Marc standards and others

An Introduction to Marc standard

MARC (machine-readable cataloging) is a standard set of digital formats for the machine-readable description of items catalogued by libraries, such as books, DVDs, and digital resources. Computerized library catalogs and library management software need to structure their catalog records as per an industry-wide standard, which is MARC, so that bibliographic information can be shared freely between computers. The structure of bibliographic records almost universally follows the MARC standard. Other standards work in conjunction with MARC, for example, Anglo-American Cataloguing Rules (AACR)/Resource Description and Access (RDA) provide guidelines on formulating bibliographic data into the MARC record structure, while the International Standard Bibliographic Description (ISBD) provides guidelines for displaying MARC records in a standard, human-readable form.

Anglo_american cataloging rules

Anglo-American Cataloguing Rules (AACR) were an international library cataloging standard. First published in 1967 and edited by C. Sumner Spalding,a second edition (AACR2) edited by Michael Gorman and Paul W. Winkler was issued in 1978, with subsequent revisions (AACR2R) appearing in 1988 and 1998; all updates ceased in 2005.

Published jointly by the American Library Association, the Canadian Library Association, and the UK Chartered Institute of Library and Information Professionals, the rules were designed for the construction of library catalogs and similar bibliographic tools. The rules cover the physical description of library resources, as well as the provision of name and title access points.

While the 2002 updates included substantial improvements to AACR's treatment of non-book materials, the proliferation of 21st century formats in a networked environment and the rise of electronic publishing signaled the necessity for significant change in the cataloging code. Plans for a third edition (AACR3) were abandoned in 2005.[2]

The international cataloging community turned its attention to drafting a completely new standard to succeed AACR. Informed by the work of the International Federation of Library Associations and Institutions (IFLA) Functional Requirements for Bibliographic Records (FRBR), the new framework was crafted to be more flexible and suitable for use in a digital environment: Resource Description and Access (RDA) was released in June 2010. The Library of Congress, National Library of Medicine, National Agricultural Library, and several national libraries of other English-speaking countries performed a formal test of RDA, resulting in a June 2011 report of findings.

RDA(Resource Description and Access)

Resource Description and Access (RDA) is a standard for descriptive cataloging initially released in June 2010,[1] providing instructions and guidelines on formulating bibliographic data. Intended for use by libraries and other cultural organizations such as museums and archives, RDA is the successor to Anglo-

American Cataloguing Rules, Second Edition (AACR2). RDA emerged from the International Conference on the Principles & Future Development of AACR held in Toronto in 1997. It is published jointly by the American Library Association, the Canadian Federation of Library Associations, and the Chartered Institute of Library and Information Professionals (CILIP) in the United Kingdom. Maintenance of RDA is the responsibility of the RDA Steering Committee (RSC). As of 2015, RSC is undergoing a transition to an international governance structure, expected to be in place in 2019.

RDA instructions and guidelines are available through RDA Toolkit, an online subscription service, and in a print format.

https://www.loc.gov/catworkshop/RDA%20training%20materials/DCatRDA/3toolkitnonlc0514.pdf

RDA training materials and texts are available online and in print.

RDA is a package of data elements, guidelines, and instructions for creating library and cultural heritage resource metadata that are well-formed according to international models for user-focused linked data applications. The underlying conceptual models for RDA are the Functional Requirements for Bibliographic Records (FRBR), Functional Requirements for Authority Data (FRAD), and Functional Requirements for Subject Authority Data (FRSAD) maintained by IFLA, and will be compliant with the Library Reference Model, the IFLA standard that consolidates them. RDA is in step with the Statement of International Cataloguing Principles published by IFLA in 2009, and updated in 2016.

https://www.loc.gov/catworkshop/RDA%20training%20materials/DCatRDA/1introconctools.pptx

https://web.archive.org/web/20131016202353/http://www.rda-jsc.org/docs/5rda-elementanalysisrev3.pdf

Functional Requirements for Bibliographic Records (FRBR)

Functional Requirements for Bibliographic Records (FRBR /ˈfɜːrbər/) is a conceptual entity–relationship model developed by the International Federation of Library Associations and Institutions (IFLA) that relates user tasks of retrieval and access in online library catalogues and bibliographic databases from a user's perspective. It represents a more holistic approach to retrieval and access as the relationships between the entities provide links to navigate through the hierarchy of relationships. The model is significant because it is separate from specific cataloguing standards such as Anglo-American Cataloguing Rules (AACR), Resource Description and Access (RDA) and International Standard Bibliographic Description (ISBD).

The ways that people can use FRBR data have been defined as follows: to find entities in a search, to identify an entity as being the correct one, to select an entity that suits the user's needs, or to obtain an entity (physical access or licensing).

FRBR comprises groups of entities:

Group 1 entities are work, expression, manifestation, and item (WEMI). They represent the products of intellectual or artistic endeavor.

Group 2 entities are person, family and corporate body, responsible for the custodianship of Group 1's intellectual or artistic endeavor.

Group 3 entities are subjects of Group 1 or Group 2's intellectual endeavor, and include concepts, objects, events, and places.

Group 1 entities

Group 1 entities are the foundation of the FRBR model:

Work is a "distinct intellectual or artistic creation." For example, Beethoven's Ninth Symphony apart from all ways of expressing it is a work. When we say, "Beethoven's Ninth is magnificent!" we generally are referring to the work.

