metadata Schema wurde von der DOI-Registrierungsagentur für Forschungsdaten DataCite entwickelt und dient dazu, die Anwender beim Anlegen von DOIs zu unterstützen. Als Benutzeroberfläche für die Erstellung und Verwaltung der DOIs wird Fabrica angeboten, das die Erstellung von DOIs und DOI-Metadatenätzen per Webformular ermöglicht.

Das DataCite-Schema besteht aus 20 »Properties« genannten Feldern, die z.T. mehrere Unterfelder enthalten. Sechs der Properties sind verpflichtend: *Identifier*, *Creator*, *Title*, *Publisher*, *PublicationYear* und *ResourceType*. Einige der Unterfelder sind ihrerseits wieder verpflichtend, z.B. das zu *ResourceType* gehörige *ResourceTypeGeneral*, dessen Wert aus einem kontrollierten Vokabular auszuwählen ist.

URL: <http://schema.datacite.org/> ; <https://doi.datacite.org/>

Technische Beschreibung: <https://support.datacite.org/docs/mds-api-guide>

5. Dewey-Dezimalklassifikation (DDC)

Die DDC (*Dewey Decimal Classification*) ist die wichtigste Klassifikation für die Sacherschließung in Bibliotheken; sie wurde in über 30 Sprachen übersetzt. Die deutsche Version wird von der Deutschen Nationalbibliothek betreut und steht unter einer Creative-Commons-Lizenz (CC-BY-NC-ND 3.0). Für Webanwendungen wurde WebDewey entwickelt, das – in der deutschen Fassung – gegenwärtig den vollständigen Inhalt der 23. Ausgabe der DDC enthält.

Die DDC ist kein Metadatenformat, allerdings verwendet WebDewey für die Datenmanipulation (Im- und Exporte etc.) ein auf MARC 21 basierendes XML-Format.

URL: <https://www.oclc.org/en/dewey.html> ; https://www.dnb.de/DE/Professionell/DDC-Deutsch/ddc-deutsch_node.html ; <https://deweysearchde.pansoft.de/webdeweysearch/>

6. Dublin Core (DC)

Das Metadatenschema der Dublin Core Metadata Initiative (DCMI) besteht aus einer Liste von 15 Kernelementen zur Beschreibung von Objekten (Dokumenten) im Internet, darunter z.B. *title*, *subject*, *description*, *publisher*. Diese Basisversion ist als »Simple DC« bekannt. Zusätzliche Felder, von denen manche aus kontrolliertes Vokabular enthalten, werden in den »DCMI Metadata Terms« (»Qualified DC«) definiert. Die Darstellung eines Datensatzes kann alternativ im RDF- oder XML-Format erfolgen.

DC ist eines der ältesten und am weitesten verbreiteten Metadaten-Schemata und eng mit der Entwicklung des World Wide Web verbunden.

URL: <https://www.dublincore.org/schemas/>

Technische Beschreibung: <https://www.dublincore.org/specifications/dublin-core/dc-xml-guidelines/2003-04-02/> ; <https://www.dublincore.org/schemas/rdfs/>

7. EPICUR (XEPICUR)

EPICUR (auch: XEPICUR) ist eine auf XML basierendes Transferformat, welches von der Deutschen Nationalbibliothek für den Datenaustausch im Rahmen ihres URN-Services verwendet wird. (Anm.: Die DNB vergibt und verwaltet Uniform Resource Names [persistente Identifikatoren] aus dem Namensraum *urn:nbn:de* und bietet einen Dienst an, der die Auflösung von URNs für Deutschland und die Schweiz ermöglicht; mehr dazu unter https://www.dnb.de/DE/Professionell/Services/URN-Service/urn-service_node.html) Die Verwaltung der URNs der teilnehmenden Einrichtungen erfolgt über eine OAI-Schnittstelle, mittels derer URNs neu registriert und die hinterlegten URLs hinzugefügt bzw. geändert werden. Das dabei zum Einsatz kommende Metadatenformat ist EPICUR.

URL: <https://wiki.dnb.de/display/URNSEVDOK/URN-Service+-+Dokumentation>

Technische Beschreibung: <https://wiki.dnb.de/display/URNSEVDOK/xepicur+-+XML-Datentransferformat+zur+Verwaltung+von+URN>

8. Europeana Data Model (EDM)

EDM ist das ab 2011 schrittweise eingeführte RDF/XML-Lieferformat für den Austausch und die Präsentation von Daten im Europeana-Portal. Es dient der Beschreibung von Objekten aller Art aus dem europäischen Kulturerbe und soll die Bedürfnisse einer breiten Schicht von Anwendern (Museen, Archive, Bibliotheken) aus allen Ländern der EU abdecken. Der Schwerpunkt liegt dabei auf dem Mapping, d.h. der Konversion bestehender Metadaten der Einrichtungen nach EDM; für die EDM-Aufbereitung von Daten nach Standards wie LIDO (für Museen) oder METS (für digitale Bibliotheken) existiert eine umfangreiche Dokumentation.

Strukturell besteht ein EDM-Datensatz aus Datenfeldern, die sich auf drei »Core Classes« für das Objekt selbst (z.B. Titel, physische Beschreibung) und vier »Contextual Classes« für zusätzliche Angaben (z.B. Aufbewahrungsort, Sammlungszuordnung) verteilen.

(Am.: Im Rahmen der oben zitierten Umfragen wurde EDM nicht genannt, nur das Vorgängerformat ESE [s.d.]; aus Aktualitätsgründen wurde es aber in diese Liste aufgenommen.)

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

Technische Beschreibung:

https://pro.europeana.eu/files/Europeana_Professional/Share_your_data/Technical_requirements/EDM_Documentation/EDM_Mapping_Guidelines_v2.4_102017.pdf ;
<https://github.com/europeana/corelib/wiki/EDMObjectTemplatesProviders>

9. Europeana Semantic Elements (ESE)

ESE ist das inzwischen obsolete Vorgängerformat zum Europeana Data Model (EDM; s.d.) und sollte nicht mehr für Metadatenlieferungen an Europeana verwendet werden. In beschränktem Umfang können ESE-Daten nach EDM konvertiert werden; Europeana rät jedoch davon ab und empfiehlt die direkte Verwendung von EDM ohne den (i.d.R. verlustbehafteten) Zwischenschritt über ESE.

URL: <https://pro.europeana.eu/page/ese-documentation>

10. Gemeinsame Normdatei (GND)

Die Gemeinsame Normdatei der Deutschen Nationalbibliothek ist eine institutionenübergreifende Normdatei für Personen (Code p), Körperschaften (k), Kongresse und Veranstaltungen (v), Geographika (g), Sachbegriffe (s) und Werktitel (w), die von sämtlichen deutschsprachigen Bibliotheksverbänden gemeinsam gepflegt wird. Sie ist Bestandteil der Verbundsysteme und wird regelmäßig über OAI-Schnittstellen aktualisiert. Die Daten stehen unter einer Creative-Commons-Lizenz und können über ein Webformular abgefragt oder als XML-Dateien in den Formaten MARC 21 oder RDF heruntergeladen werden.

URL: https://www.dnb.de/DE/Professionell/Standardisierung/GND/gnd_node.html

Technische Beschreibung:

https://www.dnb.de/DE/Professionell/Standardisierung/GND/gnd_node.html#doc58016bodyText5

11. Getty Vocabularies

Zur Zeit stellt das Getty Research Institute Thesauri zu sechs Themenbereichen zur Verfügung:

- Art & Architecture Thesaurus (AAT)
- Categories for the Description of Works of Art (CDWA)
- Cultural Objects Name Authority (CONA)
- Iconography Authority (IA)
- Thesaurus of Geographic Names (TGN)
- Union List of Artist Names (ULAN)

Die Thesauri stehen unter eine Open Data-Lizenz und können kostenfrei verwendet werden. Die allgemeine Nutzung bzw. Einbindung in lokale Anwendungen ist auf mehrere Arten möglich: über Suchoberflächen im WWW, als Linked Open Data per SPARQL-Abfrage und als XML-Downloads.

URL: <https://www.getty.edu/research/tools/vocabularies/index.html>

Technische Beschreibung:

<https://www.getty.edu/research/tools/vocabularies/obtain/index.html>

12. International Standard Archival Description (General) (ISAD(G))

ISAD(G) ist ein Standard des Internationalen Archivrats (International Council on Archives; ICA) zur Erfassung von Archivmaterialien von Körperschaften, Personen und Familien. Die Beschreibung eines Objekts besteht aus insgesamt (z.T. optionalen) 26 Elementen, die in 7 Informationsbereichen zusammengefasst sind:

- Identifikation: Signatur, Titel etc.
- Kontext: Provenienzstelle, Bestandsgeschichte etc.
- Inhalt und innere Ordnung: Form und Inhalt, Neuzugänge, Klassifikation etc.
- Zugangs- und Benutzungsbedingungen: Sprache, Reproduktionsbestimmungen etc.
- Sachverwandte Unterlagen: Aufbewahrungsort, Veröffentlichungen etc.
- Anmerkungen
- Verzeichnungskontrolle: Bearbeiterinformationen, Datum der Verzeichnung etc.