Expression is "the specific intellectual or artistic form that a work takes each time it is 'realized.'"An expression of Beethoven's Ninth might be each draft of the musical score he writes down (not the paper itself, but the music thereby expressed).

Manifestation is "the physical embodiment of an expression of a work. As an entity, manifestation represents all the physical objects that bear the same characteristics, in respect to both intellectual content and physical form." The performance the London Philharmonic made of the Ninth in 1996 is a manifestation. It was a physical embodiment even if not recorded, though of course manifestations are most frequently of interest when they are expressed in a persistent form such as a recording or printing. When we say, "The recording of the London Philharmonic's 1996 performance captured the essence of the Ninth," we are generally referring to a manifestation.

Item is "a single exemplar of a manifestation. The entity defined as item is a concrete entity." Each copy of the 1996 pressings of that 1996 recording is an item. When we say, "Both copies of the London Philharmonic's 1996 performance of the Ninth are checked out of my local library," we are generally referring to items.

Group 1 entities are not strictly hierarchical, because entities do not always inherit properties from other entities. Despite initial positive assessments of FRBR clarifying the thoughts around the conceptual underpinnings of works, there has been later disagreement about what the Group 1 entities actually mean.[5] The distinction between Works and Expressions is also unclear in many cases.

In addition to the relationships between Group 1 and Groups 2 and 3 discussed above, there are many additional relationships covering such things as digitized editions of a work to the original text, and derivative works such as adaptations and parodies, or new texts which are critical evaluations of a pre-existing text. FRBR is built upon relationships between and among entities. "Relationships serve as the vehicle for depicting the link between one entity and another, and thus as the means of assisting the user to 'navigate' the universe that is represented in a bibliography, catalogue, or bibliographic database." Examples of relationship types include, but are not limited to:

Equivalence relationships

Equivalence relationships exist between exact copies of the same manifestation of a work or between an original item and reproductions of it, so long as the intellectual content and authorship are preserved. Examples include reproductions such as copies, issues, facsimiles and reprints, photocopies, and microfilms.

Derivative relationships

Derivative relationships exist between a bibliographic work and a modification based on the work. Examples include:

Editions, versions, translations, summaries, abstracts, and digests

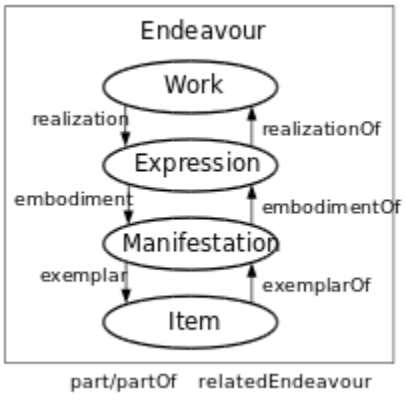
Adaptations that become new works but are based on old works

Genre changes

New works based on the style or thematic content of the work

Descriptive relationships

Descriptive relationships exist between a bibliographic entity and a description, criticism, evaluation, or review of that entity, such as between a work and a book review describing it. Descriptive relationships also includes annotated editions, casebooks, commentaries, and critiques of an existing work.



https://www.loc.gov/catworkshop/RDA%20training%20materials/LC%20RDA%20Training/FRBR_Module%203_FRBR%20&%20RDA%20&%20MARC/FRBR%20%20RDA%20%20MARC_studentversion_20120818.pdf

https://www.ifla.org/files/assets/cataloguing/frbr/frbr_2008.pdf

BIBFRAME

BIBFRAME (Bibliographic Framework) is a data model for bibliographic description. BIBFRAME was designed to replace the MARC standards, and to use linked data principles to make bibliographic data more useful both within and outside the library community.

The MARC Standards, which BIBFRAME seeks to replace, were developed by Henriette Avram at the U.S. Library of Congress during the 1960s. By 1971, MARC formats had become the national standard for dissemination of bibliographic data in the United States, and the international standard by 1973.

In a provocatively titled 2002 article, library technologist Roy Tennant argued that "MARC Must Die", noting that the standard was old; used only within the library community; and designed to be a display, rather than a storage or retrieval format. A 2008 report from the Library of Congress wrote that MARC is "based on forty-year old techniques for data management and is out of step with programming styles of today."

In 2012, the Library of Congress announced that it had contracted with Zepheira, a data management company, to develop a linked data alternative to MARC. Later that year, the library announced a new model called MARC Resources (MARCR). That November, the library released a more complete draft of the model, renamed BIBFRAME.

The Library of Congress released version 2.0 of BIBFRAME in 2016.

BIBFRAME is expressed in RDF and based on three categories of abstraction (work, instance, item), with three additional classes (agent, subject, event) that relate to the core categories. While the work entity in BIBFRAME may be "considered as the union of the disjoint work and expression entities" in IFLA's Functional Requirements for Bibliographic Records (FRBR) entity relationship model, BIBFRAME's instance entity is analogous to the FRBR manifestation entity. This represents an apparent break with FRBR and the FRBR-based Resource Description and Access (RDA) cataloging code. However, the original BIBFRAME model argues that the new model "can reflect the FRBR relationships in terms of a graph rather than as hierarchical relationships, after applying a reductionist technique." Since both FRBR and BIBFRAME have been expressed in RDF, interoperability between the two models is technically possible.