Die Library of Congress (LoC) hat einen mit ISAD(G) kompatiblen XML-Standard namens *Encoded Archival Description* (EAD) herausgegeben.

URL: <https://www.ica.org/en/isadg-general-international-standard-archival-description-second-edition> ; <https://www.loc.gov/ead/>

Technische Beschreibung:

https://www.ica.org/sites/default/files/CBPS_2000_Guidelines_ISAD%28G%29_Second-edition_EN.pdf

13. JavaScript Object Notation (JSON)

JSON ist ein von JavaScript abgeleitetes Datenformat in einfacher, für Menschen lesbarer Textform für den Datenaustausch zwischen webbasierten Systemen (vergleichbar mit XML). JavaScript ist durch die Implementierung spezieller Funktionen besonders für die Verarbeitung von JSON-Daten geeignet, aber auch andere Programmiersprachen bieten einschlägige Frameworks; d.h. JSON ist sprachunabhängig. Die Spezifizierung von JSON erfolgt durch Normen der International Electrotechnical Commission und der International Organization for Standardization (IEC/ISO).

URL: <https://www.json.org/json-de.html>

Technische Beschreibung: <https://www.iso.org/standard/71616.html> (kostenpflichtige ISO-Norm); https://www.ecma-international.org/wp-content/uploads/ECMA-404_2nd_edition_december_2017.pdf

14. LibreCat / Catmandu

LibreCat ist ein Open-Source-Projekt der Universitäten Lund, Gent und Bielefeld mit dem Zweck, Software zur Organisation und Verarbeitung von Metadaten in Bibliotheken zu entwickeln. LibreCat ist also kein Metadatenstandard, sondern ein Softwarekonsortium.

Das wichtigste von LibreCat entwickelte Werkzeug ist *Catmandu*, ein Linux-Kommandozeilen-Tool zur Konversion zahlreicher Datenformate untereinander, darunter MARC, MAB, XML, LIDO, JSON, MODS, RDF und Excel. Zur Dokumentation wird u.a. ein Online-Tutorial für Einsteiger angeboten.

URL: <https://librecat.org/>; <https://librecatproject.wordpress.com/tutorial/>
Technische Beschreibung: <https://github.com/LibreCat/Catmandu/wiki>

15. Lightweight Information Describing Objects (LIDO)

LIDO ist ein XML-Metadatenformat zur Beschreibung von Museums- und Sammlungsobjekten, welches sich teilweise am Datenschema des Getty-Thesaurus »Categories for the Description of Works of Art« (CDWA) orientiert.

Das LIDO-Schema gliedert sich in zwei Teile: einen deskriptiven (Art des Objekts, Gattung, Form, Titel etc.) und einen administrativen (Rechte, Spezifika des Datensatzes, Link zum Datensatz etc). Eine Erweiterung zur LIDO-Spezifikation ist *LIDO Terminologie*, das kontrollierte Vokabulare (Thesauri) für einzelne LIDO-Elemente definiert und sie im Sinn von Linked Open Data über URIs referenzierbar macht.

URL: <https://cidoc.mini.icom.museum/working-groups/lido/lido-overview/>
Technische Beschreibung: <https://cidoc.mini.icom.museum/working-groups/lido/lido-overview/lido-schema/>

16. Learning Object Metadata (LOM)

LOM ist ein 2002 vom Institute of Electrical and Electronics Engineers (IEEE) herausgegebener Metadatenstandard für Lernobjekte (d.h. Textbausteine, Bilder, Grafiken und Animationen) mit dem Zweck, diese in Online-Lernangeboten leichter auffindbar und zwischen verschiedenen Anwendungen austauschbar zu machen. LOM-DE, LOM-CH, LOM-FR etc. sind länderspezifische, auf den jeweiligen Bildungssektor zugeschnittene LOM-Implementationen.

Struktur eines LOM-Datensatzes (neun Kategorien mit Unterelementen):

- General: Identifier, Title, Language etc.
- Life Cycle: Version, Status, Role etc.
- Meta-Metadata: Metadata Schema, Language etc.
- Technical: Format, Size, Location etc.
- Educational: Interactivity Type, Interactivity Level etc.
- Rights: Cost, Copyright etc.
- Relation: Resource etc.
- Annotation: Date, Description etc.

- Classification: Purpose, Taxon, Keyword etc.

URL: <https://standards.ieee.org/ieee/1484.12.1/3294/> (kostenpflichtige IEEE-Norm)

Technische Beschreibung: <http://www.imsglobal.org/metadata/index.html>

17. Maschinelles Austauschformat für Bibliotheken (MAB)

MAB (1970er Jahre) bzw. MAB2 (ab 1995) sind historische Datenformate für den Austausch von bibliographischen sowie Norm- und Lokaldaten in den deutschsprachigen Bibliotheksverbänden. Die Entwicklung wurde im Jahr 2006 eingestellt, als Nachfolger kommt seither MARC 21 zum Einsatz. Als Ergänzung zu MAB2 diente MABxml, womit die Abfrage über OAI-Schnittstellen ermöglicht wurde. Außer in Altsystemen spielt MAB heute keine Rolle mehr.

URL:

https://web.archive.org/web/20181121033620/https://www.dnb.de/DE/Standardisierung/Formate/MAB/mab_node.html

18. Machine Readable Cataloging (MARC)

MARC ist ein von der Library of Congress (LoC) entwickeltes bibliographisches Datenformat, das heute (seit 1999) in der Version MARC 21 vorliegt und weltweit im Einsatz ist. Zusätzlich gibt bzw. gab es zahlreiche weitere, zumeist länderspezifisch angepasste Versionen. Der in Bibliothekskreisen besten bekannte MARC-Standard enthält zahlreiche Felder für Nummern und Codes, Klassifikationen und Signaturen, Haupteintragungen, Titel und damit verbundene Angaben etc. etc., auf die hier aus Platzgründen nicht weiter eingegangen werden soll. Von besonderer Bedeutung für die bibliothekarische Praxis ist die XML-Kodierung MARCXML, weshalb die folgenden Links speziell diese Version berücksichtigen.

URL: <https://www.loc.gov/standards/marcxml/>

Technische Beschreibung: <https://d-nb.info/1222522543/34> ;

<https://www.loc.gov/standards/marcxml/schema/MARC21slim.xsd>

19. Metadata Encoding and Transmission Standard (METS)

METS ist ein von der Library of Congress (LoC) betreuter Metadaten-Standard im XML-Format, der der Beschreibung von digitalen Sammlungen dient. Ein METS-Dokument besteht aus sieben Hauptabschnitten:

- Header: zusammenfassende Beschreibung (Erstelldatum, Autor etc.)
- Erschließungsangaben: z.B. Verweis auf einen externen OPAC-Katalogdatensatz
- Verwaltungsangaben: Herkunft, urheberrechtliche Angaben zum Dokument etc.
- Dateienabschnitt: Liste der zum Dokument gehörigen Files
- Strukturbeschreibung: innerer Aufbau des Dokuments
- Strukturverknüpfungen: Hyperlinks zwischen einzelnen Elementen im Dokument
- Verhalten: Beschreibung von Softwareeigenschaften, die nötig zur Wiedergabe bzw. Darstellung des Dokuments sind

Verweise auf externe Informationsquellen sind in den meisten Abschnitten zulässig. Die Strukturbeschreibung ist der einzige verpflichtende Abschnitt.

URL: <https://www.loc.gov/standards/mets/>

Technische Beschreibung: <https://www.loc.gov/standards/mets/mets-schemadocs.html>

20. Metadata Object Description Schema (MODS)

MODS ist ein von der Library of Congress (LoC) herausgegebener XML-Standard für bibliographische Metadaten, der einen Kompromiss zwischen der Komplexität von MARC und dem für viele Anwendungsfälle zu wenig leistungsfähigen Dublin Core bilden soll. Er enthält 20 Top-Level-Elemente zur Dokumentbeschreibung (*titleInfo, name, language, subject, abstract, accessCondition...*), denen weitere Subelemente zugeordnet sind (z.B. für *name: alternativeName, role, affiliation...*). Software der LoC und von Drittanbietern ermöglicht es, MODS-Datensätze (teilweise) nach MARC oder Dublin Core zu konvertieren oder für BibTeX und EndNote aufzubereiten.