Schema.org

Schema.org is a reference website that publishes documentation and guidelines for using structured data mark-up on web-pages (called microdata). Its main objective is to standardize HTML tags to be used by webmasters for creating rich results (displayed as visual data or infographic tables on search engine results) about a certain topic of interest. It is a part of the semantic web project, which aims to make document mark-up codes more readable and meaningful to both humans and machines.

Microdata

The following is an example of how to mark up information about a movie and its director using the Schema.org schemas and microdata. In order to mark up the data, the attribute itemtype along with the

URL of the schema is used. The attribute itemscope defines the scope of the itemtype. The kind of the current item can be defined by using the attribute itemprop.

```
<div itemscope itemtype="http://schema.org/Movie">
  <h1 itemprop="name">Avatar</h1>
  <div itemprop="director" itemscope itemtype="http://schema.org/Person">
  Director: <span itemprop="name">James Cameron</span>
(born <time itemprop="birthDate" datetime="1954-08-16">August 16, 1954</time>)
  </div>
  <span itemprop="genre">Science fiction</span>
  <a href="../movies/avatar-theatrical-trailer.html" itemprop="trailer">Trailer</a>
</div>
```

RDFa 1.1 Lite

```
<div vocab="http://schema.org/" typeof="Movie">
  <h1 property="name">Avatar</h1>
  <div property="director" typeof="Person">
  Director: <span property="name">James Cameron</span>
(born <time property="birthDate" datetime="1954-08-16">August 16, 1954</time>)
  </div>
  <span property="genre">Science fiction</span>
  <a href="../movies/avatar-theatrical-trailer.html" property="trailer">Trailer</a>
</div>
```

JSON-LD

```
<script type="application/ld+json">
{
  "@context": "http://schema.org/",
  "@type": "Movie",
```

```
 "name": "Avatar",
 "director":
  {
    "@type": "Person",
    "name": "James Cameron",
    "birthDate": "1954-08-16"
  },
 "genre": "Science fiction",
 "trailer": "../movies/avatar-theatrical-trailer.html"
}
</script>
```

Schema Types

There are a number of items that a web page can be marked up with using a Schema, with examples including:

- Article
- Breadcrumb
- Course
- Event
- FAQ
- LocalBusiness
- Logo
- Movie
- Product
- Recipe
- Review
- Video

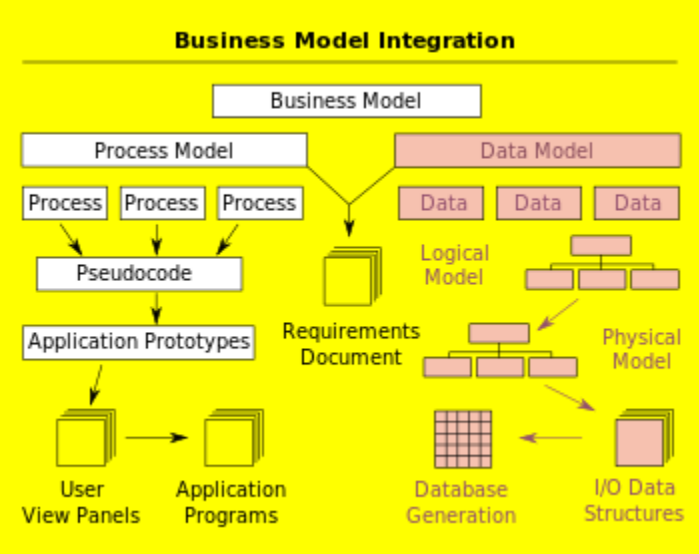
Data model

A data model is an abstract model that organizes elements of data and standardizes how they relate to one another and to the properties of real-world entities. For instance, a data model may specify that the data element representing a car be composed of a number of other elements which, in turn, represent the color and size of the car and define its owner.

The corresponding professional activity is called generally data modeling or, more specifically, database design. Data models are typically specified by a data expert, data specialist, data scientist, data librarian, or a data scholar. A data modeling language and notation are often represented in graphical form as diagrams.

A data model can sometimes be referred to as a data structure, especially in the context of programming languages. Data models are often complemented by function models, especially in the context of enterprise models.

A data model explicitly determines the structure of data; conversely, structured data is data organized according to an explicit data model or data structure. Structured data is in contrast to unstructured data and semi-structured data.



The term data model can refer to two distinct but closely related concepts. Sometimes it refers to an abstract formalization of the objects and relationships found in a particular application domain: for example the customers, products, and orders found in a manufacturing organization. At other times it refers to the set of concepts used in defining such formalizations: for example concepts such as entities, attributes, relations, or tables. So the "data model" of a banking application may be defined using the entity–relationship "data model". This article uses the term in both senses.

Managing large quantities of structured and unstructured data is a primary function of information systems. Data models describe the structure, manipulation, and integrity aspects of the data stored in data management systems such as relational databases. They may also describe data with a looser structure, such as word processing documents, email messages, pictures, digital audio, and video: XDM, for example, provides a data model for XML documents.

Functional Requirements for Authority Data

Functional Requirements for Authority Data (FRAD), formerly known as Functional Requirements for Authority Records (FRAR), is a conceptual entity-relationship model developed by the International Federation of Library Associations and Institutions (IFLA) for relating the data that are recorded in library authority records to the needs of the users of those records and facilitate and sharing of that data.

The draft was presented in 2004 at the 70th IFLA General Conference and Council in Buenos Aires by Glenn Patton. It is an extension and expansion to the FRBR model, adding numerous entities and attributes.

The conceptual work and future implementations are aimed at supporting four tasks, frequently executed by users in a library context—either the library patrons (the first three tasks), or the librarians themselves (all four tasks):

Find: Find an entity or set of entities corresponding to stated criteria;

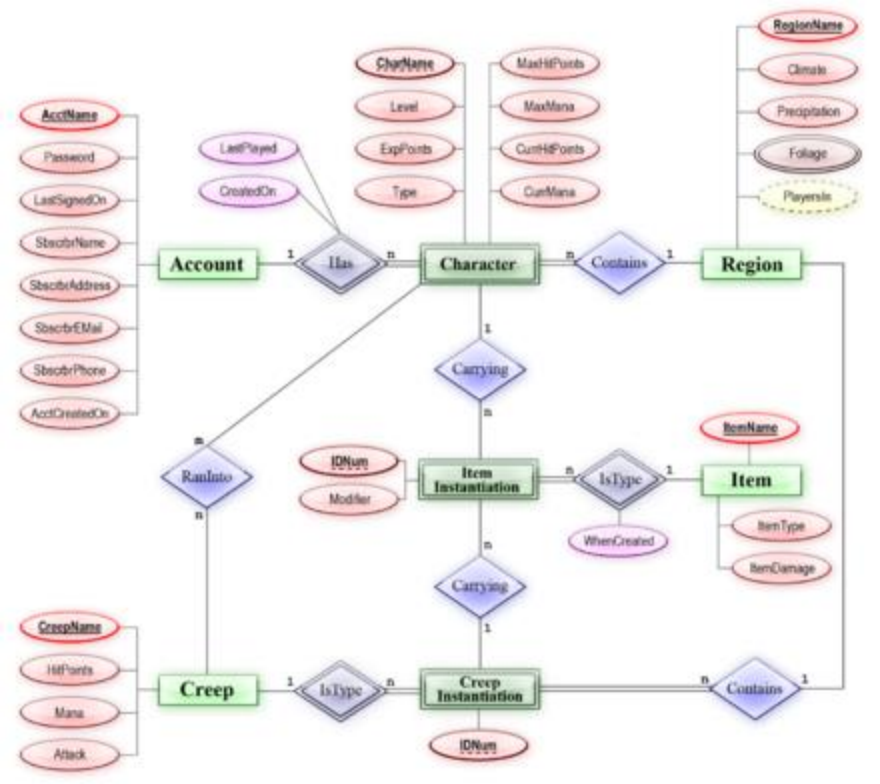
Identify: Identify an entity;

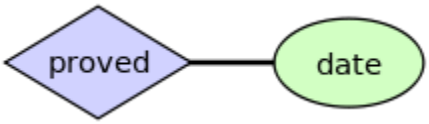
Contextualize: Place a person, corporate body, work, etc. in context;

Justify: Document the authority record creator's reason for choosing the name or form of name on which an access point is based.

Next to the development of FRAR, the Working Group on Functional Requirements and Numbering of Authority Records (FRANAR) is also charged to study the feasibility of an International Standard Authority Data Number (ISADN) and serve as the official IFLA liaison to and work with other interested groups concerning authority files.

An entity–relationship model (or ER model) describes interrelated things of interest in a specific domain of knowledge. A basic ER model is composed of entity types (which classify the things of interest) and specifies relationships that can exist between entities (instances of those entity types).



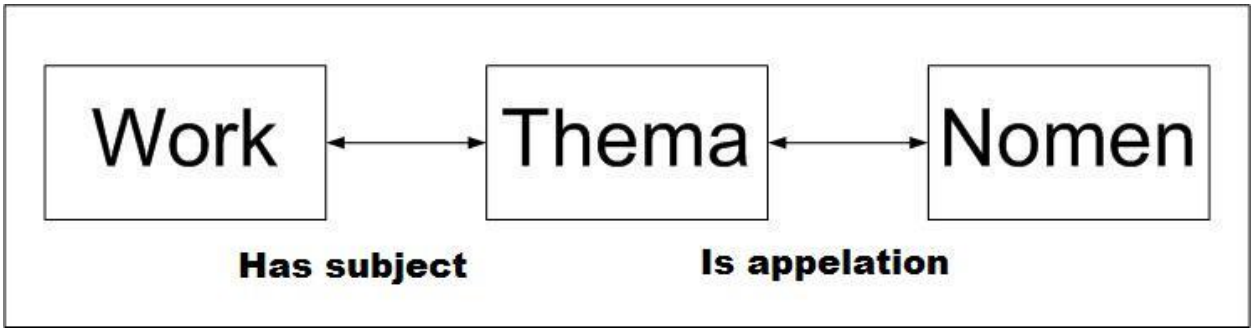


English grammar structure	ER structure
Common noun	Entity type
Proper noun	Entity
Transitive verb	Relationship type
Intransitive verb	Attribute type
Adjective	Attribute for entity
Adverb	Attribute for relationship

FRSAD

Functional Requirements for Subject Authority Data (**FRSAD**), previously known as **Functional Requirements for Subject Authority Records** (**FRSAR**), is a conceptual entity-relationship model developed by the International Federation of Library Associations and Institutions (IFLA) and published in 2010. It is a continuation of the work done on the FRBR model, detailing how "entities that serve as subjects of intellectual or artistic endeavor" can be related and controlled within the bibliographic universe. The model is intended to support global sharing and reuse of subject authority data.