URL: <https://www.loc.gov/standards/mods/>

Technische Beschreibung: <https://www.loc.gov/standards/mods/mods-schemas.html>

21. METS/MODS (auch: MODS/METS)

METS/MODS ist ein vor allem im deutschsprachigen Bibliothekswesen verbreiteter XML-Standard, der Elemente von METS und MODS miteinander kombiniert und auf die Vereinheitlichung der Beschreibung von Digitalisaten aus unterschiedlichen Projekten von Bibliotheken und Archiven abzielt. Er wird von den Praxisregeln der Deutschen Forschungsgemeinschaft unterstützt und dient vorwiegend als (z.T. verpflichtendes) Format für die Lieferung von Dokumenten an die Deutsche Digitale Bibliothek. Strukturell ist der MODS-Teil für die bibliographischen und der METS-Teil für die sonstigen Metadaten zuständig.

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

Technische Beschreibung: <https://wiki.deutsche-digitale-bibliothek.de/pages/viewpage.action?pageId=69124693>

22. Primo Normalized XML (PNX)

PNX ist das interne Metadatenformat der Suchmaschine Primo von Ex Libris. Die Vielzahl der in Primo zusammengeführten Daten aus den unterschiedlichsten Quellen werden durch die Konversion nach PNX standardisiert und für Primo optimiert. Ein PNX-Datensatz gliedert sich dabei in 13 Abschnitte mit jeweils mehreren Unterfeldern, die speziell auf die Anforderungen der Datenmanipulation in Suchmaschinen zugeschnitten sind, also z.B. die Sortierung bzw. das Ranking von Suchergebnissen, das Anlegen von Browse-Listen, die Unterdrückung von doppelt angezeigten Treffern (Deduplikation), die Darstellung von Suchergebnissen in Facetten und anderes.

URL: https://knowledge.exlibrisgroup.com/Primo/Product_Documentation/Primo

Technische Beschreibung:

https://knowledge.exlibrisgroup.com/Primo/Product_Documentation/Primo/Technical_Guide/010The_PNX_Record

23. Preservation Metadata: Implementation Strategies (PREMIS)

PREMIS ist eine Initiative der Library of Congress, die sich mit Metadaten für die digitale Langzeitarchivierung beschäftigt. Das Datenmodell von PREMIS (beschrieben im PREMIS Data Dictionary) basiert auf der Erfassung von fünf grundlegenden Entitäten und deren Beziehungen zueinander:

- Objekte: zur digitalen Archivierung vorgesehene Informationseinheiten
- Umgebungen: zur Nutzung eines Objekts nötige Technologien (Hard- und Software)
- Agenten: Personen oder Einrichtungen, die zu einem Objekt in Beziehung stehen
- Ereignisse: Aktionen, die ein Objekt oder einen zugehörigen Agenten betreffen
- Rechte: Beschreibung von Nutzungsrechten und Lizenzen für ein Objekt

Die Umsetzung des Datenmodells erfolgt üblicherweise (aber nicht zwingend) in Form eines XML-Schemas; die PREMIS-Initiative stellt ein solches Schema, das direkt auf dem Data Dictionary beruht, zur Verfügung. Es ist aber auch möglich, PREMIS innerhalb anderer Schemata zu verwenden, sofern diese die dazu nötige Komplexität aufweisen. Für die Nutzung von PREMIS in METS (s.d.) hat die Initiative eigene Richtlinien herausgegeben.

URL: <http://www.loc.gov/standards/premis/>

Technische Beschreibung: <http://www.loc.gov/standards/premis/v3/premis-3-0-final.pdf> ;
<https://www.loc.gov/standards/premis/premis.xsd>

24. Regeln für die alphabetische Katalogisierung (RAK)

Die RAK, für die österreichischen Universitätsbibliotheken von Bedeutung insbesondere in der Variante RAK-WB (= RAK für wissenschaftlichen Bibliotheken), sind kein Metadatenstandard, sondern ein seit den 1970er Jahren verwendetes Regelwerk, das noch stark am Aufbau von Zettelkatalogen ausgerichtet war. Die RAK werden nicht mehr weiterentwickelt und sind seit 2015 durch das neue Regelwerk »Resource Description and Access« (RDA; s.d.) abgelöst.