The conceptual model



Work

Work is a "distinct intellectual or artistic creation (IFLA 1998).

Thema

Is anything that can be the subject of a work. This is the abstract idea of the aboutness of a given work. Thema is independent of language and disciplines (FRSAR 2007).

Nomen

Any alphanumeric, sound, visual, or any other symbol, sign or combination of symbols by which a thema is known, referred to or addressed (FRSAR 2007). A nomen can be any expression of a thema. Ideally there will exist an authority file with every possible subject/thema. This means that it should be possible to exchange subject authority data between systems. If a user looks up a specific subject in a catalog and wants to look in other places, he or she should not have to worry about translating the query, since the system would be able to recognize the underlying thema and automatically translate it into the relevant nomen. One way to understand this is to think about how a subject can be described in different ways. For example, if one looks at a work about the city Stockholm, the capital of Sweden, there are many ways to describe Stockholm. First, one must distinguish between the city Stockholm and the many other meanings of the word; for example, Stockholm is also the name of other cities, a record label, and a syndrome. When the thema is established in an authority file, it is possible to translate across systems. The nomen for Stockholm could be anything from "Stockholm", "Stockholm (City)", "Tukholma" - the Finnish spelling of Stockholm - or a range of Zip-codes, or the longitude and latitude, or a picture of the city, or a sound. FRSAR, if implemented, enables users to perform specific and precise subject searching across multiple systems.

Relationships

Works and themas have a many-to-many relationship, meaning that any work can have more than one subject, and any subject can be expressed in one or more works. The same is true for the relationship between thema and nomen. A thema can be expressed in many different ways and a nomen can express many different themas, all depending on the given system. Besides these relationships, the workgroup has so far identified several other thema-thema and nomen-nomen relationships. Two nomens can, for example, be said to have an equivalence relationship, if they both are appellations of the same thema.

User tasks

The workgroup conducted two user studies in 2006 and 2007, and based on the results of these studies, four subject authority data user tasks were defined:

Find: to find an entity (thema or nomen) or set of entities corresponding to stated criteria

Identify: to identify an entity (thema or nomen) based on certain attributes / characteristics

Select: to select an entity (thema or nomen)

Explore: to explore any relationships between entities (thema or nomen), correlations to other subject vocabularies and structure of a subject domain

International Cataloguing Principles

The (Statement of) International Cataloguing Principles (ICP) define(s) the foundation for the creation of bibliographical cataloging rules for libraries. The ICPs are an initiative of the International Federation of Library Associations and Institutions (IFLA) to modernize and replace the old Paris Principles (PP).

The ICPs were drawn up at conferences, the IFLA Meetings of Experts on an International Cataloguing Code (IME-ICC), on four different continents: in Frankfurt am Main (2003), Buenos Aires (2004), Cairo (2005), Seoul (2006) and Pretoria (2007). First published in 2009,[1][2] they were revised again in 2014 and 2015. A new ICP edition was published in 2016.[3][4]

The ICPs are intended as a global guideline for the development of cataloging regulations. They aim to achieve uniformity in both formal and subject indexing and have been created for all types of media (not just books).

The ICPs build on Functional Requirements for Bibliographic Records (FRBRs) and other catalog traditions and, together with the FRBRs, form the basis for the Resource Description and Access (RDA) standard.

Dublin Core

The Dublin Core, also known as the Dublin Core Metadata Element Set (DCMES), is a set of fifteen main metadata items for describing digital or physical resources. The Dublin Core Metadata Initiative (DCMI) is responsible for formulating the Dublin Core; DCMI is a project of the Association for Information Science and Technology (ASIS&T), a non-profit organization.

Dublin Core has been formally standardized internationally as ISO 15836 by the International Organization for Standardization (ISO) and as IETF RFC 5013 by the Internet Engineering Task Force (IETF), as well as in the U.S. as ANSI/NISO Z39.85 by the National Information Standards Organization (NISO).

The core properties are part of a larger set of DCMI Metadata Terms. "Dublin Core" is also used as an adjective for Dublin Core metadata, a style of metadata that draws on multiple Resource Description Framework (RDF) vocabularies, packaged and constrained in Dublin Core application profiles.

The resources described using the Dublin Core may be digital resources (video, images, web pages, etc.) as well as physical resources such as books or works of art. Dublin Core metadata may be used for multiple purposes, from simple resource description to combining metadata vocabularies of different metadata standards, to providing interoperability for metadata vocabularies in the linked data cloud and Semantic Web implementations.

Dublin Core Metadata Element Set

The original DCMES Version 1.1 consists of 15 metadata elements, defined this way in the original specification:

1. Contributor – "An entity responsible for making contributions to the resource".
2. Coverage – "The spatial or temporal topic of the resource, the spatial applicability of the resource, or the jurisdiction under which the resource is relevant".
3. Creator – "An entity primarily responsible for making the resource".
4. Date – "A point or period of time associated with an event in the lifecycle of the resource".
5. Description – "An account of the resource".
6. Format – "The file format, physical medium, or dimensions of the resource".
7. Identifier – "An unambiguous reference to the resource within a given context".
8. Language – "A language of the resource".
9. Publisher – "An entity responsible for making the resource available".
10. Relation – "A related resource".

11. Rights – "Information about rights held in and over the resource".
12. Source – "A related resource from which the described resource is derived".
13. Subject – "The topic of the resource".
14. Title – "A name given to the resource".
15. Type – "The nature or genre of the resource".

Each Dublin Core element is optional and may be repeated. The DCMI has established standard ways to refine elements and encourage the use of encoding and vocabulary schemes. There is no prescribed order in Dublin Core for presenting or using the elements. The Dublin Core became a NISO standards, Z39.85, and IETF RFC 5013 in 2007, ISO 15836 standard in 2009 and is used as a base-level data element set for the description of learning resources in the ISO/IEC 19788-2 Metadata for learning resources (MLR) – Part 2: Dublin Core elements, prepared by the ISO/IEC JTC 1/SC 36.

Full information on element definitions and term relationships can be found in the Dublin Core Metadata Registry.

Encoding examples

<**meta** name="DC.Format" content="video/mpeg; 10 minutes" />

<**meta** name="DC.Language" content="en" />

<**meta** name="DC.Publisher" content="publisher-name" />

<**meta** name="DC.Title" content="HYP" />

IFLA Library Reference Model

The IFLA Library Reference Model (IFLA LRM) is a conceptual entity–relationship model developed by the International Federation of Library Associations and Institutions (IFLA) that expresses the "logical structure of bibliographic information". It unifies the models of Functional Requirements for Bibliographic Records (FRBR), Functional Requirements for Authority Data (FRAD) and Functional Requirements for Subject Authority Data (FRSAD). The IFLA LRM is intended to be used as the basis of cataloguing rules and implementing bibliographic information systems. It has Library of Congress subject heading number 2017004509.

Differences from FR-series models

IFLA LRM adds super-classes res ("thing") and agent to facilitate formal relationship definitions. Time span and place are entities rather than literal values.

It uses the same Work, Expression, Manifestation, Item (WEMI) model as FRBR Group 1 entities. The FRBR Group 2 corporate body and FRAD family are combined into a single collective agent type. FRBR Group 3 entities are deprecated. A redefined nomen type encompasses name from FRAD plus nomen, identifier, and controlled access point from FRSAD.

http://www.rda-rsc.org/sites/all/files/IFLA%20LRM%20what%20and%20why.pdf

https://www.librarianshipstudies.com/2020/04/ifla-library-reference-model-lrm.html

https://www.ifla.org/files/assets/cataloguing/frbr-lrm/transitionmapping_overview_20161207.pdf

history of marc standard

Working with the Library of Congress, American computer scientist Henriette Avram developed MARC between 1965 and 1968, making it possible to create records that could be read by computers and shared between libraries. By 1971, MARC formats had become the US national standard for dissemination of bibliographic data. Two years later, they became the international standard. There are several versions of MARC in use around the world, the most predominant being MARC 21, created in 1999 as a result of the harmonization of U.S. and Canadian MARC formats, and UNIMARC. UNIMARC is maintained by the Permanent UNIMARC Committee of the International Federation of Library Associations and Institutions (IFLA), and is widely used in Europe.

The MARC 21 family of standards now includes formats for authority records, holdings records, classification schedules, and community information, in addition to the format for bibliographic records.

Mods(Metadata Object Description Schema)

The Metadata Object Description Schema (MODS) is an XML-based bibliographic description schema developed by the United States Library of Congress' Network Development and Standards Office. MODS was designed as a compromise between the complexity of the MARC format used by libraries and the extreme simplicity of Dublin Core metadata.

History and development

The Library of Congress' Network Development and MARC Standards Office, with interested experts, developed the Metadata Object Description Schema (MODS) in 2002 for a bibliographic element set that may be used for a variety of purposes, and particularly for library applications. As an XML schema it is

intended to be able to carry selected data from existing MARC 21 records as well as to enable the creation of original resource description records. It includes a subset of MARC fields and uses language-based tags rather than numeric ones, in some cases regrouping elements from the MARC 21 bibliographic format. MODS was first announced for trial use in June 2002. Since September 16, 2022 it is at version 3.8.[1]

The number of users of MODS is unknown. Implementers are encouraged to register their uses of MODS in the implementation registry on the official MODS website. To date there are about 35 projects listed in the registry, although it is assumed that many others are making use of the standard. Users are primarily operating in the area of digital libraries, and some of the registered uses are in digital library projects at the Library of Congress.