URL: <https://d-nb.info/986402338/34>

25. Resource Description and Access (RDA)

RDA ist als Nachfolger von RAK (s.d.) das im deutschen Sprachraum (aber nicht nur dort) dominierende Regelwerk für Bibliotheken, Archive und Museen. Für die Entwicklung zuständig ist ein internationales Konsortium, dem u.a. die Library of Congress, die British Library, die Deutsche Nationalbibliothek und andere angehören. Das Grundkonzept von RDA bildet das sogenannte »Entity-Relationship-Modell«, worunter die Darstellung des Geflechts aus Entitäten (Personen, Körperschaften, Werke, Exemplare etc.), deren Merkmalen (Namen, Titel, Publikationsdaten etc.) und den zwischen ihnen herrschenden Beziehungen

zu verstehen ist. Sämtliche Informationen zum Regelwerk und dessen Anwendungsrichtlinien sowie weitere Materialien finden sich im RDA Toolkit, das allen Einrichtungen im RDA-Konsortium zur Verfügung steht.

URL: <https://wiki.dnb.de/display/RDAINFO/RDA-Info>

Technische Beschreibung: <https://access.rdatoolkit.org/> (Login erforderlich)

26. Resource Description Framework (RDF)

RDF ist ein vom World Wide Web Consortium (W3C) entworfener Standard zur Beschreibung von Metadaten und gilt als eine der Grundlagen des Semantic Web (Web 3.0). Der Zweck von RDF ist die Formalisierung von Aussagen. Jede Aussage in RDF, genannt »semantisches Tripel«, besteht aus Subjekt, Prädikat und Objekt, bildet also im Prinzip ein Analogon zu einem Satz in menschlicher Sprache. Alle Tripel, die sich auf dieselben Subjekte und Objekte beziehen, bilden ein semantisches Netz zu den beschriebenen Gegenständen, das z.B. das Ziehen von logischen Schlüssen auf maschinellem Weg ermöglicht (Beispiel: Ich arbeite an der ZBPH; die ZBPH gehört zur UBW; also arbeite ich an der UBW).

RDF per se ist ein abstraktes System, das auf verschiedene Arten dargestellt werden kann, z.B. durch geeignete XML-Schemata. Häufig verwendet werden die speziell für RDF entwickelten Sprachen Notation 3 oder Turtle.

URL: <https://www.w3.org/RDF/>

Technische Beschreibung: <https://www.w3.org/TR/rdf-schema/>

27. TEI-XML

TEI-XML ist ein besonders im geisteswissenschaftlichen Bereich verbreitetes Format der Text Encoding Initiative (TEI) für die Kodierung gedruckter und handschriftlicher Texte sowie die wissenschaftliche Textanalyse. Neben den üblichen Beschreibungselementen (Autor, Titel, Erscheinungsort etc.) ermöglicht es die Kennzeichnung von Eigenschaften der speziellen physischen Erscheinungsform eines Dokuments:

- Gliederung: Klappentext, Vor- und Nachwort, Kapitel, Glossar etc.
- Semantische Struktur: Überschriften und ihre Hierarchie, Fließtext
- Hervorhebungen durch Kursiv- oder Fettdruck
- Positionierung von Abbildungen und Formeln
- Handschriftliche Einfügungen bzw. Streichungen, Randglossen
- Auszeichnung verschiedener Handschriften in Manuskripten

Zur Abdeckung der Bedürfnisse unterschiedlicher Textsorten kann TEI-XML durch besondere Module (»Customizations«) erweitert werden. So reichert man z.B. durch Einbindung des Moduls *tei_math* TEI um die Mathematical Markup Language (MathML) an, die die Kodierung mathematischer Symbole und damit ganzer Mathematiktexte ermöglicht.

URL: <https://tei-c.org/>

Technische Beschreibung: <https://tei-c.org/release/doc/tei-p5-doc/en/html/index.html>

Weiterführende Informationen

Eine umfangreiche Quelle zu den unterschiedlichsten Klassifikationen, Thesauri und Normdateien, die teilweise über eigene Metadaten-Schemata verfügen, stellt das *Basic Register of Thesauri, Ontologies & Classifications* (BARTOC) dar. BARTOC ist eine Datenbank mit dem Ziel, eine möglichst vollständige Dokumentation von Wissensorganisationssystemen (*Knowledge Organization Systems*, abgekürzt KOS) zu erstellen und diese untereinander vergleichbar zu machen. BARTOC wurde 2013 an der Universität Basel gegründet und wird heute vom Gemeinsamen Bibliotheksverbund (GBV) betrieben.

URL: <https://bartoc.org/>

[CC BY 4 international](https://creativecommons.org/licenses/by/4.0/)
Alexander Zartl