Relationship to MARC

The MODS record has been designed to carry key data elements from the MARC record but does not define all of the MARC fields and does not use the field and subfield tagging from the MARC standard. There are data elements in MODS that are not compatible with the MARC record so there is some loss translating from MARC to MODS and from MODS to MARC. There is no commitment on the part of the Library of Congress to maintain compatibility between the two metadata formats beyond what is convenient to the community of MODS users.

The Library of Congress maintains crosswalks in XSLT format for mapping from MARC to MODS, and from MODS to MARC.[2]

Relationship to Dublin Core and qualified Dublin Core

Dublin Core is a simple schema. MODS is far more complex.

The Library of Congress maintains crosswalks in XSLT format for mapping from Dublin Core to MODS, and from MODS to Dublin Core.[3] However, no crosswalks are available for mapping between qualified Dublin Core and MODS.

Advantages

The use of MODS provides several advantages compared to other metadata schemas:[4]

High compatibility with existing resource descriptions

Less detail than MARC so various internal record element sets can be mapped to MODS

Item descriptions from outside in DC and other simpler formats can be mapped and enhanced

Maintenance board

Revisions to the schema are suggested and discussed on the MODS listserv, and approved by the MODS Editorial Committee. The MODS/MADS Editorial Committee is an international group of volunteers responsible for maintaining editorial control over MODS and MADS and their accompanying documentation as well as for the MODS and MADS XML schemas.[5] The Library of Congress carries out the application of approved changes to the schema and maintains the official web site for the standard. There is no formal standards body involved in the MODS schema at this time.

Metadata Authority Description Schema(mads)

Metadata Authority Description Schema (MADS) is an XML schema developed by the United States Library of Congress' Network Development and Standards Office that provides an authority element set to complement the Metadata Object Description Schema (MODS).

History

April 2004: Preliminary version for review

December 2004: Draft for review

April 2005: Version 1.0 published

June 2011: Version 2.0 published

September 2016: Version 2.1 published

MADS does authority control. It is a schema to define people, organizations, and geographical locations which can be involved in creating or publishing a creative work, publication, or artifact. Descriptive schemas for creative works and publications can reference MADS, with the underlying descriptive schema describing the item and referencing a MADS record which describes a creator or location. Authority control allows precise work with issues such as distinguishing multiple authors who share a name, sorting a list of books based on where each was published, and identifying all publications by scholars who are members of a specific organization or graduates of a specific college.

MADS is a schema for carrying authority control information and for describing authorities. It is not a controlled vocabulary nor a registry. It can reference controlled vocabularies or registries, such as Virtual International Authority File (VIAF), Open Researcher and Contributor ID (ORCID), or the Getty Thesaurus of Geographic Names (TGN).

MADS was originally formulated as an XML schema, but it also has an expression in Resource Description Framework (RDF), called MADS/RDF. MADS/RDF expresses and makes statements about Authorities and their Variants, which are controlled records, and distinguishes these from the real-world objects (RWOs) they describe.[4]

The Library of Congress had been representing bibliographic authority data in Simple Knowledge Organization System (SKOS) since 2009. However, they found that they couldn't present the full structure of authorities such as the LC Subject Headings (LCSH) in a general-purpose form such as SKOS. MADS/RDF is intended to complement SKOS: its classes and properties are subclasses of appropriate SKOS items. For example, madsrdf:Authority is a sub-class of skos:Concept, and madsrdf:authoritativeLabel is a sub-property of skos:prefLabel.[4]

MADS/RDF authority items use both authoritativeLabels, which are structured strings like "United States--New Jersey--Essex--Montclair", and collections of typed nodes such as {Country, State, County, City}.

Record structure and field designations

The MARC standards define three aspects of a MARC record: the field designations within each record, the structure of the record, and the actual content of the record itself.

Field designations

Each field in a MARC record provides particular information about the item the record is describing, such as the author, title, publisher, date, language, media type, etc. Since it was first developed at a time when computing power was low, and space precious, MARC uses a simple three-digit numeric code (from 001-999) to identify each field in the record. MARC defines field 100 as the primary author of a work, field 245 as the title and field 260 as the publisher, for example.

Fields above 008 are further divided into subfields using a single letter or number designation. The 260, for example, is further divided into subfield "a" for the place of publication, "b" for the name of the publisher, and "c" for the date of publication.

Record structure

MARC records are typically stored and transmitted as binary files, usually with several MARC records concatenated together into a single file. MARC uses the ISO 2709 standard to define the structure of each record. This includes a marker to indicate where each record begins and ends, as well as a set of characters at the beginning of each record that provide a directory for locating the fields and subfields within the record.

In 2002, the Library of Congress developed the MARCXML schema as an alternative record structure, allowing MARC records to be represented in XML; the fields remain the same, but those fields are

expressed in the record in XML markup. Libraries typically expose their records as MARCXML via a web service, often following the SRU or OAI-PMH standards.

Content

MARC encodes information about a bibliographic item, not information about the content of that item; this means it is a metadata transmission standard, not a content standard. The actual content that a cataloger places in each MARC field is usually governed and defined by standards outside of MARC, except for a handful of fixed fields defined by the MARC standards themselves. Resource Description and Access, for example, defines how the physical characteristics of books and other items should be expressed. The Library of Congress Subject Headings (LCSH) are a list of authorized subject terms used to describe the main subject content of the work. Other cataloging rules and classification schedules can also be used.

MARC formats

MARC formats

Name	Description
Authority records	provide information about individual names, subjects, and uniform titles. An authority record establishes an authorized form of each heading, with references as appropriate from other forms of the heading.
Bibliographic records	describe the intellectual and physical characteristics of bibliographic resources (books, sound recordings, video recordings, and so forth).
Classification records	MARC records containing classification data. For example, the Library of Congress Classification has been encoded using the MARC 21 Classification format.
Community Information records	MARC records describing a service-providing agency, such as a local homeless shelter or tax assistance provider.
Holdings records	provide copy-specific information on a library resource (call number, shelf location, volumes held, and so forth).

MARC 21

MARC 21 was designed to redefine the original MARC record format for the 21st century and to make it more accessible to the international community. MARC 21 has formats for the following five types of data: Bibliographic Format, Authority Format, Holdings Format, Community Format, and Classification

Data Format.[3] Currently MARC 21 has been implemented successfully by The British Library, the European Institutions and the major library institutions in the United States, and Canada.

MARC 21 is a result of the combination of the United States and Canadian MARC formats (USMARC and CAN/MARC). MARC 21 is based on the NISO/ANSI standard Z39.2, which allows users of different software products to communicate with each other and to exchange data.[3]

MARC 21 allows the use of two character sets, either MARC-8 or Unicode encoded as UTF-8. MARC-8 is based on ISO 2022 and allows the use of Hebrew, Cyrillic, Arabic, Greek, and East Asian scripts. MARC 21 in UTF-8 format allows all the languages supported by Unicode.

MARCXML

MARCXML is an XML schema based on the common MARC 21 standards.[5] MARCXML was developed by the Library of Congress and adopted by it and others as a means of facilitating the sharing of, and networked access to, bibliographic information.[5] Being easy to parse by various systems allows it to be used as an aggregation format, as it is in software packages such as MetaLib, though that package merges it into a wider DTD specification.

The MARCXML primary design goals included:[6]

- Simplicity of the schema
- Flexibility and extensibility
- Lossless and reversible conversion from MARC
- Data presentation through XML stylesheets
- MARC records updates and data conversions through XML transformations
- Existence of validation tools

Future

The future of the MARC formats is a matter of some debate among libraries. On the one hand, the storage formats are quite complex and are based on outdated technology. On the other, there is no alternative bibliographic format with an equivalent degree of granularity. The billions of MARC records in tens of thousands of individual libraries (including over 50,000,000 records belonging to the OCLC consortium alone) create inertia. The Library of Congress has launched the Bibliographic Framework Initiative (BIBFRAME),[7] which aims at providing a replacement for MARC that provides greater granularity and easier re-use of the data expressed in multiple catalogs.[8] Beginning in 2013, OCLC Research exposed data detailing how various MARC elements have been used by libraries in the 400 million MARC records (as of early 2018) contained in WorldCat.[9] The MARC formats are managed by the MARC Steering Group, which is advised by the MARC Advisory Committee.[10] Proposals for changes to MARC are submitted to the MARC Advisory Committee and discussed in public at the American Library Association (ALA) Midwinter and ALA Annual meetings.

Marc bibliographic standard

MACHINE READABLE CATALOGING

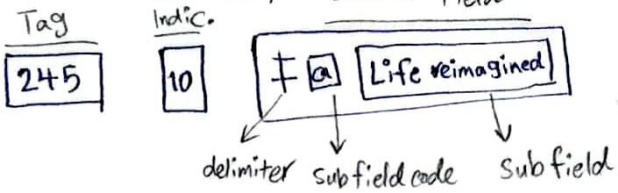
What is Marc?

- Data indexing system
- Allows computers to understand cataloging inputs

Marc anatomy 101

- Fields → containing data
- Tags → Identifing what is in the field
- Indicators → Tell computers how to work with fields
- Sub-field → sections of fields
- Sub-field Codes → identify content of subfield
- Delimiters → Mark start of subfield

A marc anatomy chart field



Important fields for copy catalogers

- 04 (cataloging source)
- 245 (title)
- 250 (edition statement)
- 264 or 260 (publication)
- 300 (physical description)
- 490 (series statement)

040 - Cataloging Source

040 | | | DLC ǂb eng ǂe rda ǂc DLC ǂd IG ǂd OCLCO ǂd YDXCP ǂVET

- ǂa who did intellectual work of original cataloging
- ǂb language of cataloging
- ǂe description conventions
- ǂc who entered the record
- ǂ who has edited the record

245 - title Statement

245 |1|0| Life reimagined : ǂb discovering your new life possibilities / ǂc Richard J.

- ǂa title
- ǂb remainder of title
- ǂc Statement of responsibility

250

250 | | | First edition

- ǂa edition Statement

264 or 260 - Publication date

264 | |1| San Francisco : ǂb berrett-koehler Publishers, Inc., ǂc [2013]

- ǂa Place of Publication
- ǂb Publisher name
- ǂc date of Publication

300-Physical Description

300|| |ix, 173 pages : ǂb illustrations ; ǂc 23 cm

- ǂa extent (#Pages, discs, etc)
- ǂb other physical details
- ǂc dimentions

490-Series Statement

490|1| |NBER working paper series ; ǂv no. 13191

- ǂa series statement